Promotor Prof. dr. Erik Weber
Vakgroep Wijsbegeerte en Moraalwetenschap
Copromotor Dr. Leen De Vreese
Vakgroep Wijsbegeerte en Moraalwetenschap

Decaan Prof. dr. Marc Boone
Rector Prof. dr. Rik Van de Walle
Juryleden Prof. dr. Henk de Regt (Vrije Universiteit Amsterdam)
Prof. dr. Theo A.F. Kuipers (Rijksuniversiteit Groningen)
Prof. dr. Maarten Van Dyck (Universiteit Gent)
dr. Dingmar Van Eck (Universiteit Gent)

Nederlandse vertaling:
Unificatie herdacht: Unificatie als een verklarende waarde in de wetenschapspraktijk en -theorie.

Kaftinformatie: The sculpture *Broken Symmetry* straddles the road to Fermilab and was created in 1978 by Fermilab's first director, Rober Wilson. The concept of symmetry breaking is very important in particle physics; it may explain why there is more matter than anti-matter in the universe. In May 2006 Nobel laureate Francois Englert gave a lecture about broken symmetry and *unification* of particle physics at Ghent University. This was my first date with my husband. (photograph taken by Steve Krave.)

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Rethinking Unification

Unification as an explanatory value in scientific practice

Merel Lefevere

Proefschrift voorgelegd tot het behalen van de graad van Doctor in de Wijsbegeerte
2018

UNIVERSITEIT
GENT

‘The truth.’ Dumbledore sighed. ‘It is a beautiful and terrible thing,
and should therefore be treated with great caution.

- J.K. Rowling

Dankwoord

Dit dankwoord dat het startpunt vormt van dit proefschrift, is eigenlijk het eindpunt van een lange, maar boeiende weg. Een weg met vele zijwegen, kruispunten en bochten, maar ook met vele supporters.

Dat ik deze weg ben opgegaan heb ik te danken aan mijn promotor prof.dr. Erik Weber. Een zomerse barbecue van de studentenkring op het einde van mijn opleiding wijsbegeerte legde de kiem voor dit proefschrift. Erik liet die avond ontvallen dat ik zou solliciteren voor de job van assistent wetenschapsfilosofie en logica. Wie Erik kent, weet dat zijn plannen heel vaak werkelijkheid worden. Wie Erik beter kent, weet ook dat hij daarbij de mensen niet uit het oog verliest. Via deze weg wil ik hem bedanken om me te gidsen in de wondere wereld van wetenschapsfilosofisch onderzoek, maar vooral ook om me tijdens dit onderzoek mezelf te laten blijven.

Ook co-promotor dr. Leen De Vreese verdient een bedanking. Niet alleen las ze dit proefschrift tot in de puntjes na en tilde ze dit proefschrift naar een hoger niveau, ze zorgde ook voor gezellige momenten, samen met Frederik, in ons bureau.

Mijn speciale dank gaat uit naar prof.dr. Theo Kuipers, prof.dr. Henk de Regt, prof.dr. Maarten Van Dyck en dr. Dingmar Van Eck die als leden van mijn doctoraatsbegeleidingscommissie en jury hun tijd en expertise aan dit proefschrift spenderen.

Werken aan een proefschrift is vaak een eenzame onderneming, maar leuke collega's maken veel goed. Bedankt aan alle collega's die deel uitmaken (of uitmaakten) van het Centre for Logic and Philosophy of Science. De vele discussiemomenten, seminars en workshops leidden tot nieuwe ideeën en inzichten. Maar ook de gezellige babbels, de barbecues, de teambuildings, het uur natuur met Raoul en andere gekke momenten hebben ongetwijfeld bijgedragen aan het voltooien van dit proefschrift. Bedankt ook aan Gitte voor de mooie kaft.

Naast iedereen die binnen de academische wereld op een of andere manier een rol heeft gespeeld bij dit onderzoek, zijn er natuurlijk ook heel wat mensen in mijn persoonlijke omgeving die een bedanking verdienen.

Mijn ouders zijn het grootste deel van mijn leven verantwoordelijk geweest voor mijn opvoeding en opleiding. Daarbij hebben ze altijd de balans proberen houden tussen mij de vrijheid geven om mijn interesses te volgen en me te blijven stimuleren om die interesses dan ook ten volle te ontwikkelen. Hun onvoorwaardelijke steun is in elke betekenis een noodzakelijke voorwaarde voor dit doctoraat. Hun hulp, en die van mijn broer Thijs, om de laatste fase van dit onderzoek te combineren met een huishouden, twee levendige zoontjes en een fulltime job was onmisbaar. Bedankt voor alles!

Alle vrienden en familieleden opnoemen die me tijdens de voorbije jaren hebben gesteund is onbegonnen werk. Ik mag me gelukkig prijzen met een diverse en ruime vriendenkring. Ze wisten wanneer ze wel of niets mochten polsen naar de vorderingen van dit proefschrift. Ze zorgden voor de nodige afleiding, die o.a. resulteerde in een uitgebreide collectie amigurumi en een topmusical. Maar bovenal zijn ze altijd blijven geloven in me.

Ik wil ook mijn leerlingen en collega's van het Klein Seminarie Roeselare bedanken. Hun interesse, (heavy) metal muziek en hun motiverende woorden waren meer dan deugddoend in de schrijffase van dit onderzoek.

De laatste vermelding in dit dankwoord gaat uit naar mijn gezin. Bedankt Joris en Korneel om me op tijd en stond vanachter mijn bureau te halen om te spelen en om me in jullie dinoverhalen alle stress te laten vergeten. Bedankt Koen, om me de tijd en ruimte te geven om het laatste jaar volop te kunnen focussen op het schrijven van dit doctoraat. Zonder jouw steun en liefde was er nooit een einde gekomen aan deze boeiende, maar intensieve weg.

Merel Lefevere

Tielt, oktober 2018

List of figures

Table of Contents

General introduction

Thanks to Philip Kitcher's seminal essay 'Explanatory Unification' (1981) the unification account became a very influential way of thinking about scientific explanations in the 1980s. It was the main rival of Wesley Salmon's causal-mechanical account, which received its full presentation in Salmon's book *Scientific Explanation and the Causal Structure of the World* (1984). In Kitcher 1989, the two approaches are compared; as can be expected, Kitcher argues that the unification account is superior.

In the 1990s, unification was still respected, but often integrated in a pluralistic view. For instance, Salmon writes in 1993:

> In this paper I have tried to show that there are at least two intellectual benefits that scientific explanations can confer upon us, namely, (1) a unified world-picture and insight into how various phenomena fit into that overall scheme, and (2) knowledge of how things in the world work, that is, of the mechanisms (often hidden) that produce the phenomena we want to understand. The first of these benefits is associated with the unification view of scientific explanation; Philip Kitcher (1989) is its present principal proponent. The second is associated with the causal/mechanical view of scientific explanation that I have advocated (Salmon, 1984). My current view is that the two accounts are by no means incompatible. (1993, p. 15)

This view is repeated in Salmon 1998 (pp. 89-90). An example of a unification account embedded in a pluralistic view is Weber 1999, which focuses on unification but explicitly acknowledges that there are other types of explanation, e.g., causal ones (see p. 480).

In the new millennium, unification became unfashionable. This was mainly due to the development of two alternative views that became very popular: the counterfactual theory of Jim Woodward (2003) and the mechanistic approach, which started with

Machamer, Darden and Craver 2000. and is further elaborated in, e.g., Bechtel & Abrahamsen 2005 and Craver 2007.

I experienced the unfashionable status of unification in the spring of 2010, when I attended a workshop about understanding and the aims of science at the Lorentz Center in Leiden. It was a very interesting event organised by Henk de Regt and James McAllister, with speakers such as William Bechtel, Michael Strevens, Lindley Darden, Sabina Leonelli, etc.[1] At the end of the workshop Henk de Regt gathered all of us for a wrap up session. One of the questions he asked was whether or not there was still anyone thinking that unification plays a part in scientific explanation. I was at the start of my research project then, and I was not confident enough to raise my hand. But Theo Kuipers did, and he was the only one.

Today the situation would not be much different. Most philosophers of science consider unification as outdated, irrelevant for explanation or infeasible. The overall aim of this dissertation is *to bring unification back into the picture* in the philosophical study of scientific explanation. After writing the dissertation, today I would certainly also raise my hand.

This dissertation starts with a concise overview of what philosophers of science have written about unification and its role in scientific explanation during the last 50 years (Chapter 1). I introduce a few key ideas, theories and thinkers before we set off for the real start. I do this partially in order to provide the reader with some background knowledge. However, at several stages later in this dissertation (especially chapters 4, 7 and 8) I use this overview to clarify how my approach and results *differ* from existing views on unification and its relation to scientific explanation and how I can avoid legitimate worries that have been raised in relation to 'traditional' (as opposed to my 'alternative') views on unification.

In order to bring unification back into the picture, I have followed two strategies, resulting respectively in Parts I and II of this dissertation. In Part I the idea of unification is used to refine and enrich the dominant causal-mechanist and causal-interventionist accounts of scientific explanation. How this is done exactly will be clarified in the Introduction to Part I and in Chapter 4. What is important now is the general underlying idea. In this part of the dissertation I grant, for the sake of argument, that explanations

---

1 Lorentz Center. 2010. *Lorentz Center International Center for workshops in the Sciences*. Available at: http://www.lorentzcenter.nl/lc/web/2010/380/program.php3?wsid=380. [Last accessed 10 May 2018].

are causal. I argue that unification is important from within this causalist perspective. In Part II I pursue an opposite strategy: I bring unification back into the picture by arguing that some legitimate scientific explanations are unificatory without being causal. This second strategy will be further clarified in the Introduction to Part II.

Part III contains two more general reflective chapters. I clarify how my approach differs from the traditional views that are described in Chapter 1 and are still being put forward by contemporary philosophers (Chapter 7). Furthermore, the results of this dissertation fit into an overall pluralistic view on both unification and on scientific explanation (Chapter 8).

In this dissertation I do not try to develop a new model of explanation and compare it to existing models. The aim is to show that there are important types of explanatory practice which cannot be properly analyzed if we neglect unification as a desideratum for explanations. What will be done here is similar in nature to what Carl Craver did in his book *Explaining the Brain* (2007). Craver's main question is "what is required of an adequate explanation in neuroscience (p. vii). He proposes norms by means of which we can distinguish good from bad explanations in neuroscience (p. viii). The mechanistic model which he develops in the book is a tool for achieving this aim. In the same vein, I discuss questions of the form "what is required of an adequate explanation if this explanation is of type X?". What these questions look like will become clear at the beginning of chapters 2, 3, 5 and 6. Existing models of explanation (e.g. the mechanistic model and the counterfactual) will be used as tools for answering such questions. And – this is of course crucial given my aim of bringing unification back into the picture – my answer to these questions includes some kind of unification as a condition of adequacy.

Chapter 1
A brief history of unification

1.1 Introduction

In this chapter the reader will get a brief historical overview of explanatory unification and its main problems. A full comprehensive history of the subject is unnecessary to understand the main arguments in this dissertation, but a few key ideas, theories and thinkers need to be introduced, before we set off for the real start. This will provide the reader with some background knowledge. But the main reason why this overview is included is that in Chapters 4, 7 and 8 I will clarify how my approach to the issue of unification in scientific explanations differs in crucial aspects from what has been done till now on this subject.

The unificationist approach of explanation has mostly been connected with Philip Kitcher, especially (1981) and (1989). An earlier attempt to develop a unificationist account of explanation can be found in the work of Michael Friedman (1974). However, both of them can be considered as belonging to the intellectual heritage of Carl Hempel's covering-law model. They retain the main idea (viz. subsumption under a covering law) and add extra requirements to solve recognised problems for this model. This is why my overview starts with Hempel's model and its problems (section 1.2).

In section 1.3 the accounts of Michael Friedman, Philip Kitcher, Gerhard Schurz and Erik Weber are summarized. These accounts can all be considered as intellectual heirs of Carl Hempel. In section 1.4 the alternative approach to unification of Uskali Mäki is briefly outlined. After this introductory chapter concrete case studies of unification will be presented.

1.2 Carl Hempel

1.2.1 The DN-model

The basic idea that Hempel proclaims is that explaining a phenomenon is the same as deriving it as an instance of a general regularity. In the deductive-nomological model, this derivation is achieved through deduction, in the inductive-statistical model this derivation is acquired through induction. I will focus on the DN (deductive-nomological) model.

The DN model imposes two conditions on genuine scientific explanations. First, the explanans must contain at least one empirical law, hence a nomological model. Second, the explanandum must be derivable from the explanans, hence a deductive model. However, the first condition of the DN model is too general. The law must be relevant for the explanation, the argument should be invalid without it. Moreover, the law by itself is not sufficient, laws only state what will happen if certain specified antecedent conditions are met. Thus, "the question "*Why* does the phenomenon happen?" is construed as meaning "according to what general laws, and by virtue of what antecedent conditions does the phenomenon occur?"" (Hempel & Oppenheim, 1948, p. 136; italics in original).

Hempel and Oppenheim (1948, p. 137 et seq.) specify a basic pattern of scientific explanations, dividing an explanation in two major constituents: the explanandum and the explanans. The explanandum describes the phenomenon for which we want an explanation. The explanans, that accounts for the explanandum, consists of certain sentences C_1, C_2,..., C_k that state the specific antecedent conditions, and of a set of sentences L_1, L_2,..., L_k which represent general laws. Furthermore there are four conditions of adequacy (Hempel & Oppenheim, 1948, p. 137):

Logical conditions of adequacy

R1 The explanandum must be a logical consequence of the explanans.

R2 The explanans must contain general laws, and these must actually be required for the derivation of the explanandum.

R3 The explanans must have empirical content, i.e., it must be capable, at least in principle, of test by experiment or observation.

Empirical condition of adequacy

R4 The sentences constituting the explanans must be true.

Arguments that fulfill conditions R1-R3 are potential explanations, only when condition R4 is fulfilled the explanation becomes an actual explanation. A DN explanation can be summarized in the following scheme:

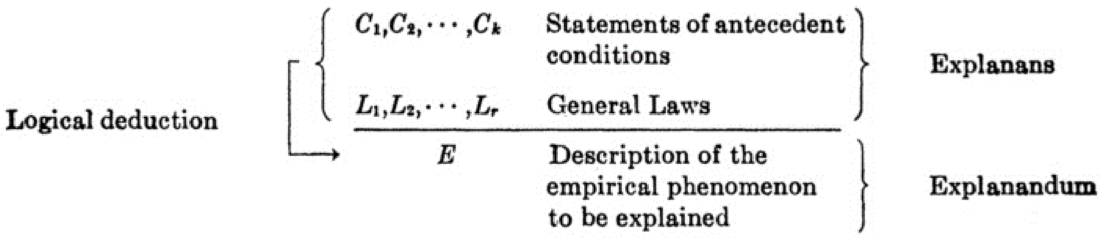

Figure 1 Scheme of DN explanation according to Hempel and Oppenheim (1948, p. 138).

It is clear from this scheme that Hempel considers DN explanations as deductive arguments. A DN explanation shows that the phenomenon is the result of a set of laws and a set of specified antecedent conditions. This argument shows that the phenomenon was to be expected, given the circumstances and the relevant laws.

Hempel and Oppenheim note themselves that the same formal analysis, including the four conditions, can be applied to prediction as well as to explanation. The difference between both is pragmatic:

> If *E* is given, i.e. if we know that the phenomenon described by *E* has occurred, and a suitable set of statements $C_1, C_2, \ldots, C_k, L_1, L_2, \ldots, L_k$ is provided afterwards, we speak of an explanation of the phenomenon in question. If the latter statements are given and *E* is derived prior to the occurrence of the phenomenon it describes, we speak of a prediction. (Hempel & Oppenheim, 1948, p. 138)

The analogy between explaining and predicting immediately illustrates the main intellectual benefit of explanations according to Hempel: understanding the world.

> ...the argument shows that, given the particular circumstances and the law in question, the occurrence of the phenomenon was to be expected; and it is in this sense that the explanation enables us to understand why the phenomenon occurred. (Hempel, 1965, p. 337)

I will end this section with a short example of the DN model constructed by Erik Weber et al. (2013, p. 2 etseq.) about light striking a mirror.

Suppose that we want to explain E:

E This reflected beam of light *a* has an angle of 45° relative to the mirror from which it bounced

According to the DN model, a potential explanans could be:

C The angle of incidence of *a* relative to the mirror was 45°.

L For all beams reflecting on mirrors: if the angle of incidence relative to the mirror is 45°, then the reflected beam also has an angle of 45° relative to the mirror.[2]

If these claims are true (condition R4), this is an explanans.

Now that Hempel's DN model has been briefly summarized, I can discuss some problems of the model and present how Michael Friedman and Philip Kitcher use unification as a solution to those problems.

1.2.2 Problems for the DN model

Hempel's classic book *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* recently celebrated its 50th anniversary[3] and the views in it are still discussed in today's philosophy of science. However, his DN model is not flawless. Three main problems emerged.

The first problem has to do with the difference between genuine laws and accidental generalisations. Let us use Hempel's own examples:

(i) All members of the Greensbury School Board for 1964 are bald.

If this is true, it is only accidentally so. Therefore, this generalisation should not be used to explain why member *n* of the 1964 Greensbury School Board is bald.

---

[2] In volume I of The Feynman Lectures on Physics Chapter 26 deals with elementary optics. There we find a simple law about mirrors: "The simplest object is a mirror, and the law for a mirror is that when the light hits the mirror, it does not continue in a straight line, but bounces off the mirror into a new straight line[.] … The light striking a mirror travels in such a way that the two angles, between each beam and the mirror, are equal. (Feynman, Leighton, & Sands, 2010, pp. 22-26)

[3] In November 2015 the Center for Logic and Philosophy of Science of Ghent University hosted a workshop called '50 shapes of explanation'. http://www.lrr.ugent.be/archive/50-shapes-of-scientific-explanation/ (accessed 25 May 2018).

(ii) All gases expand when heated under constant pressure.

This second example is a law, and thus it can be used to explain why some particular gas sample that has been heated under constant pressure has expanded. While these examples may seem obvious, the criteria to distinguish true accidental generalisations from laws are still up for philosophical discussion. Moreover, Salmon (1989, p. 15) has shown that not only truth but also modality is important for the law-like statement in a DN explanation. Consider the following two statements:

(iii) No gold sphere has a mass greater than 100 000kg.

(iv) No enriched uranium sphere has a mass greater than 100 000kg.

Statement (iii) has no modal force. It is not physically impossible to fabricate a gold sphere of mass greater than 100 000kg. It just happens to be so that no such sphere has ever been made or found in this world. If something were a sphere with mass greater than 100 000kg, we cannot not legitimately conclude from (iii) that it is not a gold sphere. We can however conclude that it is not an enriched uranium sphere, since this would be physically impossible[4]. Hempel's DN model lacks a decent account of laws to cope with such counterexamples.

A second problem has to do with explanatory irrelevancies. Kyburg's (1965) classic example shows that an argument can meet the criteria for a DN explanation and yet be defective because it contains irrelevancies:

All samples of hexed salt dissolve in water.

I have hexed this sample of salt.

This sample of salt dissolves in water.

The hexing is irrelevant to the explanation because salt dissolves in water anyway.

Another classic example has been developed by Salmon (1971, p. 34):

[4] The critical mass of enriched uranium amounts 52 kg, which means that exceeding this mass would result in a uncontrolled chain reaction and as such a nuclear explosion.

All males who take birth control pills regularly fail to get pregnant.

Jon is a male who regularly takes birth control pills.

Jon fails to get pregnant.[5]

As it was the case with the hexed salt, these premises are true and the inference fits a DN argument. However, the first premise is superfluous, males do not get pregnant, whether they take birth control pills or not. These examples show that a derivation can satisfy the DN criteria and yet fail to be explanatory.

A third problem is the asymmetry problem. Hempel and Oppenheim (1948, p. 138) have noted themselves that the same formal analysis can be applied to prediction as well as to explanation. Inferences from explanations to predictions are generally accepted, but not vice versa. Michael Scriven (1962) showed one may predict a storm when cows lie down in their fields, yet the observation that cows are lying down does not explain why the storm occurs. The most famous example to show the asymmetry problem comes from Sylvain Bromberger (1966): deducing the height of a flagpole from the length of the shadow cast by the flagpole satisfies Hempel's conditions for a DN argument, but it does not explain why the flagpole has that specific height.

The three types of counterexamples mentioned in this section show that something is missing from the DN model. In section 1.3 I will show how several philosophers follow Hempel's unofficial view by claiming that unification is the missing ingredient in explanations.

[5] At the 'Causality and Explanation in the Sciences' (CaEitS) conference in Ghent (Belgium) (September 2011), the local organizing committee came up with a couple of humoristic movies in which they capture basic reasoning mistakes on correlation, causation and explanation. They filmed both the example of the hexed salt (https://www.youtube.com/watch?v=nDFAtaVFVqk (accessed 25 May 2018)) and the example of the birth control pills (https://www.youtube.com/watch?v=-ulc7HiZoj4 (accessed 25 May 2018)).

1.3 The unificationist heirs of Hempel

In this section four intellectual heirs of Carl Hempel will be briefly summarized: Michael Friedman, Philip Kitcher, Gerhard Schurz and my promotor Erik Weber. They are all advocates of explanatory unification, albeit it derivational unification. These accounts will be contrasted with my views on unification Chapter 4.

1.3.1 Michael Friedman

An early attempt to develop a unificationist account of explanation can be found in the work of Michael Friedman (1974). According to Friedman, an explanation is an argument, a derivation from a set of premises to a conclusion. The conclusion describes the explanandum and the premises describe the explanantia. Unification has to do with the relation between the number of premises and the number of conclusions. The smaller the set of premises and the larger the set of conclusions, the more our knowledge is unified:

> "this is the essence of scientific explanation—science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given" (Friedman, 1974, p. 15).

Michael Friedman argues that Hempel's DN-model provides a "clear, precise, and simple condition that the explanation relation must satisfy", and it gives us an objective criterion for explanation. (Friedman, 1974, p. 9). But, for Friedman, the DN-model lacks a connection between explaining and understanding. In order to explain, the conjunction of the universal claim and the singular claims only need to bear the appropriate deductive relation to the explanandum. For Hempel understanding is related to expectability, if a phenomenon can be expected, it is understood. For Friedman understanding is increased when the total number of independent phenomena that we must accept as ultimate is reduced. Unifying explanatory laws or phenomena in a conjoined structure increases our understanding of the world. So, for an argument to really explain, it must not only bear the appropriate relation between the explanandum and the explanans, it must also unify the explanandum with other phenomena. Friedman uses the kinetic theory of gases as an example to show how unification increases our understanding. If the kinetic theory of gases would only show us that any collection of molecules of the sort that gases are, which obeys the laws of mechanics will also approximately obey the Boyle-Charles law, it would add nothing to our understanding (Friedman, 1974, p. 14). The kinetic theory of gases allows us to derive other gas-related phenomena, such as Graham's law of diffusion, the specific heat capacities of gases, etc. At least three brute facts can now be reduced to one brute fact.

In Friedman's account the arguments that are used to explain the gas-related phenomena fit Hempel's DN-model. To explain the Boyle-Charles law you construct an argument that shows that the gas-related behavior can be expected by deriving it from the kinetic theory of gases. Further, the kinetic theory of gases, the explaining law, must unify the explanandum with other phenomena, which indeed it does.

1.3.2 Philip Kitcher

Philip Kitcher defines an explanation as an argument, in which parts of our knowledge are derived from other parts of our knowledge (Kitcher, 1981). Our knowledge, or our set of beliefs, should be systematized to reach a maximal degree of unification. The quality of the systematization, and thus the degree of unification, is determined by the size of the conclusion set, the number of and similarity among argument patterns and the stringency of those patterns.

> Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation over and over again, and, in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute). [...] If a pattern sets conditions on instantiations that are more difficult to satisfy than those set by another pattern, then I shall say that the former pattern is more stringent than the latter. (Kitcher, 1989b, pp. 432-433, italics deleted)

Unification, according to Kitcher, is achieved by constructing arguments in which parts of our knowledge are derived from other knowledge. Take *K* to be a set of beliefs, sentences we accept. Arguments that consist of arguments and conclusions that are part of *K* are *systematisations of K*. "Science offers us a reserve of explanatory arguments, which we may tap as need arises" (Kitcher, 1981, p. 512). The explanatory store over *K*, *E(K)*, is the set of arguments that best unifies *K*. This unification is realised by argument patterns: "a theory unifies our beliefs when it provides one (or more generally, a few) pattern(s) of argument which can be used in the derivation of a large number of sentences which we accept" (1981, p. 514). The notion of an argument pattern is crucial in Kitcher's account. An argument pattern has three constitutive parts:

(i) A sequence of schematic sentences, in which the nonlogical expressions are replaced by dummy letters. This is the schematic argument.

(ii) A set of filling instructions that tell us how to replace the dummy letters in the schematic argument.

(iii) A classification for the schematic argument; this determines which sentences are premises and which are conclusions, and which inference rules are used.

Kitcher uses examples from genetics and from evolutionary biology to illustrate how explanatory unification works (1989), but I will his argument pattern 'Common Descent' (1993, p. 83). This pattern gives an answer to the question: why do the members of *G*, *G'* share *P*?

(1) *G*, *G'* are descended from a common ancestor G_0.

(2) G_0 members had *P.*

(3) *P* is heritable.

(4) No factors intervened to modify *P* along the G_0 – *G,* G_0 – *G'* sequences.

Therefore,

(5) Members of *G* and *G'* have *P.*

In this scheme there are five schematic sentences that need to be filled according to the following instructions:

G, G' and G_0 need to replaced by names of groups or organisms.

P needs to be replaced by the name of a trait of organisms.

In the scheme sentences (1) till (4) are classified as premises and sentence (5) is the conclusion, deduced from the premises.

With his account Kitcher tackles the problem of explanatory asymmetry illustrated by the flagpole example (1989, pp. 484-488).

He argues that there are at least two possible systematizations of K:

(1) The systematization S'(K) of K with the shadow-pattern. This pattern derives the length of the pole from the length of the shadow.

(2) The systematization S(K) of K that includes the origin-and-development-pattern. This pattern traces the dimensions of objects (both natural and man-made) back to "the conditions in which the object originated and the modifications that it has since undergone" (1989, p. 485). Depending on the object those derivations will be more or less complex, but the general pattern is the same in all cases.

Kitcher argues that the systematization S'(K) that includes the shadow-pattern is worse than S(K) by considering two options. First, if the origin-and-development-pattern is added in S'(K), the number of argument patterns is increased, without increasing the number of derivable conclusions. The origin-and-development-pattern is sufficient to derive the dimensions of objects. Second, if S'(K) is favored, the numbers of argument patterns stays the same. But this would lead to fewer conclusions that can be derived from the set of patterns, e.g. what about objects that are not able to drop a shadow? Kitcher concludes that therefore the origin-and-development-pattern belongs to the explanatory store, and the shadow-pattern does not. Therefore, The best systematization is S(K).

His approach to the problem of irrelevancies is similar. If the argument pattern "all hexed salt dissolves in water" is included, the set of argument patterns will be less unifying since another pattern, such as "all unhexed salt dissolves in water" will be needed too. If we remove this second pattern from the set, the number of conclusions that can be derived will be limited. If both patterns are replaced with the pattern "all salt dissolves in water", the number of argument patterns decreases (or remains the same if the pattern about unhexed salt was not included) and the number of conclusions is increased (or remains the same if the pattern about unhexed salt was included) (1989, pp. 482-484).

Kitcher's account of unification was highly influential on the literature on explanation. Both with philosophers that argued in favor of unification and with philosophers who argued against it. In 1.3.4 and 1.3.5 two philosophers, Gerhard Schurz and Erik Weber, are discussed since they refine the Kitcher-Style unification.

1.3.3 Gerhard Schurz

Besides Friedman and Kitcher, there are other philosophers who have defended unification as an epistemic goal in the construction of explanations. For Gerhard Schurz (1999) the aim of an explanatory process is the achievement of a new cognitive state, in which the premises that one relies on for the answer must be in less need of explanation than the explanandum. In his account explanation and understanding are semantically related, and one cannot understand something in terms of something else that is not understood. Unification, then, is the key to reach understanding.

Schurz states that there are four key-elements in an explanatory process: the explanation-seeking question P, the cognitive state C of the questioner, the answer A and the new cognitive state C+A. A proper answer A takes the form: “P because of the reasons Prem" in which Prem must be true and the inference Prem => P must be correct in a broad sense (deductively or probabilistically).

But this is not all, Schurz also imposes that the new cognitive state C+A must be more unified ($>_u$) than C. He motivates this constraint by referring to the being-in-need-of-explanation of question P in C, which is satisfied in C+A. This can only happen if the explanatory premises Prem are less in need of explanation than the explanandum P. Schurz justifies this condition with examples, but also in a more general way by pointing out that explanation and understanding are semantically related. An explanation can only be satisfying if and only if it provides understanding; one cannot understand something in terms of something else that is not understood. This gives us the following definition of an explanation:

> A is an explanatory satisfying answer to the question Why-P? in the cognitive state C iff (i) A claims (for some Prem) Prem & Prem => P, where Prem => P is a premise-relevant correct inference in a broad sense, and (ii) C + A >u C (Schurz, 1999, p. 98)

This definition has both a local (i) and a global (ii) condition, so that one avoids the objection that unification is a global matter while explanation is a local affair. In Schurz' view an explanation is still a local affair, but it has global effects. The major rival for the unification approach is the causality approach[6]. According to Schurz the main advantage of the unification approach is to be situated in the explanation of laws. But in the domain of singular facts he has to add asymmetry as an extra requirement for explanation, such that only inferences Prem => P are allowed that are supported by a causality theory in C. But the deeper analysis of causality brings him back to unification on the level of theories. He believes that causality is not a priori but theory-relative, since the causal theories are selected only for their unificatory virtues.

> "...think, the only scientifically acceptable answer can be this: the theoretical laws which we trust to describe the real causal mechanisms are those which unify all known empirical regularities in a superior way. And only because these theoretical laws have this overall unification power, our belief in real causal processes as opposed to non-causal correlations is rationally justified." (Schurz, 1999, p. 101)

Schurz compares different levels of unification according to their assimilation status. The representation of the cognitive state C of a cognitive agent AG is the main ingredient of unification. This cognitive state C consists of the descriptive knowledge of AG and the inferential knowledge of AG.

Together they form the state of information systems: the elements of the descriptive knowledge are information units and the inferences assimilate information. Based on those constituents Schurz makes a difference between actual assimilation, potential assimilation, dissimilated phenomena and basic phenomena.

The actual assimilation can only happen if there is predictability or nomic expectability. Potential assimilation consists of virtual or heuristic assimilation. Schurz speaks of virtual assimilation if an event is random without any further hidden cause and if there is a low probability inference Prem => Pa. Heuristic assimilation happens if phenomena are derived from theories with help of initial or boundary conditions, which are considered plausible. Dissimilated phenomena are anomalies, they cannot be assimilated within the accepted theory. These are the phenomena that are in need of explanation according to Schurz.

---

[6] This approach holds that an explanation mainly consists of uncovering a causal relation between the explanandum and the explanans. The specific nature of this relationship and of the explanans varies between different causal approaches. In Part I I will investigate if unification can be a virtue or even an extra requirement of adequacy for causal explanations. In order to do this the causal accounts of Salmon (1984, 1989, 1989), Woodward (2003) and Strevens (2008) will be mentioned.

1.3.4 Erik Weber

Erik Weber defines unification as showing that two or more different events are instances of the same (set of) law(s) (Weber, 1999, p. 481). In order to account for unificatory explanations, he uses two other types: coexistence explanations and succession explanations.

A coexistence explanation for Qa is a deductive argument in which the explanandum is an instance of a law of coexistence. A succession explanation for Qa is a deductive argument in which the explanandum is an instance from a law of succession. Those types of explanations are only valid if the used laws are valid. A coexistence law has the form $(\forall x)(Px \rightarrow Qx)$ and a succession law has the form $(\forall x)[P(x,t) \rightarrow Q\ (x,t+i)]$ in which t is time. Moreover, the laws have to fulfil four conditions:

(1) the law may not be tautological or analytical

(2) the law may not be vacuous

(3) the law may not be an accidental generalization

(4) the antecedent may not be irrelevant.

Weber can now define a unificatory explanation as an explanation consisting of (1) a coexistence or succession explanation and (2) a derivation of the law used in this explanation (Weber, 1999, p. 484)]. In order for an explanation of two or more qualitative events to be unificatory the used law must be identical.

Although Weber's account is based on Kitcher's approach there are some important differences. For Kitcher the degree of unification of a set of beliefs K is directly proportional to the quality of the systematization of K. Unification is thus best reached by iterating the same argument patterns. Weber on the other hand claims that the degree of unification of K is also related to the degree of unification of the members of the set, with how many other events is an event unified, and the cardinality of the set (Weber, 1999, p. 489). This means that for Weber unification is not only a matter of using the same patterns over and over again, but also by using patterns that contain "the same empirically adequate premises again and again" (1999, p. 490). Kitcher has no criterion to choose between two sets that are equally stringent, contain the same number of argument patterns and have an equally large number of derivable conclusions. By looking at the unification between the members of the set, Weber adds an extra criterion that urges us to use the same laws in explanations.

1.3.5 Conclusion

Kitcher refers to the covering-law model as Hempel's official view, but claims that he also holds an unofficial view:

> In contrasting scientific explanation with the idea of reducing unfamiliar phenomena to familiar phenomena, Hempel suggests this unofficial view: "What scientific explanation, especially theoretical explanation, aims at is not [an] intuitive and highly subjective kind of understanding, but an objective kind of insight that is achieved by a systematic unification, by exhibiting phenomena as manifestations of common, underlying structures and processes that conform to specific, testable basic principles". (Kitcher, 1981, p. 508)

Maybe what Kitcher says here is wishful thinking. But what is certain is that at least for the authors discussed here, Hempel's idea of subsumption was the starting point for developing a unificationist account of explanation.

I will not mention the problems for Kitcher's, Schurz' and Weber's account here, but they will be discussed later when I contrast their traditional view on unification with my proposal (chapter 4).

1.4 Alternative approaches to unification

In his 2001 article Uskali Mäki shows that although unification is usually considered a virtue in science, it can also be vice. By only looking to the explananda/explanantia ratio, as is the case in Kitcher-style unification, and thus also in top-down unification, hypotheses that are not sufficiently unifying are considered as less interesting.

> The validity of utility maximization does not depend on its being an accurate description of the behaviour of individuals. Rather, it derives from its being the underlying postulate that pulls together most of economic theory; it is the major component of a certain way of thinking, with many important and familiar implications, which have been part of economics for decades and even centuries. [...] Alternatives such as satisficing have proved next to useless in this respect. While attractive as hypotheses, there is little theory built on them; they pull

> together almost nothing; they have few interesting consequences. In judging utility maximization, we must ask not "Is it plausible?" but "What does it tie together, where does it lead?" (Aumann, 1985, p. 35)

In his 2001 article Uskali Mäki claims that the type of unification proposed by Kitcher or Friedman poses a formal constraint:

> Consider the example of self-seeking maximization in economic theory. [...] the principle of unification prescribes that the results of theoretical work be derivable from the assumption of maximization by economic agents—deviations from this rule will be proscribed as ad hoc. (Maki, 2001, p. 503)

To cope with this worry, Mäki claims that the Kitcher-style unification is not the only option. Mäki distinguishes between derivational and ontological unification (2001, p. 493 et seq.). A first kind of unification is derivational or logical unification. But, instead of basing the unification on inferential capabilities, one can focus on the referential and representational capacities of a theory. This notion of ontological unification implies that a set of phenomena is unified if they share the same ontic foundations. Mäki refers here to causes, origins and constituents. This type of unity is not only about what category the phenomena belong to, but also about how the phenomena are established. This is a kind of unification that cannot be imposed, but has to be discovered by adding the relevant entities and properties to your ontological framework (Maki, 2001, p. 498 et seq.). This ontological unification is "*a matter of redescribing apparently independent and diverse phenomena as manifestations (outcomes, phases, forms, aspects) of one and the same small number of entities, powers and processes*" (Maki, 2001, p. 498). Those phenomena are ontologically unified if they only appear to be independent from each other, while they are in fact dependent on the same deeper, more fundamental structure or process. Mäki refers to Newtonian mechanics, where unification was realized by representing a variety of sublunary and superlunary phenomena as having the same causes, namely the same forces of gravitation. Newton's theory unifies better than previous theories since there are less kinds of entities referred to in order to describe the same manifestations. For Mäki this type of unification is about showing that disconnected phenomena are only apparently disconnected, since they are "*manifestations of one and the same fundamental and relatively simple structure*" (Maki, 2001, p. 499).

In Mäki's conceptual framework there are three possibilities: purely derivational unification, purely ontological unification and a mixed unification (unification that is derivational as well as ontological). Since purely derivational unification poses formal constraints, Mäki has a clear preference for the ontological variants (purely ontological or mixed), because in that case the constraints are not merely formal.

> The point is that if there are limits to unification, they had better be ontological in character. One may hope to be able to celebrate unification as a factual discovery, while a more cautious attitude will be recommendable if it is imposed as merely a formal constraint. (Maki, 2001, p. 489)

1.5 Summary and preview

In this chapter a brief historical overview of explanatory unification was given. The key ideas of unification as presented by Friedman, Kitcher, Schurz and Weber are all part of the traditional approach to unification. In the next chapters it will become clear that this approach ignores certain scientific practices. The alternative perspective of Mäki is promising, but a deeper analysis is missing. In Chapters 3, 5 and 6 of this dissertation I will construct several case studies to elaborate the idea of unification proposed by Mäki and show how it can appear in scientific explanation. My strategy is to elaborate his vague idea by looking for cases of ontological unification in scientific practice. In this way I will be able to rethink unification. This strategy presupposes that I temporarily bracket the classical accounts of unification summarised in this chapter. They will reoccur in Chapter 4.

Part I
Unification and causal explanation

Introduction to part I

Over the last decades causation has taken a prominent role in theories of explanation. Causal-mechanist and causal-interventionist accounts have become dominant and have pushed unificationism into the background. In this first part, two lines of argument will be developed to imply that by neglecting unification altogether the baby is thrown out with the bathwater. Each line of argument relates to an important question that – rather surprisingly – no one in the philosophical literature seems to have asked.

The first question is: are there cases in which unification is an explanatory virtue, while the aim of explanation is not unification? In these cases (if they exist) the aim of explanation is different (e.g. to provide information about the causes of the explanandum or about the mechanisms that produce it), but it is better to try to reach this aim in a unified way. My answer is positive. In Chapter 2 two examples will be constructed – one about mercury thermometers and one about the ideal gas law – to demonstrate how unificatory information can have a surplus value. Unifying causal explanations can result in a higher explanatory power (this is a first possible virtue) and they can result in more cogent explanations (this is a second possible virtue).

The second question is: are there cases in which unification is a necessary condition for a satisfactory explanation? If such cases exist, these explanations have a double aim: to provide causal information *and* to unify. In Chapter 3 I construct two examples from scientific practice – one about social revolutions and one about general anaesthesia – to demonstrate that scientists are interested in similarities between particular facts or between regularities. Answers that aim to provide an adequate explanation for these resemblance questions need both causal information and unification.

In Chapter 4 I will clarify some theoretical-philosophical implications of my results. My strategy to search for cases of ontological unification in scientific practice, and to use them to rethink unification, resulted in a new possible relation between unification, causality and explanation.

Chapter 2
Unification: unnecessary but virtuous

2.1 Introduction

In this chapter[7] I will argue that unification *can* be an explanatory virtue, even in types of explanation where the aim of the explanation is *not* unification. In the literature on scientific explanation, there is a classical distinction between explanations of facts and explanations of laws. In this chapter I will focus on explanations of laws, more specifically mechanistic explanations of laws. I investigate whether providing unificatory information in mechanistic explanations of laws has a surplus value. Unificatory information – in this context – is information about how the mechanism that explains the target law (the law that is the explanandum of the explanation considered) relates to other mechanisms (which explain other, related laws). I argue that providing unificatory information can lead to explanations with more explanatory power (Jim Woodward's (2003) concept of explanatory power will be used for that) and that it may lead to more strongly supported explanations.

The structure of this chapter is as follows. In section 2.2 some basic information on laws and mechanisms is given. In section 2.3 an example is presented about the behaviour of mercury thermometers, which will be analysed in section 2.4. In Section 2.5 I present a second example: the ideal gas law. In Section 2.6 this second example will be analysed and the results of this analysis will be compared with those of section 2.4. In Section 2.7 the relation between Sections 2.3-2.6 will be discussed. The aforementioned sections provide

---

[7] This chapter is based on the published paper 'The role of unification in micro-explanations of physical laws' (Weber & Lefevere, 2014).

a positive answer for the first question of part I: are there cases in which unification is an explanatory virtue, while the aim of explanation is not unification?

2.2 Laws and mechanisms

In this section two of the key concepts used will be clarified: law and mechanistic explanation. I will proceed by explicating the aims and the method of this chapter.

I will not attempt here to give a definition of what a scientific law is. Rather, I will give a non-exhaustive list of types of laws which I intend to cover. First, there are quantitative laws, typically expressed by means of a mathematical formula. An example that I will use is the Ideal Gas Law (IGL):

> For equal quantities of gas in a container, the product of pressure P and volume V is proportional to temperature T, with a proportionality constant R (the ideal gas constant), or mathematically expressed: $PV = nRT$.

Second, there are laws in which capacities are ascribed to classes of objects, for instance:

> In mercury thermometers the mercury level first drops and then rises when they are rapidly immersed in hot water.
>
> Asters are short-day plants (i.e. they start to flower when the night length exceeds some critical value).
>
> Petunias are long-day plants (i.e. they start to flower when the night length is below some critical value).

These laws describe how the objects that belong to the class react to a certain stimulus.

Third, there are causal generalisations about populations, for instance:

> Smoking causes lung cancer.
>
> Alcohol causes liver cancer.

These claims are about a population (in these cases: humans) though this population is not always explicitly mentioned. It is claimed that in this population there is a cause-effect relation between two variables.

What is a mechanistic explanation of a law? Starting with Machamer, Darden and Craver's seminal paper 'Thinking about Mechanisms' (2000) a mechanistic tradition has developed in the philosophy of explanation. For laws, mechanistic explanations are defined as follows:

> A *mechanistic explanation* of a law is a description of the underlying mechanism.

A description of a mechanism is usually called a *model* of the mechanism. The core idea of the mechanists is that, to have explanatory value, the model has to describe the mechanism in terms of its entities, its activities and the way these entities and activities are organized. I look at some characteristic quotes. Bechtel & Abrahamsen write:

> A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena. (2005, p. 423)

Carl Craver writes:

> [M]echanisms are entities and activities organized such that they exhibit the explanandum phenomenon. (2007, p. 6, italics removed)

These quotes show that mechanists have no unique way of defining what a mechanism is. However, Phyllis Illari and Jon Williamson have constructed a definition that expresses the core consensus among philosophers in the mechanistic tradition:

> A mechanism for a phenomenon consists of entities and activities organized in such way that they are responsible for the phenomenon. (Illari & Williamson, 2012, p. 123)

This definition includes the three key terms which mechanists use: entities, activities and organization. I have to make a brief remark about the concept of organisation. This has to be interpreted liberally, so that it includes random interactions between components. In a mechanistic explanation, the behaviour of the macro-system is explained by means of component parts, the activities of these parts and the way these entities and activities interact. The term "organisation" is used by mechanists (and here) to denote this interaction, even if this interaction is "chaotic" and thus not "organised" in a strict sense.[8]

[8] This and other borderline cases of mechanisms are discussed in (McKay Illari & Williamson, 2012).

Not every explanation of a law is a mechanistic explanation. Take for instance:

General law:	All waves reflect.
Auxiliary hypothesis:	Sounds are waves.
Explanandum:	Sounds reflect.

In this explanation the explanandum law is subsumed under a more general law by means of an auxiliary hypothesis. This explanation does not refer to the (activities of the) entities of which sounds are composed, so it is not a mechanistic explanation.

The aim of this chapter is to find out what the surplus value is (if any) of providing *unificatory information* in mechanistic explanations of laws. Unificatory information is information about how the mechanism that explains the law, which is my target, relates to other mechanisms which may explain other laws. In order to make this a legitimate aim, I have to make two strongly interconnected assumptions:

1) In explanatory texts that provide a mechanistic explanation of a law, it is possible to separate the information that is required for explaining the target law from additional unificatory information.
2) Scientists, when writing explanatory texts, can opt to be minimalistic (i.e. only provide information relevant for explaining the target law) or can choose to add unificatory information.

I will look at examples of explanations of laws and I will show that these assumptions are supported. For each case, I construct two possible explanations: a minimal one (without unificatory information) and an elaborate one (with unificatory information) and then compare the relative merits of these two explanations.

2.3 Mercury thermometers

The example used here is inspired by an example of Carl Hempel (1965, p. 246). A mercury thermometer is rapidly immersed in hot water. A temporary drop of the mercury column is observed, followed by a swift rise. This event can be explained (according to Hempel's DN model) by subsuming it under the following law:

L: In all mercury thermometers that are rapidly immersed in hot water, the mercury level first drops and then rises.

Hempel uses this law as part of the explanans in his example, because he wants to give an example of an explanation of a particular fact. In my example, the law is the explanandum. Three general laws are relevant for explaining L: the law of thermal expansion of fluids, the law of thermal expansion of solids and the law of heat conduction. The law of thermal expansion of fluids is:

(1) $v = v_0 + v_0 \beta \Delta T$

(v: final volume; v_0: initial volume, β: expansion coefficient of the fluid considered; ΔT: rise (+) or fall (-) of the temperature).

The law of thermal expansion of solids is:

(2) $l = l_0 + l_0 \alpha \Delta T$

(l: final length; l_0: initial length; α: coefficient of linear expansion of the solid considered).

Finally, the law of heat conduction says that the quantity of heat dQ (joules) which, in a time period dt, flows between any two plane surfaces normal to the direction x of heat flow in a time period dt , is given by the equation

(3) $dQ = -KA\left(\frac{\delta T}{\delta x}\right) dt$

(K: thermal conductivity coefficient (joule/°C.sec.m); A: area of the surfaces; $\frac{\partial \mathrm{T}}{\partial x}$: temperature gradient).

Mercury thermometers consist of glass tubes which are partly filled with mercury. The law of thermal expansion of fluids, when implemented for mercury, gives us the equation:

(4) $v = v_0 + v_0 \times 0.00018 \times \Delta T_m$

(v: final volume of the mercury; v_0: initial volume of the mercury; 0.00018: coefficient of expansion of mercury; ΔT_m: the temperature rise of the mercury).

On the other hand, the tube of the thermometer will expand in accordance to the following equation, which is an implementation of the law of thermal expansion of solids:

(5) $r = r_0 + r_0 \times 0.000008 \times \Delta T_g$

(r: final radius of the tube; r_0: initial radius of the tube; 0.000008: coefficient of linear expansion of glass; ΔT_g: the temperature rise of the tube).

As glass is a bad thermal conductor, only the temperature of the tube will initially rise ($\Delta T_g > 0$). The temperature of the mercury will remain unchanged, $\Delta T_m = 0$ and thus also $v = v_0$. The relation between the level of the mercury and its volume is given by the geometrical formula

(6) $h = \frac{v}{\pi r^2}$

(h: height of the mercury column; v: volume of the mercury; r: radius of the tube).

From this formula and $v = v_0$ we derive that

(7) $h = \frac{v_0}{\pi r^2}$

The initial mercury level is given by

(8) $h_0 = \frac{v_0}{\pi r^2}$

Because the temperature of the glass tube rises, r is larger than r_0. The two last equations and $r > r_0$ imply that

(9) $h < h_0$

In other words, the mercury level drops in a first phase. In a second phase, the temperature of the mercury will start to rise after a few instants as a result of the heat conduction. This let the mercury expands due to the thermal expansion law of fluids. Because the expansion coefficient of mercury is large, the expansion will be greater than the extension of the tube. Consequently, the mercury level will rise.

2.4 Analysis of the mercury example

The explanation in Section 2.3 fits the aforementioned definition of a mechanistic explanation: it identifies the components (glass tube, mercury) and describes how they behave (equations (4) and (5)); and equations (6)-(9) are consequences of the organization of the parts (the most important organization-fact is that that mercury is inside the glass tube).

As suggested in section 2.2, scientists can opt to be minimalistic or they can choose to add unificatory information. In order to argue for this claim, let me separate the information that is required for explaining the target law from the additional unificatory information. Equations (1) and (2) are not really necessary for describing how mercury thermometers work. All that is needed is (4) and (5), and these can be known independently (by means of experiments with mercury and glass). Similarly, equation (3) is not used in the derivation. Instead, the instantiation is used where K is substituted for the thermal conductivity coefficient of glass.

In order to get a better grip on what is going on here, it is useful to distinguish functional laws from parametric laws (Ajdukiewicz, 1974, pp. 318-323). Let us first look at an example of a functional law:

> For all free-falling material objects x near the surface of the earth and for all numbers s en t: if s is the distance covered by x and t the time during which x is falling, then $s = \frac{9.81t^2}{2}$.

This is Galileo's law. Formally, we can write this as:

$$(\forall x)(\forall s)(\forall t)\left[(S(x) = s \wedge T(x) = t) \rightarrow s = \frac{9.81t^2}{2}\right]$$

The general form of functional laws is:

> For all objects x in domain D and all numbers $r_1, r_2, \dots, r_n$: if $M_1(x) = r_1, M_2(x) = r_2, \dots, M_n(x) = r_n$, then $F(r_1, r_2, \dots, r_n) = 0$.

An example of a parametric law is:

> For all gas samples x with fixed temperature there is a parameter k such that for all numbers p en v: if p is the pressure of x and v its volume at the same moment of time then, $p \times v = k$.

Formally, we can write this as:

$$(\forall x)(\exists k)\big[(\forall p)(\forall v)[(P(x) = p \wedge V(x) = v) \rightarrow p \times v = k]\big]$$

The general form of parametric laws is:

> For all objects x in domain D there is a parameter k such that for all numbers$r_1, r_2, \dots, r_n$: if $M_1(x) = r_1, M_2(x) = r_2, \dots, M_n(x) = r_n$, then $F(k, r_1, r_2, \dots, r_n) = 0$.

The explanation presented in Section 2.3 consists of (i) a derivation of the law to be explained from functional laws about the behaviour of the parts and knowledge about their interaction, and (ii) information (contained in parametric laws) about similar behaviour of related objects. The parametric and functional laws are logically independent. The parametric laws are not derivable from the functional ones because the parametric laws are more general, as they are about a larger class of objects; the functional laws are not derivable from the parametric laws because the functional laws are more precise, since they contain a specific value for the parameter. As a consequence, there is unificatory information which can be easily separated from the rest of the explanation. The explanation as a whole consists of a part that tells us how things work in the case at hand, and a part that connects this mechanism to potential mechanisms in similar cases. If the latter part is deleted, we have a minimal explanation. If it is added, we have an elaborate explanation. I will now compare the value of these two possible constructions.

In my evaluation I use Jim Woodward's views on explanation, introduced in his book *Making Things Happen*: "explanation is a matter of exhibiting systematic patterns of counterfactual dependence" (2003, p. 191). He discusses a classic example (2003, p. 187):

> All ravens are black.
>
> a is a raven.
>
> _______________
>
> a is black.

He claims that this is not a satisfactory explanation because it

> ... doesn't tell us about the conditions under which raven *a* would be some other colour than black. (2003, p. 193)

This fits into his general idea: the argument above tells us that we could have expected that *a* is black, given that we knew that it was a raven. But expectability is not enough for

Woodward: an explanation must also tell how the explanandum would change if the initial conditions would be different. In other words, adequate explanations ...

> ... locate their explananda within a space of alternative possibilities and show us how which of these alternatives is realized systematically depends on the conditions cited in the explanans. They do this by enabling us to see how, if these initial conditions had been different or had changed in various ways various of these alternatives would have been realized instead. (2003, p. 191)

Before I apply this idea to my case, let me give a simple example of a good and a bad explanation. According to Woodward's views on explanation, the following argument is useless:

All pendula with a length of 1 metre have a period of 2.006 seconds.

This pendulum has a length of 1 metre

This pendulum has a period of 2.006 seconds

However, the following argument is an adequate explanation:

For all pendula: $P = 2\pi\sqrt{L/g}$

All pendula with a length of 1 metre have a period of 2.006 seconds.

This pendulum has a length of 1 metre.

This pendulum has a period of 2.006 seconds.

In this second argument, the law which also occurred in the first argument is no longer a premise, it is derived from the general pendulum law. This general law is superfluous from an expectability perspective, but it is necessary in order to provide the counterfactual information. The second explanation tells us namely how the period would be different if the length of the pendulum were different.

Let me now apply Woodward's view on explanatory power to the mercury example. The minimal explanation gives some counterfactual information. We can calculate what would happen if the expansion coefficients of glass and mercury would be higher or lower than they actually are. The elaborate explanation gives the same counterfactual information, but offers something extra. It offers insight in what would happen if the thermometer would be made of different materials. What if the tube of mercury thermometers would not be made of glass, but of another transparent material? The equations (1) – (3) tell us that similar laws would apply, and that what happens if the differently composed thermometer is immersed in hot water depends on the thermal conductivity and expansion coefficient of the other material.

Woodward summarizes his idea by saying that the explanations must answer "what-if-things-had-been-different questions" and that explanations which answer more such questions are better than explanations that answer less such questions. In our mercury example, both the minimal and the elaborate explanation satisfy Woodward's minimal condition. But the elaborate explanation answers more what-if-things-had-been-different questions and is therefore better.

I can take the analysis one step further: why is it interesting to have many answers to what-if-things-had-been-different questions? One reason is practical. If the real situation, the real behaviour we are explaining, is not our ideal, the answers inform us about the possibilities of creating an alternative, more desirable situation. So the unificatory information can have a practical value.

Another reason one may give is that this is what understanding is all about: systematic knowledge of counterfactual patterns of dependence is intrinsically valuable because it provides understanding of the world. Such claim is not uncontroversial. Hempel has claimed that science serves two basic human motives: the capacity to predict and control and sheer intellectual curiosity (1965, p. 333). No problem so far. However, he equates this intellectual curiosity with the desire to show that everything could be expected by means of certain laws and initial conditions. In other words, Hempel identifies understanding with expectability. In order to argue that answers to what-if-things-had-been-different questions are intrinsically valuable, I need a different, stronger implementation of the idea of intellectual curiosity.

2.5 The ideal gas law

In this section I present as a second example the relation between the ideal gas law (IGL) and the kinetic theory of gas (KTG) to show that unification can be an explanatory value. In this case the advantage of unification is stronger than in the previous example, since it adds credibility to the overall argument.

The IGL is about standard, isolated quantities of gases in closed spaces. It claims a proportional relationship between the pressure P, the volume V and the temperature T:

IGL: $PV = nRT$ in which R is the ideal gas constant.

For an equal quantity n of gas in a container, the product of pressure P and volume V is proportional to temperature T. This proportionality constant, R, is the same for all ideal gases and is therefore called the ideal gas constant. To explain why this proportionality holds, we need to understand the underlying machinery. To derive the IGL from the KTG I will use Kuipers' account of explanations of laws (Kuipers, 2001, pp. 82-104). According to Kuipers explanations of laws have different forms, but contain only steps of five types: application steps, aggregation steps, identification steps, correlation steps, and approximation steps. As a rule, these steps occur in this order, but exceptions are possible. Not all explanations contain steps of all types, and some types may occur more than once.

KTG assumes the following basic hypotheses:

a) The ideal gas consists of molecules, and the intermolecular distance in these gases is much larger than the dimensions of the molecule itself. Therefore, we can consider those molecules as point masses.

b) The motion of this molecules is governed by Newton's laws.

c) The collisions of the molecules with each other or with the walls of the container are in correspondence with Newton's second and third law, and thus considered perfectly elastic.

The first step is to apply these laws to one molecule with a mass m and a velocity v colliding with the wall of the container. This is the application step of the KTG to the IGL. We consider a cubical box of volume V, with side length L and a wall surface A (= L^2). In that box a molecule will collide with one of the walls at a velocity $v_{x,i}$. We know that collisions between opposite walls in the x-direction on average will happen in intervals of $\Delta t_i = 2\frac{L}{v_{x,i}}$. In that collision the x-component of the momentum will be reversed from $mv_{x,i}$ to $-mv_{x,i}$.

(1) This gives us the following force on the wall: $F = \frac{2mv_{x,i}}{\Delta t_i}$

The second step is an aggregation step. By means of a statistical auxiliary hypothesis the individual law is applied to a larger number of molecules. We assume the following auxiliary hypothesis:

(a) The ideal gas is considered a homogeneous distribution of N molecules, which are in constant random motion. N is in the order of 6.022×10^{23}, the number of Avogadro N_A.

This results in the following aggregated law:

(2) $F = \sum_{i=1}^{N} \frac{2mv_{x,i}}{\Delta t_i}$.

By using $\Delta t_i = 2\frac{L}{v_{x,i}}$ as an identity hypothesis, we can perform a transformation step on (2), which gives us:

(3) $\mathrm{F} = \sum_{i=1}^{N} \frac{mv^2{}_{x,i}}{L} = \frac{2}{L}\sum_{i=1}^{N} \frac{mv^2{}_{x,i}}{2}$

To further proceed, we use the following auxiliary hypothesis:

(b) the kinetic energy for one particle in one dimension is defined as: $\mathrm{T} = \frac{1}{2}\mathrm{mv}^2$.

and apply a second aggregation step:

(4) Since there are N molecules in the container, moving in three dimensions, this gives us a total kinetic energy T_{tot} of :

$$T_{tot} = \frac{1}{2}m\sum_{i=1}^{N}\left(v^2{}_{x,i} + v^2{}_{y,i} + v^2{}_{z,i}\right) = T_x + T_y + T_z$$

Now we apply an identification step, since we assumed that the N molecules are moving in a constant random motion.

(c) Therefore, we can say that $\mathrm{T_x} = \mathrm{T_y} = \mathrm{T_z}$, and thus we get:

(5) $T_{tot} = 3T_x$

If we implement this in (3) we get:

(4) $F = \frac{2}{L}\frac{T_{tot}}{3}$

Now, another auxiliary hypothesis is added:

(d) the pressure P is defined as a force F exerted on a surface A , here the wall of the box: $P = \frac{F}{A}$

Combined with (4), we get another identification step:

(6) $P = \frac{2}{LA}\frac{T_{tot}}{3} = \frac{2}{3}\frac{T_{tot}}{V}$

By means of the statistical auxiliary hypothesis of the virial theorem, the total kinetic energy of a system is connected with his temperature T in the following formula:

(e) $T_{tot} = \frac{3}{2}NkT$ in which k is the Boltzmann constant.

We can apply this to (6) and have a third identification step, which gives us:

(7) $PV = NkT$

In another auxiliary hypothesis the Boltzmann constant k is related to the ideal gas constant R as:

(f) $\mathrm{k} = \frac{\mathrm{R}}{\mathrm{N_A}}$

As a final result we perform a last identification step on (7) and derive the Ideal Gas Law:

(8) $PV = \frac{N}{N_A}RT = nRT$

2.6 Analysis of the IGL explanation

The explanation fits the previously presented definition of a mechanistic explanation: it identifies the components (box, wall, molecules) and describes the activities (linear motion of the molecules) and how they interact (perfectly elastic collisions, no attraction between molecules, randomly distributed direction of motion of the molecules).

Like in the mercury example, I can construct here a minimal explanation and an elaborate one. The minimal explanation is the explanation obtained by taking the result of the application step (equation (1)) as a premise (i.e. by removing the application step).

Let me call the result of that move the "minimal IGL explanation" while I call the one presented in Section 2.5 the "elaborate IGL explanation".

There is an important difference between the IGL case and the thermometer case. In the thermometer case, the additional information in the elaborate explanation did not deductively imply the premises of the minimal explanation. Although the additional information results in more explanatory power, it could be removed without harming the credibility of the explanation. The reasons for accepting the explanation remain intact if this information is removed. In the case of the IGL the information that is left out in the minimal explanation provides an argument for believing equation (1): this equation can be logically derived from it.

This difference creates room for a second advantage that unificatory information may have. It provides support for premises one may otherwise doubt. Indeed, if the application step in the explanation in Section 2.5 is removed, it is unclear why one should believe that equation (1) is correct. There certainly is not a possible experiment to support it (that would require a setup where exactly one gas molecule is put in a container to measure the force it exerts on the wall of the container). The reasons for accepting this premise are theoretical rather than experimental. The elaborate IGL explanation incorporates, at least partially, the necessary theoretical reasons for believing equation (1). The minimal explanation leaves out these reasons. As a consequence, the elaborate IGL explanation is more cogent.

2.7 Conclusion

In this chapter I have shown that there are cases in which unification is an explanatory virtue, while the aim of explanation is not unification. The two examples focused on unification as a potential explanatory virtue in mechanistic explanations of laws. In both cases there was *one* target law, and I investigated whether and why unification is an explanatory virtue in such contexts.

It was demonstrated that in the context of mechanistic explanations of laws, providing unificatory information can have two advantages: it can result in a higher explanatory power and in more cogent explanations. This was done without claiming that these advantages are exhaustive: there may be other virtues attached to unification in this context. However, I think I have discovered and described two important ones (i.e. advantages that occur frequently in scientific practice).

In the next chapter the focus is on the second question: are there cases in which unification is a necessary condition for a satisfactory explanation? Are there explanations with a double aim: to provide causal information *and* to unify?

Chapter 3
Unification as a necessary ingredient of explanations

3.1 Introduction

In this chapter[9] I will argue that there are indeed explanation-seeking contexts in which unification is one of the necessary conditions for a satisfactory explanation. Resemblance questions (see section What are resemblance questions?) play a crucial role in the way I develop my argument. Although they do occur in scientific practice, they have been largely neglected in the philosophical literature. My aim in this chapter is to show that questions of this type cannot be adequately answered without unification. The first example, in section 3.3, demonstrates that resemblance questions about particular facts demand what I call causal network unification. This means that answers to such resemblance questions must give information about the causal network and at the same time explicate the unifying factors in that network. In the second half of this chapter I show that resemblance questions about regularities require what I call mechanism unification. Explanations of this type also have two conditions of adequacy: they need to provide information about the causal mechanisms, but in a unified way. In section 3.3. the general considerations will be presented, and in 3.4. they will be further clarified with an elaborate case study.

[9] This chapter is based on the published paper 'Unification, the answer to resemblance questions' (Weber & Lefevere, 2017).

3.2 What are resemblance questions?

Before we can connect unification to resemblance questions, the terminology that will be used in this chapter needs some clarification. Let us first look at explanation-seeking questions about particular facts. The simplest possible format of such questions is:

> Why is it the case that X?

Here X is a description of a particular fact. We use the label “plain questions about facts” to denote questions of this form. Examples are:

> Why did the French revolution occur in 1789?
>
> Why is Belgium a monarchy?
>
> Why did the space shuttle Challenger explode?

Bas van Fraassen (1980) drew attention to a second type of explanation-seeking question: contrastive questions. Their simplest form is “Why X rather than Y?”. Here are some examples:

> Why did John paint a portrait of the Queen, rather than a landscape?
>
> Why did John – rather than Bill – paint a portrait of the Queen?

Resemblance questions are a third type of explanation-seeking questions. Suppose we observe that both Peter and Mary have blood group A. Then we can ask the following question:

> Why do Peter and Mary both have blood group A?

Similarly, if we observe that Peter and Mary both have blue eyes, we can ask

> Why do Peter and Mary both have blue eyes?

As is clear from these examples, resemblance questions about particular facts focus on similarities between two or more facts, rather than on differences (as contrastive questions do) or on just one fact (as plain questions about facts do).

The same distinction can be made with respect to regularities. It is possible to ask plain questions, contrastive questions and resemblance questions about regularities.

Examples of plain questions are:

> Why do children of blue-eyed parents always have blue eyes?
>
> Why are all ravens black?

Examples of contrastive questions are:

> Why do pigeons have the capacity to find their way back home, while other sedentary birds do not have that capacity?
>
> Why do woodcocks migrate during the night, while pigeons cover long distances during the day?

Examples of resemblance questions are:

> Why do both humans and desk calculators have the capacity to perform exact numerical calculations?
>
> Why do bats and hedgehogs hibernate?

3.3 Causal network unification

Scientists do ask resemblance questions about particular facts, even when it is not possible to subsume them under a law. Answering them requires then what I call *causal network unification*. I first give a general characterisation of this kind of unification. Then I give an elaborate case study. Finally, I clarify how causal network unification answers a part of the question "what is required of an adequate explanation if this explanation addresses a resemblance question?"

Suppose that John and Peter both have lung cancer. When asked for an explanation for this similarity, we may give the following account:

> Smoking is a positive causal factor for lung cancer.
>
> John smokes.
>
> Peter smokes.

Giving this explanation is a very simple form of what we call causal network unification. The general scheme in which this explanation fits is:

(A) X is a positive causal factor for E.

Object *a* has X.

Object *b* has X.

When putting forward such explanation, it is claimed that the two objects about which we have a question have one causal factor in common. The explanation declares X to be a causal factor and asserts that it is present in both objects.

Causal network unification can be defined as the act of providing an explanation for a similarity between two events or more events presenting *one* causal network and applying it to *all* events involved. A causal network is a set of (possibly connected) causal factors. To apply a network to an event means to assert that all its factors are present in the causal ancestry of this event.

To clarify this, let me give a more elaborate example. Suppose one asks why John and Peter both started to smoke at the age of 15. This question can be answered by presenting the following causal network and applying it:

Parental separation (X) is a positive causal factor for smoking initiation at adolescent age (E).

Parental separation (X) is a positive causal factor for high level of rebelliousness (Y).

A high level of rebelliousness (Y) is a positive causal factor for smoking initiation at adolescent age (E).

Smoking parents (Z) are a positive causal factor for smoking initiation at adolescent age (E).

This causal network contains three causal factors. Y mediates between X and E, while Z influences E independently of Y and X.[10] To apply this network to Peter and John means to claim that their parents were divorced, that their parents smoked and that Peter and John had a high level of rebelliousness.

[10] This is not a toy example. The causal relation between parental separation and smoking initiation is investigated in Kirby (2002). He uses rebelliousness as a mediating variable and presents the smoking status of the parents as another potential causal factor. So, this network contains a selection from real scientific research and results.

Given the definitions I gave above, the general format of an explanation that results from an act of causal network unification is the following:

(B) X is a positive causal factor for E.

Y is a positive causal factor for E.

Z is a positive causal factor for E.

...

Object *a* has X, Y, Z, ...

Object b has X, Y, Z, ...

Object c has X, Y, Z ... .

...

If there are connections between the variables (e.g. mediating variables as in the example above) the explanation also contains claims specifying these connections[11].

If one assumes that the motivation behind resemblance questions about facts is curiousness about causal factors the events have in common, scheme (A) represents the minimal information an explanation has to give in order to be a minimally adequate answer. In this sense causal network unification (which results *at least* in an explanation of type (A)) is required for answering resemblance questions about facts. Scheme (B) gives a standard format in which the information provided by more elaborate explanations can be represented and compared.

The material used for the elaborate example is taken from an article of Michael Taylor on revolutionary collective action (1988) which discusses Theda Skocpol's classic *States and Social Revolutions* (1979). By using comparative methods, Skocpol has formulated a so-called "structural" explanation for three successful modern social revolutions in agrarian-bureaucratic monarchies (the French, Russian and Chinese revolution). The structural conditions that, in her view, make a revolution possible relate to the incapacitation of the central state's machineries, especially the weakening of the state's repressive capacity. The revolutions can be successfully mounted only if these structural

[11] Readers acquainted with causal graph theory (e.g. Pearl 2009 will notice that these causal networks a s we use them can be graphically represented in DAG's (directed acyclic graphs).

preconditions are met. The weakening is caused by external military pressure: because of the backward agrarian economy and the power of the landed upper class in the agrarian-bureaucratic monarchy, the attempt to increase the military power leads to a fiscal crisis. Escalating international competition has an economic (fiscal) impact and thus influences the state's repressive capacity negatively. That makes social revolutions possible. The foreign military pressure that made the respective social revolutions possible, was the following:

(1) Bourbon *France* (1787-89) was financially exhausted after the war for American independence and because of the competition with England in general.

(2) Manchu *China* (1911-16) was involved in the Sino-Japanese War (1895) and the Boxer debacle (1899-1901).

(3) Romanov *Russia* (1917) was involved in World War I.

Skocpol's theory gives an answer to the following resemblance question:

> Why was there a revolution in Bourbon France, Manchu China and in Romanov Russia?

Before I put this answer in the format given in Section 3.1, let us have a look at what Taylor adds to Skocpol's theory. According to Michael Taylor there is another causal factor which the three revolutions have in common, viz. a strong sense of community among the peasants:

> When the peasant community was sufficiently strong, then, it provided a social basis for collective action, including revolutionary collective action and rebellions and other popular mobilizations. (1988, p. 68)

Taylor shows how the participation of vast numbers of peasants in collective action could be explained by means of economic incentives and selective social incentives. Without incentives to motivate participation, collective action is unlikely to occur even when large groups of people with common interests exist. Using this account of collective action, Taylor argues that peasant collective action in revolutions was based on a strong sense of community and that this is mainly why the large numbers of people involved were able to overcome the free-rider problem familiar to students of collective action and opted for conditional cooperation. Taylor's idea that a strong sense of community is a positive causal factor for social revolutions does not contradict Skocpol's ideas: they propose different but compatible claims about factors that are causally relevant for social revolutions.

The work of Skocpol and Taylor is an illustration of the fact that scientists sometimes address resemblance questions about facts. When put into the format of causal network unification (B), Taylor's answer is:

> Large external military operations (X) are positive causal factors for social revolutions (E).
>
> Fiscal crises (Y) are positive causal factors for social revolutions (E).
>
> Weakening of the state's repressive capacity (Z) is a positive causal factor for social revolutions (E).
>
> Large external military operations (X) are positive causal factors for fiscal crises (Y).
>
> Fiscal crises (Y) are positive causal factors for weakening of the state's repressive capacity (Z).
>
> A strong sense of community among peasants (Z') is a positive causal factor for social revolutions (E).
>
> *Bourbon France, Manchu China and Romanov Russia have X.*
>
> *Bourbon France, Manchu China and Romanov Russia have Y.*
>
> *Bourbon France, Manchu China and Romanov Russia have Z.*
>
> *Bourbon France, Manchu China and Romanov Russia have Z'.*

Skocpol's answer the leaves out Z' as causal factor and the corresponding claim that it is present in the three states (the claims are in italics). Taylor's answer is richer, but the basic structure is the same: they use causal network unification as defined in earlier in this section.

The questions we are dealing with in this section have the following form:

(C) Why do objects *a*, *b*, *c*, ... have property X?

My claim is that such questions should be answered by means of an explanation that can be schematically represented by format (B) (with scheme (A) as a bottom line). Such explanations are both causal (because they identify factors that are present and causally relevant) *and* unifying (because they claim that factors *of the same type* are present in each case. So my claim is that questions of type (C) must be answered by means of a causal explanation (viz. by constructing and applying a causal network). But there is an extra, unification-imposing condition: we have to use the same causal network for all events mentioned in the question. In terms of norms for explanation (cfr. Craver 2007) simply giving a causal explanation is not sufficient, it should be of a specific type.

It is important to notice that we may fail to meet this extra condition even if we try. For instance, it is possible that we find out that no known positive causal factor for lung cancer is present in both John and Peter. In such cases, the resemblance remains unexplained: it turns out to be accidental. Whether or not the type of unification that I presented here is possible in a specific case, is a contingent matter. By trying to give an explanation of type (B), we either find an explanation of the resemblance or determine, that given the current state of knowledge, the resemblance is accidental.[12]

3.4 Resemblance questions and mechanism unification: general considerations

I now shift my attention from resemblance questions about particular facts to resemblance questions about regularities. Because the case study is longer and I need more general background (e.g. on mechanism and the mechanistic model of explanation, I split the part on mechanism unification into two sections: this one containing the general considerations and the next one containing the case study.

It is well-known that why-questions are ambiguous. If I ask: “Why are all ravens black?” this may be a question about the origin of the colour or about the function of the colour. In the first case, the why-question can be rephrased as “How is black feather colour in ravens produced?” In the second case, it can be rephrased as “What is the function of black feather colour in ravens?” In the same vein, why-questions about similarities can stem from different underlying ideas about kinds of explanations that are sought four. For instance, the example about bats and hedgehogs may be a motivated by

[12] If our knowledge of relevant causal factor increases, this may change because an adequate explanation may become possible.

a desire to know the common function(s). In that case, the question can be more precisely formulated as

> What is the common function of hibernating in bats and hedgehogs?

I will confine myself to resemblance questions that have a different type of motivation, viz. insight in the mechanisms that produce a behaviour or capacity. If this kind of insight is sought, the sample question about exact calculation can be rephrased as:

> What are the common features of the mechanisms that produce the capacity to perform exact numerical calculation in humans and in desk calculators?

This question fits into the following general format:

> (D) What are the common features of the mechanisms that produce capacity E in objects of type X, Y, Z, ... ?

Mechanisms are to be understood here as follows:

> A mechanism for a phenomenon consists of entities and activities organized in such way that they are responsible for the phenomenon. (Illari & Williamson, 2012, p. 123)

Illari and Williamson present this definition as an expression of the core consensus among philosophers in the mechanistic tradition. Many slightly different characterizations have been given in the last two decades. But they all contain a common core idea, and thus relate back to the definition used in 2.2:

> A *mechanism* is a collection of entities and activities that are organized such that they realize the macro-level behaviour of a system.

We have identified the type of questions we consider and we know what the key term (mechanism) means. But what do answers to questions of type (D) look like? Illari & Williamson characterise mechanistic explanations as follows:

> All mechanistic explanations begin with (a) the identification of a phenomenon or some phenomena to be explained, (b) proceed by decomposition into the entities and activities relevant to the phenomenon, and (c) give the organization of entities and activities by which they produce the phenomenon. (Illari & Williamson, 2012, p. 123)

This is a dynamic characterisation, which tells us how mechanistic explanations are construed. Illari & Williamson do not give a (static) definition of the end product. The following definition is complementary to what they say:

> A mechanistic explanation for a phenomenon is a description of the mechanism that produces this phenomenon.

A description of a mechanism is often called a *model* of the mechanism.

The dynamic and static characterisations cover explanations in which *one* capacity is explained (i.e. answers to plain questions). To make the step towards answers to questions about common features in mechanisms, it is important to see that models of mechanisms can vary in richness. A model of a mechanism can be very *rich*, in the sense that it provides a lot of details about the entities, activities and organisation of the mechanisms. Or it can be rather *abstract*, in the sense that it leaves out a lot of details. This richness/abstraction axis is important for us because questions of type (D) are questions about shared features of mechanism. By definition, a model that correctly describes the shared features has to omit all features that are not present in all the types X, Y, Z, ... Hence, a model that correctly describes these features will always be quite abstract.

With this background information in place, we can now define what we mean with mechanism unification. The act of mechanism unification consists in providing the richest possible mechanistic model that is correct for all types X, Y, Z,... Mechanism unification in this sense produces the most informative correct answer to questions of type (D). It results in explanations that are both mechanistic and unifying. Questions of type (D) must be answered by means of a mechanistic explanation. But there is an extra, unification-imposing condition: one should try to find the richest possible model that is valid for all types of objects included in the question. In terms of norms for explanation: simply giving a mechanistic explanation is not sufficient, it should be of a specific type. All this is analogous to Section 3, where the explanations had to be both causal and unifying.

Mechanism unification as defined here is complementary to network unification discussed in 3.3: it is an attempt to answer resemblance questions about regularities (type (D)) while causal network unification is an attempt to answer resemblance questions about particular facts (type (C)). While causal network unification results in intra-level explanations (the causal factors are situated at the same level of abstraction as the explanandum), mechanism unification results in an interlevel explanation: the objects whose capacities are explained are decomposed into smaller entities.

3.5 Resemblance questions and mechanism unification: a case study

Our aim here is to show that questions of type (D) do occur in scientific practice. The case study we use is about general anaesthesia. Our main source is the book *The Wondrous Story of Anesthesia* (Eger II, Saidman, & Westhorpe, The Wondrous Story of Anaesthesia, 2014)[13].

General anaesthesia was first practiced on October 16, 1846. On this day, Dr. William T.G. Morton, a young dentist from Boston, and Dr. John Collins Warren, the Chief of Surgery at Harvard performed the first surgery with general anaesthesia. Dr. Morton used an inhaler with volatile ether to put the patient to sleep and Dr. Warren removed a tumour in the patient's neck. It was the first time this happened: till then patients were kept in a conscious state, though e.g. nitrous oxide was often used to suppress pain. The patient, Gilbert Abbott, described his experience as a slight pressure followed by wonderful dreams (Robinson & Toledo, 2012, p. 143). With volatile ether the history of general anaesthesia took a real start. Volatile ether not only causes unconsciousness, but also amnesia, analgesia (insensibility to pain) and muscle relaxation. "General anaesthesia" refers to a state in which these four factors are present.

Only a year later the Scottish obstetrician James Simpsons discovered the anaesthetic effects of chloroform. It became the "British" anaesthetic while ether remained the "Yankee dodge" (Eger II, Saidman, & Westhorpe, The Wondrous Story of Anaesthesia, 2014, p. 28). In the first decade of the 20th century several scientists administered anaesthetics intravenously: ether, chloroform, hedonal, barbiturates and thiopental. (White, 2014, p. 630). In the 1960s etomidate and ketamine were synthesized but it was propofol that displaced thiopental as the most important intravenous anaesthetic in the 1980s. Simultaneously anaesthetic inhalation machines were developed to control the

[13] This recent book consists of two parts. First the history of anesthesia is reconstructed chronologically. In the second part the history of anesthesia is described per geographical region or per subject. I have entered the individual chapters to which I refer separately in the bibliography.

amount of volatile ether, chloroform and nitrous oxide and to make use of rebreathing to minimize the consumption. These machines enabled the use of ethylene as an anaesthetic. Another agent that was discovered in that period was cyclopropane. The new fluorine chemistry needed for the development of the atomic bomb also gave us modern anaesthetic compounds halogenated with fluorine. In much of the world, halothane was the inhaled anaesthetic of the 1960s; enflurane of the 1970s; and isoflurane of the 1980s (Eger II, Saidman, & Westhorpe, 2014, p. 132). In the 1990s desflurane and sevoflurane were released for clinical use, but sevoflurane won the popularity poll. The advances slowed down in the 2000s, when only sugammadex was released as a new anaesthetic. In today's operating rooms the use of combined intravenous and inhaled anaesthetics remains the most popular approach for general anaesthesia (White, 2014, p. 638).

The discoveries listed above lead to the following resemblance question:

> What do the mechanisms by which the chemical substances X, Y and Z induce general anaesthesia (i.e. cause unconsciousness, amnesia, analgesia and muscle relaxation) have in common?

Several scientists tackled this question. Already in 1847 von Bibra and Harless linked the phenomenon of anaesthesia to dissolving and removing lipids in the brain:

> With all its chemical attributes intact ether in the nervous system first affects the components for which it has the greatest affinity, the fatty ones. It partly dissolves them, and the solution gets taken up by the venous blood after traversing the capillaries by endosmosis. Partial dissolution of the fat by ether must of course necessarily alter the components of the nervous tissue and nullify their mode of action. (von Bibra & Harless, 1847)

In 1875, after experiments on frogs, Claude Bernard suggested that there ought to be one common mechanism for general anaesthesia for all living organisms, despite the diversity of the used agents. He suggested that a reversible coagulation of nerves underlay the production of anaesthesia by some means that applied to all forms of life – a unitary mechanism of anaesthesia. (Eger II, Saidman, & Westhorpe, 2014).

At the turn of the 20th century Meyer and Overton independently discovered a positive correlation between the lipid solubility of the anaesthetic compound in olive oil and their anaesthetic potency (Meyer, 1899 and Overton, 1901). This led to a *reversed* version of the lipid hypothesis. Contrary to von Bibra and Harless, Meyer and Overton suggested that anaesthetics dissolve *in* the lipids. In this way anaesthetics changed the state of brain cell lipids, resulting in narcosis. The lipid hypothesis served as a unitary mechanism of general anaesthesia, suggesting that when general anaesthesia occurred the concentration of the anaesthetic agent in the lipid cell membrane of the neurons would be the same for all anaesthetics. In the 1960s many researchers sought a mechanistic basis for the Meyer-Overton hypothesis. The 1970s saw a variation of lipid theories. But there were several questions that lipid theories could not answer: the effect of body temperature on the membrane lipids, the absence of anaesthesia with some lipid soluble drugs and variable anaesthetic effects with enantiomer drugs[14]. Despite the numerous papers on lipid theory, the question remained how the effect on the membrane lipid cells could be extrapolated to the membrane proteins, since these proteins were thought of as the eventual target.

The lesson we can draw from this brief survey is that from 1847 till around 1980 there were several attempts to answer the resemblance question about general anaesthesia. Lipids played a crucial role in these attempts. In the first attempts, it was posited that all anaesthetics dissolve fatty brain cell membranes. Mechanism unification results here in an abstract description of the mechanisms that leaves out many details, e.g. on how the anaesthetics reaches the brain (inhaled vs. intravenous anaesthetics). What the description does have to contain is the claim that one entity (the anaesthetic) performs an activity (dissolving) on another entity (lipid brain cell membranes). In the later attempts, again the details about how the anaesthetics enter the brain have to be left out in order to describe a unified mechanism. The entities that are present in the description of the unitary mechanism are the same, but the claim about activity is different: it is now claimed that the brain cell membranes affect the anaesthetics (they absorb them) rather than the other way around.

In the 1980s Franks and Lieb discovered that general anaesthetics can also interact in lipid-free environments (1982) (1984). They showed that agents able to inhibit two types of proteins (luciferases and cytochrome P450) have anaesthetic potencies. By the 1990s the lipid theory was largely abandoned and the focus shifted to proteins. With this shift the search for a unitary mechanism also faded into the background, since the question now became: which proteins are involved? In the last decades researchers examined the effect of anaesthetics on specific protein targets. Among those targets the $GABA_A$

[14] Enantiomers have identical chemical and physical properties except for their ability to rotate plane-polarized light. More info on enantiomers can be found on http://www.chemguide.co.uk/basicorg/isomerism/polarised.html.

receptors, glycine receptors, NMDA receptors and anaesthetic-activated potassium channels were considered important in general intravenous anaesthesia (Franks, 2014, p. 603). For inhaled anaesthetics the understanding is much less complete, since here the spinal cord and not the brain is the mediator for immobility. Identifying the specific molecular targets progresses slowly and no consensus has been reached. Currently there are two main views regarding the targets of inhaled anaesthetics. Franks & Lieb (1994) believe that only a few targets are important, while Eckenhoff (2001) and others believe that many targets contribute in a small amount to anaesthesia through inhalation. Researchers conduct studies in which they inhibit or block certain targets to explore the change in anaesthetic potency of that target. For example, a study in 2002 shows that acetylcholine receptors do not contribute to the immobility induced by inhaled anaesthetics, but may be important to amnesia (Eger II, Gong, Raines, & Flood, 2002). Many questions remain unanswered at the molecular level, and the search for how individual anaesthetic agents function still continues.

The resemblance question was left aside only for a brief period. George Mashour (2004) is strongly convinced we should not abandon the quest for a unitary hypothesis for the mechanism of general anaesthesia.

> Although general anaesthesia may be mediated by complex pharmacologic and neurologic mechanisms, we should not abandon the quest for a unitary hypothesis. Although such frameworks may be simplified, the lack of a guiding paradigm may ultimately impede the development of investigation. (Mashour, 2004, p. 428)

Mashour proposes to use the concepts of cognitive binding and unbinding to provide a unitary mechanism. Cognitive binding is the phenomenon that despite the evidence that our brain subdivides perceptual processing into different modules, we experience our perceptions as unified. To explain this phenomenon several solutions have been suggested: binding by convergence, binding by assembly and binding by synchrony. Binding by convergence claims that higher-order neurons collects the information from various lower order neurons and binds them together. Binding by assembly replaces a single binding unit by a dynamic, self-organizing cell assembly, "neurons that fire together, wire together" (Mashour, 2004, p. 429). Binding by synchrony connects neurons in the temporal dimension, such as the 40-Hz γ-band oscillations. Mashour argues that these three mechanisms probably function in complex relation with each other. His key point however is that whatever the specific mechanism or combination of mechanisms for cognitive binding is, it is cognitive unbinding that causes general anaesthesia. Studies showed that the γ-band synchronization that occurs between cortices, in the hippocampus and in the temporal lobe, in conscious state is interrupted when anesthetized (John et al., 2001, John, 2001, Ma et al., 2002 and Uchida, et al., 2000). Evidence is found that isoflurane, an anaesthetic agent, interrupts the ability of neurons - that are located in the cortical region connected to visual pattern recognition - to bind the patterns from their component parts. The neurons were still responsive to the signals from component patterns, but they were unable to bind them into an integrated representation. Anaesthesia can be connected with the disruption of cognitive binding through convergence and synchrony. Although the specific details of the process are still the subject of research, for Mashour cognitive binding/unbinding provides a common mechanism that integrates processes at the cellular, system or global scale of the brain (2004, p. 431).

Like in the explanations based on the lipid theories, acts of mechanism unification based on cognitive binding/unbinding theory must leave out all the details about how the anaesthetic reaches the brain cells. What is included in the unitary mechanism that is proposed are neurons (entities) that signal to other neurons (an activity), respond to signal of other neuron (another activity) and can bind together in the three ways described above (i.e. they can organize themselves in certain ways). The crucial claims is that the capacity of self-organisation units that allows integrated representations is impeded by all anaesthetics.

This case study shows that an actual resemblance question – what are the common features in the different mechanisms that are responsible for the anaesthetic capacity of the different anaesthetic agents? – has been addressed by scientists over a long period. Even though the first theory that claimed to provide the answer – the lipid theory – was falsified, the question itself remains present in today's science. Mashour does not ask how a specific anaesthetic agent works, he wants to know what the common features of general anaesthesia mechanisms are, and finds them in the theory of cognitive binding/unbinding. The descriptions of mechanisms leave out many details (as is normal in case of mechanism unification) but the main ingredients of mechanistic explanations are present (entities, activities and organisation).

3.6 Limited attention for resemblance questions

An added bonus of the attempts to bring unification back into the picture, is that I have drawn attention to resemblance questions. The case studies in 3.3 and 3.5 show that resemblance questions are important in scientific practice. So philosophers should pay attention to them, and they have not done that sufficiently in the past. I elaborate on this issue in this section.

After giving several examples of why-questions about particular facts, Carl Hempel writes:

> [A]nd in that case the explanatory problem can again be expressed in the form 'Why is it the case that *p*?', where the place *of 'p'* is occupied by an empirical statement specifying the explanandum. Questions of this type will be called *explanation-seeking why-questions*. (1965, p. 334)

As is clear from the examples that Hempel gives before he offers this general characterisation, "p" can be a particular fact or a generalisation:

> A scientific explanation may be regarded as an answer to a why-question such as: 'Why do the planets move in elliptical orbits with the sun at one focus?', 'Why does the moon look much larger when it is near the horizon than when it is high in the sky?', 'Why did the television apparatus on Ranger VI fail?' 'Why are children of blue-eyed parents always blue-eyed?' 'Why did Hitler go to war against Russia?'. (1965, p. 334)

So he assumes that all explanation-seeking questions about particular facts and about regularities are, in our terminology, plain questions.

This assumption has been adopted by many philosophers of explanation after Hempel. For instance, Wesley Salmon (who only analyses explanations of facts, not regularities) consistently describes the explanandum as "the fact-to-be-explained" (see e.g. Salmon W. C., 1984, p. 13 and pp.15-19). And all his examples fit the simple format "Why is it the case that *p?*". The same goes for other prominent defenders of causal approaches to explanation (e.g. Cartwright, 1983; Humphreys, 1989 and Hausman, 1998).

Kitcher also adopts this view. He describes his aim as ...

> ... to determine the conditions under which an argument whose conclusion is *S* can be used to answer the question "Why is it the case that *S*?". (1981, p. 510)

Again, only plain questions are considered.

The examples I gave in 3.3 and 3.5 show that this assumption is not justified. Scientists sometimes aim at answering resemblance questions, i.e. questions about similarities between at least two events or at least two regularities. If we do not pay explicit attention to resemblance questions, an important type of explanatory practice is neglected in our philosophical analysis.

Contrastive questions (the remaining type I distinguished in 3.2) did receive a lot of attention from philosophers of explanation. This is due to the work of Bas van Fraassen (1980), who challenged the traditional assumption in the *opposite* way (compared to what I have done here). According to van Fraassen a good why-question is always contrastive, as in the examples given in 3.2:

> Why did John paint a portrait of the Queen, rather than a landscape?
>
> Why did John rather than Bill paint a portrait of the Queen?

Van Fraassen (1980) does not deny that people ask non-contrastive why-questions. For instance, it is possible that someone asks the following question:

> Why did John paint a portrait of the Queen?

However, van Fraassen claims that such questions are inaccurate expressions of the cognitive problem the person has. The real problem is captured by a contrastive question, for instance one of the two questions above.

Van Fraassen's view entails a strong claim: *all* non-contrastive questions which scientists ask are inaccurate formulations of contrastive questions for which they really want an answer. Even if one does not agree with this, one has to admit that van Fraassen did a good job by drawing attention to the fact that many why-questions that scientists ask are contrastive in nature.

There is a specific type of resemblance questions that did get a lot of attention from philosophers of explanation: questions about extremely improbable coincidences. Salmon gives the following example:

> Two students, Adams and Baker submit essentially identical term papers in a particular course. There is, of course, the logical possibility that the two papers were produced entirely independently, and that the resemblance between them is a matter of pure chance. Given the overwhelming improbability of this sort of coincidence, no one takes this suggestion seriously. Three reasonable explanatory hypotheses are available: (1) Baker copied from Adams, (2) Adams copied from Baker, or (3) both copied from a common source. (1984, p. 207)

In the cases I have considered, there is no such extremely improbable coincidence. The fact that there has been a social revolution in France does not make revolutions in other countries improbable. The fact that Peter has lung cancer does not make the fact that John has lung cancer improbable. So, the questions I have been dealing with in 3.3 differ from the ones Salmon and other philosophers have discussed, though these are also questions about resemblances. My questions are about sets of events that, taken together, are not improbable at all but nevertheless somehow raise an interest in what they may have in common. The same goes for the questions in 3.4 and 3.5. For instance, the fact that ether is an anaesthetic does not make the existence of other anaesthetics improbable.

As Salmon rightly says in the quote above, we do not take "separate causation" seriously as explanation of extremely improbable coincidences. In order to explain the joint occurrence, we assume that there is a causal connection between the events: a causal relation form the first event to the second, from the second to the first, or a common cause. This is *not* how I explain the resemblances in 3.3; I do not claim that there is a causal connection between John's and Peter's lung cancer or between the social revolutions. The resemblance questions I have discussed are answered by showing that the causes are of the same type. Questions about extremely improbable coincidence are answered by assuming a causal connection between the events to be explained.

In sum, philosophers of explanation have paid attention to a specific subclass of resemblance questions. I have discussed another type here. The way in which they have to be answered is completely different.

3.7 Conclusion

In this chapter I have argued that resemblance questions, although they appear to be a blind spot in philosophy of science, do occur in scientific practice and that unification is required to answer them successfully. I have demonstrated this by means of two examples: one from sociology and one from anaesthesiology.

In the case of resemblance questions about particular facts causal network unification is needed in order to answer them. In this type of unification one causal network is presented and applied to two or more events to point out the common causal factors that are present in the causal ancestry of all events we consider.

When resemblance questions about regularities are asked, mechanism unification is necessary to focus on the shared elements of the underlying mechanism. The act of mechanism unification consists in providing the richest possible model that is correct for capacity E in objects of types X, Y, Z, etc.

To capture my overall views on unification, it is important to note that - contrary to Kitcher – I do not claim that unification is necessary in every explanatory context. It is crucial in the context of resemblance questions. In Chapter 7, where I discuss explanatory pluralism, I come back to this.

Moreover, resemblance questions can liberate the idea of unification from its Kitcherian connotations. My concepts of causal network and mechanism unification are not derivational, and therefore allows one to see unification as valuable even if one rejects covering law conceptions of explanation. This topic will be further discussed in the next chapter.

Chapter 4
Unification: an enrichment for causal and mechanistic accounts

4.1 Introduction

In this chapter some theoretical-philosophical implications of my results will be clarified and elaborated. My strategy to search for cases of ontological unification in scientific practice, and to use them to rethink unification, resulted in a new possible answer to the following question:

> How do unification, causality and explanation relate to one another?

In order to clarify my answer to this question, I will start with an overview of different positions that have been taken by philosophers of explanation about the relation between these three concepts:

(1) Unification is always the aim of explanation. An explanation can be causal, but this is not the aim of the explanation.

(2) Explanation always needs to be causal. Unification can be a surplus value, but it is not the aim of the explanation.

(3) Unification and causality are both explanatory, but in a different way (and therefore complementary). What is the aim of an explanation is contextually determined.

(4) Unificationist explanation and causal explanation are possible types of explanations among other types. What is the aim of an explanation is contextually determined.

Position (3) is usually called the complementarity thesis, while position (4) is a pluralistic perspective.

In section 4.2 these four positions will be further elaborated. In section 4.3 I argue that, as a consequence of my views on unification, there is another possible relation between unification and causal explanation:

(5) Unification can be embedded as an additional aim in causal accounts of explanation.

In section 4.4 I will give two suggestions about how this embedding can look like in existing philosophical theories of explanation.

4.2 Unification, causality and explanation

Several positions about the relation between unification, causality and explanation have been defended:

A first position is that explanations always need to be unifying. Often this can (accidentally) be reached by causal explanation, but unification is the aim of the explanation. Kitcher, for example, sees no reason to require that explanations are causal. Problems for Hempel, like asymmetry and irrelevance, can be solved without invoking causation (1981, pp.522-525; 1989 pp.482-487). So, if explanations are causal, they are causal by accident.

A second position defends that explanations always need to be causal. In the past there were philosophers who explicitly stated that all explanations are causal, e.g. Wesley Salmon (1984) and David Lewis (1986). Today, philosophers are more careful. James Woodward (2003), for example, does not take this strong position it in 'Making Things Happen' and Michael Strevens does not affirm it in 'Depth' (2008). Woodward acknowledges that unification is 'unquestionably intuitively appealing' (Woodward, 2003, p. 358). However, 'considerations having to do with unification do not automatically pick out those derivations that are explanatory from those that are not' (Woodward, 2003, p. 361). So, derivations can be unifying by accident, but they are explanatory because they are causal.

A third position is the complementarity thesis. In this position the aim of an explanation depends on the explanation-seeking context. In some contexts, the aim will be causation, in others it will be unification. To reach full understanding of the world, both types of explanation are necessary. Salmon, the main defender of this position, writes in 1993:

> In this paper I have tried to show that there are at least two intellectual benefits that scientific explanations can confer upon us, namely, (1) a unified world-picture and insight into how various phenomena fit into that overall scheme, and (2) knowledge of how things in the world work, that is, of the mechanisms (often hidden) that produce the phenomena we want to understand. The first of these benefits is associated with the unification view of scientific explanation; Philip Kitcher (1989) is its present principal proponent. The second is associated with the causal/mechanical view of scientific explanation that I have advocated (Salmon, 1984). My current view is that the two accounts are by no means incompatible. (1993, p. 15)

This view is repeated in Salmon 1998 (p. 89-90). Salmon argues that explanation by unification and causal explanation both produce understanding, albeit a different type of understanding (1998, p. 78). If both types can be combined, this creates an enhanced understanding of the world we live in. Salmon refers back to the formulations of Friedman (1974, p.19) of global and local understanding, and Kitcher's similar distinction between 'top-down' and 'bottom-up' approaches to explanations (1985, p. 638). Salmon argues that understanding starts as a local activity: the first aim is to understand individual facts, and generalizations or laws are built upon that individual understanding. But, for Salmon, both are needed to grasp the world we live in.

A fourth position is explanatory pluralism, in which unificationist explanation and causal explanation are possible types of explanations among other types. Schurz mentions the expectability paradigm as a third approach to explanation next to causal explanation and unificationism (2014). Weber focuses on unification but he explicitly acknowledges that there are other types of explanation, e.g. causal ones (1999, p. 480). The most influential pluralist approach to explanation is from Bas van Fraassen (1980). Van Fraassen denies that there are fixed properties of explanation. What counts as a satisfactory answer to an explanation-seeking question in science depends on the context. An explanation is "salient to a given person because of his orientation, his interests, and various other peculiarities in the way he approaches or comes to know the problem – contextual factors" (van Fraasen, 1980, p. 125). There are no general criteria for explanations, it all depends on contextual factors.

Unification in the aforementioned positions is usually considered in its classical sense. I will argue in 4.3. that my views on unification lead to a fifth position about the relation between unification, causality and explanation.

4.3 My views on unification

In order to argue that my views on unification lead to a new position about the relation between unification, causation and explanation, I take two preparatory steps. First, I show how my results from Chapter 3 answer some questions about ontological unification that Mäki (2001) leaves unanswered (section 4.3.1). Second, I will show that my cases from the previous chapter constitute purely ontological unification, which is different from unification as conceived in the classical accounts presented in Chapter 1 (4.3.2).

4.3.1 Ontological unification

The aim of this dissertation is to bring unification back into the picture. In this first part I tried to do this by searching for examples of ontological unification in scientific practice. For the sake of clarity, I repeat Mäki's characterisation of ontological unification as seen in Chapter 1:

> Ontological unification is a matter of redescribing apparently independent and diverse phenomena as manifestations (outcomes, phases, forms, aspects) of one and the same small number of entities, powers, and processes. Those phenomena are thereby revealed to be only apparently independent; as a matter of actual fact, they are dependent on the same underlying structure of entities, forces, and processes. (2001, p. 498)

Although this characterisation is vague (cf. the many ontological categories occurring in it), the concept of ontological unification is a useful tool to draw attention to a broad range of possible unification-related scientific practices. This is, in my view, the main merit of Mäki's paper: drawing attention to the fact that unifying phenomena may mean that we show that they share the same ontic foundations (causes, origins, constituents).

Mäki's proposal leaves a number of important questions unanswered:

(a) What does ontological unification look like in scientific practice? What types of ontological unification are there? How do these types work?

(b) Why is ontological unification important? What kinds of questions does it address?

Causal network unification and mechanism unification are specific instances of ontological unification. The ontic foundations were respectively the common causal factors and the common features in causal mechanisms. These types of unification provide a partial answer to the questions mentioned in (a). This is not a complete answer, because other types of ontological unification will be presented in Part II of this dissertation. The fact that unification is a necessary ingredient for, at least, resemblance questions provides an answer to the questions in (b).

4.3.2 Derivational versus bottom-up unification

As we have seen in 1.4.2, Mäki characterises derivational unification as follows:

> Combining the two ideas that explanation is a matter of inference or derivation and that explanation involves unification of phenomena gives us the notion of derivational unification. (2001, p. 493)

Causal network unification does not presuppose that we derive the explananda that we want to unify from overarching premises. Neither does mechanism unification. So, the forms of unification that I have proposed in Chapter 3 are *purely ontological*: they are ontological and non-derivational. I will call this type of unification also *bottom-up* unification.

Because they are purely ontological, the types of unification I propose are different from what is proposed in the accounts of unification presented in Chapter 1:

(a) in the account of Weber unification is mixed (ontological plus derivational),

(b) in the accounts of Kitcher and Schurz unification is clearly derivational (while it is unclear whether they require an ontological component).

I will now elaborate on these two differences. In Weber's account (1999) unification happens if one can show that two (or more) events are instances of the same (set of) law(s). The idea of subsumption is very much present in this account. Unification is reached by constructing arguments in which the explananda are subsumed under the same law. So, we have derivational unification.

The unification act is based on the use of laws. This use of laws implies that unification has an ontological aspect. So, the proposal in Weber 1999 promotes a form of mixed unification: derivational and ontological at the same time.

As we have seen in Chapter 1, a main idea of Kitcher (which he shares with Hempel) is that, while all explanations are arguments, the converse is not true. He uses argument patterns to distinguish explanations from non-explanatory arguments. For an individual with knowledge *K,* an argument *A* can only be an explanation if it is acceptable relative to *K* (i.e. if the premises of *A* are members of *K*). But not all acceptable arguments are explanations: an acceptable argument is an explanation if and only if it instantiates an argument pattern that belongs to a privileged set of argument patterns. This set of argument patterns is privileged because it has a higher unifying power with respect to *K* than any other conceivable set of argument patterns. The unifying power of a set of argument patterns is determined by four factors: (i) it varies directly with the number of accepted sentences (i.e. the number of elements of *K*) that can be derived by means of acceptable arguments that instantiate a pattern in the set; (ii) it varies conversely with the number of patterns in the set; (iii) it varies directly with the stringency of its members; and (iv) it varies directly with the degree of similarity of its members.

I already mentioned a similarity between Hempel and Kitcher: explanations are arguments but not vice versa. Relatedly, there is a second similarity between Hempel and Kitcher: the basic act one has to perform to explain something is to construct an argument which shows that the explanandum is to be expected. The main difference is that, from Kitcher's perspective, there is a good and a bad way to do this. The good way is to use only arguments that belong to the explanatory store (the set of arguments that instantiates an argument pattern that belongs to the privileged set of argument patterns). The bad way is to use these plus other Hempelian DN explanations. The first strategy results in a unified knowledge system (according to Kitcher's account of unification) while, in Kitcher's view, the second strategy results in an uninteresting kind of expectability. In Kitcher's account, there are no ontological constraints imposed, so this may be purely derivational unification.

Unification in the account of Schurz (1994, 1999) may also be purely derivational. In Schurz' account unification is not an all-or-nothing characteristic, but is gradual. He compares different levels of unification according to their assimilation status. As I will argue, neither type of assimilation or non-assimilation can be categorized in the class of bottom-up unification, they are all instances of derivational unification.

An explanation for Schurz means that a cognitive state is reached in which the explanandum is inferred by a set of premises that are in less need of explanation, because they are unified. This unification is gradual because a set of premises can be more or less assimilated with a given theory. The actual assimilation can only happen if there is predictability or nomic expectability. This is exactly what happens in derivational unification. Potential assimilation consists of virtual or heuristic assimilation. Schurz speaks of virtual assimilation if an event is random without any further hidden cause and if there is a low probability inference Prem => Pa. Heuristic assimilation happens if phenomena are derived from theories with help of initial or boundary conditions, which are considered plausible. Both virtual and heuristic assimilation have a top-down direction. Dissimilated phenomena are anomalies, they cannot be assimilated within the accepted theory. These are the phenomena that are in need of explanation according to Schurz. The last assimilation status is the class of basic phenomena. According to Schurz they are the fundamental theories with which every fact or empirical law will assimilate or dissimilate.

Schurz cannot exclude merely formal constraints. Both in the actual and in potential (virtual or low probability) inference assimilation some sort of subsumption is needed, thus there is no room for non-derivational unification, such as bottom-up unification. One can be tempted to think that dissimilated phenomena are eligible for bottom-up unification, since they are not derivable from a theory. However, the goal for Schurz is to find a theory (or adapt a theory) so that the dissimilated phenomena are no longer dissimilated. In bottom-up unification we do not presuppose an accepted theory in which the phenomena have to fit, there is no theory presupposed at all.

The main difference between the accounts of Kitcher and Schurz on the one hand and my views on unification on the other hand is that, for them, the activity that is required for unification is deduction: an explanandum has to be derived from an argument pattern (Kitcher) or from a theory (Schurz). However, this does not necessarily mean that unification is purely derivational. The result of classical unification can be mixed: both derivational and ontological, but this is not a requirement for Kitcher or Schurz.

A major advantage of my view is that it allows us to think about the value of unification in contexts where the explanation does not have the form of an argument. The explanations that involve causal network or mechanism unification are not derivational. That is important, because it means that they are interesting even if one rejects covering law conceptions of explanation, such as Hempel's and Kitcher's. I developed tools that adherents of non-derivational views on explanation (such as the causal and mechanistic account) can use to think about the value of unification. This idea will be further elaborated in Section 4.4.

4.3.3 Consequences for the relation between explanatory unification and causal explanations

In 4.2 I showed that there are four positions taken about possible relations between unification, causality and explanation. Note that these positions all consider unification as *derivational* unification. Section 4.3 enables me to further specify these positions:

(1) Explanation always needs to be unifying. An explanation can be causal, but this is not the aim of the explanation.

(2) Explanation always needs to be causal. Unification can be a surplus value, but it is not the aim of the explanation.

(3) Unification and causality are both explanatory, but in a different way (and therefore complementary). What is the aim of an explanation is contextually determined.

(4) The pluralistic position: unification and causal explanation are possible types of explanations among other types. What is the aim of an explanation is contextually determined.

A consequence of considering unification as purely *ontological*, viz. non-derivational and based upon shared underlying causal processes and structures, is that another relation between unification and causal explanation is possible:

(5) Embedding: unification can be embedded as an additional aim in causal accounts of explanation.

Because causal accounts are ontological, such embedding is only possible if unification is also ontological. This is why I had to show that my views on unification are ontological

(4.3.1.). Furthermore, because causal explanations are non-derivational, the embedding of unification in those explanations is only possible if unification is considered non-derivationally (4.3.2).

4.4 Embedding unification in causal accounts of explanation

I will now make two concrete suggestions about how existing causal accounts of explanation can be optimized for answering resemblance questions by using unification as a second aim of a satisfactory explanation. In this way I can clarify how embedding will work in scientific practice.

4.4.1 Unification as reverse difference-maker in causal accounts

Michael Strevens' book *Depth. An Account of Scientific Explanation* (2008) contains an elaborate account of explanatory relevance. Strevens considers and rejects what he calls the "minimal causal account of event explanation". According to this account ...

> [A]n event is explained by whatever other events causally influence it, together with the laws and background conditions in virtue of which they do so. (Strevens, 2008, p. 41)

His argument against this view is:

> The most obvious difficulty facing the minimal causal account is the apparently unreasonable vastness of a complete causal explanation. As I pointed out above, in a quasi-Newtonian world like our own, an event's minimal explanation ought in principle to mention anything that has ever exerted a gravitational force on the objects involved in the event, anything that had previously exerted a force on these exerters, and so on. But all scientific explanations, even the most well regarded, describe much less than the complete causal history of the explanandum. (Strevens, 2008, p. 43)

Thus, the question arises: how do we decide which causal information to include in the explanation and which information to exclude? In order to answer this question, Strevens develops an optimizing procedure which is the core of his kairetic account in chapters 3 and 4 of his book.

> The kairetic theory provides a method for determining the aspects of a causal process that made a difference to the occurrence of a particular event. The essence of the theory is a procedure that does the following: given as input a causal model *M* for the production of an event *e*, the procedure yields as output another causal model for *e* that contains only elements in *M* that made a difference to the production of *e*. A model that contains explanatory irrelevancies is then "distilled" so that it contains only explanatorily relevant factors. (Strevens, 2008, p. 69)

The need for such a procedure has been recognised by many philosophers before Strevens, e.g. Peter Lipton:

> Suppose that my car is belching thick, black smoke. Wishing to correct the situation, I naturally ask why it is happening. Now imagine that God (or perhaps an evil genius) presents me with a full Deductive-Nomological explanation of the smoke. This may not be much help. The problem is that many of the causes of the smoke are also causes of the car's normal operation. Were I to eliminate one of these, I might only succeed in making the engine inoperable. By contrast, an explanation of why the car is smoking rather than running normally is far more likely to meet my diagnostic needs. (Lipton, 1993, p. 53)

Like Strevens, Lipton claims that a good explanation is one in which explanatorily irrelevant details are weeded out. Strevens provides a procedure for doing this in a systematic way[15].

An obvious move in optimisation is that we leave out factors that did not make a difference. Strevens calls this *elimination* of causal factors. However, Strevens also discusses the possibility of *abstraction*:

> I throw a cannonball at a window, and the window breaks. Does the fact that the cannonball weighs exactly 10 kg makes a difference to the window's breaking? The natural answer to this question is no. The fact that the cannonball is rather heavy

[15] Strevens claims that the criterion for identifying a model containing only the necessary difference-makers resembles unificationism's criterion for identifying the most unifying theory, such as abstractness, cohesion and generality (2004).

> made a difference, but the fact that it weighed in at exactly 10 kg did not. (Strevens, 2008, p. 96)

In such cases, the optimizing procedure ensures that the proposition "The ball's mass was 10 kg" is not in the explanation, but that e.g. the proposition "The ball's mass was greater than 1 kg" is in it. This is what Strevens calls abstraction (2008, p. 97). Abstraction is more refined than elimination because it is gradual, while elimination is all or nothing.

The desideratum of keeping only causal factors that make a difference is certainly plausible in cases where we want to answer a contrastive question, such as in Lipton's example. This is also what Strevens has in mind in his 2008 book, though he does not formulate the questions in a contrastive way.

In order to see how I can improve on this, we have to look at the issue of optimisation from a different perspective, viz. resemblance questions. What is needed for such questions is an *opposite* optimizing procedure: one that throws out the difference-makers. The social revolutions example of Chapter 3 (3.3) can illustrate this. The complete causal story about any of the three revolutions contains claims that are irrelevant for explaining the other revolutions because they point at differences rather than at common causal factors. For instance, what is usually called *the* Russian Revolution (October 1917) was preceded by a revolution in February which ended the regime of the tsars but was not a big social revolution. This information is irrelevant for answering the resemblance question. In order to explain the resemblance, the focus needs to be on the factors that were present in all three cases: external military/economic pressure and a strong sense of community. Other factors, the difference makers, have to be removed from the complete causal history.

In sum, my critique on Strevens is that his optimizing procedure is not universally valid. A second, opposite optimizing procedure is needed for contexts in which his procedure leads to the wrong result.

By embedding unification in Strevens' causal account, and using it as a guiding principle for a second optimizing procedure, his account can be applied to the context of resemblance questions.

4.4.2 Unification and optimising mechanistic explanations

Unification can also be embedded in a mechanistic account: it can be the guiding principle behind a procedure for optimising mechanistic explanations in certain contexts (viz. resemblance questions).

According to this account a mechanistic explanation for a phenomenon is a description of the mechanism that produces this phenomenon. Mechanisms are to be understood as follows:

> A mechanism for a phenomenon consists of entities and activities organized in such way that they are responsible for the phenomenon. (Illari & Williamson, 2012, p. 123)

Mechanisms not only explain why a phenomenon occurs, but also, by giving information about the entities, activities and organization, how the phenomenon occurs.

Within this mechanistic view on explanation, there is a group of philosophers who regard models that are more complete and more specific as superior to mechanistic models that are more abstract (e.g. Machamar, Darden, & Craver, 2000; Darden, 2006 and Craver, 2007). According to Machamer et al. (2000) abstract mechanistic models that omit details are not explanations, they are merely schemas that need to be filled with descriptions of parts and activities. Darden (2006) agrees and sees such abstract models as templates for explanations. Craver sees such a 'mechanisms sketch' as an incomplete model of a mechanism (2007, p.13).

Arnon Levy and William Bechtel are not convinced of this and focus on the merit of abstraction in mechanistic modelling (Bechtel, 2009 and Levy & Bechtel, 2013). Levy and Bechtel conclude that

> It is always possible and, we argue, often desirable to overlook the more concrete aspects of a system and represent its organization abstractly as a set of interconnections among its elements. Oftentimes such a detail-poor representation will be well suited for the explanatory purposes at hand. (2013, p. 255)

This is because

> biologists expect similarities among the mechanisms responsible for the same or similar phenomena in related organisms. (2009, p. 763)

Bechtel constructs an example about circadian rhythms to illustrate this (2009). The idea is that there is a 24-hour rhythm that affects body temperature, metabolism functions, mental functions and sexual activities in almost all living organisms. Instead of constructing a different mechanistic model for every type of organism, biologists assume that certain parts of the mechanism are conserved throughout different species.

From Chapter 3 it should be clear that I agree with Levy and Bechtel. The analysis of the case study about general anaesthesia provides an argument in favour of their claim that mechanisms which omit details can still be explanatory.

My results from Chapter 3 allow me to add two important elements to Levy and Bechtel's claims about the value of abstraction in mechanistic explanations.

First, the value of abstraction depends on the type of explanation-seeking questions. In their example, biologists ask a resemblance question, since they focus on similarities between different species. Levy and Bechtel argue convincingly that sometimes abstraction is necessary, but they do not specify why and under which conditions. My analysis of mechanism unification implies that abstraction is necessary when answering questions of the following form:

> What are the common features of the mechanisms that produce capacity E in objects of type X, Y, Z, ... ?

The question in the example of Levy and Bechtel can be rephrased as such a resemblance question:

> What are the common parts of the mechanisms that produce capacity C in species A and B?

Second, by clarifying the type of question that is asked in their example, it becomes clear why abstraction is a necessary requirement for the model to be explanatory. The explanation has to provide a mechanistic model and it has to be unifying. In order to meet this second condition of unification, some degree of abstraction is necessary.

In explanation-seeking contexts that demand both abstraction and a mechanist model, unification can be the guiding principle on the completeness/abstraction axis. The most optimal mechanistic explanation is the one that provides the richest possible mechanistic model that is correct for all types X, Y, Z,... Mechanism unification produces the most informative correct answer to such resemblance questions.

In this sense, the idea of unification can be embedded as an optimization strategy in accounts of mechanistic explanation.

4.5 Conclusion

By rethinking unification ontologically I have shown that there is a new position possible about the relation between unification, causation and explanation. Unification can be embedded in causal accounts of explanations. In 4.4 I have given two suggestions how unification can be an enrichment for causal accounts of explanations. My ontological conception of unification is not only closer to scientific practice, it can also help to improve philosophical theories on causal explanation.

This finalizes the first part of this dissertation: bringing unification back into the picture by searching for cases of unification in causal explanations.

Part II
Unification and non-causal explanations?

Introduction to part II

In the second part of my dissertation I continue my strategy of digging into scientific practice to find cases of ontological unification. In order to do so, I will again bracket the classical accounts of unification in Chapter 5 and Chapter 6. Of course, once the case studies of ontological unification are developed, a comparison with the philosophical literature becomes useful; I will do this in part III.

In Part II I add a second set of brackets. I distance myself from the dominant literature that all explanations must be causal. I will investigate whether there are cases of ontological unification in non-causal explanations.

In Chapter 5 optimality explanations in biology will be investigated. These are highly generalized explanations about population traits. Since a good optimality model needs to be applicable to the whole population and to future populations of a species, or even to a set of populations, the unifying aspect will not be difficult to uncover. What is at stake here is whether or not this is an example of a causal unifying explanation or a non-causal unifying explanation. In the case of optimality models, it will be argued that the explanation itself is not causal, but is based on physical dependency relations. These relations supervene on causal relations in the world. I call this physical dependency unification.

In Chapter 6 I will use an example of chemical systematization, better known as the periodic table, to show that something unifying is happening in the absence of causal relations. It will be argued that the QM model for the systematization of chemical elements can be explanatory, even in the absence of causal explanations. Furthermore, if one accepts the possibility of such a non-causal explanation, then a new kind of ontological unification becomes possible. I will call this *structural unification*.

Chapter 5
Unification in biological optimality explanations

5.1 Introduction

In this chapter I will focus on a type of explanation that is often used in evolutionary biology to *"investigate adaptations by representing the evolution of a particular phenotype as a function of the relative fitnesses of a set of possible trait values"* (Rice, 2012, p. 685). Explanations of this type are usually called optimality explanations. Elliott Sober has characterized them as equilibrium explanations (1983). Contrary to dynamic models, optimality models do not present the history of individual changes of a species that produced the equilibrium of a population. Instead, an optimality model gives an explanation as to why the current phenotypic trait is the optimal one as compared to other possible trait values. The idea is that natural selection will eliminate the non-optimal phenotypes from the population.

It is quite clear that a good optimality model needs to be applicable to the whole population and to future populations of that species. The phenotypic traits are considered at population level, the individual trajectory of how that trait is developed in individuals within that population is not relevant. In case of highly generalized patterns the model needs to be applicable to multiple populations now and in the future. In section 5.2 two examples of optimality models are given. The first example is about a system-specific resemblance question, the second example is about a highly generalized pattern. I will show that optimality models provide answers to resemblance questions. Similar to the resemblance questions in previous chapters, here too ontological unification will be required in order for such an answer to be satisfactory. My aim in this chapter is to explicate what kind of ontological unification is present in optimality models.

First, I will argue that optimality explanations are not an instance of previously discussed types of ontological unification in section 5.3. To give a positive argumentation for the claim that optimality models provide another type of ontological unification, some difficult issues must be tackled. In section 5.4 I will argue that optimality models are not causal explanations. In section 5.5 I argue that, despite the fact that they are non-causal, they do provide some kind of ontological (and derivational) unification. I will call this kind *physical dependency unification.* In section 5.6, I will show that these non-causal physical dependency relations used in optimality models supervene on existing causal relations.

5.2 Optimality explanations as answers to resemblance questions

In this section it will be clarified how an optimality explanation works by presenting two examples, one about dung flies (5.2.1) and one about the equilibrium of sex ratios (5.2.2). Although the term 'unification' is not mentioned in papers about optimality models I will show that they form answers to resemblance questions and therefore presuppose a form of ontological unification. Furthermore, the two examples illustrate two types of optimality models: models about system-specific phenomena and models about highly generalized patterns that are system-transcendent.

5.2.1 Dung flies: system-specific resemblance question

In evolutionary biology scientists use highly idealized mathematical models to explain the current state of populations. They do so by showing that the current state is the result of an optimization strategy, favored by natural selection. In this model optimization theory, a mathematical technique, is used to determine control variables and design variables. An optimality model will connect possible strategies to optimize the design variables, within the context-specific constraints and tradeoffs. After identifying all the components of the optimality model, the optimal strategy of the design variables can be deduced. The model shows why a certain state, or a certain variable is to be expected in a certain population or set of populations. (Rice, 2015)

In order to clarify how optimality models work, the example from Parker and Stuart on copulating time in dung flies (1976) will be used, as it is summarized by Collin Rice (2012). Apparently dung flies *(Scatophaga stercoraria)* copulate for 36 minutes on average (design value). Through experiments and observations Parker and Stuart found out that female dung flies mate with multiple males (context-specific constraint). If this occurs, the second male fertilizes around 80% of the eggs, while the first male only fertilizes around 20%. This explains why a male dung fly guards the female for some time before flying off to another mate. The average time a male dung fly spent on searching a mate and guarding the female after copulating was 156 minutes. The total cycle is 156 + c minutes, where c is the time spent copulating. Experiments showed that an increased copulating time resulted in a higher average number of fertilized eggs. However, there is a tradeoff: the time spent copulating one female, decreases the available time the male has to search other mates. Moreover, the return (number of fertilized eggs) gained from longer copulating time does not increase linearly, but becomes smaller and smaller over time, this is another tradeoff. Parker and Stuart use an optimization strategy to determine the maximal tradeoff between copulating time and the rate of fertilized eggs across several cycles. They calculated that the optimal copulation time is 41 minutes, close to the observed value of 36 minutes.

The model that Parker and Stuart built provides accurate predictions based on empirically observed parameters. According to Elliot Sober it captures the major constraints and tradeoffs involved in the shaping by natural selection of this behavior. It is an adaptationist explanation of a single characteristic in a single species, since the model assumes that natural selection will maximize the optimization criterion. (Sober E. , 2000, p. 135). Sober sees two positive features in their model. First, Parker and Stuart did not just use 'fitness' as a criterion, but they stated a very specific criterion of optimality: maximizing the number of eggs fertilized per time unit. Second, they measured exactly what the investment returns are and were able to make accurate predictions. Rice emphasizes another feature of the model: the explanation uses a number of idealizing assumptions, such as an infinite population, randomized mating, the inheritance of copulating strategies. By idealizing these factors, it is assumed that other evolutionary factors would not keep the population of dung flies from reaching this optimal strategy (2012, pp. 688-689). The strategy that optimizes the criterion, here the amount for fertilized eggs, is the equilibrium point of the evolving population. Rice calls it an equilibrium explanation that shows why a particular strategy is the best available solution:

> Once the strategy set and the optimization criterion have been identified, an optimality model describes an objective function, which connects each possible strategy to values of the design variable(s) to be optimized. [...] Certain trade-offs and context-specific limitations will constrain the optimal design. [...] Once these components are specified, one can deduce which of the available strategies will yield the optimal value(s) of the design variable(s). The strategy that optimizes the model's criterion, in light of various constraints and trade-offs, is deemed the optimal strategy. By mathematically representing the important constraints and trade-offs, an optimality model can demonstrate why a particular strategy is the best available solution. (Rice, 2015, pp. 591-592)

Now, let us take a closer look at the explanation-seeking question in the example of the dung flies. Starting from the observation that dung flies copulate for approximately 36 minutes, we can ask contrastive questions and a resemblance question. An example of a contrastive question is:

> Why do dung flies copulate for approximately 36 minutes instead of 5 minutes or 1 hour?

The resemblance question is:

> Why do *all* dung flies copulate for approximately 36 minutes?

This resemblance question asks why there is so little variation. What is at stake is why do they all behave *more or less the same*?

In the model of Parker and Stuart both the contrastive and the resemblance question are answered. My focus will be on the resemblance question. A satisfactory answer to those questions needs to show that phenomena are dependent on the same deeper, more fundamental structure or process, in this case the biological and environmental factors that influence the copulating behavior of dung flies. This means that ontological unification will be required to provide a good answer to this resemblance question.

In section 5.3 of this chapter I will investigate whether the ontological unification here is similar to the previously discussed types of ontological unification. If not, another conception of unification needs to be constructed for this type of explanations.

5.2.2 Sex ratios: resemblance questions about highly generalized patterns

The example used in section 2.1 is an optimality explanation of a system-specific phenomenon. Optimality models are also used to explain highly general patterns. In order to show how optimality models work on this higher level, Fisher's model of equilibrium sex ratios (1930) is used as summarized by Rice (2015).

Fisher asked a very basic biological question: why is the sex ratio in populations often 1:1? For my argument it is not necessary to elaborate on all the specific, mathematical procedures that are used in the construction of this model. The optimizing criterion in Fisher's model is the amount of mating opportunities, thus he assumes that natural selection will favor strategies that optimize this criterion. If there is an imbalance between sex ratios, producing offspring of the minority sex is the best way to propagate one's genes. There will be an advantage for parents who produce the minority sex, since their children have more mating opportunities, and thus more expected offspring. The births of the minority sex will become more common in the population, until there is a sex ratio of 1:1 and the fitness advantage fades. The reasoning applies both to female and male sex as the minority sex. A 1:1 sex ratio is the stable equilibrium state of a population, since it is the only state in which selection does not favor the production of the minority sex and thus is an evolutionarily stable strategy. In this model it is important to see that there is a perfectly linear tradeoff between the ability and the cost to produce sons and daughters. The resource cost of raising a daughter is the same as raising a son. One fewer son, means one more daughter, and vice versa.

As in the previous example this model also makes idealizing assumptions such as randomized mating, an infinite population, equal access to resources etc. Furthermore, the model assumes that other evolutionary factors will not influence the population from reaching this equilibrium, namely a 1:1 sex ratio.

Starting from the observed 1:1 sex ratios, several contrastive questions can be asked. For instance:

> Why is the sex ratio in most populations 1:1 instead of 1:2?

The explanation-seeking question that Fisher addressed is a resemblance question:

> Why do *all* populations have a 1:1 sex ratio?

In other words: why is the sex ratio *the same* for all populations, regardless of the species to which they belong?

This type of explanation-seeking question focuses on similarities between two or more populations, rather than on differences (as contrastive questions do). Just as in the example with the dung flies the optimality model here answers the resemblance question as well as contrastive questions. The difference is that while the optimality model about dung flies is a question about particular facts, the optimality model about sex ratio is an attempt to answer a higher-level resemblance question about regularities.

5.3 Optimality models and previously discussed types of ontological unification

If optimality models would be instances of causal network unification or mechanism unification, my task of explicating the kind of ontological unification that is involved here, would be rather easy. However, as will be shown in sections 5.3.1 and 5.3.2, this easy route does not work.

5.3.1 No causal network unification

Consider the following resemblance questions:

i. Why do *all* dung flies copulate for approximately 36 minutes?

ii. Why do *all* populations have a 1:1 sex ratio?

The general format of these questions is the following:

I. Why do all objects of population *P* have property X?

II. Why do all populations have property X?

In order to provide causal network unification (as in Chapter 3), positive or negative causal factors need to be determined that are present in the causal ancestry of all objects (organisms or populations). Remember the general format of an explanation that results from an act of causal network unification:

X is a positive causal factor for E.

Y is a positive causal factor for E.

Z is a positive causal factor for E.

...

Object a has X, Y, Z, ...

Object b has X, Y, Z, ...

Object c has X, Y, Z ...

...

Thus, objects a, b and c have property E.

The causal factors X, Y and Z are present in the causal ancestry of the objects. This means that these factors make a difference and if one intervenes on a cause X it would lead to a change in the effect Y. In order to show that these models are no causal network unification, I turn to the argument of Collin Rice.

Rice (2012, 2015) firmly disagrees with Angela Potochnik (2007, 2010) and other philosophers who see optimality explanations as a special type of causal explanations. He gives two arguments for his position, a bad one and a good one. I will use both, but not in the same way. The first argument is useful for my claim that optimality models are not causal network unification. The second argument will be used in section 5.4. to show that they are not causal.

Rice's first argument states that optimality explanations are not causal explanations because they are synchronic instead of diachronic. Optimality models are a separate kind of explanations since they are based on optimization theory, a mathematical technique to "determine what values of some control variable(s) will – given a set of tradeoffs and constraints – optimize the value of some design variable(s)" (2012, pp. 695-696). Optimality models do not focus on specific parts of causal processes, but they are based on a completely different set of relationships: population-level constraints and tradeoffs. To create such a model a lot of causal information is censored: "all information about step-by-step dynamics of the evolving system" is omitted (2012, p. 698). Besides censoring, there is another feature of optimality models: the essential use of idealization. Removing idealizations, such as infinite populations, randomized mating, a constant payoff structure and the exact copying of parent's phenotypic strategies by their offspring, would eliminate the explanatory value of the optimality models. These models focus on population-level relations that are synchronically represented and leave all the dynamics out of the explanation. Therefore, Rice argues, they are non-causal explanations.

> In representing a trade-off, a biological optimality model does not reference any causal processes of the biological population or any events within the population's causal history [...] Nowhere does the model describe a causal process (or causal trajectory) *that unfolds over time or any events that occur prior to the explanandum*. (Rice, 2012, p. 699; italics added)

Rice is too quick to classify optimality explanations as non-causal because of the synchronic relations they model. His argument rests on the arbitrary demand about the temporal asymmetry: in a causal explanation the explanans must occur before the explanandum. Let us compare this to Woodward's notion of explanations:

> I suggest below that the distinguishing feature of causal explanations, so conceived, is that they are explanations that furnish information that is potentially relevant to manipulation and control: they tell us how, if we were able to change the value of one or more variables, we could change the value of other variables. (2003, p. 6)

Nowhere does Woodward demand that causal explanations need to be dynamic. Causal explanations are causal because they provide information about the conditions under which the phenomenon would be different. So, Rice's first argument is not cogent with respect to the claim that optimality models are non-causal explanations. I come back to this in section 5.4.

However, Rice's argument is useful in determining whether or not optimality models are an instance of causal network unification. The fact that optimality models describe synchronic relations implies that they do not give information about the causal ancestry of the phenomenon. No causal process or causal trajectory is described. Causal network unification was characterized in a way that excludes synchronicity: in causal network unification "one causal network is presented and applied to two or more events to point out the common causal factors that are present in the causal ancestry of all events we consider" (section 3.7). Because of the synchronic nature of optimality models, they do not meet this requirement and are not an instance of causal network unification. The notion of causal ancestry is crucial in the latter.

5.3.2 No mechanism unification

Let us look back at the resemblance questions from 5.2 (and rephrased in 5.3.1). If we treat these questions as a request for mechanism unification, we get:

> What are the common features of the mechanism which makes dung flies copulate for approximately 36 minutes?
>
> What are the common features of the mechanism which makes sure that in almost all populations the sex ratio is 1:1?

In order to decide if the components of an optimality model can function as features of a mechanism, I use Stuart Glennan's distinction between causal production and causal relevance. This difference arises in the discussion whether or not natural selection is a cause, and thus eligible to be used as a causal explanation.

Optimality explanations do not focus on population-genetic mechanisms but on the determination of evolutionary stable phenotypes. Orzack and Sober define this evolutionary stable strategy (ESS) as:

> A phenotype of an individual is optimal (relative to a variety of alternatives) because it outperforms the other phenotypes and thereby results in a higher fitness. [...] As a result, other phenotypes are eliminated from the population (or nearly so) or prevented from invading. (1994, p. 3)

The idea is that natural selection must lead to the occurrence of an evolutionary stable strategy (ESS) observable in a stable phenotype, such as a 1:1 sex ratio or an average copulation time of 36 minutes in the dung fly population. But does natural selection really 'cause' this ESS?

There are two opponent views in this matter. The dynamical view sees selection as a force that causes evolutionary change. In this view population level causation is defended. The statistical view claims that natural selection is epiphenomenal and merely a statistical outcome of individual causal processes. Glennan tackles the discussion by bringing up a distinction between a causal process and causal relevance (2009, p. 327). Causal productivity is a relation between events. Glennan illustrates how this definition of causality is used in the literature of mechanisms and causal processes. In this meaning causality is about events producing another event, about objects doing something. He gives some examples:

The bowling ball knocked over the pin.

The explosion made Edward deaf.

The firing of neuron A caused the firing of neuron B. (2009, p. 327)

Causal relevance is a relation of dependence between a fact (or an event) and an event, and it is closely connected to counterfactual approaches to causation as can be found in Woodward 2003. Causal relevance solves problems with causation by omission or with causality based on properties. For instance:

If I had turned off the shower faucet, the bathroom would not be flooded.

The fact that the soccer ball weighs around 450g is relevant for it breaking the window.

With this distinction in mind, Glennan argues that natural selection does not cause the higher frequency of a trait in a population, since a population or a population level property is not causally productive. Populations do consist of parts, but that is not enough to consider them as objects or entities. The criteria to be considered as an entity is to have parts with a stable structure that can engage in an activity as a unified entity. Glennan illustrates this with a fish and a water bug. “When a fish kills a water bug, it kills the whole water bug, it cannot kill its legs but not its body. On the other hand, when a fish kills a water bug, it does not kill the whole population of water bugs.” (2009, p.333).

Glennan saves natural selection from epiphenomenalism by showing that population level traits can be causally relevant without being causally productive. When biologists invoke fitness to explain the higher frequency of a specific phenotype in a population, it is not the reproductive rate of the phenotype in itself that is causally relevant, but the ratio of the reproductive rate of that phenotype compared to that of variants. This is a population level property; thus, causal relevance is not limited to the individuals that make up the population.

In my view Glennan gives good reasons to accept natural selection and population level properties as causally relevant. But what is important for me here, is that in his line of reasoning he also gives me an argument not to accept population-level properties as being eligible as explananda for mechanistic explanations. Glennan gives the following definition of a mechanism, which is very similar to the one Illari and Williamson have given:

> A mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon. (2017, p. 17)

From this definition it follows that only higher-level properties of systems that are produced by the activities and interactions of lower-level entities can be given a mechanistic explanation. In Glennan's view biological populations do not satisfy this condition:

> This is because populations are not typically entities that enter into productive causal relations. (Glennan, 2009, p.335).

A population is not a system, it is a logical aggregation. Therefore, there is no mechanism that produces the properties of a population:

> If we manipulate a population level property like the relative frequency of a frequency dependent trait, we will have a causal influence on selective outcomes. Because a population is a logical aggregation of rather than a causal product of the individuals of which it is composed, when one manipulates a property of the population one ipso facto manipulates the properties of individuals within the population. (Glennan, 2009, p. 336)

In sum, optimality models operate at the population level, which cannot be seen as a system in the 'mechanism' sense; thus, these models are not instances of mechanism unification because they do not constitute mechanistic explanations.

5.3.3 Preview

In this section I argued that optimality models are not instances of causal network or mechanism unification. There is no easy way to clarify how ontological unification operates in these optimality models. In the remainder of this chapter I take three steps in order to get grip on the kind of unification used here. First, I claim that the explanations provided by optimality models are not causal (5.4). Second, I investigate whether they are explanatory at all. I will claim that they provide explanations because they describe shared physical dependency relations. These relations make ontological unification possible in these models. I will call this *physical dependency unification* (5.5). Finally, I will further clarify this type of unification by arguing that it supervenes on causal relations (5.6).

5.4 Are optimality models causal?

There has been a debate about whether or not optimality models are causal explanations between Collin Rice (2012, 2015) and Angela Potochnik (2007, 2010). This debate will be a thread throughout this and the following sections. In a sense, they are both right and wrong. This will become clear after showing how unification operates in optimality models. But first things first, are these models causal or not?

Rice gives two arguments to claim that they are not. I have discussed the first argument in section 5.3.1 and showed that it is not cogent to argue that optimality models are not causal, but it was sufficient to show that they do not use causal network unification.

His second argument is more convincing: trade-offs are not causes; thus, they cannot provide causal explanations. Let us take a step back and first show why these trade-offs are so important. In 5.3.1. we saw that one of the key features of optimality models is their degree of idealization and generalization. A lot of causal information is censored. The core of the model is not the representing of causal processes but representing population-level constraints and tradeoffs. It is the relation between trade-offs that explains why the current state of a population is as it is. In the case of the dung flies there were two trade-offs: the time spent copulating one female decreases the available time the male has to search other mates and the number of fertilized eggs gained from longer copulating time does not increase linearly, but becomes smaller and smaller over time. This is also the case in the example of the sex ratios. The particular initial conditions and the causal trajectory of the system are not important in the optimality model. The causal trajectory of how a parent produces a child is irrelevant here. It is equally unimportant what causal factors play a role in choosing a specific mate. What does matter is that the cost for raising a child is independent of its sex, that mating opportunities are better for offspring of the minority sex, that members of the minority sex will produce more offspring of that same sex since they have better mating opportunities. By mathematically representing these trade-offs and context-specific constraints, it is possible to deduce from an optimality model why a particular strategy is the best available solution (Rice, 2015, p. 592).

It is difficult to accept trade-offs as causes. First, the representation of tradeoffs is inaccurate if they are compared with the actual mechanisms that are active *within* the population. For instance, in the model it is assumed that trade-offs are constant. Second, trade-offs are relationships between values. It is indeed difficult to see how 'average fertilization rates' cause 'copulating time' or how the trade-off between population-level averages is a cause for the equilibrium state that is the explanandum. That the relation between trade-offs and the equilibrium state is non-causal can be related to Glennan's claim that populations are logical aggregations. Yes, it is true that trade-off values can be different, but it is not possible to manipulate a trade-off directly. Therefore, optimality models cannot be captured by an interventionist account. The idea that an optimality model is modular is only an illusion according to Rice:

> I maintain that the apparent modularity of these relationships within the idealized model is merely an illusion created by the use of several idealizations and abstractions that eliminate the complexity of the causal networks of real-world biological systems. Modular relationships within an idealized mathematical model cannot establish modularity of causal relationships in the model's target system(s). (Rice, 2015, p. 605)

The only manipulation that is possible is on the level of the individual that make up the population. In sum, I agree with Rice that the core relations that make up the optimality model are non-causal, even if they are in a way counterfactual. However, I disagree with his second conclusion: "the explanatory claim and the causal claim are independent of one another" (Rice, 2015, p. 605). I will explain this in section 5.6.

An obvious question that arises now is: are optimality models explanatory at all given that they are not causal. I will investigate this in the next section.

5.5 Optimality explanations: shared physical dependency relations

That optimality models do not use causes is not an argument in itself that they are non-causal explanations: one can also deny that they are adequate explanations at all. For instance, Michael Strevens (2008, p.288) argues that optimality models give us minimal causal information: the actual causal history must be within the set of possible causal trajectories, limited by context-specific constraints and trade-offs. But the actual causal information is 'black-boxed'. This implies that optimality models are at best only partial explanations:

> Because a model that secretes some mechanisms in its explanatory framework does not confer what I called [...] 'deep' of 'full' understanding of the phenomenon that constitutes its target, the black-boxing model is limited in its explanatory power. A deep explanation of the ecosystem's stability must flesh out the model's black boxes rather than leaving the causal details in the framework. (Strevens, 2008, p. 159)

The problem for Strevens is that the features of optimality models are discretely multiply realizable kinds. There is no single, stable explanation across all populations, even if there is a single high-level structure that is shared among the populations. His conclusion is even more fare-reaching:

> The sense in which the black-boxing model explains the stability of a wide range of ecosystems is at best partial, then: the model does not itself explain stability in each such system; it rather provides the schema for the individual, case-by-case explanations. (Strevens, 2008, p. 160)

Strevens' position contradicts my approach. My analysis starts by considering how these models are used by scientists: as explanations. Optimality models are not temporary explanations or steps towards an explanation. They are the best-suited explanations for the task at hand: providing highly idealized equilibrium explanations, or in my terms: providing answers to resemblance questions.

Woodward's view is more in line with scientific practice here. His notion of explaining is broader than causal explanations:

> An explanation must answer a what-if-things-had-been-different question, or exhibit information about a pattern of dependency. (2003, p. 201)

This means that if I can show that optimality models exhibit information about dependency relations, they are explanations. Woodward uses the criterion of mirroring physical dependency relations to distinguish explanatory derivations from non-explanatory derivations:

> The idea is that these derivations trace or mirror the relations of physical dependency that hold between the explanans conditions and the explananda phenomena-relations that would be revealed if, for example, we were to physically intervene to alter the explanans conditions. (2003, p. 201)

If I can show that optimality models mirror such physical dependency relations, they can be classified as explanatory according to Woodward's view.

In order to do so, I will use Weber, Van Eck and Mennes' interesting analysis of the epistemic value of biological ascriptions, or functional explanations:

> Biological advantage ascriptions are valuable because they provide the means for answering questions of the following form: "What would happen if (due to mutation) in some individuals of species *s* item i' would have a different property e' (while the habitat remains unchanged)?" (201+)

Their conclusion is that functional explanations provide answers to questions of the form: *what would happen if* a disturbance occurred. If such a disruption would occur, similar causal processes and mechanisms would be activated in the individuals that make up the population (or set of populations).

Let me explain this by comparing it, as Weber, Van Eck and Mennes do, with Woodward's contrastive account (201+).

For Woodward an explanation must show how the explanandum would change if the initial conditions were different. In other words, adequate explanations

> … locate their explananda within a space of alternative possibilities and show us how which of these alternatives is realized systematically depends on the conditions cited in the explanans. They do this by enabling us to see how, if these initial conditions had been different or had changed in various ways, various of these alternatives would have been realized instead (2003, p. 191).

In Woodward's view an explanation must give an answer to the question: what if things had been different? What if we could intervene in the relevant causal factors of the explanandum and change one or more factors, would the explanandum be different or not? In order to give a satisfactory answer to that question, counterfactual dependence must be established.

Functional explanations not only look back or focus on the present, but they are also prospective. They give information about what might happen in the future if the current state of affairs would change. Biological ascriptions formulate reasons to prefer one theoretical possibility above others. In this sense, functional explanations are not causal explanations, since they do not rely on reconstructing causal processes or causal relevance from the past to the present. Instead functional explanations use theoretical possibilities to argue how future states of affairs might be if the current state is disrupted. Optimality models work in a similar way. In order to explain the current state they do not reconstruct a causal ancestry, but they show that future states may be different if trade-offs are changed due to genetic mutations or environmental factors.

The explanation in optimality models is based on structural relationships between trade-offs and constraints. Optimality explanations tell us *what would happen if* the relationship between those tradeoffs and constraints changes, e.g. because of a change in the habitat of the population or because of a genetic mutation. By focusing on these structural relationships optimality models can answer forward-looking resemblance questions of the type:

> What would happen to phenotypic trait X of a population (or a set of populations) if the current state of affairs (environmental or genetic) would change? What would happen if a disruption occurs that disturbs the current evolutionary equilibrium?

An optimality model can answer such questions because it exhibits the dependency relations between context-specific constraints, trade-offs and the state of a population, even if these relations are not causal. The deductive structure of the optimality model mirrors counterfactual physical dependency relations, and so, they qualify as explanations. The explanatory unification that occurs in these models will be called *physical dependency unification*.

5.6 Physical dependence unification supervenes on causal relations in the world

In this final section I will do something that I have tried to avoid until now in this dissertation: making metaphysical claims. This is where Angela Potochnik's role in the debate about the status of optimality models comes in handy (2007, 2010). Her perspective on the causal aspect of optimality explanations is interesting, but she uses it in a wrong way. I will show that Potochnik and Rice are arguing about different issues. In their debate epistemological and metaphysical claims are intertwined with one another. This will be important for the further characterisation of my notion of physical dependency unification.

5.6.1 Ontological supervenience

Rice argued that optimality models are not-causal, since they are purely structural, highly idealized explanations. However, there is a difference between idealizing and denying the existence of causal trajectories. Rice too speaks of causal information:

> The actual strategies are, however, causally relevant to the evolutionary process that occurred. This causal information, however, is explicitly not included in the explanation provided by the optimality model. (2015, p. 599)

To clarify this, Potochnik's argument that optimality models are causal explanations can help. It is true that optimality models use highly simplified assumptions instead of complex dynamics , but that does not mean that they can use any simple assumption:

> Optimality models are epistemically dependent on unrepresented features of genetic transmission. The success of the simplifying assumptions that optimality models use to stand in for genetic dynamics must be established on a case-by-case basis. (Potochnik, 2010, p. 226)

An optimality model can only be successful if the model parameters and background assumptions are confirmed by other research. If this were not the case, an optimality model would be an example of purely derivational unification.

My views on the metaphysical underpinnings of optimality models is as follows. It is true that the explanandum is *derived* from the optimality model, but the dependency relations in the model (the trade-offs) are based on causal relations in the world. This implies that optimality models do not presuppose some 'spooky' metaphysics in which there would be physical dependency relations in the world that do not supervene on causal relations.

The causal relations on which the physical dependency relations in the model supervene do not explain the equilibrium, but they are relevant for the model. The non-causal physical dependency relations between populations, population level traits, the environment, fitness and selection trail supervene on existing causal relations in the world. Epistemologically speaking, optimality models are non-causal, in the sense that their explanans does not contain causes of the explanandum (5.4). Ontologically speaking, they are causal, because the physical dependency relations that make the model explanatory, supervene on causal relations in the world.

Potochnik attempts to construct an example in support of her claim that optimality models provide causal explanations. I will now show that her example is in fact an argument for my views on the metaphysical underpinnings of optimality models.

5.6.2 An example: the reproduction of bacteriophages

Potochnik's example is about the reproduction of bacteriophages by lysis. Viruses that infect bacteria (bacteriophages) reproduce by a process, called lysis, whereby the virus multiplies itself inside the infected bacterium. In a next stage the viral particles burst out into the environment and kill the host bacterium in the process. Wang et al. (1996) and Bull et al. (2004) examined the different stages of phage lysis and found a fitness trade-off between early and late lysis. When bursting out early the new viral particles can start reproducing and the time between generations is shorter. When the lysis happens later more particles are accumulated in the host cell. The optimal lysis time depends on the length of the different stages in the life cycle. The optimality model of Wang et al. (1996) and Bull et al. (2004) breaks down the life of a bacteriophage in three phases:

1) Dispersal phase: when a viral particle is being released from the host until it infects a new host.
2) Juvenile phase: when the bacteriophage infects a new host until the first progenies have matured inside the host.
3) Adult phase: when the first progenies have matured until the host is lysed.

The longer phases (1) and (2) take, the longer lysis should be delayed, according to the optimality model. This model is an instance of a highly generalized pattern: organisms with longer life cycles benefit from delaying the age at which reproduction begins (Potochnik, 2010, p. 219). There are different dynamics of lysis timing in different phage populations, but the selection pressure is the same, therefore

> any model of lysis timing that incorporated genetic dynamics would apply only to a narrow range of bacteriophage populations. Such an explanation would obscure the key determiner of bacteriophage lysis timing: the time it takes to reproduce is balanced against the number of progeny in a way that maximizes reproductive potential. (Potochnik, 2010, p. 220)

The specific causal trajectories are left out of the optimality model, since it focuses on a different set of relationships: population-level constraints and tradeoffs.

This does not mean that the causal information is completely irrelevant. It black-boxed. For instance, optimality explanations determine traits as heritable without characterizing the actual genetic causes. But the success of these optimality models is dependent on what is inside the black-box. Potochnik calls this epistemic interdependence (Potochnik, 2010, p. 226 etseq.). When Bull et al. (2004) assess the model developed by Wang et al. (1996), they do so by using genetic information. There are two assumptions in the model: first, the only reproductive option for phages is lysis and second, lysis timing can vary without changing other traits of the bacteriophage. Both assumptions are tested and corroborated by digging deeper into the genetic system of the bacteriophages.[16] So, even though the optimality model in itself is not a causal explanation, it is dependent on genetics and its causal relations.

This analysis is in accordance with Glennan's position. Optimality explanations do not refer to detailed biographies of organisms, but

> while these biographies may be unnecessary for our explanation, it is these organisms and their life histories that produce change. Without the individual organisms, there are no populations, and without the activities and interactions of these individual organisms, there are no changes in populations. (2009, p. 338)

This is exactly what is needed to make an optimality model not 'spooky': the population level traits supervene on existing causal relations at the level of the individual organisms. While Rice focusses on the non-causal features of optimality models on the epistemological level, the arguments that Potochnik gives are relevant at the ontological level. In their papers they intertwine epistemological and ontological claims.

[16] For more detail, see Bull et al. (2004) and Potochnik (2010).

5.7 Conclusion

Now I have all the elements to characterize the kind of ontological unification that is present in optimality explanations. I have called this *physical dependency unification.* This kind of unification has three features:

(1) It does not constitute a causal explanation.
(2) It explains by mirroring physical dependency relations in the world.
(3) It is not metaphysically spooky, since the physical dependency relations supervene on causal relations in the world.

As in the other types, physical dependency unification offers a way to answer a resemblance question. Contrary to the other types, it does not figure in a causal explanation. The difficulty here was that populations, population level traits etc. could not be treated as causes. However, optimality models provide information about physical dependency relations (albeit at a highly generalized level) and thus they are explanatory: they tell us how tradeoffs relate to one another and how they result in an evolutionary stable state in a population. There are no references to causal processes *in the model.* The model is constructed by applying mathematical optimization procedures to physical dependency relations. Comparable to functional explanations, optimality models do not rely on reconstructing causal processes or causal relevance from past to present. On the other hand, the unifying power of optimality models is not purely derivational, it still is ontological unification because it refers back to physical dependency relations that exist in the world. How those relations arise, is black-boxed in the model.

Thus, epistemologically speaking, optimality models form non-causal explanations, in the sense that their explanans does not contain causes of the explanandum. Nevertheless, the success of the model is based on an epistemic dependency on ontological causal factors such as genetic information, environmental information etc. Throughout this dissertation I have taken on Woodward's view that an explanation is about physical dependency relations. This commits me to a certain form of realism about explanatory relationships: explaining is more than exposing a logical or deductive relation:

> Derivational relations do not have a role to play in explanation that is independent or prior to such dependency relations, but rather matter only insofar as (or to the extent that) they correctly represent such relationships. (Woodward, 2003, p. 202)

Metaphysically speaking, this kind of unification is still causal, because the physical dependency relations that make the unification explanatory, supervene on causal relations in the world. This means that no 'spooky' metaphysics are presupposed.

Therefore, physical dependency unification is ontological in two ways: (1) it is not purely derivational, because it is based on physical dependency relations; (2) it commits me to a minimal realistic position about the ontology of explanations.

In the next chapter I will investigate whether there exists a kind of ontological unification that does not supervene on causal relations in the world.

Chapter 6
Systematization of chemical elements

6.1 Introduction

The periodic table is one of the best-known systems of classification in science. A periodic table is a diagrammatic representation of all the known chemical elements ordered according to their atomic numbers. The Russian chemist Dimitri Mendeleev stated in a presentation to the Russian Chemical Society in 1869 that there was a dependency between the chemical properties and the atomic weights of the elements. If the chemical elements are ordered according to the atomic weight the chemical properties of elements appeared to reoccur after certain definite intervals: the properties exhibit a periodicity. Since 1869, over 700 different versions of the periodic table were published (Scerri E. , 2007, p. 20).

As we will see further in this introductory section, contemporary versions of the periodic table differ in important respects from those in the 19th century. However, they share an important property: they encode information about which elements display similar chemical behavior.

My first aim in this chapter is to show that this property of all periodic tables implies that they suggest explanation-seeking questions of a type that is familiar from Part I and Chapter 5 of this dissertation: resemblance questions (section 6.2).

Quantum mechanical models of the behavior of electrons (which describe an electron configuration for each kind of atom) can provide explanations in response to these resemblance questions. These explanations are presented in section 6.3. Given what we have established in Part I, it makes sense to ask the following questions:

a) Do these explanations provide causal network unification as discussed in Chapter 3?

b) Do these explanations provide mechanism unification as discussed in Chapter 3?

The second and third aim in this chapter is to argue that the answers to these questions is negative (sections 6.4 and 6.5).

Given these negative answers, the question arises: are the QM models explanatory if they are non-causal? In section 6.6 I will argue that they are.

The fourth aim is to argue that there may be a type of ontological unification that is based on physical dependency relations that do not supervene on causal relations. I will call this structural unification (section 6.7).

These specific aims fit into the overall goal of this chapter: to explicate what is going on in the quantum mechanical explanations. As a whole, this chapter is an additional argument in favor of the thesis that explanatory unification without causation is possible. I want to convince the reader that, like in Chapter 5, we have non-causal explanations that provide ontological unification. The difference is that no underlying causal relations are present here.

6.1.1 18th and 19th century

It is beyond the scope of this dissertation to give a full account of the discovery of the periodic law. However, I will provide some context based on the book of Eric Scerri 'The periodic table: its story and its significance' (2007). Antoine Lavoisier, the French chemist, used his own fortune to make one of the finest balances of his day and he started weighing different reacting substances. Next to his great discovery of the oxygen theory of combustion and the law of conservation of mass, his quantitative approach paved the way for laws of chemical combination and an empirical approach to chemistry. At the end of the 18th century Jeremias Benjamin Richter did measurement experiments that led to the concept of equivalent weights. Now properties of elements could be compared on a numerical scale. In 1803 John Dalton, famous for the atomic theory, published what could be seen as the first list of atomic weights. In the next 60 years the atomic weight would become the main criterion to arrange chemical elements. By means of numerical relations such as triads[17] predecessors of Mendeleev's table were created, that actually were successful in capturing the correct classification of the main group elements known at that time. What is missing here is the periodicity of types of elements at regular intervals. For this we had to wait until the 1860s. Several factors contributed to the rapid discoveries of many periodic tables in that decade. In 1860 a first international meeting of chemists was held in Karlsruhe. Here one set of atomic weights was accepted, instead of many systems before. Another factor is the development of the new technique of spectroscopy that led to the discovery of new chemical elements. Spectroscopy led to more elements and thus fewer gaps in systematizations and now every element had its own unique spectral fingerprint, which led to a better understanding of chemical elements. Scerri (2007) examines the six discoverers of periodicity in his book, but the leading discoverer is Dimitri Mendeleev. His version of the periodic system had by far the greatest influence on chemistry at his time and still today.

The main organizational principle of his periodic table was the atomic weight of elements. Stephen G. Brush formulates the core idea as follows:

> "The periodic law (as formulated in the nineteenth century) states that when the elements are listed in order of atomic weight, properties such as valence will recur periodically for example, after seven elements." (Brush, 1996)

This central organizational principle enabled Mendeleev to design a two-dimensional classification of the elements known at his time. Ordered according to atomic weight the elements Li, Be, B, C, N, O and F were put in the first column. Since the next known element Na was chemically similar to Li, Mendeleev put Na in a new column, next to Li (see figure 2).

[17] Johann Wolfgang Döbereiner (1829) ordered some chemical elements by their physical properties in triads. For instance, he observed that strontium had an atomic number that was intermediate to that of calcium and barium, and these three elements had similar properties.

Typische Elemente							
			K = 39	Rb = 85	Cs = 133	—	—
			Ca = 40	Sr = 87	Ba = 137	—	—
			—	?Yt = 88?	?Di = 138?	Er = 178?	—
			Ti = 48?	Zr = 90	Co = 140?	?La = 180?	Tb = 231
			V = 51	Nb = 94	—	Ta = 182	—
			Cr = 52	Mo = 96	—	W = 184	U = 240
			Mn = 55	—	—	—	—
			Fe = 56	Ru = 104	—	Os = 195?	—
			Co = 59	Rh = 104	—	Ir = 197	—
			Ni = 59	Pd = 106	—	Pt = 198?	—
H = 1	Li = 7	Na = 23	Cu = 63	Ag = 108	—	A·. = 199?	—
	Be = 9,4	Mg = 24	Zn = 65	Cd = 112	—	Hg = 200	—
	B = 11	Al = 27,3	—	In = 113	—	Tl = 204	—
	C = 12	Si = 28	—	Sn = 118	—	Pb = 207	—
	N = 14	P = 31	As = 75	Sb = 122	—	Bi = 208	—
	O = 16	S = 32	So = 78	Te = 125?	—	—	—
	F = 19	Cl = 35,5	Br = 80	J = 127	—	—	—

Figure 2 Mendeleev's published periodic system, of 1869.
(Mendeleev, 1869, p. 70)

Mendeleev ends his first paper with eight explicitly stated points, the first of which being the periodic law and the last being the prediction that many yet unknown elements are to be discovered to fill the gaps. In 1871 he published a third paper with many different formats of the periodic table and detailed predictions about these elements that are yet to be discovered.

At the end of the 19^{th} century, the noble gas Argon was discovered by Lord Rayleigh and William Ramsay. The discovery of helium, krypton and xenon followed soon thereafter. Suddenly a whole group of elements was discovered without being predicted by the periodic table. The noble gases formed a serious threat for Mendeleev's system. At first, Mendeleev denied that argon and helium were new elements, as he would not accept elements to be competely inert. In 1900 Ramsay suggested to Mendeleev that argon and the other noble gases should be added as a new group, between the halogens and the alkali-metals. As such, each period would be extended by one element. Mendeleev wrote as an answer to this suggestion:

> "This was extremely important for him [Ramsay] as an affirmation of the position of the newly discovered elements, and for me as a glorious confirmation of the general applicability of the periodic law."[18]

The periodic table passed the test and this led to the version of the periodic table that is familiar to many people: the one with eight groups. (Scerri E. , 2007, p. 151 etseq.)

6.1.2 20th and 21st century

Since Henry Mosely, the periodic table is no longer ordened according to atomic weight, but by atomic number. Moseley concluded from spectrometric research that certain radioactive characteristics correlated with the atomic number, but not the atomic mass. Therefore he proposed the atomic number as a more basal property of an element than the atomic mass (Moseley, 1913). In the Rutherford atomic model an atom has a nucleus with a positive charge equal to the total number of positive charged protons on the nucleus (the neutrons in the nucleus are neutral). The atom as a whole is electrically neutral, thus the number of positively charged protons (*p)* is equal to the number of negativily charged electrons revolving around the nucleus (*e*). Thus, the atomic number $z = p = e$. This new ordering principle solved certain problems with Mendeleev's table and provided a more natural ordering principle.

Let us continue with the periodic law itself. Hettema & Kuipers call the core idea in such tables Mendeleev's Periodic Law (MPL), and formulate it as follows:

$$\text{MPL} \quad e \sim e' \qquad \text{iff} \qquad |\, z(e')|\ \textit{is a multiple of}\ 8$$

in which *z* orders the elements according to mass, with the requirement that there can be only one element with a specific atomic number, thus if $z(e) = z(e')$, then $e = e'$ (1988, p. 396)[19] MPL generates at least two predictions: the number of elements in between two chemically similar elements and the existence of similar elements at certain distances from a given element. This is why Mendeleev's table had empty spaces for

[18] As cited by J.R. Smith, *Persistence and Periodicity*, unpublished Ph.D. thesis, University of London, 1975, p. 460. Also see D. Mendeleev, *An Attempt Towards a Chemical Conception of the Ether*. This statement appears in the Russian edition of 1902 as a footnote.

[19] Isotopes form an exception to this rule: isotopes are elements with a small difference in atomic weight, but with identical chemical properties. This is solved when the elements are no longer ordered according to atomic weight but according to their atom number. See below.

elements that were not yet discovered. To summarize, Mendeleev's periodic law has four properties:

(i) Ordered according to atomic weight, there is a repetition of chemical properties.

(ii) There is a constant period length.

(iii) Every period contains 8 elements.

(iv) If there is an empty space, this predicts a yet unknown element.

Scerri compares this to a musical scale: a returning note, denoted by the same letter, sounds like the original note, but is not identical to it, being an octave higher or lower (2007, p. 18).

Empirical evidence shows that the periodicity, unlike notes on a Western musical scale, is neither constant nor exact. The length of a period varies and the elements within any column of the periodic table are not exact recurrences of each other. The periodic table as it is standardized by IUPAC[20] is organized in 18 vertical groups and 7 horizontal periods. This systematization follows the ordering according to increasing atomic number. The first period contains two elements, the second and third each contain eight, the fourth and fifth contain eighteen and so on. This means that the original properties (ii) and (iii) are no longer applicable: the period length is no longer eight, and no longer constant. Moreover, even though there is a neat ordering according to atomic number, the visual representation shows a certain discontinuity: in the first, second and third period there are no similar elements for elements in group 3 to 12.

This means that also the original property (iv) is no longer valid. Only the original idea of (i) remains, albeit in a more approximate nature[21].

[20] See https://www.iupac.org/cms/wp-content/uploads/2015/07/IUPAC_Periodic_Table-28Nov16.jpg (last accessed on 1 July 2018).

[21] The approximate nature of the periodic law could question the use of the term 'law'. But this is beyond the scope of this dissertation.

6.1.4 Summary

In the previous paragraphs we have shown that the notion of periodicity no longer has the same meaning as it did in the early days of the periodic law. The main thing that has been preserved is the idea of similarity in chemical properties between groups of elements. Now that we have clarified this, we can proceed to the next step: show that there are resemblance questions involved.

6.2 Raising explanation-seeking questions

In Mendeleev's first table, elements with similar properties were arranged in the same horizontal rows. In the IUPAC periodic table these elements are vertically ordered and are numbered from 1 through 18. Horizontally, the atomic number *z* of elements increases from one period or row to the next. The arrangement of the elements is a visualization of the fact that the chemical elements exhibit a certain repetition in their properties, such as:

- Elements of group 1in solid state are good electrical conductors.[22]
- Elements of group 3 through 12 are very hard and have very high melting and boiling points, are good electrical conductors and are malleable.
- Elements of group 18 do not form compounds and have very low boiling and melting points.[23]
- Elements of group 17 easily form compounds with elements of group 1 and 2.[24]

This systematization of chemical elements in the table has several practical uses. Fernelius describes the table as "the font of much chemical information" (1986, p. 263). It

---

[22] The alkali-metals include hydrogen (H), lithium (Li), sodium (Na), potassium (K), rubidium (Rb), caesium (Cs) and francium (Fr).

[23] Group 18 contains the so-called "inert gases" or "noble gases": helium (He), neon (Ne), krypton (Kr), xenon (Xe) and radon (Rn).

[24] Group 17 contains the so-called "Halogens": fluorine (F), chlorine (Cl), bromine (Br), iodine (I), astatine (At) and tennessine (Ts)

is a visual representation of certain trends of element properties useful for solving chemical problems.

The arrangement in the table provides information about common chemical characteristics, and immediately generates important explanation-seeking questions. For example, all the elements in group 1, also known as the alkali metals (Li, Na, K, Rb, Cs and Fr) have common properties, such as

- they are good electrical conductors,
- they have relatively low melting points,
- they have a high reactivity, even violently with atmospheric oxygen and water,
- they have a low density,
- their compounds are white or colourless and
- they are soft and cannot withstand force.

This raises several resemblance questions, e.g.:

Why are lithium, natrium, potassium, rubidium, caesium and francium good electrical conductors?

Why do lithium, natrium, potassium, rubidium, caesium and francium have low melting points?

Similar questions arise in other groups, e.g.:

Why are helium, neon, krypton, xenon and radon not found in compounds?

Why do titanium, zirconium, hafnium and rutherfordium have very high melting points?

These explanation-seeking questions focus on the common factors in certain regularities, chemical properties in this context. These resemblance questions fit into the following general format:

Why do the elements in group X have the properties X_1, X_2, X_3,....X_n?

Before we can investigate what kind of unification is present here (cfr. second till fourth aim of this chapter) we have to know what the explanations look like.

6.3 How to answer the explanation-seeking questions raised by the systematization?

6.3.1 The electron and the quantum mechanical explanation

The successful systematization of the chemical elements in the periodic table begs for an explanation:

> Although the periodic table was an outgrowth of Mendeleev's periodic law formulated in 1869, this early grouping of elements with similar properties was done empirically, just by seeing how the material behaved. The odd shape of the periodic table, however, simply begs for an explanation, and quantum mechanics has given it. (Kelter, Mosher, & Scott, 2009, p. 250)

Although Mendeleev remained opposed to an explanation in terms of atomic structure, the traditional explanation for the periodic system invokes electron configurations. J.J. Thomson's discovery of the electron in 1887 gave him a place in the annals of physics, but he also put forward the first explanation of the periodic table in terms of electrons. Even today, most physicists and chemists believe that the electron holds the key to understanding the periodic table. In fact, it is not the odd shape that is explained by quantum mechanics[25]. The atom models explain the many regularities that have been empirically discovered for the groups (vertical columns) and series (horizontal rows) of the table (see the examples in section 6.2).

A core idea of the contemporary model of atoms is that they consist of a dense positively charged nucleus and electrons that are clustered around this nucleus. We owe this idea to Ernest Rutherford's work in 1909-1911. Another core idea of Rutherford, viz. that the electrons revolve around the nucleus like planets around the sun, did not survive. In the quantum mechanical models, the behavior of electrons is characterized mainly in terms of *orbitals*. Given the indeterministic nature of quantum mechanics, it is impossible to predict the exact location of an electron. However, one can determine a spatial region in which the electron is present with a certain probability (e.g. a region in which the

---

[25] For a detailed technical discussion on the quantum mechanical explanation, see Philips 2003 chapters 9 to 11).

electron is present with 0.9 probability, i.e. a region in which the electron is present 90% of the time). Such regions are called orbitals. Each electron occupies an orbital and there can be no more than two electrons in the same orbital.

Orbitals have certain properties. Thus, by saying that an electron occupies an orbital, we say something about where the electron is:

- Each orbital has an energy level, determined by the principal quantum number (the energy level is the most probable distance of the electron(s) in the orbital to the nucleus);
- Each orbital also has a shape determined by the angular momentum quantum number;
- Each orbital also has a spatial orientation, determined by the magnetic quantum number.

Energy levels are denoted by natural numbers (1,2, 3,...) while the shapes of the orbitals are denoted as *s, p, d* and *f.*

According to the quantum mechanical picture, atoms have a shell structure. The first shell is at energy level 1 and has only one, *s*-shaped orbital. So the first shell can contain no more than two electrons. The second shell is at energy level 2; it has one *s*-shaped orbital and three *p*-shaped orbitals. The latter have a different spatial orientation (determined by the third quantum number). It can contain 8 electrons. Each atom has a so-called electron configuration. This configuration tells us which shells are (partially or fully) filled and which shells are empty. And it also tells us how the shells are filled (which orbitals). For instance, the configuration of hydrogen atoms is 1s, which mean that they have one electron in the *s*-shaped orbital of level 1 and nothing at higher levels. The configuration of helium atoms is $1s^2$, which means that they have two electrons in that orbital and nothing elsewhere. The configuration of lithium is $1s^22s$, which means that the first shell is filled (like helium), that there is one electron in the s-orbital of energy level 2, and nothing in the p-shaped orbitals of level 2 or at higher levels.

The electron configuration obeys the following principles:

a) the *Aufbau principle* states that as protons are added to the nucleus, electrons are successively added to orbitals of increasing energy, beginning with the lowest-energy orbitals. This rule is often accompanied by a diagram like the one shown in figure 3, which represents the Madelung or n + l rule.

b) *Hund's rule* which states that when electrons fill orbitals of equal energies, they occupy as many different orbitals as possible.

c) the *Pauli exclusion principle* which states that only two electrons can occupy a single orbital, and if they do so, they have opposite spins (spin up and spin down respectively, the fourth quantum number).

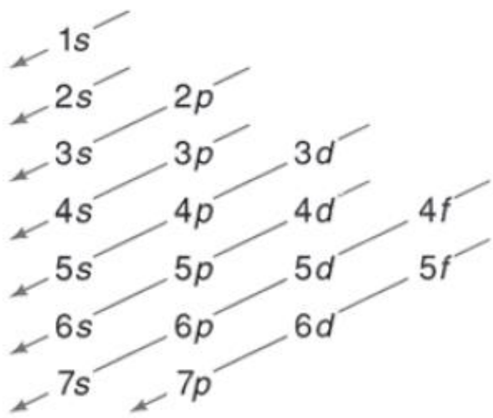

Figure 3 Madelung (or n + l rule) for the order of filling of orbitals (Chang, 2000, p. 601).

The electron configuration allows us to explain some of the regularities mentioned in Section 6.2:

- The elements of group 1 have a single electron in the highest energy level; the rest of the configuration is like that of the preceding inert gas. This outlying electron is relatively free to move about. This explains their conductivity.
- The elements of group 18 are atoms with a closed-shell structure: their outer shell is completely filled. As a result, these atoms are difficult to excite, they are non-reactive.
- The elements of group 17 need only one more electron to form a closed shell. Since the acquisition of an extra electron is energetically favourable, these atoms are very reactive. And they can obtain the required electron easily from elements of group 1, since these have outlying electrons.

6.3.2 Critique on QM explanation

There are several issues with the quantum mechanical explanation of the systematization of chemical elements. First, Eric Scerri argues that probability-based orbitals do not exist according to quantum mechanics (1997, p. 549). The explanation in terms of electronic configuration and the number of outer-shell electrons an atom possesses is at best an approximate explanation. Second, in the most-used formats of the periodic table, every period ends with a noble gas, an atom with a closed shell structure. There is no strict quantum mechanical explanation for this. The reason why a period closes at the known noble gases is an empirical matter. In chapter 9 of his book Scerri demonstrates that Niels Bohr did not deduce electron configurations from physical principles, but he was working backward from chemical and spectroscopic facts. The claim that the periodic table is deductively explained by quantum mechanics is therefore problematic:

> A feature that seems to generally go unnoticed is the need to assume the empirical order of shell filling rather than trying to derive it from the theory. The order in which orbitals are occupied with electrons is not derived from first principles. It is justified post facto and by some complex calculations. (Scerri, 2007, p. 237)

The three principles for electronic configuration as mentioned in 6.3.1. are essentially empirical, they are not derived from the principles of quantum mechanics. These principles summarize the knowledge from empirical data on atomic spectra. There are plenty of examples where the electronic configuration of an atom does not strictly follow the principles. For instance the element potassium. The configuration of the preceding element argon is $1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^6$. The expected configuration for potassium would thus be $1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^6$, $3d^1$ according to the pattern that the next electron should be added at the next available orbital at increasing distances from the nucleus (Hund's rule). Empirical evidence shows a different configuration for potassium: $1s^2$, $2s^2$, $2p^6$, $3s^2$, $4s^1$. In the next element calcium the new electron also occupies a space in the 4s orbital, but in element 21, scandium, the orbital energies have reversed, giving the 3d orbital a lower energy. (Scerri, 2007, pp. 234-235).

Furthermore, having a particular electronic configuration is neither a sufficient nor a necessary condition to belong to the same group within the periodic system. Nickel, palladium and platinum belong to group 10 but each shows a different outer-shell configuration. They are grouped together because of similar chemical properties. Helium, beryllium and magnesium on the other hand do share the same outer-shell structure of two electrons, yet they do not fall into the same group.

Nevertheless, the QM explanation of the systematization of chemical elements is considered as an explanation in science (e.g. Ostrovsky 2001 and 2006, Schwarz & Rich 2010), even by its critics (e.g. Scerri 1998, Anderson, Gomatam, & Behera, 2014).

6.3.3 Summary and preview

In section 6.2 we showed that the systematization of chemical properties in the periodic table raises the type of explanation-seeking questions that came into the focus of this dissertation. In this section we clarified what the QM explanations look like. We are now ready to start the philosophical analysis. It will be argued that these explanations are not an instance of causal network unification (6.4) nor of mechanism unification (6.5). Since the mechanistic and the interventionist account of causal explanations do not fit what happens in the QM explanation, the question arises whether QM provides any causal explanation at all. In section 6.6 I will argue that QM models provide structural explanations. If structural explanations exist, then a kind of ontological unification that is based on those structural explanations becomes a possibility. I will call this structural unification (6.7).

6.4 The QM explanations do not use causal network unification

Let us investigate whether the QM explanations are an instance of causal network unification.

Consider the following resemblance question:

> Why are helium, neon, krypton, xenon and radon not found in compounds?

The general format of this question is the following:

> (A) Why do objects of type *A*, *B*, *C*, ... have property X?

This format is different from the resemblance question from Chapter 3 about the three social revolutions.

There we had:

> (B) Why do objects *a*, *b*, *c*, ... have property X?

Although these questions appear to be very similar, they are not. In question (A) the explanandum is a similarity about a capacity, e.g. conductivity, being a compound (or not), etc. In question (B), as seen in Chapter 3, the explanandum is a similarity between events.

There is a second difference. To provide causal network unification (as in Chapter 3), positive or negative causal factors need to be determined that are present in the causal ancestry of all objects (organisms or populations). Remember the general format of an explanation that results from an act of causal network unification:

> X is a positive causal factor for E.
>
> Y is a positive causal factor for E.
>
> Z is a positive causal factor for E.
>
> ...
>
> Object a has X, Y, Z, ...
>
> Object b has X, Y, Z, ...
>
> Object c has X, Y, Z ...
>
> ...

Thus, objects a, b and c have property E.

The causal factors X, Y and Z are present in the causal ancestry of the objects. This means that these factors make a difference and if one intervenes on a cause X it would lead to a change in the effect Y.

QM models provide synchronic explanations. The state of a compound (e.g. conductivity of an atom) at time t is explained by the state of the constituents (electron configuration) at the same time t. The fact that QM models describe synchronic relations implies that they do not give information about the causal ancestry of the phenomenon. No causal process or causal trajectory is described. Causal network unification was characterized in a way that excludes synchronicity: in causal network unification "one causal network is presented and applied to two or more events to point out the common causal factors that are present in the causal ancestry of all events we consider" (section 3.7). Because of the synchronic nature of QM models, they do not meet this requirement and are not an instance of causal network unification. The notion of causal ancestry is crucial in the latter.

One could argue that the criterion of causal ancestry is too strict, and that there is some other type of causal unification. But what could function as a causal factor in this case? The interventionist account of causation, as proposed by Woodward (2003) maintains that causes are factors that make a difference. This means that if intervenes on a cause X, it would lead to a change in the effect Y. This also means that a cause X should be able to take at least two values (theoretically). Furthermore, it must be clear how cause X is connected to effect Y, the causal relation needs to be identified.

Lauren Ross (2017) claims that the QM explanation of the periodicity and group trends are causal explanations. The atomic features of an element serve as causes, according to Ross, since "changes in these features produce changes in these properties" (Ross, 2017, p. 11). If one were to intervene in the electron configuration of an element, this would produce changes in the properties of that element.

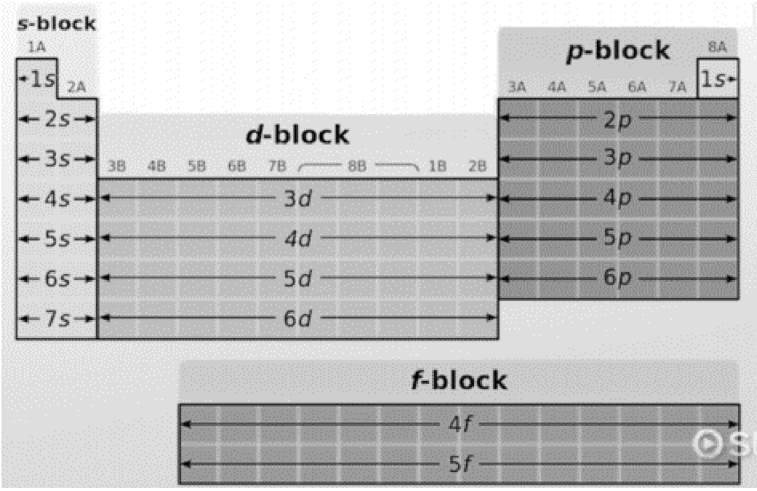

Figure 4 Blocks of the periodic table (Ross, 2017, p. 12)

Based on the work of Charles Janet the chemical elements in the periodic table can be seen in blocks. Each element in such a block has the same atomic orbital type. The ordering by increasing atomic number also shows a periodic pattern in valence electron configuration according to the Mandelung rule (figure 3). Elements in the s-block have valence electrons in s-orbitals, elements in the p-block have valence electrons in p-orbitals etc. Each orbital is progressively filled across a period (vertical column), and the position in the block tells us how many electrons occupy the respective orbital. For instance, an element in the first position of the p-block, has a single valence electron in the p-orbital.

If the electron configuration is the cause of certain chemical properties of an element, experiments should be able to explicate the causal relation between both. It is indeed so that experiments show that atomic alterations can cause changes in the chemical and physical properties of a substance, and result in the transmutation of one element into another. Rutherford succeeded in converting nitrogen into oxygen through alpha particle bombardment (Rutherford and Soddy 1903). These bombarding techniques led to the synthesis of new elements (lantanides and actinides).

Now, this all seems to add up for individual elements, but what about the initial resemblance question about groups in the periodic table:

> Why do objects of type *A, B, C*, ... have property X?

Is there a common causal structure that can be invoked to explain the similarity of the elements? As seen in section 6.3.2, there are several problems with the QM explanation, especially when it is treated as a causal, interventionist explanation. If the electron configuration functions as a causal explanation, it should be possible to predict which configurations relate to particular properties and to predict where chemical properties recur in the periodic table. In 6.3.2. it was made clear that the electron configurations and the periodicity were often discovered empirically, not by uncovering a causal structure.

Neither is it the case that resemblance questions as mentioned before can be answered by uncovering a common causal structure in terms of electron configuration. The Aufbau principle has no less than 20 exceptions (Scerri, 1998). Having a specific electronic configuration is neither a sufficient nor a necessary condition to belong to the same group within the periodic system (see examples in 6.3.2.).

Even if the electron configuration is the key to understanding the properties of an individual chemical element, it can be doubted if this happens in a causal way. Furthermore, it does not explain the periodicity of those properties. Therefore, a satisfactory answer to the resemblance question at hand is not a type of causal network unification.

6.5 The QM explanations do not use not mechanism unification

6.5.1 Introduction

Consider the following question:

> Why are lithium, natrium, potassium, rubidium, caesium and francium good electrical conductors?

If we treat this question as a request for mechanism unification, we get:

> What are the common features of the mechanisms which respectively make lithium, natrium, potassium, rubidium, caesium and francium good electrical conductors?

The main point I want to make in this section is that this is not a viable interpretation, because there are no such mechanisms. One cannot give a mechanistic explanation of the regularity that makes Lithium a good electrical conductor, because there is no known mechanism that produces this conductivity. The same holds for Natrium and the other elements of Group 1. Since there are no mechanisms, there can be no mechanism unification.

In Sections 6.5.2 till 6.5.3 I argue that, while the quantum mechanical explanations are clearly bottom-up explanations, they are not mechanistic explanations in the way the latter are usually understood. Bottom-up explanations are explanations that invoke entities at a lower level (compared to the level at which the explanandum phenomenon is situated). The quantum mechanical explanations described in Section 6.3 are clearly bottom-up explanations in this sense: they involve decomposing atoms into nuclei and electrons, and thus invoke entities at a lower level. However, as we will see in Section 6.5.2, mechanistic explanations are usually thought of as a species of causal explanations. In Section 6.5.3 I show that this implies that the quantum mechanistic explanations of chemical properties cannot be seen as mechanistic explanations.

6.5.2 Mechanistic explanations as a species of causal explanations

The causal nature of mechanistic explanation is clearly explicated by Stuart Glennan – one of the leading philosophers favouring a mechanistic approach to scientific explanations – in a recent joint paper with Meinhard Kuhlman. They write:

> One way to summarize the mechanistic consensus is this: Mechanisms consist of parts (entities, components) that are so organized that the activities and interactions of these entities are productive of a phenomenon. (Kuhlman & Glennan 2014, p. 339)

They do not define what a mechanistic explanation is. Like most mechanistic philosophers, they focus on defining and clarifying what a mechanism is: for my purposes, it is good to also have a definition of what a mechanistic explanation is. The definition I propose is:

> A mechanistic explanation for a phenomenon is an explanation in which the explanans describes the mechanism that produces this phenomenon.

This is a static characterisation. From a dynamic point of view, mechanistic explanations can be characterised as follows:

> All mechanistic explanations begin with (a) the identification of a phenomenon or some phenomena to be explained, (b) proceed by decomposition into the entities

> and activities relevant to the phenomenon, and (c) give the organization of entities and activities by which they produce the phenomenon. (Illari & Williamson, 2012, p. 123)

This is a dynamic characterization which tells us how mechanistic explanations are construed. It fits perfectly with the definition of the end product and the definition of mechanisms given by Kuhlman & Glennan.

Kuhlman & Glennan discuss four important features of mechanisms and mechanistic explanations. The second feature is important for us here:

> A second feature of the New Mechanist consensus is the idea that phenomena exhibited by the mechanisms are produced by the activities and interactions of parts. The terms 'activity', 'interaction', and 'produce' are all transparently causal. If the activities and interactions are not genuinely causal, then mechanism can't produce anything. Mechanistic explanation is a species of causal explanation and the legitimacy of mechanistic explanation depends upon the interactions between parts being genuinely causal. (2014, p. 339; italics in original)

I will use this feature in 6.5.3. The third feature they mention is also important for me:

> A third feature of the New Mechanist consensus is its focus on organization. It is the organization of the entities (and their activities) that allows the mechanism to produce the phenomenon that it does. A pile of lawnmower parts does not [make] a lawnmower. While mechanists emphasize the importance of spatial and temporal organization, it is ultimately the causal organization upon which the productive capacities of the mechanism depend. (2014, pp. 340-341)

This feature will also be used in 6.5.3.

6.5.3 Application

The quantum mechanical explanations invoke entities at a lower level: they involve decomposing atoms into nuclei and electrons. These elements have a strict, orderly organisation: the characteristic electron configuration of each type of atom.

A first potential trouble lies with the activities. Being located in an orbital is a property but it is debatable whether it is an activity: nothing is 'done'. 'Being there' is, it seems, a passive property rather than an activity.

Secondly, in a mechanistic explanation the causal organization is essential to the explanation:

> To clarify the role or organization in mechanisms, consider as a brief example the mechanism for starting a lawnmower engine. The engine is started by rapidly pulling a cord while the throttle is set to an appropriate level. The cord is attached to a flywheel which in turn engages a clutch which causes the crank shaft to move, which in turn moves the piston, allowing air and fuel into the cylinder. The flywheel is also connected to a magneto—a device which uses the rotation of magnets to generate a voltage. The magneto is attached to the sparkplug which produces the spark that ignites the fuel-air mixture in the piston. The production of the phenomenon (namely the starting of the mower) *depends essentially on organization*. The parts must be spatially organized so that the same part—the flywheel—may simultaneously engage the clutch and turn the magneto. Timing is also essential here. The parts must be so organized that the spark generated by the spark plug enters the cylinder at the correct time in the piston's cycle*. These spatial and temporal arrangements determine the causal organization of the system*. (Kuhlman & Glennan, 2014, p. 341; emphasis added)

Even though the chemical elements have a strict, orderly organization, i.e.: the characteristic electron configuration of each type of atom, this is not a 'causal organization'. The spatial organization is not connected to any causal interaction that there may be between the electrons. For instance, it is certainly not the case that only electrons within the same orbital can have an influence on each other's behaviour.

6.6 Why the QM-models are explanatory

Since the mechanistic, the causal-interventionist and the causal ancestry account of explanations do not fit what happens in the QM explanation, the question arises whether QM provides any causal explanation at all. In the literature it appears that there is no consensus about whether or not quantum mechanics is compatible with our views on causality[26]. Instead of arguing that QM-models are causal explanations or non-causal explanation, I would like to start from the question: if QM models are not causal, can they still be explanatory?

In order to answer this question, I focus on plain explanation-seeking questions, for instance:

> Why is lithium a good electrical conductor?

I will come back to the resemblance questions in 6.7.

The QM-model answers this question by referring to the electron configuration:

> The good electric conductivity of lithium physically depends on its electron configuration, viz. $1s^22s$.

One might argue that this is an example of purely derivational unification, but it is not. On the contrary, I will argue that it is purely ontological.

First, the chemical properties cannot be directly derived from the electron configuration. As seen in section 6.3.:

> There is a problem with the claim that the periodic table is deductively explained by quantum mechanics. A feature that seems to generally go unnoticed is the need to assume the empirical order of shell filling rather than trying to derive it from the theory. The order in which orbitals are occupied with electrons is not derived from first principles. It is justified post facto and by some complex calculations. (Scerri, 2007, p. 237)

That the QM models do not provide a full deductive account, might be a problem for physics, but it does not need to be a problem for the explanatory status of QM-models. Deductive information can be relevant in certain explanations, but it is not what makes an explanation explanatory. As I have written in Chapter 5, it is the mirroring of deductive relations to physical dependency relations that make a derivation explanatory. Dependencies always guide us towards explanations, derivations do not. (Woodward, 2003, p. 203).

[26] For some interesting papers about this topic I refer to Brukner, 2014; Plotnitsky, 2010; Riggs, 2009; Ringbauer et al 2016 and Winter, 2017.

> The idea is that these derivations trace or mirror the relations of physical dependency that hold between the explanans conditions and the explananda phenomena-relations that would be revealed if, for example, we were to physically intervene to alter the explanans conditions. (2003, p. 201)

Throughout the dissertation I have used this criterion of representing physical dependency relations as the demarcation between an explanation and a description that is non-explanatory. Nowhere in Woodward's criterion is the condition that these relations need to be causal:

> The *underlying* or unifying idea in the notion of causal explanation is the idea that an explanation must answer a what-if-things-had-been-different question, *or exhibit information about a pattern of dependency*. (2003, p. 201; italics added)

Causality is not a criterion for the representation of physical dependency relations, it is the other way around. If we accept that the QM models mirror such physical dependency relations, they can be classified as explanatory according to Woodward's view. More generally, in order to accept QM derivations as explanations we have to accept claims of the following format:

> Chemical property X of atoms of type Y physically depends on their typical electron configuration.

The constituents that form the base of the QM explanation of the periodic law, the electrons, obey three organizational principles: the Aufbau principle (Mandelung rule), Hund's rule and the Pauli Exclusion Principle. These principles can be seen as pure mathematical rules. For instance, the Pauli Exclusion Principle can be formulated as: no two electrons can have four identical quantum numbers. Usually the principle is formulated making also an ontological claim such as: no two electrons can simultaneously occupy the same quantum state with respect to orbital and spin. Concepts such as 'occupy', 'quantum state' and 'orbital spin' suggest more than pure mathematical constraints of behavior. They tell us something about how electrons can behave, about certain characteristics they may have. Quantum mechanics not only provides useful mathematical tools but also makes ontological claims:

> Quantum mechanics neatly accommodates the existence of particles that are indistinguishable in principle: we simply construct a wave function that is noncommittal as to which particle is in which state. [...] This is the famous Pauli

exclusion principle. It is not (as you may have been led to believe) a weird ad hoc assumption applying only to electrons, but rather a consequence of the rules for constructing two-particle wave functions, applying to all identical fermions. (Griffiths, 2005, p. 204)

Quantum mechanics does tell us something about relations of physical dependency. However, section 6.4 and 6.5 suggest that if there is such a physical dependency, it is a bit spooky: it is not causal dependence (unlike what is the case in causal network unification and mechanism unification) and it does not supervene on causal relation (unlike what is the case in optimality explanations). So, in line with the idea that physical dependencies can be non-causal, I define structural explanations as follows:

> An explanation is structural if and only if the explanandum physically depends, in a non-causal way, on the micro-structure described in the explanans.

With this definition in place, the derivations from the QM models may be called structural explanations.

Some people may say that structural explanations as I define them cannot exist, because they reject the idea of "physical dependency in a non-causal way" as spooky metaphysics. In their view all physical dependency relations are causal or supervene on causal relations. If the QM-models are non-causal, they would consider them as non-explanatory.

6.7 The QM explanations provide structural unification

This section is meant for those readers who accept the idea that physical dependency relations that do not supervene on causal relations exist. If structural explanations exist, then a kind of ontological unification that is based on those structural explanations becomes possible.

I define structural unification as follows:

> Structural unification consists in the act of answering a resemblance question by identifying crucial common features in a given set of structural explanations.

Note that this presupposes that you have a set of structural explanations. Hence, if the QM models do not provide structural explanations, they do not provide structural unification. The notion of structural explanations means that the explanandum physically depends on the micro-structure described in the explanans, in a non-causal way.

Consider the following example:

> Why are lithium, natrium, potassium, rubidium, caesium and francium good electrical conductors?

The answer is:

> They all have a single electron in the highest energy level; the rest of the configuration is like that of the preceding inert gas.

The unifying explanation in this answer refers back to common features in the set of structural explanations of those elements: they all have a single electron in the highest energy level.

The notion of structural unification means that a resemblance question is answered by identifying common features in a given set of structural explanations. The idea that there is a more fundamental structure showing how disconnected phenomena are only apparently disconnected, but are manifestations of one and the same structure is structural unification. This is not purely derivational unification, but a matter of successfully representing how things are physically dependent upon one another in the world. In this section I claimed it is possible to have an explanation that is based on physical dependency relations that do not supervene on causal relations. Furthermore, answering resemblance questions by referring to the similarities in those non-causal physical dependency relations, is a new kind of ontological unification. This kind of non-causal ontological unification is structural unification.

6.8 Conclusion

In this sixth chapter a form of unification that is non-causal was found by studying the QM explanation of the periodic law. In sections 6.1, 6.2 and 6.3 the necessary background information was given: the development of the periodic law, how this systematization is both an explanans and an explanandum and how the periodic law is explained by referring to quantum mechanics and electron configurations. As an explanandum the periodic law raises several resemblance questions and thus some form of unification is required to answer those questions.

In sections 6.4 and 6.5 previous types of unification were considered, but neither causal network unification nor mechanism unification was applicable. In 6.6 it was argued that if QM-models are not causal, it can still be accepted that they are explanatory, since they provide structural explanations. In 6.7 I argued that if such structural explanations are possible, there is also a kind of ontological unification that is based on the common features in these structures. In my view the QM explanation for the resemblance questions is a genuine explanation based on structural unification. This form of unification is still ontological, because it is based on physical dependencies, even if those dependencies do not supervene on causal relations.

Part III
Further reflections and conclusions

Introduction to Part III

This third and final part of my dissertation contains some further reflections and conclusions. During the course of this research I have experienced how the philosophy of explanation is dominated by accounts of causal explanation. Explanatory unification is often considered old-fashioned. However, last year, there was one person, Sorin Bangu, who wrote a paper about unification. The goal of his paper is similar to this dissertation: rethinking unification and bringing it back into the picture. In Chapter 7 I will compare his approach to mine.

In the final chapter I will formulate my primary results, and I will elaborate on their implications for thinking about unification and explanation. The reader may have noticed that my approach to explanation is as disunified as my approach to unification itself. The different forms of ontological unification were quite diverse. This relates to the method I have used. Throughout this dissertation the types of unification that were discussed emerged from digging into scientific practice. During the research for this dissertation I could explore topics from a variety of scientific fields. This philosophy-of-science-in-practice approach steered me towards a pluralistic view on unification and on explanation.

Chapter 7
Unification in the current philosophical literature

In 2017 the *European Journal for Philosophy of Science* published a paper by Sorin Bangu: 'Scientific explanation and understanding: unificationism reconsidered'. At first, I was enthusiastic to read that the aim of the paper was to revive unificationism by revision of its doctrines. Finally, I found someone with a similar research interest.

In section 7.1 I will explain how Bangu rethinks unification by adding an ontological constraint. In section 7.2 I will clarify why my enthusiasm was short-lived.

7.1 Unificationism reconsidered by Sorin Bangu

Bangu rightly diagnoses that unification in its classical form is flawed (2017, p. 104). By making the notion of unification more precise and replacing it with what he calls 'ontological-reductive' unification, he hopes to reconnect unification with explanation and understanding (2017, p. 104). Bangu uncovers two assumptions about unification: explanations provide understanding, and understanding is achieved through unification. The nature of understanding and the relation between explanations and understanding does not need to be problematized for the purpose of this section[27]. It is the relation between unification and explanation that deserves our attention. Bangu states that

---

[27] For more on the nature of understanding and the historical variation of criteria for understanding in science I refer to the work of Henk de Regt (2009) (2017).

explanation is an ordered triple [EXS, d, EXD] consisting in the phenomenon to explain (i.e., the explanandum, abbreviated as EXD), the assumptions made when explaining (the explanans, EXS), and a dependence relation d, which makes explicit how EXS and EXD are connected (2017, pp. 105-106). Schematically this looks like this:

$$\text{EXS} \longmapsto \text{EXD}$$

In order for the explanandum to be explained, the explanans needs to be identified and the relation d is specified (2017, p. 106). One way to fill in the dependence relation is Hempel's deductive-nomological model. But as seen in 1.2.2 this led to several problems. Bangu analyzes unification approaches to explaining by distinguishing two strategies.

The first strategy adds a third condition to the criteria for an explanandum to be explained. This third condition constrains the explanans. Schurz and Lambert fit this strategy, since they require that the explanans is in less need of an explanation than the explanandum (Schurz, 1999, p. 97) (Schurz & Lambert, 1994, p. 105). Another way to fill in this strategy condition is the requirement that the explanans is the cause of the explanandum. Also Friedman's account of unification (1974) fits this first strategy, according to Bangu, since the explanans must be *"more comprehensive"* (Friedman, 1974, p. 19; italics in original). The explanans should not only enable us to derive the explanandum but other phenomena as well. The requirement that is added by Friedman is that the explanans needs to be a unifier (Bangu calls this UEXS).

A second strategy does not constrain the explanans but the appropriate dependence relations. This is what Kitcher (1981) does by claiming an explanandum is explained if it can be derived from an argument pattern. This dependence relation, or argument pattern in Kitcher's terminology, is further limited with requirements such as the size of the conclusion set, the number of and similarity of argument patterns and the stringency of those patterns. Just as with Friedman's strategy this is an objective criterion, because it is measurable.

Bangu schematizes Friedman's and Kitcher's views on unification as follows:

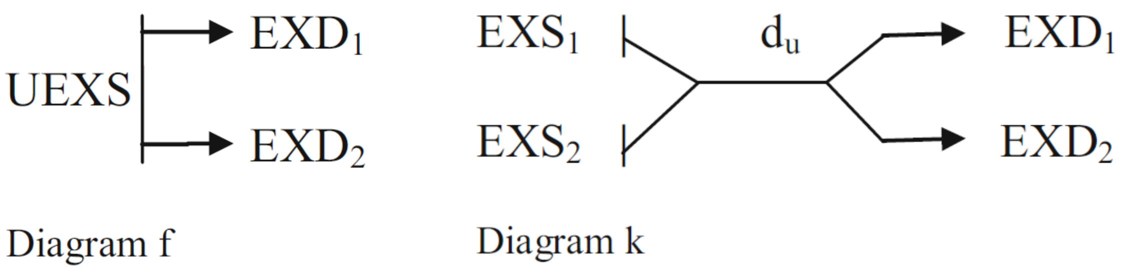

Figure 5 Schematic representation of two strategies of unification by Bangu (2017, p.111 and p. 114). Diagram f presents unification according to Friedman (1974), diagram k presents the Kitcher-style unification (1981).

According to Bangu both types of unification have their merits. Friedman-style unification increases understanding by reducing the number of brute phenomena. Kitcher-style unification reduces the number of dependence relations. But both approaches see understanding as a global feature. According to Bangu neither approach provides a satisfactory answer to how our understanding of individual phenomena is increased when we derive it from a unifier or by using an argument pattern (2017, pp. 113, 116). Bangu tries to do better. He considers the Newtonian mechanical framework and its capacity to derive Kepler's and Galileo's mechanical results and the combined gas law to show how unificationist accounts can enable us to give better descriptions of individual explananda. The unifying descriptions are better because the explanandum is conceptually enhanced:

> That is, the unifier UEXS, once identified, does not leave the descriptions of the phenomena to be unified unchanged, but unifies while re-describing (and/or correcting). Thus, the key-question to which the standard unificationist didn't have an answer—how is our understanding of an individual phenomenon increased when we derive it from a unifier?—now receives one: we can claim a better understanding of an individual explanandum because of our capacity to provide a conceptually enhanced description of it. (Bangu, 2017, p. 118)

Unifying explanations redescribe the phenomena that were in need of explanation by introducing a different ontology than that of the explanandum before the unifying explanation. For instance, an atom in group I of the periodic table is no longer just an atom, but a conglomerate of neurons, protons and electrons that behave in a specific way.

Bangu concludes that genuine unification is ontological-reductive: it assumes a sparser ontology than the ontology of the phenomenon/-a derived from it.

Summarized, for Bangu a phenomenon is explained if and only if:

(i) an ontologically-reductive unifier explanans is identified, and

(ii) the way in which this explanans ensures the phenomenon arises is presented. (2017, p. 123)

This means that an explanation needs to unify the explanandum and the explanans, by showing that the explanandum can be derived from the explanans. (This limits the dependence relation between EXD and EXS.) Furthermore, the unification needs to be ontological-reductive: the explanandum needs to be redescribed in a sparser ontology than before the explanantion. (This limits the EXS.)

Unfortunately, there is a high price to pay: genuine unification is difficult and hard to get[28]. But, if it is reached, this genuine unification is superior to causal explanations, according to Bangu: it provides both local and global understanding (2017, p. 124).

7.2 Problems for Bangu's proposal

There are several reasons why my enthusiasm of Bangu's paper was premature. In Bangu's definition an explanation is triple of an explandum, an explanans and a dependence relation between both of them.

My first problem relates to the first part of his strategy to grasp genuine explanations: constraining the explanans. Bangu argues that an explanans has to be an ontological-reductive unifier. However, this position ignores cases where the explanans does not have to be a unifier. Scientists solve problems in different explanation-seeking contexts. There are contexts in which unification is completely irrelevant for explanation. Weber & Van Bouwel 2009 et al. argue convincingly that there are explanation-seeking contexts in

[28] For instance, in a footnote on page 123 Bangu acknowledges that as a consequence of his view Maxwell's unification of electromagnetism should not be considered as genuine unification.

which the sole aim of the explanation is uncovering causal relations between the explanans and the explanandum.

A second problem is related to the dependence relation between the explanans and the explanandum. Bangu remains vague about the dependence relation, as mentioned in his second criterion of explanation*: "the way in which this explanans ensures the phenomenon arises is presented."* (2017, p.123). According to a footnote on p.110 the nature of the dependence relation is ultimately left open. But how open can a dependence relation really be if ontological *reduction* is required? It seems to me that this requirement unnecessarily constrains both the nature of unification and the nature of explanation.

Even if we restrict ourselves to contexts where unification is the aim of an explanation his criterion is still too limited. Let us compare two examples to show this.

Remember these resemblance questions:

A. Why was there a revolution in Bourbon France, Manchu China and in Romanov Russia? (Chapter 3)
B. Why do *all* dung flies copulate for approximately 36 minutes? Why is there so little variation? Why do they all behave more or less the same? (Chapter 5)

Question A was answered by showing that all three revolutions had common causal factors. Both Bourbon France, Manchu China and Romanov Russia state's repressive capacity was weakened due to a fiscal crisis because of large military expenses. They all had a strong sense of community among peasants. Nowhere is it claimed that the phenomena can be *reduced* to these common causal factors. Although the revolutions were similar, they are not the same. The causal ancestries of the three revolutions are not identical. The revolutions cannot be reduced to their common causal factors. The dependence relation between the explanandum (the three revolutions) and the explanans (the common causal factors) is one of causal network unification: ontological, causal but non-reductive.

Question B was answered by constructing an optimality model to show how copulating time of 36 minutes is an optimal strategy to maximise the number of eggs fertilized per time unit. In order to work optimality models need to use a number of idealizing assumptions such as an infinite population, randomized mating, the inheritance of copulating strategies. All the dynamics, thus the causal factors that make dung flies copulate for 36 minutes, are left out of the model. Tradeoffs are not causal relationships between values: ‘average fertilization rates’ does not cause ‘copulating time’. Similarly, the explanandum is not ontologically reduced to the explanans. An optimality model does not give a new ontological description of the phenomenon so that it can be reduced to this new level, but it exhibits a physical dependency between populations, population level traits, the environment, fitness and selection. The dependence relation between the explanandum (copulating time) and the explanans (optimality model) is one of physical dependency unification: this is a form of ontological and non-reductive unification[29].

The two of the examples used earlier in this dissertation show that unification often is non-reductive in the sense that Bangu requires.

7.3 Conclusion

The value of Bangu’s paper is that he attempts to bring unification back into the picture and that he does this in an ‘ontological way’. However, they are problems, as clarified in 7.2. The first problem that I have mentioned there relates to Bangu’s monism about explanation (all explanations have to be unificatory). The second problem relates to Bangu’s monism about unification (unification is always ontological-reductive). In Chapter 8, I discuss the implications of the view that I have developed throughout this dissertation. I am a pluralist about explanation and a pluralist about unification. This is another way to understand my disagreement with Bangu.

[29] The nature of unification in optimality models is also derivational, the equilibrium state is mathematically derived from the trade-off relations and the constraints. But the equilibrium state cannot be reduced to those trade-offs and constraints.

Chapter 8
General conclusions

8.1 Introduction

In this final chapter, it is time to take stock and to summarize the main results of this dissertation. The main goal defined in the general introduction was to bring unification back into the picture by rethinking the concept. To reach this aim I have followed a new strategy: looking for ontological unification in scientific practice.

The primary results that I have reached by means of this new strategy are:

(1) *Causal network unification* exists. I have analyzed how it works and shown why it is important.

(2) *Mechanism unification* exists. I have analyzed how it works and shown why it is important.

(3) *Physical dependency* unification exists. I have analyzed how it works and shown why it is important.

(4) *Structural unification* exists. I have analyzed how it works and shown why it is important.

These primary results have two important implications:

1) I am a pluralist about the nature of unification.

2) I am a pluralist about the nature of explanation.

I will discuss these implications in sections 8.2 and 8.3. In section 8.4 I will make some suggestions for future research.

8.2 Pluralism about unification

The primary results summarized 8.1 have four main implications for my views on unification:

1) Unification in adequate explanations is sometimes, but not always derivational.
2) Unification in adequate explanations is always ontological.
3) Unification plays a variety of roles in explanations.
4) Ontological unification can take on different forms.

Each implication marks a difference with all or most traditional accounts discussed in Chapter 1. And each implications is an aspect of my 'pluralism about unification'. I discuss each of them below.

First, unifying explanations, in the traditional accounts, are derivational and require subsumption. In order to explain something, one has to construct an argument to show that the explanans is to be expected and/or is an instance of a (set of) law(s). In my view unification does not need to be derivational. Causal network unification, mechanism unification and structural unification do not presuppose that the explananda are derived from overarching premises. In the case of the physical dependency unification in optimality models unification is mixed, it is both derivational and ontological. So unification is sometimes derivational, and sometimes not.

Second, unification is always ontological. Mäki proposed this broad idea in 2001, but elaborate examples from scientific practice are not mentioned in the paper (he briefly mentions the Newtonian unification of mechanics). This notion of ontological unification implies that a set of phenomena is unified if they share the same ontic foundations. In the case study of the social revolutions the ontic unification is based on a network of common causal factors, which I called *causal network unification*. In the example of general anaesthesia, a common causal mechanism was found that explains why all those different substances induced general anaesthesia. This was called *mechanism unification*. In Chapter 5 optimality explanations were analysed. This taught us that there is a type of unification that is based on physical dependency relations that supervene on causal relations in the world. This was called *physical dependency unification*. In the QM explanation of the periodic phenomena the physical relations do not supervene on causal relations in the world, but there still is a deeper ontological structure that forms the basis for the unification. This was labelled *structural unification*. In all examples the unification was not purely derivational, but was based on a deeper, underlying common ontological structure. In this concluding chapter, I will take the liberty to make an extra claim: derivational unification is not explanatory. This is a consequence of the criterion that I have used throughout this dissertation to classify descriptions as explanations: representing physical dependency relations (Woodward, 2003, p. 201). If unificatory derivations mirror those physical relations, in other words if there is also ontological unification, then they are explanatory, otherwise they are not. This is why I cannot adopt the accounts of Kitcher and Schurz, because they allow for purely derivational explanations.

Third, unification plays a variety of roles in explanations. Contrary to Kitcher (1981, 1989) I do not claim that unification always is a necessary ingredient of explanation. A causal explanation without unification can be adequate (cf. the example in the following section). However, Chapter 2 shows that unification can be an explanatory virtue of causal explanations. The example of the mercury thermometers and the ideal gas law demonstrate how adding unification to a causal explanation can improve that explanation in at least two ways: it can result in a higher explanatory power and in more cogent explanations. In other circumstances (cf; chapters 3, 5 and 6) unification is a necessary condition of adequacy for explanations. This is connected then to the fact that one tries to answer resemblance questions.

Fourth, if unification is a necessary aspect of an explanation, it can take on different forms. I have shown that at least four different kinds of ontological unification exist (see chapters 3, 5 and 6).

These four implications result in a pluralist perspective on unification. This pluralism is the indirect result of the strategy used in this dissertation. Starting from scientific practice made sure I kept an open mind on what an explanation is and what unification is. From thee diverse case studies, a new framework for thinking about unification emerged. In a way, giving up unification as traditionally presented by philosophy of science enabled me to embrace new pluralist ideas of unification conveyed by scientific practice.

8.3 Pluralism about explanation

The third result in the previous section states that unification can take on a variety of roles in explanation. This pluralist view on unification, does not automatically entail a pluralist view on explanation. I this section I will briefly explain why I adopt pluralism about unification but also about explanation.

The pluralistic perspective on explanation, is related to the liberal pragmatic approach Van Fraassen proposes:

> So scientific explanation is not (pure) science but an application of science. It is a use of science to satisfy certain of our desires; and these desires are quite specific in a specific context, but they are always desires for descriptive information. ... The exact content of the desire, and the evaluation of how well it is satisfied, varies from context to context. It is not a single desire, the same in all cases, for a special sort of thing, but rather, in each case, a different desire for something of a quite familiar sort. (1980, p. 156)

The key idea of explanatory pluralism is that, depending on the structure of the explanation-seeking question and the motivation behind it, the answer will have different conditions of adequacy.

After reading the previous chapters it should be clear that I am at least an explanatory pluralist in the following senses:

> Some adequate explanations are mechanistic, others are not;
>
> Some adequate explanations require causation, others do not; and
>
> Some adequate explanations require derivation[30], others do not.

Despite the pluralism that I am committed to, one could argue that it is possible for me to uphold the following monistic thesis:

> Unification always plays a role in explanation, either as an aim or as a virtue.

However, my view is that besides being a virtue or an aim, unification can be completely irrelevant for an explanation. Let me make a brief detour about Van Fraassen's pragmatism to elaborate this. Van Fraassen proposed a simple but very effective

[30] This is not derivation in the sense of Kitcher (1981, 1989) or Hempel (1965).

perspective on explanation: "an explanation is not the same as a proposition, or an argument, or a list of propositions; it is an *answer*." (1980, p. 137). Explanations are answers to why-questions. Those questions typically take the form 'Why *P*?', but, according to Van Fraassen, they all have the underlying structure: "Why *P* in contrast to *X*?" (1980, p. 127). The contrast-class is determined by the context, which in turn is defined by a set of background assumptions. Van Fraassen's own example (1980, p.127) makes this clear:

Why P: (a) Why did Adam eat the apple?

In the simplest form "Why X rather than Y?" contrastive questions have two important features: the topic (in this case X) which is taken to be true and the foil or contrast-class (in this case Y) which is taken to be false. In the case of (a) the question can be interpreted in at least three different explanation-seeking questions:

(a) Why was it *Adam* who ate the apple, rather than Eve?

(b) Why was it *the apple* Adam ate, rather than a pear?

(c) Why did Adam *eat* the apple, rather than throwing it away?

Depending on the context, the plain why-question (a) will have the underlying structure of (b), (c) or (d).

I partially adopt these pragmatist ideas, but I also modified them. I accept the idea that explanations are answers to why-questions. But contrary to Van Fraassen, I do not accept the strong claim that all explanation-seeking questions are contrastive questions. In most case studies the explanations that were analyzed were answers to resemblance questions. These questions, that are present in scientific practice, focus on similarities instead of differences. So, I do accept an adapted claim: there are different kinds of explanation-seeking questions, and explanations need to provide a satisfactory answer to those questions. The focus on resemblance questions in this dissertation does not stem from a believe that these are the only kind of explanation-seeking questions. The focus on these questions is directly related to the aim of this dissertation: bringing unification back into the picture.

Several philosophers present convincing case studies to argue for explanation-seeking contexts where unification is irrelevant. Weber et al. have constructed the fictitious example of two neighboring cities, Koch City and Miasma City to illustrate how some explanation-seeking questions require only causal information (Weber, Van Bouwel, &

Vanderbeeken, 2005; Weber & Van Bouwel, 2007). This implies that on top of the pluralist claims already mentioned, I also adopt the following view:

> Depending on the kind of explanation-seeking question, ontological unification can be a virtue, a necessary aim or completely irrelevant for a satisfactory explanation.

A complete account of explanatory pluralism is beyond the scope of this dissertation, for this I can refer to the book 'Scientific Explanation' of Weber, Van Bouwel & De Vreese for some interesting perspectives (2013)[31]. Remember that pluralism was one of the four possible positions in the literature about the relation between unification, causation and explanation. At the end of this dissertation, it becomes clear that I accept an adapted version of position (4), the pluralist approach:

> (4') Ontologically unificationist explanations and causal explanations (with or without embedding of ontological unification) are possible types of explanations among other types. What is the aim of an explanation is contextually determined.

Critics of this 'pragmatic explanatory pluralism' could claim that this approach can only lead to scattered, fragmented results, since they are so context-specific. Indeed, my approach to unification seems very disunifying. The different case studies result in different types of unification. However, they are all instances of ontological unification. Without assuming that science is uniform in its use of unification in explanatory contexts, the result of this research is more than mere particular results. It shows that a general idea, namely that of rethinking unification as an ontological strategy rather than a derivational tool is much more fruitful to get a grip on different scientific practices.

---

[31] I do not completely follow their views on explanation, since they still accept purely derivational unification as explanatory.

8.4 Prospects for future research

At the end of a dissertation one would expect to feel some relief that the work is done. The reality is quite the opposite, as this is just a beginning. With every insight, new questions emerge. So, how could this project to rethink unification be further developed?

One obvious strategy is to increase the inventory of case studies from scientific practice. I have tried to include examples from different scientific disciplines but time is limited. Scientific practice was my guide to find different types of ontological unification. I am convinced that the types brought forward in this dissertation are not the only instances of ontological unification. Exploring other scientific disciplines could result in a wider range of examples and could refine our overall picture of unification.

Another strategy could be to further develop the pragmatic approach to explanation and unification. Acknowledging that science is a human activity and that humans have many different interests paves the way for a pluralist approach to explanation. But there is another feature of pragmatism that remains unexplored here. Andrea Woody (2014) points out that retreating from abstractions clears the decks for questions about the social nature of scientific practice. One line of research could explore how the concept of unification is generated and transmitted, how it is entangled with issues of authority, expertise and trust. To shift the focus from how an individual scientist, or a group of scientists, use unification in their practice to how unification is treated within (philosophy of) science. It could be interesting to investigate how unification was treated as a virtue and/or a vice over the course of history in specific disciplines.

A third strategy could be to investigate the role of unification in other scientific activities. This is the position of Halonen and Hintikka, who claim that unification has nothing to do with explaining, but it is a great criterion for theory choice (1999). Also Margaret Morrison argues that even in the most impressive cases of unification in science, this unification is not explanatory (2000).

Finally, and at least one of my supervisors will frown at me for adding this, one thing that always intrigued me is why human beings are so prone to unification. Even the most passionate defenders of purely causal approaches to explanation admire the unifying power of Newtonian mechanics, Einstein's special relativity, Maxwell's theory of electromagnetism, Darwin's evolution theory etc. Despite all the literature on causal explanations, it is still one of the main aims of fundamental physics to find one unifying theory. Science transcends the mere practical concerns as Hempel points out:

> The second basic motive for man's scientific quest is independent of such practical concerns; it lies in his sheer intellectual curiosity, his deep and persistent desire to know and to understand himself and his world. (1965, p. 333)

But why do we want a unified worldview? Is this urge for unification something that can be found beyond the boundaries of science? Another broader line of philosophical research could start from the human condition itself.

As a final note and future promise, I solemnly swear that next time I am asked if anyone still believes in explanatory unification, I will definitely raise my hand.

Epilogue

The scientific process is less apparent than it sometimes seems. Research does not always start from a clear-cut hypothesis, it is an organic process with ideas, mistakes, doubts and new insights. Often it is only at the end of the process -the writing and rewriting of the dissertation- that everything falls into place. Claiming that this dissertation is a chronological reconstruction of the actual research process would be misleading. And yet, I still remember the moment where the seed for this was planted. It was during a lecture at EPSA 2009 in Amsterdam. The lecture was not particularly interesting (for me), so I started reading Mäki's article from 2001 on explanatory unification. On my notepad I drew two simple schemes. They looked a bit like this:

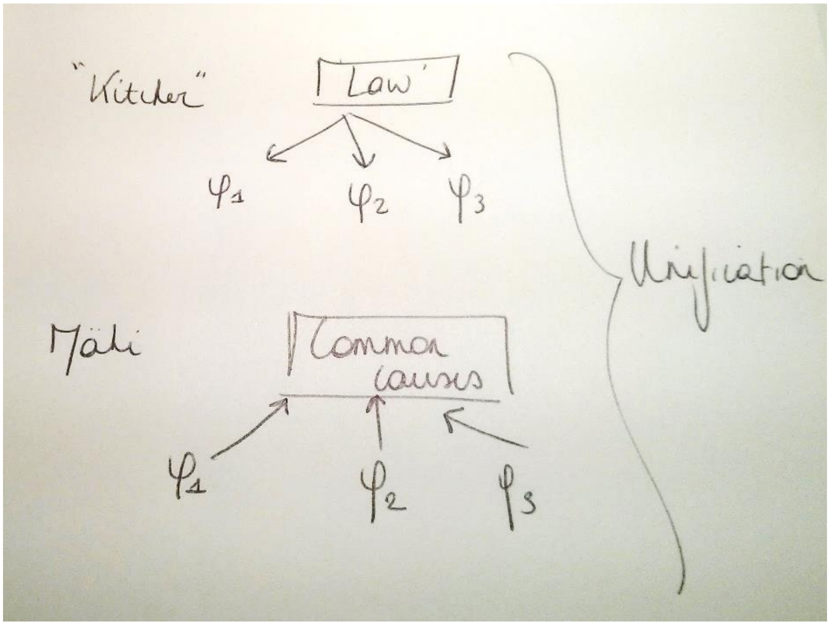

Figure 6 Reconstruction of notes made during EPSA 2009 while reading Mäki's article on explanantory unification (2001) for the first time.

At that moment, my first hypothesis was born: what if the classical direction of unification was turned upside down? A first presentation of the idea was published in the proceedings of EPSA2009 (Weber et al., 2012), there it was presented as bottom-up unification, in contrast to Kitcher-style derivational unification. From that moment on I was triggered to explore scientific practice to find diverse examples of ontological unification. The names I used for what is now called ontological unification would change more than once during my research. But whatever it is called, it led me to an interesting journey along physics, sociology, chemistry, biology and even anesthesia.

Summary in English

Thanks to Philip Kitcher's seminal essay 'Explanatory Unification' (1981) the unification account became a very influential way of thinking about scientific explanations in the 1980s. It was the main rival of Wesley Salmon's causal-mechanical account, which received its full presentation in Salmon's book *Scientific Explanation and the Causal Structure of the World* (1984). In Kitcher 1989, the two approaches are compared; as can be expected, Kitcher argues that the unification account is superior.

In the 1990s, unification was still respected, but often integrated in a pluralistic view. In the new millennium, unification became unfashionable. This was mainly due to the development of two alternative views that became very popular: the counterfactual theory of Jim Woodward (2003) and the mechanistic approach, which started with Machamer, Darden and Craver 2000. and is further elaborated in, e.g., Bechtel & Abrahamsen 2005 and Craver 2007.

Today most philosophers of science consider unification as outdated, irrelevant for explanation or infeasible. The overall aim of this dissertation is *to bring unification back into the picture* in the philosophical study of scientific explanation.

This dissertation starts with a concise overview of what philosophers of science have written about unification and its role in scientific explanation during the last 50 years to provide the reader with some background knowledge (Chapter 1).

In order to bring unification back into the picture, I have followed two strategies, resulting respectively in Parts I and II of this dissertation.

In Part I the idea of unification is used to refine and enrich the dominant causal-mechanist and causal-interventionist accounts of scientific explanation. In this part of the dissertation I bracket the classical ideas about unification: deduction and derivation. I do grant, for the sake of argument, that explanations are causal and argue that unification is important from within this causalist perspective.

In Chapter 2 two examples will be constructed – one about mercury thermometers and one about the ideal gas law – to demonstrate how unificatory information can have a surplus value. Unifying causal explanations can result in a higher explanatory power (this is a first possible virtue) and they can result in more cogent explanations (this is a second possible virtue).

In Chapter 3 I construct two examples from scientific practice – one about social revolutions and one about general anaesthesia – to demonstrate that scientists are interested in similarities between particular facts or between regularities. Answers that aim to provide an adequate explanation for these resemblance questions need both causal information and unification. In this chapter I define two kinds of ontological, non-derivational unification: *causal network unification* and *mechanism unification*.

In Chapter 4 I will clarify some theoretical-philosophical implications of my results. My strategy to search for cases of ontological unification in scientific practice, and to use them to rethink unification, resulted in a new possible relation between unification, causality and explanation: embedding unification in causal accounts of explanation.

In Part II I continue my strategy of digging into scientific practice to find cases of ontological unification. But here I distance myself from the dominant literature that all explanations must be causal. I will investigate whether explanatory unification is possible in non-causal explanations.

In Chapter 5 optimality explanations in biology will be investigated. These are highly generalized explanations about population traits. Since a good optimality model needs to be applicable to the whole population and to future populations of a species, or even to a set of populations, the unifying aspect will not be difficult to uncover. What is at stake here is whether or not this is an example of a causal unifying explanation or a non-causal unifying explanation. In the case of optimality models, it will be argued that the explanation is not causal, but is based on physical dependency relations. These relations supervene on causal relations in the world. I call this *physical dependency unification*.

In Chapter 6 I will use an example of chemical systematization, better known as the periodic table, to show that something unifying is happening in the absence of causal relations. It will be argued that the QM model for the systematization of chemical elements can be explanatory, even in the absence of causal explanations. Furthermore, if one accepts the possibility of non-causal structural explanations, then *structural unification* becomes possible.

The upshot of Chapters 5 and 6 is that some legitimate scientific explanations are unifying without being causal. This is a second way in which I want to bring unification back into the picture.

Part III contains some further reflections and conclusions. During the course of this research I have experienced how the philosophy of explanation is dominated by accounts of causal explanation dominate the literature. Explanatory unification is often considered old-fashioned. However, last year, there was one person, Sorin Bangu, who wrote a paper about unification (2017). The goal of his paper is similar to this dissertation: rethinking unification and bringing it back into the picture. In Chapter 7 I will compare his approach to mine.

In the final chapter I will formulate my primary results, and I will elaborate on their implications for thinking about unification and explanation. The different forms of ontological unification were quite diverse. This relates to the method I have used. Throughout this dissertation the types of unification that were discussed emerged from digging into scientific practice. This philosophy-of-science-practice approach steered me towards a pluralistic view on unification and on explanation.

In this dissertation I do not try to develop a new model of explanation and compare it to existing models. The aim is to show that there are important types of explanatory practice which cannot be properly analyzed if we neglect unification as a desideratum for explanations.

Nederlandse samenvatting

Dankzij Philip Kitchers baanbrekende essay ‘Explanatory Unification’ (1981) werd unificatie een heel invloedrijk model voor wetenschappelijke verklaringen in de jaren 1980. Het was de belangrijkste rivaal van het causaal-mechanistische model van Wesley Salmon, uiteengezet in Salmon’s book *Scientific Explanation and the Causal Structure of the World* (1984). In Kitcher 1989 werden de twee modellen vergeleken, en zoals verwacht, argumenteert Kitcher dat het unificatiemodel superieur is.

In de jaren 1990 werd unificatie nog steeds waardevol geacht, maar het werd meestal geïntegreerd in een pluralistische visie op verklaring. In het nieuwe millennium werd unificatie beschouwd als voorbijgestreefd. Dit is vooral te wijten aan de ontwikkeling van twee alternatieve modellen die enorm populair werden: de tegenfeitelijke theorie van Jim Woodward (2003) en de mechanistische benadering die begon met Machamar, Darden en Craver 2000 en verder uitgewerkt werd door o.a. Bechtel & Abrahamsen 2005 en Craver 2007.

Vandaag beschouwen de meeste wetenschapsfilosofen unificatie als ouderwets, irrelevant of onbereikbaar. Het hoofddoel van dit proefschrift is om unificatie terug onder de aandacht te brengen in het filosofisch onderzoek naar wetenschappelijke verklaringen.

Dit proefschrift start met een beknopt overzicht van hoe wetenschapsfilosofen unificatie en haar rol in wetenschappelijke verklaringen benaderd hebben tijdens de laatste 50 jaar.

Om unificatie terug onder de aandacht te brengen heb ik twee strategieën gevolgd, die resulteren in respectievelijk Deel I en Deel II van dit proefschrift.

In Deel I werd het idee van unificatie gebruikt om de dominante verklaringsmodellen, namelijk de mechanistische en causaal-interventionistische benaderingen, te verfijnen en te verrijken. In dit deel neem ik (voorlopig) aan dat alle verklaringen causaal zijn. Het hoofdargument van dit deel is bijgevolg dat unificatie ook binnen deze causale modellen belangrijk is. Wat ik echter verwerp zijn is het traditionele idee dat unificatie deductief en derivationeel is.

In Hoofdstuk 2 worden twee voorbeelden uitgewerkt – een over kwikthermometers en een over de ideale gaswet – om aan te tonen hoe unificatie een meerwaarde kan bieden. Het unificeren van causale verklaringen kan bijdragen tot een hogere verklaringskracht en het kan zorgen voor meer overtuigende verklaringen.

In Hoofdstuk 3 worden twee voorbeelden uit de wetenschapspraktijk ontwikkeld – een over sociale revoluties en een over algemene verdoving – om aan te tonen dat wetenschappers geïnteresseerd zijn in gelijkenissen tussen feiten of regelmatigheden. Dit soort vragen worden gelijkenisvragen genoemd. Antwoorden die een adequate verklaring vormen voor deze vragen dienen zowel causale informatie als unificatie te bevatten. In dit hoofdstuk definieer ik twee ontologische, niet-derivationele vormen van unificatie: *causale netwerk unificatie* en *mechanisme unificatie*.

In Hoofdstuk 4 worden enkele theoretisch-filosofische implicaties van de resultaten uit de vorige hoofdstukken toegelicht. De gehanteerde strategie, namelijk zoeken naar voorbeelden van ontologische unificatie in de wetenschapspraktijk, en deze gebruiken om unificatie te heroverwegen, resulteert in een nieuwe mogelijke relatie tussen unificatie, causaliteit en verklaring: unificatie inbedden in causale benaderingen van verklaringen.

In Deel II blijf ik zoeken naar voorbeelden van ontologische unificatie in de wetenschapspraktijk. Hier distantieer ik me niet alleen van de klassieke visie op unificatie, maar ook van de dominante visie in de literatuur dat alle verklaringen causaal moeten zijn. Ik zal onderzoeken of verklarende unificatie mogelijk is bij niet-causale verklaringen.

In Hoofdstuk 5 worden optimaliteitsverklaringen in de biologie onderzocht. Dit zijn sterk gegeneraliseerde verklaringen over populatiekenmerken. Een goed optimaliteitsmodel moet toepasbaar zijn op de hele populatie en op toekomstige populaties van een soort, of verzameling van soorten. Het unificerende aspect van dergelijke verklaringsmodellen is dus niet moeilijk te ontdekken. Wat hier op het spel staat is of deze modellen ook causale verklaringen zijn. In het geval van optimaliteitsmodellen wordt er geargumenteerd dat de verklaring zelf niet causaal is, maar gebaseerd is op fysieke afhankelijkheidsrelaties. Deze relaties superveniëren evenwel op causale relaties in de wereld. Deze vorm van unificatie zal ik *fysieke afhankelijkheidsunificatie* noemen.

In Hoofdstuk 6 gebruik ik de systematisering van de chemische elementen, beter bekent als de periodieke tabel, om aan te tonen dat er iets unificerends gebeurt in de afwezigheid van causale relaties. Er wordt geargumenteerd dat kwantummechanische modellen van de periodieke tabel verklarend kunnen zijn, zonder dat ze causaal zijn. Als deze structurele verklaringen aanvaardbaar zijn, dan wordt een vierde soort ontologische unificatie mogelijk: *structurele unificatie*.

Het resultaat van Hoofdstuk 5 en 6 is dat er legitieme wetenschappelijke verklaringen bestaan die unificeren zonder causaal te zijn. Dit is een tweede manier om unificatie terug in beeld te brengen.

Deel III bevat verdere reflecties en conclusies. Tijdens het onderzoek heb ik ervaren hoe dominant de causale benaderingen van verklaringen zijn in de wetenschapsfilosofische literatuur. Unificatie wordt vaak gezien als ouderwets. Toch was er vorig jaar een persoon, Sorin Bangu, die een paper publiceerde over unificatie (2017). Het doel van zijn paper is gelijkaardig aan het doel van dit proefschrift: unificatie heroverwegen en terug onder de aandacht brengen. In Hoofdstuk 7 vergelijk ik zijn benadering van unificatie met mijn visie.

In het finale hoofdstuk formuleer ik de primaire resultaten en werk ik hun implicaties voor het benaderen van unificatie en verklaring uit. De vier vormen van ontologische unificatie waren nogal divers. Dit kan toegeschreven worden aan de gehanteerde methode. De verschillende soorten unificatie die besproken werden in dit proefschrift zijn het resultaat van onderzoek gestuurd door bevindingen uit de wetenschapspraktijk zelf. Deze filosofie-van-de-wetenschapspraktijk-benadering stuurde me naar een pluralistische visie op unificatie en op verklaring.

In dit proefschrift heb ik niet geprobeerd om een nieuw verklaringsmodel te ontwikkelen en dit te vergelijken met bestaande modellen. Het doel was om aan te tonen dat er belangrijke types van verklarende praktijken bestaan die niet naar behoren geanalyseerd kunnen worden als we unificatie negeren als een desideratum voor verklaringen.

Bibliography

Ajdukiewicz, K. (1974). *Pragmatic Logic.* Dordrecht: D.Reidel Publishing Company.

Anderson, G., Gomatam, R., & Behera, L. (2014). Contradictions in the quantum mechanical explanation of the periodic table. *2nd International Conference on Mathematical Modeling in Physical Sciences 2013 Journal of Physics: Conference Series*, *490.*

Aumann, R. (1985). What is game theory trying to accomplish? In K. Arrow, & S. Honkapohja, *Frontiers of economics.* Oxford: Blackwell.

Bangu, S. (2017). Scientific explanation and understanding: unificationism. *European Journal for Philosophy of Science*(7), 103-126.

Bechtel, W. (2009). Generalization and disovery by assuming conserved mechanisms: cross-species research on circadian oscillators. *Philosophy of Science, 76*, 762-773.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: a mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 421-441.

Brock, W. (1992). *The Norton History of Chemistry.* New York: Norton.

Bromberger, S. (1966). Why questions. In R. G. Colodny, *Mind and cosmos.* Pittsburgh, PA: University of Pittsburgh Press.

Brukner, C. (2014). Quantum causality. *Nature Physics, 10*, 259-263.

Brush, S. G. (1996). The reception of Mendeleev's periodic law in America en Britain. *Isis, Vol.87*(No.4), 595-628.

Bull, J., Pfennig, D. W., & Wang, I. N. (2004). Genetic details, optimization and phage life histories. *Trends in Ecology and Evolution, 19*(2), 76-82.

Cartwright, N. (1983). *How the laws of physics lie.* Oxford: Clarendon Press.

Chang, R. (2000). *Physical Chemistery for the Chemical and Biological Sciences.* Sausalito, CA: University Science Books.

Chang, R. (2007). *Chemistry* (10th ed.). McGrawHill Education.

Craver, C. F. (2007). *Explaining the brain.* Oxford: Clarendon Press.

Darden, L. (2006). *Reasoningin in Biological Discoveries.* Cambridge: Cambridge University Press.

de Regt, H. W. (2017). *Understanding Scientific Understanding.* Oxford: Oxford University Press.

de Regt, H. W., & Dieks, D. (2005). A contextual approach to scientific understanding. *Synthese, 144*, 137-170.

de Regt, H. W., Leonelli, S., & Eigner, K. (Eds.). (2009). *Scientific Understanding: philosophical perspectives.* Pittsburg: University of Pittsburg Press.

Döbereiner, J. W. (1829). Versuch zu einer Gruppirung der elementaren Stoffe nach ihrer Analogie. *Annalen der Physik und Chemie, 15*, 301-307.

Eckenhoff, R. G. (2001). Promiscuous ligands and attractive cavities: how do the inhaled anaesthetics work? *Molecular Interventions, 1*(5), 258-268.

Eger II, E. I., Gong, D. X., Raines, D., & Flood, P. (2002). Acetylcholine receptors and thresholds for convulsions from flurothyl and 1,2-dichlorohexafluorocyclobutane. *Anesthesia & Analgesia, 95*, 1611-1615.

Eger II, E. I., Westhorpe, R., & Saidman, L. (2014). 1846-1860: Following the discovery of anaesthesia. In E. I. Eger II, R. Westhorpe, & L. Saidman (Eds.), *The Wondrous Story of Anaesthesia* (pp. 27-35). New York: Springer.

Eger II, E., Saidman, L., & Westhorpe, R. (2014). 1860–1910: The Specialty of Anesthesia Develops Slowly. In E. I. Eger II, L. Saidman, & R. Westhorpe (Eds.), *The Wondrous Story of Anesthesia* (pp. 37-49). New York: Springer.

Eger II, E., Saidman, L., & Westhorpe, R. (Eds.). (2014). *The Wondrous Story of Anaesthesia.* New York: Springer.

Fernelius, W. C. (1986). Some reflections on the periodic table and its use. *Journal of chemical education, Vol.63*(No.3), 263-266.

Feynman, R., Leighton, R., & Sands, M. (2010). *The Feynman lectures on physics (the new millennium edition).* Reading: Addison-Wesley Publishing Company.

Fisher, R. (1930). *The genetical theory of natural selection.* Oxford: Clarendon Press.

Franks, N. P. (2014). The unfolding story of how general anaesthetics act. In E. I. Eger II, L. Saidman, & R. Westhorpe (Eds.), *The Wondrous Story of Anaesthesia* (pp. 597-608). New York: Springer.

Franks, N. P., & Lieb, W. R. (1982). Molecular mechanisms of general anaesthesia. *Nature, 300*, 487-493.

Franks, N. P., & Lieb, W. R. (1984). Do general anaesthetics act by competitive binding to specific receptors? *Nature, 310*, 599-601.

Franks, N. P., & Lieb, W. R. (1994). Molecular and cellular mechanisms of general anaesthesia. *Nature, 367*, 607-614.

Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy, 71*(1), 5-19.

Glennan, S. (2009). Productivity, relevance and natural selection. *Biology and Philosophy, 24*, 325-339.

Glennan, S. (2017). *The new mechanical philosophy.* Oxford: Oxford University Press.

Griffiths, D. J. (2005). *Introduction to quantum mechanics* (2nd ed.). Upper Sadle River, NJ: Pearson Education International.

Halonen, I., & Hintikka, J. (1999). Unification - it's magnificent but is it explanation? *Synthese, 120*, 27-47.

Hausman, D. M. (1998). *Causal asymmetries.* Cambridge: Cambridge University Press.

Hempel, C. G. (1962). Deductive-nomological vs. statistical explanation. In H. Feigl, & G. Maxwell (Eds.), *Minnesota Studies in the Philoosphy of Science vol.3* (pp. 98-169). Minneapolis: University of Minnesota Press.

Hempel, C. G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: Free Press.

Hempel, C. G. (1966). *Philosophy of Natural Science.* NJ: Prentice Hall.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science, 15*, 135-175.

Hettema, H., & Kuipers, T. A. (1988). The periodic table: its formalization, status and relation to atomic theory. *Erkenntnis, Vol.28*(No.3), 387-408.

Hitchcock, C., & Woodward, J. (2003). Explanatory generalizations, Part II: plumbing explanatory depth. *Noûs, 37*(2), 181-199.

Humphreys, P. (1989). *The chances of explanation.* New York: Princeton University Press.

Ihde, A. (1969). Foreword. In J. van Spronson, *The periodic system of chemical elements: a history of the first hundred years.* Amsterdam: Elsevier.

Illari, P., & Williamson, J. (2012). What is a mechanism? Thinking about Mechanisms across the Sciences. *European Journal for Philosophy of Science, 2*, 119-135.

John, E. R. (2001). The neurophysics of consciousness. *Consciousness and Cognition, 10*, 184-213.

John, E. R., Prichep, L. S., Kox, W., Valdés-Sosa, P., Bosch-Bayard, J., Aubert, E., . . . Gugino, L. D. (2001). Invariant reversible QEEG effects of anaesthetics. *Consciousness and Cognition, 10*(2), 165-183.

Kelter, P., Mosher, M., & Scott, A. (2009). *Chemistry: the practical science.* Boston: Houghton Mifflin Company.

Kirby, J. (2002). The influence of parental seperation on smoking initiation in adolescents. *Journal of Health and Social Behavior, 43*, 56-71.

Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science, 48*, 507-531.

Kitcher, P. (1985). Two approaches to explanation. *Journal of Philosophy, 82*, 632-639.

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher, & W. Salmon, *Scientific Explanation* (pp. 410-505). Minneapolis: University of Minnesota Press.

Kitcher, P. (1993). *The Advancement of Science: Science Without Legend, Objectivity Without Illusions.* New York: Oxford University Press.

Kitcher, P., & Salmon, W. (1989). *Scientific Explanation* (Vol. Minnesota studies in the Philosophy of Science Voclume XIII). (R. N. Giere, Ed.) Minnesota: University of Minnesota Press.

Kuhlman, M., & Glennan, S. (2014). On the relation between quantum mechanical and neo-mechanistic ontologies and explanatory strategies. *European Journal for Philosophy of Science, 4*, 337-359.

Kuipers, T. (2001). *Structures in science.* Dordrecht: Kluwer Academic Publishers.

Kyburg Jr., H. E. (1965). Salmon's paper. *Philosophy of science, 32*(2), 147-151.

Lange, M. (2017). *Because without cause.* New York: Oxford University Press.

Leach, M. R. (2018, 05 21). *1869 Mendeleev's Tabelle I.* Retrieved from The Internet Database of Periodic Tables: https://www.meta-synthesis.com/webbook/35_pt/pt_database.php?PT_id=9

Levy, A., & Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of Science, 80*(2), 241-261.

Lewis, D. (1986). Causal explanations. In D. Lewis, *Philosophical Papers Volume II* (pp. 214-240). Oxford: Blackwell Publishers.

Lipton, P. (1993). Making a difference. *Philosophica, 51*, 39-54.

Ma, J., Shen, B., Stewart, L. S., Herrick, I. A., & Leung, L. S. (2002). The septohippocampal system participates in general anesthesia. *The Journal of Neuroscience, 22*(2), 1-6.

Machamar, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*(1), 1-25.

Maki, U. (2001). Explanatory unification: double and doubtful. *Philosophy of the social sciences, 31*(4), 488-506.

Mashour, G. A. (2004). Consciousness unbound. Towards a paradigm of anaesthesia. *Anesthesiology, 100*, 428-433.

McKay Illari, P., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciencess. *European Journal for Philosophy of Science, 2*(1), 119-135.

Mendeleev, D. I. (1869). Sootnoshenie svoistv s atomnym vesom elementov. *Zhurnal Russkeo Fiziko-Khimicheskoe Obshchestv, 1*, 60-77.

Meyer, H. (1899). Welche eigenschaft der anasthetica bedingt inre Narkotische wirkung? *Archiv for Experimentelle Pathologie und Pharmakologie, 42*(2-4), 109-118.

Meyers, R. (2003). *The basics of chemistry.* Westport: Greenwood Press.

Morrison, M. (2000). *Unifying Scientific Theories.* Cambridge: Cambridge University Press.

Moseley, H. (1913). The attainment of high potentials by the use of Radium. *Proceedings of the Royal Society, 88*(605), 471-476.

Nagel, E. (1961). *The structure of science.* Indianapolis: Hackett.

Newton-Smith, W. H. (2000). Explanation. In W. H. Newton-Smith, *A compantion to the philosophy of science* (pp. 127-133). Oxford: Blackwell.

Orzack, S., & Sober, E. (1994). Optimality models and the test of adaptionism. *The American Naturalist, 143*(3), 361.

Ostrovsky, V. N. (2001). What and how physics contributes to understanding the periodic law. *Foundations of Chemistry, 3*, 145-182.

Ostrovsky, V. N. (2006). Physical explanation of the periodic table. *Annals of the New York Academy of Sciences, 988*(1), 182-192.

Overton, C. E. (1901). *Studien über die Narkose zugleich ein Beitrag zur allgemeinen Pharmakologie.* Meisterschwanden: Gustav Fischer.

Parker, G., & Stuart, R. (1976). Animal behavior as a strategy optimizer: evolution of resource assessment strategies and optimal emigration thresholds. *The American NAturalist, 110*(976), 1055-1076.

Pearl, J. (2009). *Causality* (2nd ed.). New York: Cambridge University Press.

Pearl, J. (2009). *Causality: Models, Reasoning and Inference* (2nd ed.). Cambridge: Cambridge University Press.

Perrin, J. (1901). Les hypothèses moléculaires. *Revue Scientifique 4th series, 15*, 449-461.

Philips, A. C. (2003). *Introduction to Quantum Mechanics.* Hoboken (New Jersey): Wiley.

Plotnitsky, A. (2010). On physical and mathematical causality in quantum mechanics. *Physica E: Low-dimensional Systems and Nanostructures, 42*(3), 279-286.

Potochnik, A. (2007). Optimality modeling and explanatory generality. In C. Bicchieri, & J. Alexander (Ed.), *Proceedings of the 2006 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers. 74 (5)*, pp. 680-691. The University of Chicago Press.

Potochnik, A. (2010). Explanatory independence and epistemic interdependence: a case study of the optimality approach. *The British Journal for Philosophy of Science, 61*(1), 213-233.

Rice, C. (2012). Optimality explanations: a plea for an alternative approach. *Biology and Philosophy, 27*(5), 685-703.

Rice, C. (2015). Moving beyond causes: optimality models and scientific explanation. *Noûs, 49*(3), 589-615.

Riggs, P. J. (2009). The causal theory of quantum mechaics. In *Quantum Causality. Studies in History and Philosophy of Science.* Dordrecht: Springer.

Ringbauer, M., Giarmatzi, C., Chaves, R., Costa, F., White, A. G., & Fedrizzi, A. (2016). Experimental test of nonlocal causality. *Science Advances, 2*(8). doi:10.1126/sciadv.1600162

Robinson, D., & Toledo, A. (2012). Historical development of modern anaesthesia. *Journal of Investigative Surgery*, 141-149.

Ross, L. (2017). Finding causal structure. preprint. Retrieved 05 22, 2018, from http://philsci-archive.pitt.edu/13258/

Salmon, W. (1989). Four decades of scientific explanation. In P. Kitcher, & W. Salmon (Eds.), *Scientific Explanation: Minnesota Studies in the Philosophy of Science (Vol. XIII)* (pp. 3-219). Minneapolis: University of Minnesota Press.

Salmon, W. C. (1971). Statistical explanation. In W. C. Salmon, *Statistical explanation and statistical relevance* (pp. 29-87). Pittsburgh: Pittsburgh University Press.

Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World.* Princeton, NY: Princeton University Press.

Salmon, W. C. (1993). The value of scientific understanding. *Philosophica, 51*, 9-19.

Salmon, W. C. (1998). *Causality and Explanation.* Oxford: Oxford University Press.

Scerri, E. (1997). The periodic table and the electron. *American Scientist*, 546-553.

Scerri, E. (1998). How Good Is the Quantum Mechanical Explanation. *Journal of Chemical Education, 75*(11), 1384-1385.

Scerri, E. (2007). *The periodic table: its story and its significance.* New York: Oxford University Press.

Scerri, E. (2012). Mendeleev's Periodic Table Is Finally Completed and What To Do about Group 3? *Chemistry International, 34*(4), 28-31.

Schurz, G. (1999). Explanation as unification. *Synthese, 120*, 95-114.

Schurz, G. (2014). Unification and explanation: unification as a protoype concept. *Theoria, 79*, 57-70.

Schurz, G., & Lambert, K. (1994). Outline of a theory of scientific understanding. *Synthese, 101*(1), 65-120.

Schwarz, W., & Rich, R. L. (2010). Theoretical basis and correct explanation of the periodic system: review and update. *Journal of Chemical Education, 87*(4), 435-443.

Scriven, M. (1962). Comments on professor Grunbaum's remarks at the Wesleyan Meeting. *Philosophy of Science, 29*(2), 171-174.

Skipper, R. A. (1999). Selection and the extent of explanatory unification. *Philosophy of Science, 66*, S196-S209.

Skocpol, T. (1979). *States and Social Revolutions.* Cambridge: Cambridge University Press.

Sober, E. (1983). Equilibrium explanation. *Philosophical Studies, 43*, 201-210.

Sober, E. (1995). Natural selection and distributive explanation: a reply to Neander. *British Journal for the Philosophy of Science, 46*, 384-397.

Sober, E. (2000). *Philosophy of Biology* (2nd ed.). Boulder: Westview Press.

Strevens, M. (2004). The Causal and Unification Approaches to Explanation Unified—Causally. *Noûs*, 154-176.

Strevens, M. (2008). *Depth: an account of scientific explanation.* Cambridge: Harvard University Press.

Taylor, M. (1988). Rationality and revolutionary collective action. In M. Taylor, *Rationality and Revolution* (pp. 63-97). Cambridge: Cambridge University Press.

Uchida, S., Nakayama, H., Maehara, T., Hirai, N., Arakaki, H., & Nakamura, M. (2000). .Suppressionof gamma activity in the human medial temporal lobe by sevoflurane anaesthesia. *NeuroReport, 11*(1), 39-42.

van Fraasen, B. C. (1980). *The scientific image.* Oxford: Clarendon Press.

Van Fraassen, B. (1980). The pragmatic theory of explanation. In B. Van Fraasse, *The Scientific Image* (pp. 136-155). Oxford: Oxford University Press.

von Bibra, E., & Harless, E. (1847). Chemical psychiological theory of the action of inhaled ether (excperts). In *Die Werkung des Schwefelaethers in chemischer und physiologischer Beziehung* (B. Fink, Trans., pp. 157-165). Erlangen: Verlag von Carl Heyder.

Wang, I. N., Dykhuizen, D. E., & Slobodkin, L. B. (1996). The evolution of phage lysis timing. *Evolutionary Ecology, 10*, 545-558.

Weber, E. (1999). Unification: what is it, how do we reach and why do we want it? *Synthese*, 479-499.

Weber, E., & Lefevere, M. (2014). The role of unification in micro-explanations of physical laws. *THEORIA. An International Journal for Theory, History and Foundations of Science,, 29*(1), 41-56.

Weber, E., & Lefevere, M. (2017). Unification, the answer to resemblance questions. *Synthese, 194*(9), 3501 - 3521.

Weber, E., & Van Bouwel, J. (2007). Assessing the explanatory power of causal explanations. In J. Persson, & P. Ylikoski (Eds.), *Rethinking explanation* (pp. 109-118). Springer.

Weber, E., & Van Bouwel, J. (2009). Causation, unification and the adequacy of singular explanations. *Theoria, 24*(3), 301-320.

Weber, E., Van Bouwel, J., & De Vreese, L. (2013). *Scientific Explanation.* Dordrecht: Springer.

Weber, E., Van Bouwel, J., & Lefevere, M. (2012). The Role of Unification in Explanations of Facts. In W. H. de Regt, S. Hartmann, & S. Okasha (Ed.), *EPSA Philosophy of Science: Amsterdam 2009* (pp. 403-413). Dordrecht: Springer Netherlands.

Weber, E., Van Bouwel, J., & Vanderbeeken, R. (2005). Forms of causal explanation. *Foundations of science, 10*(4), 437-454.

Weber, E., van Eck, D., & Mennes , J. (201+). On the structure and epistemic value of function ascriptions in biology and engineering sciences. *in review.*

Weller, M., Overton, T., Rourke, J., & Armstrong, F. (2014). *Inorganic chemistry* (6th ed.). Oxford: Oxford University Press.

White, P. F. (2014). A history of intravenous anaesthesia. In E. Eger II, R. Westhorpe, & L. Saidman (Eds.), *The Wondrous Story of Anaesthesia* (pp. 629-639). New York: Springer.

Winter, B. K. (2017, 05 17). *Causality and quantum theory.* Retrieved from Arvix.org: arXiv:1705.07201

Woodward, J. (2003). *Making things happen: a theory of causal explanation.* Oxford: Oxford University Press.

Woodward, J., & Hitchcock, C. (2003). Explanatory generalizations, Part I: a counterfactual account. *Noûs, 37*(1), 1-24.

Woody, A. I. (2014). Chemistry's Periodic Law. Rethinking representation and explanation after the turn to practice. In L. Soler, S. Zwart, M. Lynch, & V. Israel-Jost (Eds.), *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science* (pp. 123-150). New York: Routledge.