Analogical reasoning in science education
– connections to semantics and scientific modelling in thermodynamics

Jesper Haglund
Studies in Science and Technology Education (FontD)

Distributed by:
The Swedish National Graduate School in Science and Technology Education (FontD)
The Department of Social and Welfare Studies (ISV)
Linköping University
S-60174 Norrköping
Sweden

Jesper Haglund (2012)

Analogical reasoning in science education – connections to semantics and scientific modelling in thermodynamics.

Cover image: First-grader Lisa’s self-generated analogies for heat.

ISSN: 1652-5051


Copyright: Jesper Haglund

Printed by: Liu-Tryck, Linköping University, Linköping, Sweden, 2012
Abstract

Analogical reasoning is a central cognitive ability that is used in our everyday lives, as well as in formal settings, such as in research and teaching. This dissertation concerns how analogies and analogical reasoning, attention to semantics and insight into scientific modelling may be recruited in order to come to terms with challenges in science education, in particular within the field of thermodynamics. In addition, it provides a theoretical framework of how analogy relates to semantics and the practice of scientific modelling, three fields of study which all strive to map correspondences between two different domains. In particular, the dissertation addresses the following research questions: To what degree is analogy involved in connecting different representations of a phenomenon to each other and to the represented phenomenon? How do students’ self-generated analogies relate to the practice of scientific modelling?

The dissertation comprises four published journal articles and a ‘cover story’. The first article is a semantic investigation of the word ‘entropy’, the second article is an empirical study of the view on scientific modelling in different traditions of knowledge, and the third and fourth articles are empirical studies of self-generated analogies for thermal phenomena among preservice physics teachers and first-graders, respectively. From a methodological point of view, the empirical studies were conducted in a primarily qualitative tradition, where central lines of reasoning are exemplified by analysis of dialogue excerpts. The two studies on self-generated analogies provided the participants with extensive scaffolding in the form of social interaction among peers, interaction with physical phenomena and discussion of their representations of the phenomena. The theoretical framework is developed in the cover story, which provides a background to the individual studies and reanalyses of the findings.

A key claim of the dissertation is that any phenomenon can be represented in many different ways, all potentially adequate and useful in different contexts, emphasising different aspects of the phenomenon. Applied to the field of analogical reasoning, it is argued that students can generate several analogies themselves in order to get a richer, complementary view of a phenomenon, as opposed to be provided with a presumed best analogy. As for scientific models, many different representations or models may bring across different aspects of a phenomenon at varying degrees of idealisation and within different traditions of knowledge. Finally, in semantics, one word may correspond to several distinct, yet related, meanings: the phenomenon of polysemy. These three perspectives may provide constructivist approaches to conceptual development in science teaching, in which students are encouraged to connect to and enrich their everyday understanding of encountered concepts and phenomena in dialogue, rather than merely abandoning them for one single, supposedly correct, scientific concept. In addition, science education research can come quite far with structural approaches to analysing analogical reasoning and scientific modelling, establishing correspondences between entities in different domains, ultimately striving for isomorphism, perfect matches. However, other dimensions, such as the perceptual, embodied nature of our cognition, the pragmatic, contextual circumstances in which any act of reasoning is performed, and the specificities of language, should be taken into account for a fuller view.
Preface

This dissertation has been written as part of the research programme *Significance of the Representational Forms for Learning Science*, at Campus Norrköping, Linköping University, within the Swedish National Graduate School in Science and Technology Education (FontD).

One can approach the endeavour of postgraduate studies and the project of writing a doctorate dissertation in a number of ways. Högskoleverket (the Swedish National Agency for Higher Education) (2012) provides the following two broad alternative routes:

Public debate about the scope of a dissertation has a long history, and has been full of vicissitudes. Should it represent a life’s work or be part of a programme of training and a first relatively comprehensive research assignment? The latter view has come to dominate the discussion and the recent reform in postgraduate training emphasizes that what is involved is a programme of education that should be completed within a relatively limited period of time. The PhD is a kind of journeyman’s certificate, evidence that the postgraduate has the capacity to conduct research.

I have predominantly adopted the latter view, in which the doctoral study is a process of enculturation or socialisation, of taking part gradually more actively and preparing to become a member in a culture, in this case the community of science education research. In line with this view, my ambition has been to experience as many facets as possible of the professional life of research. Arguably, the most important aspect of academic life is to publish one’s findings and therefore I have strived to enter the competitive scene of publishing in academic journals.

They say that the proper response to ‘how to eat an elephant?’ is ‘one bite at a time’. In this vein, apart from the appeal to graduate studies as a process of socialisation into the research community, the PhD by publication approach has also provided a means to chop up the writing of the dissertation into a set of standalone manuscripts. This approach, however, comes along with two connected challenges: Making the dissertation a coherent whole; and, reaching a sufficient theoretical depth. In the present dissertation, the introductory ‘cover story’ (‘kappa’ in Swedish, meaning ‘coat’) serves the overall purpose of framing the included articles against the background of fundamental theories and previous research, showing how the articles relate to each other and constitute a larger whole. As a way to confront the two challenges, the ambition in my cover story is to go beyond a comprehensive summary of the articles. First, I have striven towards making deeper investigations of the key concepts and how they are related to each other than what is permitted in the typical journal article format. This may be characterised as an extended and updated literature review, but the resulting theoretical framework, depicted in Figure 2, also aspires to be a synthesis, as applied to the field of science education. Next, following the approach of my fellow doctorate student (now PhD) Karin Stolpe, particularly rich empirical data have been selected from the four articles and reanalysed as examples from all three perspectives in the theoretical framework. The revisiting of the content in the individual articles also serves the function of an exegesis – in adding a new layer of interpretation to the findings in these studies – bearing in mind the time that has passed since the studies were conducted and the manuscripts were written.

Science education can be seen as positioned in the intersection of many academic disciplines and aspects of the society at large, including natural sciences proper, the history, philosophy and language of science, educational psychology, theories of learning and teaching, school as an institution. In my graduate studies, I have come to take an interest in a broad range of subjects, where thermodynamics, semantics, analogical reasoning and scientific modelling may be seen as keywords that span the studies in this dissertation (see Figures 1 and 2). I saw the potential of thermodynamics in science teaching when working as an upper secondary teacher in physics prior to my graduate studies, since it stood out as
particularly well suited for integrated teaching across physics and chemistry, but also reaching out to the development of technology and the society as a whole. This interest fitted well with the research in thermodynamics teaching already initiated by Helge Strömdahl and Fredrik Jeppsson when I became a graduate student. An attraction to the particular concept ‘entropy’ has followed along the way, partly originating from reading Tor Nørretrander’s book *Märk världen* in the mid-1990s, which showed how entropy can be applied to different disciplines within and outside the natural sciences.\(^1\) I have nourished a general interest in language for long, and it has been deepened, particularly as applied to thermodynamics, after I started the graduate studies. The road into the world of analogies opened with Richard Hirsch’s introduction to cognitive linguistics, in which analogy may be characterised as the cognitive substrate of metaphor, while scientific models, particularly as expressed in visual external representations, entered the scene with the research of Lena Tibell and Konrad Schönborn in the field.

Apart from the studies included in the dissertation, I have also contributed to research in other related areas. This includes studies on deliberate instructional metaphors for entropy (Jeppsson, Haglund, & Strömdahl, 2011), implicit conceptual metaphors used in relation to entropy in textbooks (Amin, Jeppsson, Haglund, & Strömdahl, 2012) and in problem solving exercises (Jeppsson, Haglund, Amin, & Strömdahl, 2012), conceptual understanding of thermal phenomena as expressed with self-generated analogies among physics teacher students (Haglund & Jeppsson, in progress) and among first-graders (Haglund, Jeppsson, & Andersson, in progress), further investigation of the issue of reference of thermodynamics concepts (Strömdahl, Haglund, & Jeppsson, in progress), and the use of thermoimaging for secondary teaching in thermodynamics (Schönborn, Haglund, & Xie, in review). Some of these themes are touched upon briefly in the cover story.

**Acknowledgements**

Alison Lee (2010) has looked upon the Scandinavian phenomenon of ‘PhD by publication’, i.e. dissertations in the form of a compilation of journal articles, also in the social sciences, by interviewing a Swedish PhD student and her supervisor throughout her doctoral studies, from the perspective of how to get the required quality in the research and subsequent reporting of the findings. Lee finds two institutionalised activities to be of utmost importance: *supervision*, and *seminars*, ranging from informal work-in-progress sessions to more formal toll-gate presentations. Lee and the interviewed supervisor particularly emphasise the role of the work-in-progress seminars, in which you are forced to formulate your ideas and open up for comments, necessary critique and proposals for further development. Having identified the importance of supervision and seminars with the research group, I would like to take the opportunity to thank my main supervisor and academic father Helge Strömdahl. Central lessons from you have been to emphasise the particularities of science content and semantics, to dare diving into challenging philosophical issues, to chisel out the core message of each study and to reach out to an international audience. Further I would like to thank my co-supervisors Roland Kjellander at the Department of Chemistry, Gothenburg University, Shu-Nu Chang Rundgren at the Department of Chemistry and Biomedical Sciences, Karlstad University and Konrad Schönborn at the Department of Science and Technology, Linköping University. Roland, as an academic uncle, you have been an inspiring example in focusing on getting the facts right, producing quality research rather than spending time on the tactics of how to reach out with it, and seeing teaching as an integral part of academic life. Shu-Nu, together with Carl-Johan Rundgren, you have served the role of my academic older siblings,

---

\(^1\) Published as *The user illusion: cutting consciousness down to size* (Nørretranders, 1998) in English.
always willing to give a helping hand in times of frustration in the life of academia. Konrad, your suggestions on how to structure texts, attention to detail and wise research advice have been much appreciated throughout my doctorate studies. Of the colleagues at the research group within FontD and later TekNaD, towards whom I am collectively grateful, Fredrik Jeppsson, my brother in arms, has played an essential role in the work towards this dissertation as a close collaborator and being there to bounce ideas pretty much on an hourly basis, and Johanna Andersson brought in much sought-after experience in science teaching among younger children. In addition, Anna Ericson, with your helpful and flexible administrative support, you have been the steady rock of our group in times of change.

While I share Lee’s (2010) view of the importance of the local environment, the wider networks within the research community should not be forgotten. In this regard, my dissertation work has greatly benefited from the three courses and the virtual network of the national graduate school, FontD. The courses – organised and run by Lena Tibell, Konrad Schönborn, and Shu-Nu Chang Rundgren – were very practical in orientation and geared towards supporting our first modest stabs at conducting research in a constructive environment. As for the network, interaction with the Karlstad University contingent has been particularly rewarding, not least with Michal Dreschler, who reviewed my work at the 60 % tollgate, and Margareta Enghag, who helped us with the ownership construct. The collaboration with other graduate schools in Finland, Germany and the Netherlands has also given the opportunity to present and receive feedback on my research in a semi-formal setting, where input from Ismo Koponen at the University of Helsinki has been much appreciated. Finally, the scene of international conferences has brought the opportunity to test ideas in the international community, and, most importantly, to make contacts with researchers in the field. A very fruitful collaboration with Tamer Amin at the Lebanese American University in Beirut – bringing along a wealth of learnedness in science education and related areas, relentless scrutiny of manuscripts and a good spirit – including reviewing this dissertation at the 90 % tollgate, came about through such conference encounters, and the contact with Risto Leinonen at the University of Eastern Finland, Joensuu, contributed to homing in on the concept of entropy. Taking graduate courses has also been a way to new worlds of knowledge and to put the research on a more rigorous footing. Here, Richard Hirsch and Fredrik Stjernberg at the Department of Culture and Communication, Linköping University contributed in the fields of linguistics and the philosophy of science, respectively, and Thord Silverbark at the Department of Literature and History of Ideas, Stockholm University, was helpful with regards to the history of scientific ideas.

Last, but not least, I would like to take the opportunity to recognise the anonymous children, students, teachers, headmasters, researchers and textbook authors who have partaken in the empirical studies that my colleagues and I have conducted.

Thank you all for making this journey possible!
# Table of contents

**Abstract** ............................................................................................................................ 1

**Preface** ............................................................................................................................... 2

  Acknowledgements ................................................................................................................... 3

**Table of contents** ................................................................................................................. 5

1. **Introduction** ..................................................................................................................... 7

   1.1. Individual studies of the dissertation ............................................................................. 7
   1.2. Structure of the dissertation .......................................................................................... 9
   1.3. Purpose of the dissertation ............................................................................................. 10

2. **Thermodynamics and thermodynamics education** ......................................................... 12

   2.1. Thermodynamics and statistical mechanics ............................................................... 12
   2.2. Thermodynamics education ......................................................................................... 14
   2.3. Confronting challenges in thermodynamics and science education ......................... 16
         2.3.1. Approaches to analysing and inducing conceptual change .................................. 16
         2.3.2. Teaching science with analogies and analogical reasoning ................................. 18
         2.3.3. Using scientific models and modelling to teach science ..................................... 22
         2.3.4. Exploiting language and representations in science education ........................... 23

3. **Perspectives of the theoretical framework** ....................................................................... 28

   3.1. Analogical reasoning – within cognitive psychology ..................................................... 28
         3.1.1. Analogical reasoning .............................................................................................. 28
         3.1.2. Concepts and mental models ................................................................................. 34
   3.2. Semantics – the study of the meaning of language .......................................................... 36
         3.2.1. Semantics – meaning and reference ...................................................................... 36
         3.2.2. Polysemy and homonymy ..................................................................................... 39
         3.2.3. Metaphor .............................................................................................................. 39
         3.2.4. Extension of ‘language’ – Semiotics and representations ....................................... 40
         3.2.5. Pragmatics – language in context ....................................................................... 42
   3.3. Scientific modelling – within philosophy of science ....................................................... 42
         3.3.1. Scientific modelling ............................................................................................... 42

4. **Synthesis of the theoretical framework** ........................................................................... 44

   4.1. Analogical reasoning and semantics .............................................................................. 44
         4.1.1. Analogy and relational language ......................................................................... 45
         4.1.2. Analogy and metaphor ......................................................................................... 45
         4.1.3. Cognitive linguistics and conceptual metaphor ..................................................... 48
         4.1.4. Cognitive linguistics and polysemy ....................................................................... 49
         4.1.5. Metaphor in discourse ......................................................................................... 51
   4.2. Analogical reasoning and scientific modelling ............................................................... 51
         4.2.1. Recruitment of analogical reasoning in scientific modelling .............................. 52
   4.3. Scientific modelling and semantics ............................................................................... 54
1. Introduction

This dissertation investigates ways in which analogical reasoning, a focus on scientific models and modelling, and attention to semantics may be taken advantage of in science education and science education research, particularly as applied to teaching and learning of thermodynamics. In this introductory chapter, the individual articles, which constitute part of the dissertation, are briefly presented and the structure of the dissertation is outlined. In addition, the purpose of the dissertation is presented, together with the auxiliary questions and research questions, which have guided the research.

1.1. Individual studies of the dissertation

The dissertation comprises a cover story and four individual articles, reprinted with kind permission from the journals in which they have been published:


This is a study of the different senses of the word ‘entropy’ from a semantics and science education perspective, by use of principled polysemy (Evans, 2005; Tyler & Evans, 2001) and the two-dimensional semantic/semiotic analysing schema (2-D SAS) (Strömdahl, 2012). The different senses are found to relate to each other in a radial structure and we propose that keeping a focus on the referents of the senses is important in thermodynamics teaching.


The perspective on scientific modelling among teachers and students representing theoretical and practical traditions of knowledge (Molander, 2002) was studied by asking them to interpret computer animations of combustion engines. Vehicle mechanics teachers are sceptical towards using idealised models of engines in their teaching and prefer more realistic representations to simplifying abstractions based on ideal Otto engines.


Preservice physics teacher students were asked to create self-generated analogies (Blanchette & Dunbar, 2000) for two thermodynamic processes. Through analysis with the structure-mapping theory (Gentner, 1983), it was found that the students elaborated self-generated analogies to a greater depth than analogies recalled from teaching, attributed to the students assuming ownership (Enghag & Niedderer, 2008) for their own analogies.


First-graders were introduced to analogies in terms of “things that work in the same way”, interacted with physical phenomena and were asked to come up with analogies for them. The children grasped the structural aspect of analogies in familiar domains and some of them managed to come up with analogies also for the more abstract natural phenomena, supporting previous research on young children’s capacity for analogical reasoning (Goswami, 1992).
In his framework *Didactical Transposition*, Chevallard (1989) characterises science education research in terms of a set of subfields, standing roughly in a causal relationship with each other:

1. As science education researchers, we have to grasp the historical and current *science content*, at an appropriate level of depth. This means that we have to be in constant rapport with the rapidly developing progress in the natural sciences.
2. Selected aspects of this science content have to be reformulated in terms of suitable *science education content*, feeding into curricula and syllabi in the educational system.
3. Next, the ‘what’ question of curricular statements has to be interpreted into the ‘how’ question of concrete *teaching approaches*.
4. Finally, in implementation or enactment of the teaching, the ultimate goal is to achieve *student understanding* of the taught content and the development of abilities.

<table>
<thead>
<tr>
<th>Science content</th>
<th>Science education content</th>
<th>Teaching approaches</th>
<th>Student understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article 1 – Senses of entropy</td>
<td></td>
<td>Article 2 – Otto engine animations</td>
<td></td>
</tr>
<tr>
<td>Article 3 – Preservice teachers’ analogies</td>
<td></td>
<td>Article 4 – First-graders’ analogies</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. Categorisation of the articles in the dissertation as ‘Didactical Transposition’ (Chevallard, 1989).*

The roles of the studies included in this dissertation can be classified roughly by use of the Didactical Transposition framework, as depicted in Figure 1. First, the science content related to thermal phenomena is put to the foreground in all four studies. However, this is most pronounced in article I, which, as mentioned, presents an investigation of the different senses of entropy and their connections, within different fields of science and outside. Although the investigation is done with educational implications in mind, it can be classified as pertaining to science content, *per se*. Article II deals with conceptualisations of models among teachers and students and the role of the models in the representation of a technical artefact. It therefore straddles the categories of teaching approaches and student understanding. Articles III and IV can be interpreted as explorations of the teaching approach of asking students to come up with their own analogies, although more scaffolding would have to be provided if applying the approach in regular education.

Duit (2007) classifies different strands of science education research in the *Model of Educational Reconstruction*, based on the German ‘Didaktik’ tradition, according to:
1. *Analysis of content structure*, in which the subject matter is clarified and its educational significance is assessed. This analysis can further be subdivided in: *elementarization*, a process in which key ideas of a field of study are singled out for inspection; and, *construction of content structure for instruction* with the intended age group in mind, going beyond mere reduction or simplification of the content.

2. *Research on teaching and learning*, i.e. findings from general pedagogy and science education with regards to learning theories, teaching approaches, teachers’ views and conceptions, etc.


In the Model of Educational Reconstruction, article I fits well with the first category with its focus on the content structure, both from the within science point of view and regarding educational implications. The second category, however, brings to the fore the influence of advances in fields outside natural science. Here, article II deals with conceptions of modelling, while articles III and IV look at participants’ analogical reasoning with different kinds of scaffolding. In addition, relating to the third category of real teaching, the studies resulting in articles III and IV were carried out as instructional events, although in a research setting.

**1.2. Structure of the dissertation**

Chapter 2, Thermodynamics and thermodynamics education, follows this introduction of the dissertation. Here, the scientific field of thermodynamics is presented briefly together with an account of the character and challenges of thermodynamics education. In addition, as a justification of the dissertation as a whole, the three perspectives comprising the theoretical framework – analogical reasoning, semantics and scientific modelling – are brought up as possible approaches to come to terms with the challenges in thermodynamics education and science education in general. The schema of the theoretical framework is shown in Figure 2.

![Figure 2. Schema of the theoretical framework. Relationship between the three theoretical perspectives and their application in thermodynamics education.](image-url)
The three theoretical perspectives represent different academic fields of study:

- **Cognitive psychology**, within which analogical reasoning is a cognitive ability.
- **Philosophy of science**, with a particular focus on scientific models and modelling.
- **Semantics**, as a part of the more general fields of linguistics and semiotics, where signs represent or stand for, on the one hand, concepts in our minds and, on the other, referents in the world.

The three theoretical perspectives will first be analysed in their own right in chapter 3, *Theoretical background*, and in chapter 4, *Synthesis of the theoretical framework*, it is investigated how the components of the framework relate to one another. For instance, how is analogical reasoning related to the issue of scientific modelling? A full-fledged synthesis of, for example, the philosophy of science and the philosophy of language, would be a quite daunting endeavour and clearly beyond the scope of this dissertation. However, some interesting commonalities will be pointed out, but also areas where the approaches lead to contrasting or complementing views.

As guidance to the reader, chapters 3 and 4 provide an in-depth general investigation of previous research relating to the three perspectives of the theoretical framework. These chapters are not essential reading, should you be primarily interested in the findings and implications of the articles comprising the dissertation. You may consider proceeding to the concluding section 4.4, with an overview of the main lines of argument in chapters 3 and 4.

Next, after a presentation and discussion of the methodological choices made in chapter 5, *Methodological framework*, overviews of the articles and a reanalysis of selected data from them are provided in chapter 6, *Summary, reanalysis and discussion of the individual studies*. The ambition has been not only to summarise the findings of the individual articles in this cover story, but to carry out reanalyses with regards to the research questions of the overall dissertation, and to use the data in the individual studies as cases in point.

Finally, in chapter 7, *Conclusions and implications*, the research questions are revisited and implications for educational research and the practice of science teaching are drawn.

Throughout the dissertation, the focus of a particular section will be pointed out by highlighting the relevant parts and connections in the schema of the theoretical framework (Figure 2).

### 1.3. Purpose of the dissertation

The overall purpose of this dissertation is to explore how advances in the fields of analogical reasoning, scientific modelling and semantics may contribute to science education research and science teaching, particularly regarding thermodynamics.

With regards to doctoral dissertations in mathematics education, and hopefully applicable also to science education, Niss (2010) characterises two types of questions that inform the research: First, there are the *research questions*, which are genuine and non-trivial in the respect that the answer to them is not already known before the research is carried out. However, before posing such research questions proper, the doctorate student typically has to pave the way by settling a set of *auxiliary questions*, in terms of establishing what is already known in a particular field of research. Answering such auxiliary questions is an important step towards formulating and also answering the research questions.

Accordingly, the research accounted for in this dissertation was guided by the following *auxiliary questions*:
A1. How have the fields of analogical reasoning, semantics and scientific modelling been used in the study and development of thermodynamics and thermodynamics education?

A2. How are the fields of analogical reasoning, scientific modelling and semantics related to one another?

The ambition is to provide answers to these auxiliary questions through the development of the theoretical framework of the dissertation in chapter 3, Theoretical background, and chapter 4, Synthesis of the theoretical framework.

Next, in chapters 5, Methodological framework, and 6, Summary, reanalysis and discussion of the individual studies, the focus is on the original research that was reported upon in the articles and the reanalysis in this cover story. The reanalysis was informed by the following research questions:

R1. To what degree is analogy involved in connecting different representations of a phenomenon to each other and to the represented phenomenon?

R2. How do students’ self-generated analogies relate to the practice of scientific modelling?

Note that these research questions of the dissertation as a whole are not identical to the research questions of the individual articles, which are provided in the summaries of the articles in chapter 7. Instead, the research questions of the dissertation are intended to relate the studies to one another, so that they can provide illustrating examples of the points made in the development of the theoretical framework.

We start the development of the theoretical framework with its three perspectives in chapter 2 by giving a brief background to thermodynamics as a scientific discipline and how the three perspectives have been suggested in science education research in order to come to terms with challenges in thermodynamics education and science education in general. Chapter 2 chiefly addresses auxiliary question 1.
2. Thermodynamics and thermodynamics education

In the following chapter, the scientific field of thermodynamics is introduced and issues related to the teaching and learning of thermodynamics are brought up. In addition, an overview is given of how analogies, scientific modelling and attention to semantics have been proposed in science education research as means to come to terms with such issues.

2.1. Thermodynamics and statistical mechanics

The field of thermodynamics may be broadly characterised as dealing with phenomena related to the transfer and transformation of energy between systems and their surroundings. Traditional thermodynamics deals with phenomena in or close to thermal equilibrium, where change in measurable, macroscopic quantities does not occur, or occurs sufficiently slowly.

The **zeroth law of thermodynamics** states that if two thermodynamic systems are separately in thermal equilibrium with a third system, they are also in equilibrium with each other. The third system may be considered as a thermometer, a measurement devise for temperature, $T$, which is an intensive physical quantity, i.e. a quantity that does not depend on the size of a system, and a state function, i.e. a quantity that depends on the state of a system but not on how the system has come to that state.

The **first law of thermodynamics**, $dU = δQ + δW$, states that the infinitesimal change of the internal energy of a system, $dU$, is equal to the sum of the heat transferred to the system, $δQ$, and the work performed on the system, $δW$, and implies the conservation of energy; that energy cannot be created or destroyed, but only transformed. Work and heat are process variables involved in energy transfer between systems, but not state functions. In other words, they depend on how changes to systems occur, the paths changes take, and not only on the state of the systems. Work is such energy change that is due to variation in external parameters, such as the volume or number of particles of a system, and heat may be defined as the remaining part of the energy change (Kjellander, 2009). Heat can be exchanged between systems with three different kinds of mechanisms: heat conduction involves transfer of energy due to a temperature difference within a solid or from one solid to another solid in thermal contact, for example by means of propagation of vibrations between neighbouring microscopic particles; convection is the transfer of energy due to the movement of fluids, i.e. gases and liquids; and radiation means energy transfer by means of electromagnetic radiation, which is emitted from all matter at temperatures above absolute zero.

The **second law of thermodynamics** relates to the tendency of energy to disperse and may be formulated in the way that heat cannot spontaneously flow from a body of lower
temperature to a body of higher temperature. By introducing the extensive state function **entropy** with the inequality: \(dS \geq \delta Q/T\), where \(dS\) is an infinitesimal entropy change of a system, the second law of thermodynamics may be formulated as the tendency of the total entropy of a system and its surroundings to increase in any **irreversible processes**, i.e. such processes that cannot run backwards in time spontaneously. **Reversible processes** that are symmetric with regards to time imply constant entropy and equality in the expression: \(dS = \delta Q/T\). The other way around, having introduced entropy, temperature may be defined as:

\[
\frac{1}{T} = (\frac{\partial S}{\partial U})_{V,N},
\]

i.e. the inverse of the partial derivative of the entropy with regards to the energy, given constant volume \(V\) and number of particles \(N\).

Apart from the laws of thermodynamics, **equations of state** specify the relationships between state functions of a system. Among these, the **ideal gas law**: \(pV = nRT\), states a relationship for ideal gases between the pressure, \(p\), the volume \(V\), the amount of substance \(n\), and the absolute temperature \(T\), where \(R\) is the ideal gas constant. Ideal gases are such gases that are assumed to consist of randomly-moving particles that interact exclusively through exchange of energy during collisions and occupy a negligible part of the system’s volume.

Thermodynamic systems can undergo change in many different ways involving exchange of heat and work with the surroundings. Such **thermodynamic processes** are described with regards to their influence on the involved state functions. **Isothermal processes** occur at constant temperature: \(\Delta T = 0\), and typically involve thermal contact so that heat can be exchanged between the considered system and a heat bath, a large system with constant temperature and ideally infinite heat capacity: \(C = Q/\Delta T\). For ideal gases with a constant amount of substance, the pressure of isothermal processes is inversely proportional to the volume: \(p = k/V\), where \(k\) is constant. **Isobaric processes** occur at constant pressure: \(\Delta p = 0\), and **isochoric processes** conserve the volume: \(\Delta V = 0\), which means that no pressure-volume work is performed by or on the considered system. Finally, in **adiabatic processes**, no heat is exchanged between the system and its surroundings: \(Q = 0\), and for ideal gases, the change adheres to the following relation: \(pV^\gamma = k\), where \(k\) and \(\gamma\) are constants, the latter of which depending on the heat capacity of the particular gas. Complex processes, comprising several steps that start and end in the same state are **cyclic processes**.

Thermodynamic processes may be illustrated in \(pV\) graphs, depicting the pressure against the volume, such as that of an ideal Otto cycle shown in Figure 3. When a mixture of fuel and air in a cylinder of a combustion engine ignites at its minimum volume, the pressure is assumed to increase instantaneously, corresponding to the vertical line at the left. During the Power stroke (1), the gas expands adiabatically, by exerting a force on a piston and performing work on the surrounding, resulting in decreasing pressure. Next, the resulting exhaust fumes are let out by opening the exhaust valve at maximum volume, which decreases the pressure. In the Exhaust stroke (2), the exhaust gases are pushed out of the cylinder through the exhaust valve, and in the Intake stroke (3), a new fresh mixture of fuel and air comes into the cylinder through the intake valve, both under the assumption of constant pressure. Finally, during the Compression stroke (4), the mixture is compressed by the piston adiabatically, so that the pressure increases as the volume decreases, preparing for a new ignition, which closes the cycle.

---

3 For the sake of completion, the **third law of thermodynamics** states that the entropy of perfectly ordered substances tends to zero at zero absolute temperature.
The ideal Otto cycle may be used to represent physical combustion engines, both of the Otto design, used for example in petrol engines, and diesel engines. However, many simplifications and idealisations are made in the modelling, such as the assumption of instantaneous combustion of the fuel-air mixture.

While classical thermodynamics deals with macroscopic, measurable properties involved in thermal phenomena, microscopic accounts deal with motions of and interaction between the systems’ constituting particles, such as atoms and molecules. The Boltzmann-Maxwell distribution details how the velocities and kinetic energies of particles are distributed in a system described by classical mechanics. In the kinetic theory of gases, the temperature of an ideal gas can be shown to be proportional to the average kinetic energy of the particles and thereby the average of the square of the velocities. Statistical mechanics is a microscopic theory that applies to a broad range of phenomena. In statistical mechanics, for an isolated system, i.e. a system that does not exchange energy or particles with the surroundings, the entropy may be defined as: \( S = k_B \ln \Omega \), where \( \Omega \) is the number of microstates, and \( k_B \) is Boltzmann’s constant, assuming the equal a priori hypothesis that all available microstates have equal probabilities. The number of microstates is the number of ways that the energy and particles can be distributed or configured microscopically, corresponding to one macroscopic thermodynamic state. The entropy may be generalised also for systems exchanging energy and particles with the surroundings as:

\[
S = -k_B \sum p_i \ln p_i,
\]

where \( p_i \) is the probability that the system is in microstate \( i \) in Gibbs’ formulation.

### 2.2. Thermodynamics education

The field of thermodynamics is a central domain in science that helps us understand fundamental aspects of the character and development of the natural world. Still, it is a domain of human knowledge that is primarily exclusive to people who have deliberately
chosen to include natural sciences as a part of their trade of life. As C. P. Snow (1993/1959, pp. 14-15) famously put it:

A good many times I have been present at gatherings of people who, by the standards of the traditional culture, are thought highly educated and who have with considerable gusto been expressing their incredulity at the illiteracy of scientists. Once or twice I have been provoked and have asked the company how many of them could describe the Second Law of Thermodynamics. The response was cold: it was also negative. Yet I was asking something which is about the scientific equivalent of: Have you read a work of Shakespeare’s?

The exclusivity of thermodynamics is unfortunate from many perspectives. Apart from offering answers to fundamental science questions such as ‘why do processes tend to happen in one direction in time, but not the other?’, the field of thermodynamics holds the potential to give insights into the nature of science, including grasping its unitary character overarching the disciplines of physics, chemistry and biology. In addition, in times where global warming as a consequence of overexploitation of fossil fuels is central to our inability to ensure sustainable development, understanding of thermodynamics may contribute to the development of crucial technology to the benefit of society as a whole.

Given the wealth of natural phenomena to which thermodynamics may contribute a deeper understanding, one may wonder why it has not become a more influential part of mainstream culture as implied by Snow (1993/1959) above. One of the obstacles to taking on thermodynamics, undoubtedly, is its abstract character. Piaget and Garcia (1977) identified that children’s ability to understand heat phenomena is delayed by one developmental stage, compared to understanding mechanical phenomena, because they cannot coordinate seeing and manipulating them in interaction. Subsequently, it has been found that it is also difficult for university students to grasp and apply fundamental concepts within thermodynamics, such as the ideal gas law (Kautz, Heron, Loverude, & McDermott, 2005; Kautz, Heron, Shaffer, & McDermott, 2005), the first law of thermodynamics (e.g. Loverude, Kautz, & Heron, 2002; Meltzer, 2004), and the second law of thermodynamics (e.g. Christensen, Meltzer, & Ogilvie, 2009; Cochran & Heron, 2006). In particular, students have been found to be prone to apply the ideal gas law in cases where other approaches are required and misapply microscopic models of thermal phenomena (Loverude, et al., 2002). For instance, students were found to see collisions of particles as the cause of increased temperature or that energy was released as a consequence of the collisions (Leinonen, Räsänen, Asikainen, & Hirvonen, 2009).

Within thermodynamics teaching, the central concept of entropy has been found particularly difficult to grasp (e.g. Carson & Watson, 2002; Christensen, et al., 2009; Sözbilir & Bennett, 2007). One reason for this is that it is not directly measurable – there is no entropy-meter – but entropy is derived from other quantities. It is also a highly theoretical construal, without obvious connections to our everyday language, experiences or physical senses. As a consequence, after a basic physics course on thermodynamics, a natural response by a student to the question what entropy is might be: ‘it’s $S$, a letter’, with a focus on the algebraic formalism, or ‘disorder’. If approached in physical chemistry, the student would get used to working with entropy, for instance by looking up the entropy change of particular reactions in tables, but would not necessarily reach an in-depth conceptual understanding. In contrast, a microscopic introduction in reference to microstates gives the opportunity to develop such more fundamental ideas, but it may be hard to relate this view of entropy to macroscopic phenomena, such as the functioning of heat engines, from which entropy once originated. Reif (1999) argues in favour of a microscopic atomistic approach to teaching of thermal physics and emphasises the need to understand the underlying mechanisms of the involved phenomena. In addition, he points out the difficulty among students to build visualizable mental models with macroscopic approaches. In contrast, Loverude, et al. (2002)
propose that since students tend to apply microscopic models in inadequate ways, the concepts have to be firmly understood in macroscopic contexts first, using for example bicycle pumps, before microscopic explanations can be introduced.

2.3. Confronting challenges in thermodynamics and science education

Against the background of the centrality of thermodynamics in physics and in science in general, in combination with the difficulty in attaining its fundamental concepts, there is a challenge in thermodynamics education in developing teaching approaches that are conducive to learning. The present dissertation ultimately aims to contribute to this endeavour. In this section, we bring forward approaches that have been suggested to analyse and come to terms with challenges in thermodynamics education and science education at large.

2.3.1. Approaches to analysing and inducing conceptual change

Posner, Strike, Hewson and Gertzog (1982) introduced the notion of conceptual change in science education, following in the Piagetian tradition in reference to the process of accommodation – radical reorganisation of concepts, but also in analogy to Kuhn’s (1962) account of revolutions in the history and sociology of science as one scientific theory is replaced by another. Research on children’s and adolescents’ understanding of scientific concepts had revealed that many of them held misconceptions, ideas which are not in line with the sanctioned view in science, also after teaching of the subjects. Posner, et al. put forward the view that students should be induced to realise the cognitive conflict between their conceptions and the corresponding scientific concept, and be convinced to replace the former with the latter, due to the science account being more intelligible and plausible.

The conceptual change tradition of research in science education has been very fruitful, refined along the way and has branched off in many directions. However, the basic premises of Posner, et al. (1982) have also received criticism. Greiffenhagen and Sherman (2008) argue that the underlying analogy between conceptual change in the history of science and in the process of learning of the individual is invalid; the learner possesses no such thing as a stable theory of a phenomenon that can be replaced as part of teaching through confrontation with a supposedly more convincing theory. Siding with this view and partly by use of techniques within neurocognition such as fMRI (functional magnetic resonance imaging), Dunbar, Fugelsang and Stein (2007) found that learners are reluctant to abandon their misconceptions and even have difficulties taking in the conflicting science account when it is presented to them; their conscious mental processing simply shuts off as a response to the conflicting perceptions. Similarly, Smith, diSessa and Roschelle (1993) argue that the abandonment of one theory and adoption of a new one, proposed as part of teaching, does not square well with a basic assumption of constructivism, that we can only build on what we already know. Instead, they propose a resource perspective, where learning implies coordination in new ways of what we already have available, i.e. claiming continuity in learning, rather than the more disruptive abandonment of deficient naïve theories, as put forward by Posner, et al.

These ideas further build on diSessa’s (1983) theory of the learner’s conceptual change as a matter of development and coordination of phenomenological primitives (p-prims), described as basic intuitive knowledge elements at a more fine-grained level than the typical naïve theory or conception.

Another line of criticism against the original conceptual change perspective has focused on the assumption that we typically hold only one conception of a phenomenon. In his challenge of the conceptual change approach, Linder (1993) sees two fundamental perspectives on conceptions: On the one hand, conceptions may be seen as something isolated in our heads, in line with the mental model tradition adopted by Posner, et al. (1982). On the other hand, conceptions can be characterised in terms of interaction between a person and the
world, bringing in the particularities of the contexts into the conceptualisation. From this point of view, learning is not a matter of exchanging one conception for another, but coming to recognise appropriate conceptions in different contexts, through a process of conceptual dispersion. Linder builds his argument with examples in physics. For instance, within physics many different conceptualisations of fundamental concepts are also available, such as matter and time, building on Newtonian, quantum or relativistic theories, etc. In addition, at home, even trained physicists happily talk about vacuum cleaning as ‘sucking’. All of these conceptions may be useful and appropriate, but in different contexts. Similarly, based on examples from biology teaching – the function of the body among younger children and Darwinian evolution at upper secondary school – Caravita and Halldén (1994, p. 106) argue that “the aim of learning, science for example, is not to abandon old ideas in favour of new ones, but rather to extend our repertoire of ideas about the physical and cultural world, to refine their organization and coherence”. They also point out that learning often takes place at an epistemological meta-level: knowing that causal explanations are called for in accounting for the theory of evolution, but not necessarily when reflecting on why there is life on Earth.

Mortimer (1995) argues that we may have several different, potentially complementary conceptions and that learning may imply a conceptual profile change, a change in the set-up of conceptions and in which circumstances a certain conception is recruited, rather than conceptual change, as such. Mortimer, Scott and El-Hani (2012) have shown how the conceptual profile change approach may be used in conjunction with classroom discourse analysis to follow how groups of students come to enrich their conceptualization of scientific concepts by adding new ways of thinking about the concepts, construed as new ‘zones’. In particular, Amaral and Mortimer (2004) studied a series of three upper secondary school lectures on the second law of thermodynamics, involving the concepts of spontaneity and entropy. They identified four different zones comprising the students’ conceptual profiles, each tending to be recruited at different points of progression of the teaching and in different discursive contexts. First, the students were invited by the teacher to express their view of thermal phenomena, where they exposed an everyday understanding of spontaneity in terms of episodes that tend to happen by themselves, within a perceptual/intuitive zone. As the teaching progressed, involving more authoritative dialogue based on the textbook and lecturing, the students came to appropriate a formalist zone, involving more complex concepts such as free energy, and a rationalist zone, where spontaneity is connected to the microscopic distribution of energy across particles in a system. Interestingly, along the way, the students used an empirical zone, where increasing entropy of a process was connected to increasing disorder, as an intermediary communicative way of connecting their intuitive thinking about spontaneity with more scientifically adequate accounts. This empirical zone was recruited throughout the entire teaching sequence, in conjunction with the other increasingly advanced zones. In a similar vein, Petri and Niedderer (1998) describe how a student goes through a learning pathway and comes to develop three distinct conceptions of the structure of the ‘atom’, with different perceived scientific value to the student and different strength in terms of how likely they are to be triggered throughout the teaching sequence and Taber (2000) introduces the related notion of multiple frameworks. Furthermore, in response to the view of Chi and colleagues (e.g. Chi, Slotta, & De Leeuw, 1994) that an obstacle to conceptual understanding is that we tend to make errors in ontological categorisation (e.g. heat or electric current belong to the ‘process’ category according to science, but are often classified as ‘objects’ by novices), diSessa (1993) and Gupta, Hammer and Redish (2010) argue that also our ontological categorisations may be very flexible and context-dependent.

The present dissertation aspires to contribute to the search of ways to induce conceptual change among learners. In this regard, I sympathise with the view that an individual may embrace several parallel conceptions of a phenomenon, all potentially appropriate and
applicable in different contexts, and of learning as expanding the repertoire of ideas. I also adhere to the resource perspective in that we should make best use of what the students already know and their reasoning capabilities in the introduction of new subject matters. Section 3.1.2, Concepts and mental models, provides an overview of different views of what concepts are, and in chapter 7, Conclusions and implications, we come back to a discussion of how parallel conceptions relates to the development of several scientific models or analogies for a phenomenon and to the matter of polysemy in semantics.

**2.3.2. Teaching science with analogies and analogical reasoning**

Analogical reasoning – in essence, coming to understand an unknown, abstract domain by comparison to another, more familiar and/or concrete domain – has been put forward as “the core of cognition” (Hofstadter, 2001, p. 499) and central to learning (Gentner, 2003). In Gentner’s (1983) *structure-mapping theory*, which will be analysed in depth in section 3.1.1, analogy is based on structural similarity between the compared domains.

From a developmental psychology perspective, laboratory tests of analogical reasoning are often based on solving tasks of the structure A:B::C:D, corresponding to the question of what relates to C in the same way as A relates to B. For instance, when presented to ‘cat is to kitten as dog is to …?’, a respondent is supposed to answer ‘puppy’. In fact, Piaget, Montangero and Billeter (1977/2001) argue that analogical reasoning can be reduced to the mathematical capacity to reason by proportionality, with the same structure. They conclude that analogical reasoning is not fully developed until Piaget’s formal-operational stage, typically entered at the age of 11-12 years, from a study where participants of different ages were asked to pick out the most appropriate picture representing the ‘D’ answer from a selection of cards. Subsequent studies, however, have found that substantially younger children are capable of performing analogical reasoning, provided that the source domain is properly understood and that they have grasped the task of coming up with analogies (Brown, 1989; Goswami, 1992). Goswami (1992) even conjectured that analogical reasoning is an innate ability of infants, a view which later has been supported by Gentner (2003), but also criticised by Richland, Morrison and Holyoak (2006), who argue that analogical reasoning is dependent on the gradual development of the working memory capacity to manipulate complex information.

Analogies are used extensively in science education, in textbooks and the teaching practice (Auibusson, Harrison, & Ritchie, 2006; Duit, 1991), but presenting science by means of analogy is not without risk, as pointed out by Glynn (1989, p. 198):

> Teachers and authors should explain to students that analogies are double-edged swords. An analog can be used to correctly explain and even predict some aspects of the target concept. At some point, however, every analogy breaks down. At that point, miscomprehension and misdirection can begin.

Spiro, Feltovich, Coulson and Anderson (1989) propose that one way of avoiding a focus on idiosyncratic aspects – potentially misleading – of each analogy, is to make use of several complementary *multiple analogies* of a natural phenomenon as accounted for in science. As a parallel to how scientists use a series of closely related *bridging analogies* in problem solving, J. Clement (1993) puts forward that a similar approach can be adopted in science teaching. He gives the example of how the counterintuitive concept of a ‘normal force’ being exerted onto a book by a table on which it lies can be conveyed to students by a series of analogies. The difficulty lies partly in the fact that there is no visible motion or displacement in the interaction between the table and the book. To make this more salient, Clement comes up with the analogy of a book that compresses a spring by its weight. However, due to difficulties in seeing how the spring corresponds to the table, Clement bridges the gap by introducing the
idea of placing the book on top of a flexible board or a foam mattress, arguing that the table yields in the same way, but at the microscopic invisible level.

The potential drawbacks of using analogies in teaching go beyond the intrinsic weakness of each analogy. In fact, even though analogies are used spontaneously by scientists and laypeople (Blanchette & Dunbar, 2000; J. Clement, 1987, 1988; Dunbar, 1995) in their everyday activities, it has been found difficult for participants in psychological tests to recruit analogical reasoning across tasks, which according to the researchers were considered to be structurally similar (Gick & Holyoak, 1980). This is the more intriguing against the background of analogical reasoning as a possibly innate capability (Gentner, 2003; Goswami, 1992). Under what circumstances are we prone to focus on relational structure, the tell-tale sign of analogical reasoning (Gentner, 1983)? Blanchette and Dunbar (2000) argue that one reason for the meagre results in experimental settings is the double challenge that the participants are confronted with: first, they are supposed to have a thorough understanding of the source domain; second, they are supposed to map it to the less familiar target domain, and to do so in the way the researchers have intended. One way to reduce the challenge is to give explicit hints to the participants that a new phenomenon is structurally similar to a previously encountered phenomenon, which has been found to lead to substantially more successful problem solving compared to when hints are not given (Gick & Holyoak, 1980). A more radical approach, with the promise of avoiding the challenge altogether is to ask the participants to come up with their own analogies for a phenomenon, by use of self-generated analogies (e.g. Blanchette & Dunbar, 2000; Wong, 1993a, 1993b); you are likely to only come to think of domains that you are in command of and that you yourself find structurally similar to the topic to be learnt. Similarly, there is an opportunity in making use of students’ spontaneous analogies in the classroom, which according to May, Hammer and Roy (2006) has not received enough attention, particularly in elementary science. Subsequently, in relation to energy, Lancor (2012) has found that different self-generated metaphors or analogies are given by students in different subject traditions: biology, chemistry and physics.

Another dimension of analogical reasoning reflects the overall setting in which it is performed. Wilbers and Duit (2006) propose a fundamental difference between the processes of generating post-festum analogies, which is done for communicative purposes once you have understood the compared domains and you see them as connected, and of heuristic analogies, which are carried out in an iterative fashion when you gradually explore the ins and outs of a studied phenomenon. The post-festum analogy is created by the scientist, textbook author or teacher who want to convey their ideas to an audience, while the process of heuristic analogy may be applied by the researcher when carrying out research in exploration of a new field, but also by the student who wants to bring structure to a field of study.

More recently, social aspects of generation and interpretation of analogies in discourse have received increasing attention in science education research. Aubusson and Fogwill (2006) invited secondary chemistry students to collaborate in the development of a role play as an analogy of a process involving formation and dissolution of chemical bonds, where the students acted as atoms and ions. The approach offered an opportunity for the students to express, negotiate and develop their understanding in conversation as the analogy role play developed. Similarly, Bellocchi and Ritchie (2011) investigated the use of analogy role play and analogy generation in small group work among secondary chemistry students. They found such exercises to stimulate productive hybridisation between the students’ own experiences from their everyday life and the scientific accounts of the studied phenomena. In contrast, Yerrick, Doster, Nugent, Parke and Crawley (2003) studied preservice physics teachers working with analogies for electric circuits in small group exercises within a guided-inquiry approach and found that self-generated analogies, corresponding to the students’ personal
theories, were found to divert them from the focus on the studied phenomena, lead to incorrect conclusions, and unconstructive group dynamics.

Another recent trend in science education has been to investigate the use of conceptual metaphor in language involving abstract physical quantities, such as energy or entropy (Amin, 2009; Amin, et al., 2012; Brookes & Etkina, 2007; Jeppsson, et al., 2012). One pattern in language is that we have come to talk about such abstract state functions as substances or locations, and change of state in terms of a flow or movement along a path.

Regarding analogies for thermal phenomena, the idea of seeing heat as if it were a fluid that may be contained in warm objects and flow from one object to another has been powerful throughout history and is still dominating everyday language in our time (Amin, 2001). For a long time it was a productive idea in science, and a foundation for the caloric theory of heat (Müller, 2007) and – less commonly recognised – also for the preceding set of theories based on phlogiston as a ‘matter of fire’ (Lewowicz, 2011). The caloric theory of heat was abandoned in the mid-19th century with the development of modern thermodynamics, since it was discovered that heat is not a conserved quantity. Consequently, in science education, a substance-based view of heat is often regarded as a misconception and an obstacle to a modern understanding of heat in terms of energy transfer (e.g. Erickson & Tiberghien, 1985).

The caloric theory of heat was still in full swing when it was applied by Sadi Carnot in his investigation of the theoretical limits of the efficiency of heat engines. In this endeavour, he introduced the ‘waterfall analogy’, comparing the mechanism of a heat engine in generating ‘motive power’, i.e. the modern work, during transfer of heat from high temperature to lower temperature on the one hand to the flow of water in a waterfall from a high to a lower level on the other. Gentner and Jeziorski (1993) put forward Carnot’s way of introducing heat flow in reference to the waterfall analogy as a typical case of the modern use of analogy with a focus on coherent structures. In particular, Carnot does not allude to surface similarities or causal relations across the domains, such as bringing in the temperature of the flowing water into the account. First, Carnot puts the phenomena side by side and establishes that there is a limit to the power that may be generated in both processes. Next, he points out explicit counterparts: the quantity of caloric, i.e. heat, corresponds to the quantity of water; the difference in temperature is mapped to the height of the fall; and, at a higher level of reasoning, the power generated depends on the quantity and height/temperature difference. Having established these coherent correspondences, he transfers an inference from the waterfall to the heat engine in the form of a hypothesis (which later is confirmed) that the power of the heat engine is proportional to the temperature difference.

Carnot’s waterfall analogy has been used and elaborated in science teaching. One such approach is Karlsruhe Physics, which makes explicit use of a shared quantitative structure across different subfields of physics and physical chemistry: a potential difference in an intensive quantity is coupled with a flow of a ‘substance-like’ extensive quantity, resulting in an energy flow or transformation (Herrmann, 2000). For instance, in thermodynamics, a temperature difference is coupled with a flow in entropy (corresponding to the ‘everyday sense of heat’ in this model) and in electric circuits, voltage is coupled with electric current. Karlsruhe Physics is a novel approach to continuity models of physics and is appealing for secondary teaching and engineering dynamics (Fuchs, 1996). However, its idiosyncratic style has provoked doubts as to whether it provides a productive foundation for further physics studies (Strnad, 2000). An exclusive focus on macroscopic accounts of thermodynamics and entropy in particular may hinder the attainment of microscopic accounts, such as those provided in statistical mechanics. Kaper and Goedhart (2005) have also developed a teaching sequence for undergraduate chemistry where they introduced entropy by reference to the waterfall analogy. Before introducing steam engines, they first worked out the correspondences between a waterfall and a piston performing pressure-volume work on the
surroundings at expansion, where height is mapped to pressure and the mass of the water to the volume of the gas in the cylinder, and derived expressions for the efficiency. Next, they asked the students to identify an expression for the efficiency of a heat engine and the energy transformed, stimulating them to recognise that a yet unknown extensive quantity, 'X', should be coupled with temperature, later identified as the entropy.

Arnold and Millar (1996) present a teaching sequence about thermal phenomena, particularly heat transfer, for secondary physics, by focusing on a story about the interrelationship between the concepts ‘temperature’, ‘heat’ and ‘thermal equilibrium’. They argue that the issue of separating the intensive quantity temperature from the extensive quantities heat and energy, i.e. following in the historical footsteps of James Black, is one of the main challenges of secondary teaching of thermal physics. In addition, due to the difficulty for students to see heat in terms of energy transfer, Arnold and Millar adopted the approach of using the term ‘heat’ as something that is stored in and flows between warm objects in line with the caloric theory of heat. The intended story is introduced by showing the analogy between heating water with a candle and controlling the water level in a glass container with input and output valves, where the temperature of the heated water corresponds to the water level in the container. The candle corresponds to the input valve and the heat being transferred can be regulated by adjusting the distance. Heat dispersal to the surroundings is mapped to the output valve. After some time of heating with the candle, the temperature of the water reaches a stable level, which is explained by equal amounts of heat being added to and dispersing from the water, i.e. a case of a steady state. This corresponded to a stable water level, due to equal input and output of water via the valves.

If the waterfall analogy may be useful in providing a macroscopic understanding of heat, completely different ways of thinking are required for microscopic thermodynamics. As we will see below, Hesse (1966) uses the comparison between molecules in a gas and interacting billiard balls as a paradigmatic case of the use of analogical reasoning in scientific modelling. Indeed, this type of comparison was important in the early development of the kinetic theory of gases. Achinstein (1987) relates that Maxwell set out to develop a theory of gases on the assumptions that: gases are composed of minute particles in rapid motion; the particles are perfectly elastic spheres; the particles interact only during impact; and, their motion is subject to the principles of Newtonian mechanics. As Achinstein points out, these basic assumptions “are highly speculative, involving as they do the postulation of unobservable particles exhibiting unobserved motion” (p. 410). Actually, the assumptions involved a set of idealising simplifications in relation to what was known by science at the time. For instance, Clausius had already recognised the need to attend to rotational motion of molecules, which was not incorporated into Maxwell’s original model. From these and a few more basic assumptions, however, Maxwell created a probabilistic model for the distribution of velocities, derived known macroscopic relations, such as the ideal-gas law: \( pV = NRT \), for which the kinetic model was argued to provide causal explanations, and later investigated the internal structure of gases, e.g. with regards to viscosity, with gradually refined assumptions.

Niaz (2000) argues that a number of lessons can be drawn from the historical development of the kinetic theory of gases. First, the kinetic theory of gases was controversial for decades, initially due to competition from caloric theorists and later from scientists of an empiricist leaning, such as Mach and Ostwald, who disregarded a theory that postulated the existence of invisible entities, i.e. atoms and molecules. This conveys the message that theory change is rarely unproblematic. Second, the kinetic theory of gases was elaborated gradually by removing Maxwell’s simplifying assumptions one by one. Here, van der Waals’ approach to encompassing particle attraction in dense gases and Boltzmann’s contributions to statistical mechanics should be noted. Finally, students’ reasoning about changes involving gases may be broadly categorised as making use of an ‘algorithmic mode’, primarily by application of
the ideal gas law, or as formation of a ‘conceptual gestalt’ in the form of microscopic modelling providing causal explanations to macroscopic phenomena. Exposing the students to the way the kinetic theory of gases developed in science may induce them to use the advantageous ‘conceptual gestalt’ way of reasoning about these matters. In the development of the theoretical framework in chapters 3 and 4 and the methodological framework in chapter 5, different theoretical approaches to analogies and analogical reasoning will be brought up and analysed, in order to ascertain how to make best use of them in science teaching and science education research.

2.3.3. Using scientific models and modelling to teach science

Gilbert (2004) argues that a well thought out use of models and modelling may provide a route to more authentic science education. It would be characterised by the representation of processes in science, showing its element of creativity and giving insight into the nature of satisfactory explanations of phenomena in the world-as-experienced. As part of his extensive work on how the scientific revolution can be used in science education and particularly Galileo’s study of the pendulum, Matthews (2004) argues that abstraction and distancing from the real world through idealisation were at the core of Galileo’s approach. He notes:

Despite people seeing swinging pendulums for thousands of years, no one, not even the great Leonardo da Vinci who studied pendulum motion, saw what Galileo ‘saw’ (p. 690).

In Matthews’ view, the sole dependence on common-sense observation was an obstacle to understanding the underlying physical laws in the Aristotelian and medieval traditions.

Lehrer and Schauble (2006) discuss how model-based reasoning can be promoted in science education and propose a typology of models:

- **Physical microcosms** are physical scale models, which relate to the represented phenomena by visual similarity, but as recognised by e.g. Vosniadou (1989) also invite to investigation of relational structures.
- **Representational systems** are “half-in and half-removed from the world” (p. 377). Here the challenge of modelling a phenomenon conflates with representation of the phenomenon. When children enter school, they have already developed a capability for representation, letting something stand for another, as in their drawings or in pretend play, which may be recruited for conceptual development. We will briefly come back to the relation between models and representations when discussing the connection between modelling and semantics in section 4.3.3.
- **Syntactic models** reflect the underlying structure of a phenomenon, often of a quantitative kind, and are thus further removed from the world.
- Finally, with **emergent models**, such as billiard ball models of a gas, relations and interaction between the constituent entities produce phenomena not evident from descriptions of the entities or their relations alone, but appear at a level of higher complexity. Chi (2005) argues that such emergent phenomena are particularly challenging to understand in science education.

Following Hesse (1966), whose ideas in this regard are expanded upon in section 4.2.1, Lehrer and Schauble (2006) argue that models essentially are analogies, and that modelling may be introduced in science teaching as analogical reasoning skills develop. However, they also claim that successful introduction of model-based instruction hinges on adopting an inquiry approach to learning, in which students get the opportunity to pose questions in relation to investigated phenomena and create their own models of them. These models
further develop in a cyclical fashion, involving presenting and defending the model against any criticism from peers and teachers, followed by revision and refinement, generating different types of representations.

Even though model-based instruction is an attractive approach to science teaching, with its possibilities to strengthen understanding of core concepts and the nature of science as a whole, it does not come easily. Leach, Driver, Millar and Scott (1997) studied the views of the nature of science among 9-, 12- and 16-year-old students. They found that the students often have an unproblematic view of empirical enquiry, where observation is followed by careful description, without an intermediate modelling or interpreting step. No students consistently reach the level of model-based reasoning, where:

explanation is viewed as involving coherent systems of theoretical entities, some of which may be conjectural. Multiple possibilities for explanation are thus entertained. Empirical enquiry involves evaluation of theory with respect to evidence, and is acknowledged as problematic (p. 159).

Grosslight, et al. (1991) performed an interview study on the conception of models and their use in science among experts and 7th-, and honours 11th-grade students. While the 7th-graders see models as toys or copies of reality, some of the 11th-graders realise that somebody has created the model and that this modeller has had to make conscious decisions in simplification and high-lighting of certain characteristics. Only the experts, however, acknowledge that models serve to develop and test ideas, rather than being copies of reality, and that modelling involves evaluation of several designs, with regards to how well they serve the intended purpose of the model. Justi and Gilbert (2002) studied views on the nature of models among 39 science teachers at different levels. They found that the teachers were convinced that an understanding of the purpose of why a model is being produced is a prerequisite for successful modelling. Another criterion for successful modelling is experience of the phenomenon being represented and related phenomena. The teachers recognised the particular advantages of different modes of representation. Harrison and Treagust (2000, p. 1011) see practical obstacles to the use of models in science education:

Modelling is the essence of thinking and working scientifically. But how do secondary students view science models? Usually as toys or miniatures of real-life objects with few students actually understanding why scientists use multiple models to explain concepts.

Black (1962, p. 219) calls for humility when approaching the everyday view of models:

To speak of “models” in connection with a scientific theory already smacks of the metaphorical. Were we called upon to provide a perfectly clear and uncontroversial example of a model, in the literal sense of the word, none of us, I imagine, would think of offering Bohr’s model of the atom, or a Keynesian model of an economic system.

Examples such as Carnot’s waterfall analogy and the kinetic theory of gases show how close the practice of scientific modelling is to analogical reasoning, a central theme of section 4.1 below. In addition, in section 4.3, a deeper investigation is made of the nature of the relationship between phenomena and representations of them in modelling.

2.3.4. Exploiting language and representations in science education

In science education research, the role of language has been recognised particularly within socio-cultural perspectives on learning. With the notion of a zone of proximal development (ZPD), Vygotsky (1962) related to the difference between what an individual is capable of doing or understanding by herself and what she can achieve if assisted by somebody more knowledgeable serving as an example, primarily in communication by use of language.
Similarly, although within a more cognitive framework, Bruner and colleagues introduced the notion of scaffolding, describing the way learning can be enhanced in interaction with another person (Wood, Bruner, & Ross, 1976).

In his criticism of ‘mentalist’ approaches to education, Lemke (1990) proposes social semiotics as a tool to analyse “the dynamic development of a trajectory of meaningful action, socially shared and jointly constructed by teacher and students… this progressive sequence of things we say, diagrams we draw, equations we write, experiments we perform” (Lemke, 1998). He even promotes a wider view on language itself and suggests that part of the objective of science teaching is to make students capable of ‘talking science’ (Lemke, 1990):

I want us to recognize that in addition to scientific English, or Spanish, or Catalan, there are also other essential ‘languages’, in the sense of cultural systems of semiotic resources in science: the languages of visual representation, the languages of mathematical symbolism, and the languages of experimental operations. The goal of science education, I want to argue, ought to be to empower students to use all of these languages in meaningful and appropriate ways, and, above all, to be able to functionally integrate them in the conduct of scientific activity (Lemke, 1998).

Roth, Tobin and Shaw (1997) use Latour’s (1987) idea of a cascade of inscriptions, where one representation is succeeded by others in a sequential way, in their account of how a physics teacher conveys the mechanics of a rolling ball by means of tables and graphs of the distance, velocity and acceleration as a function of the time, and a money analogy for the phenomenon. They show that the task of translating between such different representations is complex, but argue that the connection between representations is taken to be self-evident and rarely attended to in teaching. The importance of being able to move from one representation to another has received particular attention in mathematics education research. Duval (2006) introduces the notion of representation registers, such as spoken or written natural language, algebraic notation, or a graphic representation of an object and claims: “Changing representation register is the threshold of mathematical comprehension for learners at each stage of the curriculum” (p. 128). If children do not manage to make such changes of representations, their mathematical knowledge remains compartmentalised and fragmented. Duval’s concluding statement “the content of a representation depends more on the register of the representation than on the object represented” (p. 114, italics in the original) reminisces of McLuhan’s (1964) dictum: ‘The medium is the message’. Similarly, within the field of science education, Airey and Linder (2009) argue that science students are expected to engage in disciplinary discourse, in the form of “the complex of representations, tools and activities of a discipline” (p. 29). In turn, different modes of discourse, e.g. spoken language, images or mathematics, is another dimension of points of entry to representation within the discipline. One ambition of science teaching is for students to achieve discursive fluency, “a process through which handling a mode of disciplinary discourse with respect to a given disciplinary way of knowing in a given context becomes unproblematic, almost second-nature” (p. 33), within one mode of discourse, but also in combinations of several modes. Further, adopting Gibson’s (1979) notion of affordance, Fredlund, Airey and Linder (2012, p. 658) introduce disciplinary affordances as “the inherent potential of [a] representation to provide access to disciplinary knowledge”. Where Duval (2006) sees translation of representation between different registers as challenging, adopting a disciplinary affordance view takes yet another step: the disciplinary knowledge that is accessed by means of one representation may in fact be completely inaccessible through another representation.

Interpretation and production of visual representations have received particular attention in science education research, and apart from the need to develop scientific literacy, a parallel call for development of visual literacy has been expressed (Gilbert, 2005). DiSessa (2004) expresses the need for students to develop metarepresentational competence, extending the
typical ambition that students should learn to interpret and use representations that are sanctioned in science. Such competence also encompasses creating your own representations, being able to communicate and assess the quality of representations, and knowing what representations may be suitable in a particular situation, in short: the problematic of representation, resonating with the view that “scientists are designers of representation” (p. 326). Particularly encouraging is his finding that students spontaneously recruit a wealth of resources in coming up with new representations, and even classes of representations, of phenomena and concepts, although to the extent that it might hinder adopting conventionalised representational forms. In addition, students as young as sixth-graders bring into class a developed sense of quality criteria for representations, such as completeness, precision and parsimony, remiscing of typical quality criteria for scientific theories in general. Similarly, Ainsworth, Prain and Tytler (2011) emphasise the importance of making your own representations for learning and particularly argue that making your own drawings improves engagement and learning in science education.

Apart from the primarily socio-cultural interest in language and representational systems as a means of human interaction, some science education research has focused specifically on semantics, the meaning of words and sentences. For instance, Williams (1999) brings forward several aspects in which semantics has an effect on physics communication and learning. First, there is a tension between words that have a precise meaning in physics, but other, more vague meanings in everyday contexts, such as ‘work’ or ‘charge’. According to Williams, when introducing students to the scientific meanings of such words, we cannot afford to make use of the laboratory vernacular used in communication between fellow scientists, but have to adopt a more formal language, a science equivalent of ‘hochdeutsch’. Second, and possibly more challenging, many words used primarily within science have different meanings in its different sub-disciplines. For instance, a single atom of helium may be seen as a valid representative of a ‘molecule’ to a chemist, focusing on its behaviour in interaction with other particles, but not to a physicist, due its lack of a molecular spectrum. Against this background, Williams calls for carefulness in formulation in physics textbooks, but also precision in the language we use in the classroom.

Brookes and Etkina (2009) suggest that attention to different meanings of words may provide an alternative to the view of students holding misconceptions of physical phenomena. For instance, a student may be well aware that an object will move with constant velocity if all resistance is removed, but still talk of the force of the object. In such cases, the necessary change would not have to involve reconceptualisation of the phenomenon, but merely to reassign appropriate words to the involved physical quantities. What the student has called ‘force’ may correspond to a physicist’s notion of ‘kinetic energy’ or ‘momentum’ and learning would amount to disambiguation between many different senses of ‘force’, in current and historical science, and in everyday language. In my view, telling students that physicists use common words in a specific, slightly different way than what they have encountered in everyday language is a more convincing, constructive and sympathetic approach than claiming that they have misunderstood basic physical phenomena. Further, Brookes (2006) argues that teachers should use a language that is consistent with the correct ontological category (Chi, et al., 1994). Particularly, with regard to ‘heat’, Brookes proposes the use of “energy flows from A to B by heating” as improvement of the typical “bad sentence” “heat flows from A to B”, which does not convey the idea of heat as a process variable. In contrast, in our research on the recruitment of conceptual metaphor in problem solving (Jeppsson, et al., 2012), we argue that it would be unfortunate for teachers to restrict themselves to such formulaic language. Instead, there has to be flexibility in the formulation of each individual sentence, making use of for example metaphorical language, extemporisation against the background of a deep conceptual understanding of the scientific theories at hand.
Focusing on younger students, Wiser and Amin (2001) argue that coming to see heat in terms of energy transfer is “the core stumbling block” (p. 332) in science teaching in the thermal domain, due to the fact that the students have a preconception of heat as ‘hotness’. The scientific view and the students’ preconceptions cannot be reconciled from an ontological point of view: “Since ‘energy is not hot,’ it is impossible to believe that heat is energy with the student conceptualization” (p. 338). In order to make students adopt the view of heat as energy transfer, they developed a microscopic computer simulation model of conduction where energy is exchanged between a hotplate and a piece of metal by means of molecular interaction. In one of the modes of depiction, units of energy were represented as letters “E” in motion in the system, the amount of heat was conveyed as the number of Es in transfer and the temperature as their crowdedness. One crucial component to the teaching sequence was to put forward learning of the scientific meaning of ‘heat’ as learning another sense of the word, in addition to previously known everyday senses of the word, based on perceived hotness:

To help students accept the scientific theory, we make the two conceptualizations compatible and integrate them. We explain how ‘science heat’ produces the sensation of hotness… /…/ Thus the everyday view is not replaced by the science view, nor does it simply coexist with it in a relativistic fashion, but is accounted for by the science view (p. 334).

Such focus on two related senses of a word, one explaining the other, makes explicit use of the semantic phenomenon of polysemy – that one word may have several related senses – which we return to in section 3.2.2. Interestingly, it provides yet another approach to conceptual change as a widening of repertoires (Caravita & Hallén, 1994).

Adolescents have been found to have difficulties discerning heat from temperature (e.g. Erickson & Tiberghien, 1985). However, based on the occurrences of the words ‘heat’ (as a noun and verb) and ‘temperature’ in a large corpus of language predominantly from non-scientific contexts, Amin (2001) found that the linguistic properties of heat as a noun and temperature differ markedly. While heat can either be conceived as a spatially located cause of ‘hotness’ or the state of hotness itself, temperature is taken to be a location on a vertical scale, representing a level of hotness. Amin suggests an underlying conceptual structure, a ‘minimal cognitive model’, in which heat (as a noun and verb) and temperature represent different aspects.

While ‘heat’ and ‘temperature’ are introduced early in teaching about thermal phenomena and part of everyday English, “entropy” has retained its technical and advanced character even though a broad range of different senses have developed. Grad (1961) identifies a large set of different meanings of entropy within science and argues that this plethora has contributed to the difficulty in grasping the concept. As will be expanded upon in section 6.1, Sklar (1993) critiques the view that thermodynamics can be ‘reduced’ to statistical mechanics and puts forward entropy as particularly problematic. As we have identified, several metaphors have been proposed for teaching entropy (Jeppsson, et al., 2011). The most common of these in teaching is probably ‘entropy is disorder’, typically made more explicit by use of the analogy between a thermodynamics system and a messy room. Entropy as disorder has been criticised, among other things, for its static ‘snapshot view’ of the concept (Lambert, 2002) and students tend to focus on the spatial configuration and ignore the aspect of energy (Brosseau & Viard, 1992). However, as we saw above, Amaral and Mortemer (2004) found that students may recruit the conception of entropy as disorder as a way to connect their intuitive way of reasoning about spontaneity with scientific accounts of concepts in the thermal domain. Other proposed metaphors have been to see entropy in terms of ‘freedom’ in conjunction with ‘information’ (Brissaud, 2005) or ‘spreading’ (Leff, 1996).

Within our research group, Strömdahl (2012) has developed a two-dimensional semiotic/semantic analysing schema (2-D SAS), detailed in Figure 4, as a tool for analysis of
the different senses of terms used in science, particularly physical quantities. Along the horizontal axis of the diagram, there are three kinds of semiotic elements, adopted from the semiotic triangle put forward by Ogden and Richard (1923), which is presented in section 3.2.4. Here, the meaning of a representational, purely linguistic entity, a word, is established by identifying what it corresponds to in our minds, a concept, and in the world, a referent. Along the vertical axis, there are the different, related senses of the word. The non-formal sense relates to the use of the word in everyday language. For instance, the word ‘heat’ evokes ideas of things and situations that feel warm, i.e. the referent is things that appear warm to our physical senses. The scientific qualitative model and physical quantity both relate to heat as interpreted in scientific modelling. In the case of heat, these senses refer to a model of energy exchange between thermodynamic systems due to a temperature difference. Finally, the empirical quantification is decided by determining units of measurement and measurement procedures, in the case of heat within the field of calorimetry. As will be detailed in the following, the 2-D SAS was used and elaborated upon in order to analyse the different senses of the term ‘entropy’ in article I of this dissertation (Haglund, Jeppsson, & Strömdahl, 2010), in particular with regards to the way the existence of several different senses influences teaching and learning of entropy. 2-D SAS has also been used in our investigation of change of concepts and referents of terms in thermodynamics (Strömdahl, et al., in progress).

Figure 4. Two-dimensional semiotic/semantic analysing schema (2-D SAS), applied to the word ‘heat’, adapted from Strömdahl (2012).

The significance of language as a means for teachers and more knowledgeable peers to provide ‘scaffolding’ or a zone of proximal development to the individual in his or her learning is expanded upon in the methodological framework in chapter 5. The connections between polysemy – that words may have many related senses – and our potential to embrace several different conceptions, models or analogies for an experienced phenomenon are discussed in the synthesis of the theoretical framework in chapter 4. Such connections are brought up again and exemplified in the reanalysis of the individual studies in chapter 6, with a particular emphasis on the learners’ own representations of the phenomena.
3. Perspectives of the theoretical framework

In this chapter, the three perspectives of the theoretical framework are presented one by one within their different academic fields of study, in order to be able to compare them and position them in the following synthesis of the theoretical framework.

3.1. Analogical reasoning – within cognitive psychology

The first perspective in the theoretical framework is analogical reasoning, as indicated in the schema of the theoretical framework to the right. It is studied primarily within cognitive psychology, an academic discipline that focuses on mental processes, such as perception, memory, thought and problem solving. Apart from analogical reasoning, in this section ‘the stuff’ of which our thoughts are made, e.g. mental models or concepts, are also brought up.

3.1.1. Analogical reasoning

Constructivism as a perspective on learning relies on the assumption that you can only learn new things by relating them to what you already know. Ausubel (1968, p. vi) asserted:

*If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly.*

However, as pointed out by Plato (1976, 80e) in *Meno*, there lies a paradox in this venture:

*[A] man cannot search either for what he knows or what he does not know. He cannot search for what he knows – since he knows it, there is no need to search – nor for what he does not know, for he does not know what to look for.*

What has become known as *Meno’s paradox* brings across the genuinely puzzling issue of how learning can come about against the background of the dichotomy between things that we know and those that we do not know. The volume *Similarity and analogical reasoning*, edited by Vosniadou and Ortony (1989), focuses on the idea that one approach to linking the new to what is known is by similarity: seeing a phenomenon as similar to a previously encountered phenomenon, or more specifically, being able to classify a new entity as belonging to an already known category. However, it is also recognised that similarity is a notoriously vague and subjective notion; commonalities can be established between any two entities in the world. One way of delimiting the focus is to analyse only such similarities that have the structure of analogy, involving the cognitive process of analogical reasoning.

Mary B. Hesse (1966) describes analogies between two analogues, which may be entities or phenomena, etc., in terms of two dimensions. She uses the phenomena of sound and light as an example. First, the different properties of each of the analogues stand in relations to each other, which typically are in some way causal; these are *vertical relations*. For instance, regarding sound, such causally related properties can be loudness, pitch and the fact that it is...
detected by our ears. Next, properties of one of the analogues are mapped to properties of the other analogue, creating horizontal relations, which are characterised by similarity in some respect. In some cases, the similarities of properties that are mapped are restricted to filling structurally similar roles in the vertical relations. Using the example of sound and light, loudness may be seen as related to brightness, while pitch is related to colour. In line with Vosniadou and Ortony (1989), Hesse (1966, p. 76) admits that comparison by similarity is a “notoriously inaccurate, incomplete, and inconclusive procedure”. However, this is not fatal to reasoning by analogy in a science context, since it is primarily used to generate hypotheses, not to test them in a rigorous, logical way. We return to Hesse’s view of how analogies are used in science in section 4.2, connecting analogue reasoning and scientific models.

Hesse (1966) also notes a connection between analogies and mathematical proportions. For instance, the phrase ‘pitch is to sound what colour is to light’ may be written in the form:

\[
\frac{\text{pitch}}{\text{sound}} :: \frac{\text{colour}}{\text{light}}.
\]

According to Hesse, analogies and proportionalities are both reflexive, since an entity is (trivially) analogous to itself, and symmetrical. In contrast to Piaget, et al. (1977/2001), however, Hesse acknowledges that analogies cannot be reduced to proportions. Unlike proportions, analogies do not uniquely identify the missing forth entity. For instance, a fish is to a fishtail as a bird is to ‘?’ could be complemented with a bird’s legs or its tail, depending on in what respect similarities are identified. In addition, analogies are not transitive, i.e. \(a:b::c:d\) and \(c:d::e:f\) does not always yield \(a:b::e:f\), since the first relations may consist of different similarities.

Tversky (1977) challenges the view of similarity as a symmetrical relation:

Similarity has been viewed by both philosophers and psychologists as a prime example of a symmetric relation. Indeed, the assumption of symmetry underlies essentially all theoretical treatments of similarity. Contrary to this tradition, the present paper provides empirical evidence for asymmetrical similarities and argues that similarity should not be treated as a symmetric relation. /…/ We say ‘Turks fight like tigers’ and not ‘tigers fight like Turks’. Since the tiger is renowned for its fighting spirit, it is used as the referent rather than the subject of the simile (p. 328).

The issue of asymmetry has become central to the investigation of analogy and similarity. In order to account for asymmetrical similarity, Tversky (1977) presents an alternative set-theoretical contrast model where the perceived similarity between two objects is taken to be a weighted function of attributes that are shared between the compared entities and attributes that are distinctive to one of them. Tversky provides evidence for his view from psychological studies of perceived similarity between different types of entities, such as statements of the type ‘country \(a\) is similar to country \(b\)’, where:

most of the subjects chose the phrase in which the less prominent country served as the subject and the more prominent country as the referent. For example, 66 subjects selected the phrase ‘North Korea is similar to Red China’ and only 3 selected the phrase ‘Red China is similar to North Korea’. These results demonstrate the presence of marked asymmetries in the choice of similarity statements, whose direction coincides with the relative prominence of the stimuli (Tversky, 1977, p. 334).

As mentioned, Gentner (1983) has developed the structure-mapping theory as an account of the structure of analogy and the process of analogical reasoning, and it was used to classify the generated comparisons, including analogies, in articles III and IV of the current

---

4 Etymologically, analogy stems from classical Greek ‘ἄναλογία’ (anologia), meaning proportion.
dissertation. She describes analogy as encompassing the mapping from a known base or source domain to an unfamiliar target domain, where domains are psychologically interpreted as systems of objects, attributes of the objects and relations between the objects. Furthermore, the analogy differs from other comparisons between domains, in that it emphasises mappings between structural relationships in the two domains, but not on object attributes. Gentner exemplifies with the Rutherford-Bohr analogy between the structure of the atom and the solar system, where the central massive atomic nucleus corresponds to the sun and stands in a relation to smaller orbiting electrons, the counterpart of the planets in the solar system. The analogy goes beyond surface similarities, i.e. any shared attributes of the constituting objects: a focus on the sun being warm and yellow in fact would give a misleading idea of the atom. In the structure-mapping theory, Gentner further proposes a systematicity principle, in that we favour mappings that are coherent and of high structural order, where a first-order relation is a relation between objects, a second-order relation a relation between first-order relations, etc. Cases where comparisons are based on matches on both attributes and structural relations are not analogies, but classified as literal similarities, for instance when comparing our solar system to that of another star. The structure-mapping theory has been operationalised in the Structural Mapping Engine (SME) (Falkenhainer, Forbus, & Gentner, 1989), where mapping of two domains and analogical inference can be simulated by a computer program. The structures of the domains are modelled by entering entities and predicates (attributes, relations and functions) in predicate calculus format, and the model has an exclusively structural focus, aiming ultimately for isomorphism between the compared domains, i.e. a perfect match of objects and their relations across the domains. However, Bowdle and Gentner (1997) acknowledge Tversky’s (1977) view of comparison by similarity as an asymmetrical phenomenon. Isomorphism is rarely achieved and in psychological experiments, subjects tend to make and prefer comparisons in the direction from a more structured and coherent source domain to a less structured and coherent target domain.

Holyoak and Thagard (1989) argue that in addition to such structural constraints on our thought processes involved in mapping, there are also semantic and pragmatic constraints. The semantic dimension involves the realisation that individual entities and attributes in the two domains may be similar in some respect. Such attribute similarity is typically most important in retrieval of the domains and their relevant features – the first step of analogical reasoning where the domains to be compared and what aspects of them to focus on are identified. However, attribute similarity may also strengthen or weaken the analogical mapping across the domains, depending on how the surface similarity matches the structural pattern. Pragmatic constraints involve for instance goal-driven activity or realisation that a certain entity or relation is particularly important to map. Holyoak and Thagard incorporated such constraints in their predicate calculus model ACME, where semantic similarity between predicates across the domains is input by assigning numerical weights. Similarly, pragmatic concerns were catered for by giving weights of importance to the elements and to hypotheses of which predicates would match across the domains. By running a range of known analogical mappings, e.g. Gentner’s (1983) atom-solar system comparison, Holyoak and Thagard indeed found such semantic and pragmatic constraints to have an influence on the model’s ability to establish the intended mapping between domains, reproducing the patterns of psychological experiments of the difficulty of the analogies among adults and children under different semantic and pragmatic conditions. For instance, they tried the underlying analogical structure of the metaphor ‘Socrates is a midwife of ideas’ in three versions. First, they entered only relevant structural features of the two domains, which yielded a good structural match. Next, when they added predicates that were irrelevant and potentially distracting, such as ‘Socrates is a father’ or ‘Socrates drinks hemlock’, the appropriate match was not established. However, by further adding the pragmatic constraints that ‘Socrates’ ought to be mapped to ‘midwife’
and that ‘ideas’ should be mapped to something, in other words being important – reasonable assumptions in interpreting the metaphor – the intended pattern again was obtained.

C. A. Clement and Gentner (1991) designed a series of psychological experiments to test the degree to which structural systematicity can account for the perceived aptness of alternative facts in a fictitious story, bearing in mind the potential role of pragmatic and semantic constraints. They found that pragmatic constraints, in the form of the extrinsic goal of finding causal relations in the stories, were not sufficient to determine mapping. On the contrary, mappings could be established if the facts were structurally systematic across the domains, even in the absence of such extrinsic goals. However, they also suggest that their view is not incompatible to that of Holyoak and Thagard (1989). For instance, pragmatic goal considerations might be formulated in terms of higher-order structural relations, and used across the domains. In addition, Clement and Gentner argue against semantic similarity alone, thereby criticising Tversky’s (1977) contrast model focusing on sets of individual features, rather than relations. Although admitting that some static pragmatic constraints may be interpreted in terms of structural or semantic constraints to the mapping in line with Clement and Gentner, Spellman and Holyoak (1996) argue that goals may play a more distinct pragmatic role in guiding the mapping process, without forming a part of the mapping itself.

Chalmers, French and Hofstadter (1992) argue against the popular approach within artificial intelligence (AI) research to model cognitive processing, such as analogical mapping, in isolation, without acknowledging its interrelation with ‘high-level perception’, i.e. the act of representing sensory input. Particularly, they attack the structure-mapping theory (Gentner, 1983) and SME (Falkenhainer, et al., 1989) in that the process of representation is bypassed and its assumed result is entered into the system. Further, Chalmers, et al. argue that AI research has adopted an ‘objectivist’ stance, criticising the way our thoughts are treated as objects or statements, here in the form of predicate calculus. They mean that such approaches are bound to fail, since they do not properly represent the flexibility of our high-level perception. In a reply to Chalmers, et al., Forbus, Gentner, Markman and Furguson (1998) admit that much more remains to be done in order to incorporate representational aspects of structure-mapping. Nonetheless, they point to the success of the overall endeavour of investigating the cognitive functioning of analogy as a ‘domain-general process’: “In few other domains has the connection between computational and psychological work been as close and as fruitful as in this one” (p. 231). Particularly, SME has been successful in modelling, explaining and predicting patterns found in experimental psychological studies. In addition, the representational issue has been addressed by using other software that can interpret data closer to the phenomena in terms of physical theories, which provide input to SME. Forbus, et al. point to a dividing line in the different perspectives in that Hofstadter and colleagues are primarily interested in the creativity of exceptional scientists, while the Gentner tradition has focused on more mundane everyday cognitive phenomena. However, what ultimately is at stake is the degree to which it is useful to analyse cognition and mind as limited to the brain, and where complexity calls for taking into account the entire body, but also the environment, in terms of the extended mind (Clark & Chalmers, 1998) or distributed cognition (Hutchins, 1995).

Another line of research has investigated how expert scientists and laypeople make use of analogies in more authentic settings than the psychology laboratory (e.g. J. Clement, 1988, 2009; Dunbar, 1995; Nersessian, 2008a). Overall, J. Clement (2009) frames analogical reasoning as one of several examples of non-formal, imagistic reasoning, which previously has not received the attention it deserves as a tool in science problem solving and theory development. For instance, Clement (1988) asked expert scientists (including a physics Nobel laureate) to decide whether a metal spring with twice the diameter of another spring, but otherwise identical (with regards to number of coils, length, material, etc.), would stretch
more, less or the same when the same weight is hung on them. In solving this problem, the participants spontaneously made extensive use of analogical reasoning and Clement gives the example case of subject S2, which has become a minor classic in the field. S2 first came up with the analogy between a spring and a straight metal rod, which would bend more the longer the rod. However, S2 was dissatisfied with this comparison, since it did not account for the homogenous stretching throughout the length of the spring. Now, S2 imagined segmenting the spring, first in square coils and then in hexagonal coils, which in the limit would behave like the original circular coils. When reasoning about such hexagonal coil, S2 got the critical insight: stretching of a spring is due not only to vertical bending, but also to horizontal torsion. Clement (2009) has later argued that analogies in this and subsequent studies on problem solving were drawn in ways quite different from what previously had been found in psychological experiments. First, inferences are rarely drawn straight off by recall or association from the source domains, but only after a process of transformation. This leads to a series of analogies which get aligned and evaluated through different cognitive mechanisms, including structure mapping, as suggested by Gentner (1983), but also: conserving transformation, where the subjects made sure that the compared sources and targets behave in the same way in crucial aspects, while certain features are changed; imagistic alignment, where the domains are compared before the mind’s eye through mental simulation; generation of intermediate bridging analogies where the domains are seen as too distanced. The ubiquity and usefulness of such gradual transformations in leading to insight led Clement to argue for a more inclusive view on analogy than Gentner, who would probably dismiss several of these examples as literal similarities. On a more fundamental note, Clement sees that with Gentner’s propositional, structural approach to analogy, she has not enough taken into account imagistic reasoning, which we will come back to in section 4.3.2.

Nersessian (2008a) has investigated how Maxwell used insight into mechanics and machine design in drawing analogies in his development of a theory of electromagnetism. She has also reanalysed Clement’s (1989) case study of subject S2. One conclusion is that as opposed to in laboratory experiments, analogical reasoning rarely comes alone. Analogies are rather used in conjunction with other non-formal ways of reasoning, such as imagery and thought experiments. In addition, the level of complexity is much higher in these more authentic examples than in the typical psychology laboratory set-ups. First, there is no clear-cut solution to the target problem to be discovered in the source domain. Instead, Nersessian describes an iterative process of bootstrapping where:

- each domain supplies constraints that can be looked at as one of the straps, the intermediary hybrid models are strap crossings, and each crossing supports or contributes to further model building and enhanced target understanding (Nersessian, 2008a, p. 133, italics in the original).

Nersessian further argues that “the interpretation of the target constraints and the goals of the target problems guide the selection of the salient constraints and candidates for transfer within the source domains” (2008a, p. 150). Although acknowledging Gentner’s (1983) focus on relational structure and the systematicity principle as particularly important for scientific reasoning, she lends support to the view of Holyoak and Thagard (1989) in that pragmatic and semantic constraints have to be considered also in the mapping process. In addition, Nersessian claims that the practice of using analogical problems with an available solution in the source domain in psychological experiments has led to neglect of the process of building the model representations. Reminiscent of the notion of heuristics analogies (Wilbers & Duit, 2006), novel features emerge through combinations of source and target constraints:

Casting the net wider to include self-generated analogies in problem-solving, especially in science, however, reveals the need for more empirical studies aimed at understanding a range of representation-
building processes. Although our cases might be considered extraordinarily creative, my intuition is that if spontaneous use of analogy in protocol experiments or "in the wild" were to be investigated systematically, building intermediary hybrid models would be seen to be a significant dimension of mundane usage. In sum, analogy comprises many processes, and the focus of current research leaves much of what is creative about such reasoning unaccounted for (Nersessian, 2008a, p. 153).

Nersessian (2008a) also points out that there is no limitation to one source domain only when approaching a target phenomenon in such authentic analogical reasoning. Maxwell, for instance, started by investigating the connections between magnetic phenomena and mechanics by modelling the rotation of one vortex in isolation. However, investigation of electromagnetic phenomena required modelling interaction between several vortices, which would cause jamming due to friction. Maxwell therefore recruited insight into engineering mechanics in seeing the turning vortices in terms of cog-wheels, some of which are idle.

Dunbar (1995; Dunbar & Blanchette, 2001) suggests that a combination of in vivo studies of scientific practice and in vitro studies in the psychology laboratory is more fruitful than using either of the approaches in isolation when investigating aspects of cognition, such as analogical reasoning. With the ambition of finding out 'how scientists really reason' in in vitro settings, Dunbar (1995) proposes that the typical "arbitrary task that has a tenuous relationship to real science" (p. 366) could be replaced by providing subjects with more authentic science cases, giving the example of how Monod and Jacob discovered that there are regulator genes that control the activity of other genes. However, in order to see how scientists use cognitive resources in representing and solving authentic research problems, typically in social contexts, Dunbar (1995) followed the activities in four molecular biology laboratories for one year. Intriguingly, Dunbar participated in a meeting and audio recorded the communication when an important discovery – how certain cells proliferate in certain regions of the body – was made in one of the projects he followed. Analogies were indeed found to be an important feature of the scientists' reasoning. First, local analogies within the same domain were often used, particularly in reference to similar experiments in practical troubleshooting. In addition, regional analogies between two domains belonging to a common class, for example phage viruses and retro viruses, were used in elaboration of theories, generation of hypotheses and experiment design. Finally, long-distance analogies between domains far apart were used to highlight salient features of the research when introducing new staff members to the work. However, in contrast to previous claims of the utility of such long-distance analogies, they were not involved in the phase of making new discoveries. Analogical reasoning was also found to be cultural-bound, as it was not used at all in one of the laboratories, where other troubleshooting routines had developed. Dunbar and Blanchette (2001) conclude from these and follow-up in vivo studies that the majority of comparisons are based on superficial similarities but nevertheless often useful for fixing practical experimental problems, while the remaining analogies focusing on structural relations are used for hypothesis generation. Subsequently, such in vivo findings have been used to design controlled in vitro experiments for the psychology laboratory, particularly with regards to self-generated analogies (Blanchette & Dunbar, 2000; Dunbar & Blanchette, 2001).

In conclusion, in my view, Gentner’s (1983) structure-mapping theory has become a kind of ‘touchstone’ in research of analogical reasoning, in the sense that other researchers have developed and tried their ideas, in arguing against it or identifying ways in which it could be refined. It is fair to say that Gentner and colleagues have been able to defend the overall view that relational structure is at the core of analog and have shown that the structure-mapping theory is useful in accounting for a wide range of psychological phenomena related to analogical reasoning. However, it is the entirety of the scientific debate, with attention to issues such as the particularities of human perception, contextual circumstances and cognition ‘in the wild’, which has brought our understanding of analogical reasoning forward.
3.1.2. Concepts and mental models

As a parallel to Meno’s paradox of learning, another problem looms large in the entire field of human understanding: *we do not know what other people think*, at least not directly. At despairing moments, it is even tempting to assume a solipsist philosophy: that a person cannot even know if anything exists, apart from the awareness of her own thoughts. Still, paradoxically, as children we develop a *theory of mind*, which enables us to attribute mental states and thoughts to ourselves and to other people (Goswami, 2007).

Different theories of learning have adopted different research approaches to the problem of our inability to get insight into our fellow human beings’ thoughts. With a *behaviourist* stance, the mind is treated as a black box, the inner intricacies of which we can never tell. We can only control and study the input we give it and the output we get from it in terms of the behaviour of a person. Within *sociocultural* perspectives, the interest is directed primarily towards the social interaction between individuals as they participate in a culture. In other words, as characterised by Sfard (1998), this line of research has been conducted against the background of a rather recent metaphor, where knowing or learning is characterised as *participation*. From this perspective, the sociocultural and behaviourist traditions have in common that there is no need for an elaborate account of what happens inside the head of individual human beings. *Cognitive* and *constructivist* traditions, on the other hand, take an explicit interest in the mind and thoughts of the individual, and assume the challenge of understanding how thoughts are structured in our minds and how we develop them, albeit in an indirect way due to the nature of the problem. One possible starting point is to assume that our personal thoughts are built up of building blocks, *conceptions*, which are similar to notions that have developed in the shared culture of individuals, *concepts*, including those that are used in specialist communities such as science. Such focus on the objects of knowledge fits with an *acquisition* metaphor, where coming to know something is seen as acquiring or gaining possession of it, either through transmission from somebody more knowledgeable or, from the constructivist perspective, through active construction by the learner (Sfard, 1998).

Taking the constructivist’s interest in what is inside our minds, the next question is: what are these concepts or conceptions? Carey (2009) argues that a fully worked-out theory of concepts should be able to account for a list of phenomena collated from the rather different perspectives of a theoretical philosopher and a cognitive psychologist. The philosopher would focus on: (1) reference, how concepts refer to the world; (2) distinguishing objective concepts from our subjective beliefs about them, conceptions; (3) epistemological warrant, justification of belief. Starting in this tradition, Carey accounts for the classical view of the ‘British empiricists’ such as John Locke, who categorised concepts as either primitive or complex. First, primitive concepts are sensory and intrinsically meaningful, and refer directly by being activated in specifiable ideal circumstances. For instance, the primitive concept ‘round’ is accessed by vision or touch and activated in interaction with round objects. Complex concepts, in turn, are defined from a set of primitive concepts by means of necessary and sufficient conditions and refer by the causal connection to them. These basic premises were elaborated within the logical positivism in the early 1900s, but have since received substantial criticism, some of which has emanated from the late Wittgenstein’s (1968/1953) notion of *family resemblance*. Take the complex concept ‘game’, for instance. Wittgenstein argues that you cannot determine its extension by listing a set of necessary and sufficient conditions. In football, you play collectively against another team, one of which will win to the demise of the other, typically for recreational purposes, while a game of ‘solitaire’ or playing charades share some but not all of these characteristics. The psychologist’s list of phenomena would be: (1) prototypicality in categorisation; (2) the inferential role of cognitive processing of mental representations, that meaning is not attributable to reference or extension in the world alone; and (3) concept acquisition, learning. Leaning on Wittgenstein, cognitive psychologist
Eleanor Rosch (1973) established the effect of prototypicality for categories such as ‘bird’, where ‘robin’ is seen as of higher ‘exemplariness’ than ‘chicken’ or ‘ostrich’ among participants in psychology experiments. A psychological effect of graded prototypicality appears in spite of all these three entities being perfectly valid representatives of the avian class of animals, hence fulfilling any rational necessary or sufficient conditions for what a bird is. The other way around, a rubber duck would be readily recognised as a bird, albeit of slightly less prototypicality than the robin, without fulfilling such criteria. On the other hand, Keil (1989) found that against a background of transformation so that a raccoon comes to look exactly like a skunk (plastic surgery, fur dying, smelly gland implant, etc.) and thus fulfilling any conceivable descriptions of a skunk, we still claim that it is a raccoon, showing an essentialist commitment which contradicts both the empiricist view and that of prototypicality. In conclusion, Carey (2009) argues in favour of a theory-theory of concepts, i.e. a theory that “many everyday concepts are terms in intuitive theories” (p. 502).3

Another tack on concepts is to identify them as mental models, a notion that was introduced by Craik (1943) and became influential in cognitive science by the early 1980s. Nersessian (2008b) reviews the heterogeneous theoretical framework that has developed around mental models, particularly from the perspective of conceptual change in science and science learning. In her interpretation, the core of the argument is that we have a capacity to make up ‘small-scale models’ of physical phenomena in our minds, for the purpose of being able to simulate potential outcomes of action. Nersessian (2008b, p. 393) gives the everyday example of how we solve the problem of getting a large sofa through a door-way:

The default approach to solving the problem is usually to imagine moving a mental token approximating the shape of the sofa through various rotations constrained by the boundaries of a doorway-like shape. In solving the problem, people do not customarily resort to formulating a series of propositions and applying logic or doing trigonometric calculations.

These representations are perception-based, but figurative and schematic and in this respect not like running a ‘movie in the head’ with realistic interacting entities. In Nersessian’s (2008b) view, such ‘model-based reasoning’ is made use of also in science problem solving (much more of which is to follow), although it is complemented by a wide range of specialist activities, such as physical experiments and mathematical calculations.

Nersessian (2008b) sees a fundamental dividing line in cognitive science between researchers that claim that all mental representations are of a logical, language-like, ‘propositional’ nature, which refer to corresponding physical phenomena by description, and those that hold the idea that there are also perceptual, imagistic, ‘iconic’ representations, which refer by means of demonstration or similarity. A further distinction may be made between ‘modal’ representations, which retain some perceptual similarity to what is represented, and arbitrary ‘amodal’ representations, which do not. Nersessian sees Craik’s original account as an example of the increasingly dominating view of mental models as of a primarily iconic and modal nature. These developments are clearly parallel to the debate about analogical reasoning with regards to its focus on structure only.

In the seminal book Mental models, Johnson-Laird (1983) puts forward the view that we typically do not make inferences by following formal logic, but by reliance on mental models, each representing a distinct possibility. For instance, Johnson-Laird (2006) brings forward evidence that we are slower at conceptualising inclusive disjunctions (A or B, i.e. the union

---

3 As we saw above, however, within science education research the view of concepts as part of intuitive theories has been criticised by diSessa (1983), claiming that children’s intuitive ideas are far more fragmented than ‘theories’, and coordinated in a flexible context-dependent way.
$A \cup B$ in set formalism) than corresponding exclusive disjunctions (A or B, but not both, i.e. the symmetric difference $(A \cap B') \cup (A' \cap B)$), in spite of their simpler logical structure. The reason, according to Johnson-Laird, is that conceptualising an inclusive disjunction requires creating three different mental models (only A; only B; A and B), whereas the exclusive disjunction is represented by only two mental models, and thereby faster to process by our minds. In addition, as opposed to logical deduction, our problem solving abilities are dependent on the familiarity and degree of abstraction of the context. This conclusion is supported with evidence from psychological experiments with the selection task, where a subject is invited to decide the minimum amount of evidence that is required to test if a rule of the form ‘if $p$ then $q$’ is violated. If the rule is abstract and out of context, such as: “If a card has a vowel on one side then it has an even number on the other side” (Johnson-Laird, 1983, p. 30), then we are poor at picking out what evidence is required to establish rule violation. On the other hand, concrete contextualised problems of the same logical structure, e.g. “Every time I go to Manchester I travel by train” (p. 31), are easier to solve. The gist of Johnson-Laird’s argument goes along the lines of: a) we have exceptional cognitive abilities; b) we are poor at making logical deductions (at least in certain contexts); c) hence, our cognitive abilities are not (exclusively) attributable to logical deductions.

Since the introduction of the notion of mental models, it has been recognised that cognition is not an isolated phenomenon in a person’s mind. Rather, human cognition is both ‘embodied’ in the sense that it is grounded in our perceptions and bodily experiences (e.g. Barsalou, 2008; Johnson, 1987) and ‘distributed’ in that one person’s cognition is coordinated with external representations and the cognition of others (Hutchins, 1995; Zhang & Norman, 1994). For instance, in Cognition in the wild, Hutchins (1995) describes the interplay between people and equipment required in navigating a fleet into a harbour. We return to the issue of distributed cognition in relation to representational systems in section 3.2.4.

### 3.2. Semantics – the study of the meaning of language

We now turn to the second perspective of the theoretical framework, semantics, the study of the meaning of words and compounds of words. Semantics may be seen as a subdiscipline within the larger fields of linguistics and philosophy of science, which are touched upon in this section. In addition, from a semiotic perspective, apart from spoken and written language, we also encompass other signs and representations, particularly visual external representations.

#### 3.2.1. Semantics – meaning and reference

Linguistics, the study of language can be seen as roughly (but, as we will see, not unproblematically) divided in different sub-disciplines, including (Saeed, 1997):

- **Syntax**, the grammatical structure of language, how sentences are formed by words.
- **Semantics**, the study of meanings of words and sentences.
- **Pragmatics**, the study of the use of language in context, in terms of communicative acts.
In semantics, there are fundamentally two different theoretical approaches to how words and sentences come to mean something. On the one hand, with referential approaches, meaning of words is determined by what they refer to or denote in the world. On the other hand, with representational approaches, meaning derives from our conceptual structures (Saeed, 1997).

Both are instantly attractive and appealing to common sense, but are problematic when analysed further within linguistics (Saeed, 1997) and the philosophy of language (Lycan, 2000). In referential approaches to semantics, “linguistic expressions have the meanings they do because they stand for things; what they mean is what they stand for” (Lycan, 2000, p. 3).

For instance, ‘cat’ means what it does in reference to the animals roaming about in the world. According to Lycan (2000), two of the main problems with a referential approach are:

- There are words that are meaningful, but without referring to anything in the world, e.g. ‘Pegasus’ or ‘if’. (In fact a sentence like ‘Pegasus never existed’ is true, without implying reference.)
- There are co-referring words, differing in meaning, e.g. ‘the Pope’ and ‘Joseph Ratzinger’.

In response to the weaknesses of a purely referential approach to meaning, Frege (1948 [1892]) distinguished a word’s referent from its sense. In Lycan’s (2000, p. 32) words:

> For Frege, the ‘sense’ was, roughly, a particular ‘way of presenting’ the term’s putative referent. Though itself an abstract entity, rather than a mental or psychological one, the sense reflects a person’s conception or way of thinking of the referent.

With a representational approach, Russell (1905) developed a Theory of Descriptions, in which the meaning of ‘definite descriptions’ (such as ‘the author of Waverley’), but also ‘proper names’ (the author ‘Sir Walter Scott’), are determined by a set of necessary and sufficient conditions or logical propositions.

Kripke (1980) criticised Russell’s view that reference fixing may be established by a set of necessary and sufficient descriptions. Indeed, all known descriptions of an entity may eventually prove to be wrong, but we still refer to the same thing! Instead, Kripke offered the view that the reference to an entity is fixed by an imagined moment of baptizing in history, connecting the entity to a word:

> Someone, let’s say a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain (p. 91).

This historical causal chain of communication gave rise to the Causal Theory of Reference, developed by Putnam (1973) and Kripke, where reference of a word is established without resorting to tentative descriptions. Apart from accounting for reference of proper names such as that of a person, the theory was elaborated to account also for natural kinds, e.g. ‘tiger’ and ‘gold’. For such terms, necessary and sufficient conditions were particularly problematic, since what is seen as fundamental and a necessary condition in one scientific theory, e.g. a DNA sequence of a species or the atomic number of an element, may be abandoned by the next theory, due to the intrinsic tentative character of the scientific endeavour.

Not surprisingly, weaknesses were found also in the Causal Theory of Reference as an account of word meaning. Devitt and Sterelny (1999) put forward the quia problem, based on the fact that an entity may belong to several categories. If you point to an entity and say ‘that is a tiger!’, you may intend any category in which it is included, among these ‘animals’, ‘a particular species of large felines’ and ‘striped, furry things’. In order to discriminate between
this wealth of possible readings, some kind of description of what the name means is needed. In addition, the Causal Theory of Reference could not account for how certain words have been abandoned as a consequence of losing their reference. For instance, accounting for how ‘phlogiston’ lost its reference requires it to be embedded in a theory of its assumed characteristics. The challenge is to strike a balance between referential and representational approaches, for instance by means of causal descriptivism, where a causal chain is complemented by a set of specifying descriptions (e.g. Kroon, 1985).

If we, after all, conclude – with Frege (1948 [1892]) – that the sense of a word is required in order to establish its meaning (although also a referent is needed), the next question would be: what is the sense? One approach is that senses are mental items, such as mental images or concepts. The idea of senses as mental images has been criticized, since images are too specific. For instance, there is no one suitable image for ‘triangle’ or ‘horse’ which reflects the width and flexibility in the meaning of these words. It fares even worse with abstract words, such as ‘democracy’. However, the concept approach has its fair share of issues:

The most usual modification of the image theory is to hypothesize that the sense of some words, while mental, in not visual but a more abstract element: a concept. …/… This seems reasonable enough but the problem for many linguists is that psychologists are still very involved in investigating what concepts might be like. Unless we have a good idea of what a concept is, we are left with rather empty definitions like ‘the sense of the word dog is the concept dog’ (Saeed, 1997, p. 33).

In other words, seeing senses as identical to concepts hands over the problem to cognitive psychology and the question of what a concept is. An alternative approach, followed by for example Russell (1905) above, is to see senses as a set of abstract propositions, with a truth value independent of people’s minds and natural languages, and ideally eternal. In this view, propositions are objects of mental states; you may think about a proposition or believe that it is true, but the proposition and mental states are not identical.

One problem with propositional approaches is the abstract character of propositions; they do not stand in a causal relation to our actions or what happens in the world, i.e. the referential connection that Kripke (1980) tried to establish with the Causal Theory of Reference. In contrast, the late Wittgenstein (1968/1953) argued that words and sentences should not be seen as ideal, abstract entities. In Lycan’s (2000, p. 76) interpretation, words and sentences should rather be seen as ‘game pieces or tokens, used to make moves in rule-governed conventional social practices. A ‘meaning’ is not an abstract object; meaning is a matter of the role an expression plays in human social behavior”. In addition, Strawson (1950) shows that Russell runs into problems with his abstract approach of interpretation of the meaning of a word or a sentence, since he ignores the context – the concrete communicational practice – in which it is uttered, the influence of pragmatics on semantics, if you will. One approach to make the distinction between semantics and pragmatics clearer is to use different notions for, on the one hand, different stable meanings of words in our shared language, corresponding to entries in a dictionary, and on the other, subjective, context-dependent intentions of a person’s language in use. In cognitive linguistics, more of which to come in sections 4.1.2 and 4.1.3, it is common to use the notion ‘sense’ for the former, stable semantics of words, and apply ‘meaning’ to the more subjective, pragmatic intentions of words (Evans & Green, 2006), a distinction which is adopted in general throughout our research work and in this dissertation.6

6 Somewhat confusingly, in English translations of Vygotsky’s oeuvre (e.g. Vygotsky, 1962), the terms are swapped, so that stable ‘meanings’ of a word may correspond to several, subjective ‘senses’ unique to particular communicative situations. This convention is not adopted here.
3.2.2. Polysemy and homonymy

Polysemy is the semantic phenomenon where one word has two or more distinct, yet related senses. For instance, ‘paper’ may denote: a material made of wood pulp; a sheet of said material; a journal composed of several such sheets; its electronic counterpart, etc. This should not be confused with homonymy, where one word has several unrelated senses, as in ‘bat’: nocturnal, flying mammal vs. piece of equipment in baseball and cricket.

3.2.3. Metaphor

As a figure of speech described in ancient Greece, a metaphor has the structure ‘X is Y’, but goes beyond the literal truth of the sentence. For example, with the statement ‘my boss is a pig’, an employee does not primarily intend to convey that his immediate superior is a porcine mammal – which a literal interpretation would yield – but rather allude to his questionable behaviour and moral character or possibly his personal hygiene; hence a metaphor. The meaning of ‘metaphor’ is often widened to represent other forms of figurative language, not strictly following the form ‘X is Y’, such as ‘one giant leap for mankind’.

Saeed (1997) distinguishes between two fundamentally different traditional perspectives on metaphor. First, there is the classical view on metaphor, ascribed to philosophers such as Aristotle. Here, metaphor and other figures of speech are seen as a merely decorative feature of language. Figurative expressions can and preferably should be replaced by more literal counterparts, carrying the same meaning. In contrast, there is the romantic view on metaphor, in which figurative speech is indispensable and permeating all language, represented for example by the romantic 19th-century poet Coleridge.

From the point of view of philosophy of language, metaphor is intrinsically problematic, and Lycan (2000) refer to them as coming from ‘the dark side’, since we cannot establish an answer to the pertinent question ‘what do metaphors really mean?’ or ascertain whether a metaphor is true or not. Hence, the patronising classical view has long dominated the field. As Max Black introduces in the essay Metaphor, in his influential Models and Metaphors:

To draw attention to a philosopher’s metaphors is to belittle him – like praising a logician for his beautiful handwriting (Black, 1962, p. 25).

Black argues that, at his time, the dominating perspective was a substitution view of metaphor, in line with the classical view above, where a metaphorical expression is seen as a substitute for a literal expression, which would have expressed the same meaning. In this view, metaphor may achieve effectiveness of language. In cases where the appropriate word is missing: “Metaphor plugs the gaps in the literal vocabulary (or at least, supplies the want of convenient abbreviations)” (pp. 32-33). However, often metaphor is seen to serve merely a stylistic or decorative purpose. In the slightly elaborated comparison view, metaphors are seen as based on underlying comparisons, in the form of simile or analogy. The difference between the metaphor ‘my boss is a pig’ and the corresponding simile ‘my boss is like a pig [in the respect that he acts disrespectfully]’, is that of form and that the metaphor implies a comparison that the simile states explicitly. However, this fails to bring across the novel, surprising effect of a good metaphor. Rather: “It would be more illuminating in some of these cases to say that the metaphor creates the similarity than to say that it formulates some similarity antecedently existing” (p. 37). Consequently, Black puts forward his preferred interaction view of metaphor. Here, in the interpretation of ‘man is a wolf’, ‘man’ and ‘wolf’

---

7 However, Kittay (1987) claims that Aristotle has been ascribed a more negative and naïve view on metaphor than what he embraced.
interact as seen through a filter. A ‘system of associated commonplaces’, typically attributed to wolfs, is evoked, but in a non-trivial way. “The wolf-metaphor suppresses some details, emphasizes others – in short, organizes our view of man” (p. 41), but also our view of the wolf. In this view, a metaphor cannot be replaced by a literal counterpart without losing its cognitive content and potential for insight. Black (1993) further distinguishes metaphorical statements from metaphor-themes, where the first relates to specific sentences or sets of sentences as parts of communication or expression, while the letter is an abstraction of the format ‘the metaphor of A as B’, available for repeated use and elaboration.

3.2.4. Extension of ‘language’ – Semiotics and representations

A word may be seen as corresponding to a referent in the physical world and a sense, the intention of the speaker. The field of semiotics – the study of interpretation of signs – can be seen as an extension of semantics; in addition to interpreting the meaning of spoken and written language, the focus is now widened to include other representations of the world and our thoughts, such as images, graphs, diagrams, actions and gestures, etc. Peirce (1985) proposed a typology of different kinds of signs, where the three overarching categories are: an icon, which is in some way similar to the object it stands for; an index, which has some kind of causal relation to the object, for instance a bullet-hole being an indication of a shot, and; a symbol, which is arbitrary and conventional, e.g. words. In addition, he brought forward a semiotic triangle in order to display the relationship between three types of entities (p. 5):

A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the interpretant of the first sign. The sign stands for something, its object.

Ogden and Richards (1923) present an alternative semiotic triangle, where a symbol is causally connected to a thought or reference, which, in turn, is causally connected to a referent. In this framework, however, there is typically not a causal relation between the symbol and the referent, due to the arbitrary nature of the symbol. As mentioned, this framework was adapted in the development of 2-D SAS (Strömdahl, 2012) and article I of this dissertation and in the following, this broad interpretation of ‘symbol’ will generally be adopted, roughly equivalent to Peirce’s ‘sign’.

Zhang and Norman (Zhang, 1997; Zhang & Norman, 1994) have investigated the effect of different external representations on our problem-solving abilities in psychological experiments, and developed a theoretical framework of the interaction between such external representations and internal representations in our minds in terms of distributed cognition (Hutchins, 1995). Zhang (1997) defines external representations as:

the knowledge and structure in the environment, as physical symbols, objects, or dimensions (e.g., written symbols, beads of abacuses, dimensions of a graph, etc.), and as external rules, constraints, or relations embedded in physical configurations (e.g., spatial relations of written digits, visual and spatial layouts of diagrams, physical constraints in abacuses, etc.) (p. 180).

Internal representations, on the other hand, are the knowledge and structure in our minds and memory. Zhang argues that the interaction between external and internal representations has not received sufficient attention. For instance, artificial intelligence research has typically seen external representations as perceptual input for subsequent cognitive processing. Within

---

8 Ogden and Richards note that exceptional non-arbitrary symbols and referents exist, such as onomatopoeic words or pictures. Ideographic written languages, such as Mandarin, is another case.
the situated cognition or socio-cultural traditions, on the other hand, the need to analyse internal cognitive models of the environment are not focused upon. In contrast to both these traditions, the distributed cognition approach “explores how cognitive activity is distributed across internal human minds, external cognitive artifacts, and groups of people, and across space and time” (p. 182).

Zhang and Norman (1994) made psychological experiments of how the representational structure influences the ability to solve the Tower of Hanoi (TOH) problem, where three disks with holes in them are to be moved from one pole to another, and other isomorphic problems, i.e. with the same relational structure. One difference was that some rules of the game could be represented either externally or internally. For instance, in the standard version of the TOH problem, disks are stacked onto each other, providing the physical constraint that you can only pick the top disk, if you pick them one by one. In contrast, if the same problem is framed as picking one of three oranges on a plate, you have to remember the rule that you can only pick, for instance, the largest orange. Other factors were also varied in the experiments, such as if there is an ordinal dimension (e.g. big, medium, small) or arbitrary categories (e.g. red, green, yellow). In psychological experiments, they found that the ability and time to solve the problems varied considerably. Although the problems are isomorphic from the designer’s or analyst’s point of view, they become completely different problems to the problem solver confronted with the task. In conclusion, the study provides evidence for a representational effect: “…the phenomenon that different isomorphic representations of a common formal structure can cause dramatically different cognitive behaviors” (p. 88).

Similarly, Zhang (1997) investigated how the nature of the external representations influences problem solving of isomorphic versions of Tic-Tac-Toe, where subjects were invited to play against a computer, where the task was to find a strategy that achieves draws with the computer. For instance, in the ‘Line’ version of the problem, subjects could be assisted by the external representation of winning triplets arranged in straight lines, as opposed to the ‘Number’ version where such winning triplets, adding up to 15, would have to be represented internally. Another difference was the influence of biases or heuristic rules, where the ‘Number’ problem was disadvantaged, since following a common ‘larger-is-better’ rule would favour ‘9’, instead of the optimum ‘5’ as a starting move. Such differences in external representations had a large effect on the success in finding the appropriate strategy. External representations are seen as providing affordances (Gibson, 1979) to certain interpretations in our interaction with the environment. In conclusion, in cases where the underlying abstract structures are not known, “the format of a representation can determine what information can be perceived, what processes can be activated, and what structures can be discovered from the specific representation. This is called representational determinism” (Zhang, 1997, p. 213). Zhang finds representational determinism particularly pronounced in the impact on our cognition of the historical introduction of writing, and when comparing the Arabic position numeration system to more cumbersome predecessors. As we have seen above, the influence of representational forms has been recognised also in mathematics and science education (Duval, 2006; Fredlund, et al., 2012; Roth, et al., 1997).

Vessey (1991) has developed the idea of cognitive fit as an approach to assess whether information presented in a table vs. in a graph is most effective for decision-making performance. Rather than pointing out one type of representation as categorically superior, she argues that it is a matter of fitting the representation to the task at hand. Graphical representation presents spatial relations and therefore emphasises information about relationships in the underlying data set and provides an overall view, whereas tables represent specific, discrete data values. Therefore, if a task involves making comparisons or seeing trends, graphs are more efficient than tables. Conversely, if it is a matter of lookup of a particular value, tables are to be preferred.
### 3.2.5. Pragmatics – language in context

When the importance of language is attended to in educational research, it is typically from a socio-cultural perspective of language in discourse and on language as a means of providing ‘scaffolding’ for our learning (Wood, et al., 1976), i.e. with a focus on pragmatic aspects. Such considerations have impacted the methodological approaches to the two last studies included in this dissertation, particularly with regards to peer collaboration and adult-child interaction. We come back to these issues in the methodological framework in chapter 5.

### 3.3. Scientific modelling – within philosophy of science

The third perspective of the theoretical framework is scientific modelling. It is studied primarily in the philosophy of science, which takes an interest in the characteristics of the scientific method and issues of epistemology, our knowledge of the physical world and how we justify the knowledge (Ladyman, 2002). Here, we focus on cognitive and historical aspects of scientific modelling. Although interesting and relevant for the field of study, the issue of scientific realism vs. different types of relativism is largely excluded in this dissertation.

#### 3.3.1. Scientific modelling

Models can be seen as bridges between scientific theory and the world-as-experienced. However, they come in all shapes and sizes and can be used for a variety of purposes. They can be simplifications, providing explanations of experienced real-world phenomena, but also be generated from theories, in order to make predictions and compare with the real world through experiments. Models may be resized representations of objects, such as scale models of viruses or galaxies, or depicting abstract concepts, such as processes, mathematical relationships or systems with their interconnected components (Gilbert, 2004). One broad distinction may be made between internal, cognitive mental models, related to above, which we use in imagining situations and phenomena, and external models, used to convey and communicate our ideas through some kind of medium, the primary focus in this section.

Black (1962) categorises scientific models in different types, ranging from the concrete to the more abstract and theoretical:

- **Scale models**, preserving the relative proportions of an original. “We try to bring out the remote and unknown to our own level of middle-sized existence” (Black, 1962, p. 227).
- **Analogue models**, which through a change of medium reproduces the structure of relationships of the original. The aim is to establish correspondences or, ultimately, an isomorphism between the original and the analogue, regardless of irrelevant surface characteristics, which inevitably will differ between the media.
- Through the use of **mathematical models**, a phenomenon is represented in mathematical terms. This involves the identification of relevant variables, formulation of hypotheses, simplifications to make calculations manageable, solving the formulated mathematical equations and formulation of predictions, suitable for empirical testing.
• Black uses Maxwell’s representation of the electrical field as an incompressible fluid as the paradigm of a theoretical model. Initially, Maxwell used the model as an imaginary aid for thought, but came to regard it as an increasingly realistic representation. In this way, theoretical models can be classified by the way of reasoning. Either the electrical field is handled as if it were filled by an incompressible medium, or alternatively it can be taken as being such a medium. While the first way of reasoning is seen as a ‘heuristic fiction’, the other aspires to an ontological categorisation. The main advantage of theoretical models is that they can be a speculative instrument in the exploration of a new field, by analogy with a better known field. In the same way as metaphors, the theoretical model selects and magnifies certain aspects that could not be seen otherwise.

• Finally, with archetypes, Black (1962) refers to “a systematic repertoire of ideas by means of which a given thinker describes, by analogical extension, some domain to which those ideas do not immediately and literally apply” (p. 241), and gives the example of using expressions from physics, e.g. ‘tension’ or ‘forces’, in accounting for psychological phenomena. Black argues that we often use an archetype without being aware of it, since it is “an implicit or submerged model operating in a writer’s thought” (p. 239).

Hesse (1966) accounts for the role of models and analogies in science by referring to proponents of two very different views on the character of physics. First, Duhem (1991/1914) distinguishes between a French/German tradition, using an abstract mathematical approach, and an English tradition, making use of concrete, material models. As noted by Hesse (1966, pp. 70-71), Duhem uses the example of electrostatics in a devastatingly sarcastic way:

This whole theory of electrostatics constitutes a group of abstract ideas and general propositions, formulated in the clear and precise language of geometry and algebra, and connected with one another by the rules of strict logic. This whole fully satisfies the reason of a French physicist and his taste for clarity, simplicity, and order. /.../ Here is a book [by Oliver Lodge] intended to expound the modern theories of electricity and to expound a new theory. In it there are nothing but strings which move around pulleys, which roll around drums, which go through pearl beads, which carry weights. /.../ We thought we were entering the tranquil and neatly ordered abode of reason, but we find ourselves in a factory.

Duhem concludes that models may be useful psychological tools in approaching scientific theories (although less often than generally supposed), but they are not essential elements of the theories. In contrast, Campbell (1920) argues for the opposite: that models are an essential part of theories. Without models, theories cannot provide intelligible explanations of the phenomena they account for and making predictions is possible only by use of models. Hesse (1966) formulates her position as a hypothetical debate between modern disciples of the two antagonists. She argues that although advances in modern physics have made the field increasingly abstract in favour of Duhem’s position, “an element of truth remains in Campbell’s insistence that without models theories cannot fulfil all the functions traditionally required of them, and in particular that they cannot be genuinely predictive” (1966, p. 5). We look closer on Hesse’s view on scientific models and analogies in section 4.2.1.

One characteristic of modelling is that several – possibly mutually inconsistent – models may be used to investigate and illustrate different aspects of one phenomenon. The paradigmatic case is the wave-particle duality in quantum physics, epitomised by the double-slit experiment, in which it can be shown that any particle has wave-like characteristics. Particle-based and wave-based descriptions cannot be reconciled logically, but, according to Bohr (2011/1929), contribute to a more complex understanding of phenomena in a complementary way, taking into account also the observer of the phenomena. Using multiple models of a phenomenon is not limited to the individual’s multiple conceptualisations, but is an intrinsic characteristic of the science practice.
4. Synthesis of the theoretical framework

Up to this point, I have given a theoretical background of the three theoretical perspectives of the dissertation in isolation. I now proceed to develop a synthesis of the theoretical framework by investigating how the individual theoretical perspectives relate to one another.

One overall characteristic that is shared between analogical reasoning, scientific modelling and reference in semantics is that they all involve the process of mapping across entities or representing one entity in terms of another, as shown in Figure 5.

![Figure 5. Shared structure of analogical reasoning, scientific modelling and reference in terms of mapping or representation.](image)

Although the structure of mapping is a common feature, an important difference is the direction and degree of symmetry of the relations. In scientific modelling, the emphasis is on getting to know more about the world by means of devising a model for it (Ladyman, 2002). In semantics, on the contrary, our focus is instead on understanding what words and expressions mean by establishing what they refer to (Saeed, 1997). Analogical reasoning, finally, may be seen as the more generic of the three and is recruited to make inferences about a target domain by invoking a structurally similar source domain, i.e. typically an asymmetric relation (Gentner, 1983). However, analogical reasoning may also be used symmetrically in mutual alignment or ‘bootstrapping’ when approaching two unknown domains of potentially similar structure (Kurtz, Miao, & Gentner, 2001).

4.1. Analogical reasoning and semantics

We first put side by side the realms of analogical reasoning on one hand, and semantics and semiotics on the other, and discuss a number of touching points. In particular, we analyse the debate whether or not interpretation of metaphors is connected to analogical reasoning.
4.1.1. Analogy and relational language

Gentner (2003) addresses the ancient fundamental question of “why we’re so smart”, i.e. how come the human species has such outstanding cognitive and learning capabilities in comparison to other species. The short answer is ‘our knack for analogy and language’:

[What makes humans smart is (1) our exceptional ability to learn by analogy, (2) the possession of symbol systems such as language and mathematics, and (3) a relation of mutual causation between them whereby our analogical prowess is multiplied by the possession of relational language. /.../ In short, analogy is the key to conceptual learning, and relational language is the key to analogy (pp. 196-197).

Relational terms are terms that establish relations between entities. There are relational nouns. For instance, ‘parent’ implies a relation between an adult and his or her child, and ‘mammal’ stands in a superordinate type/token relation to a class of more prototypical cats, bats and other beasts of that ilk. There are also relational verbs, such as ‘cause’ and ‘prevent’, prepositions, e.g. ‘on’, etc. Gentner (2003) argues that relational terms are generally more abstract and difficult to learn, in comparison to words denoting objects, and more sensitive to idiosyncrasies of a particular language or culture. The upshot, however, is that they serve a number of important functions in learning and cognition. By hearing and remembering a relational term, initial registration, we become alert to the relational aspect of the phenomenon, even before we have fully grasped its meaning. In addition, assigning a name to a relational pattern facilitates abstraction out of a particular context. Next, relational terms imply reification of relational patterns, so that we can use them as entities in higher-order reasoning in a recursive manner. Finally, coming to use a stable set of relational terms promotes uniform relational coding: we are more likely to see new situations in terms of the relational structure provided by relational language, i.e. transfer of knowledge.

4.1.2. Analogy and metaphor

One immediate point of connection between analogical reasoning and semantics is the issue of the relationship between analogies and metaphors and how analogical reasoning relates to generation and interpretation of metaphors. It is tempting to use the terms interchangeably and they are often lumped together in the fixed phrase ‘analogies and metaphors’, but there are subtleties to the matter. One possible approach is to see ‘metaphor’ and ‘analogy’ as two concepts – from different disciplines or at different levels – that typically can be used to describe the same phenomenon and often come together. A metaphor may be seen as a language phenomenon, an explicit expression that invites to seeing something in terms of another, e.g. ‘man is a wolf’. In order to interpret the metaphor, we carry out a structural mapping, i.e. use analogical reasoning, belonging to the realm of psychology. Another dimension along which metaphor and analogy can be discriminated is that of elaboration. While the metaphor typically only alludes to a connection between two domains, the analogy spells out the connection explicitly. Duit (1991, p. 650) sees this difference as crucial: “It appears to be the very essence of a metaphor that the grounds of the comparison are hidden. Metaphors always have some aspect of surprise; they provoke anomaly.” In this perspective, what starts as a metaphor, leaving the interpretation to the listener, may form into an analogy by pointing out the relevant correspondences across domains. However, while fleshing out the details of the analogy increases the informational content, the metaphor loses something which may have been stimulating an interest: the suspense and suggestion of a novel idea.

Ortony (1979) regards interpretation of metaphors (A is B) and similes (A is like B) as cases of nonliteral similarity. He acknowledges Tversky’s (1977) account of the asymmetry of literal similarities, but claims that it cannot sufficiently account for the radically asymmetrical relation of such nonliteral similarity statements. However, by recognising a
salience imbalance of matching attributes of the two domains as integral to metaphors, he offers a modified version of Tversky’s model, which focuses on the difference in salience of matching attributes between the target and source domains. If the objects which are compared share a lot of high-salience attributes, as in the example ‘billboards are like placards’, it is a literal comparison. If the source has primarily high-salient attributes that are either matched with low-salient attributes of the target or not matched at all, as in ‘billboards are like warts’, the match is non-literal – in this case a simile. The other combinations (low salience in source matched with low/high salience in the target) are anomalous. For instance, compare the anomalous ‘sleeping pills are like sermons*’ to the corresponding ‘sermons are like sleeping pills’, where the shared ‘anaesthetic’ quality has higher salience in the sleeping pills. The recognition of the salience level of matching attributes relative to the source is particularly powerful in accounting for the quite different interpretation of inverted similes, e.g. when going from ‘butchers are like surgeons’ to ‘surgeons are like butchers’. In addition, Ortony argues that it a common mistake is to assume that a simile invariably would be more literal than a corresponding metaphor. For instance, ‘encyclopaedias are like gold mines’ is no less figurative than the corresponding ‘encyclopaedias are gold mines’. Instead, “the locus of metaphoricity lies not in the surface structure of a statement (e.g., the presence or absence of ‘like’) but in the underlying comparison itself” (Ortony, 1979, p. 175). Similarly, Gentner brought forward the view that metaphors are typically based on analogical reasoning when introducing the structure-mapping theory:

Many (perhaps most) metaphors are predominantly relational comparisons, and are thus essentially analogies. …/… Although most metaphors are relationally focused, some are predominantly attribute matches. …/… Finally, for metaphors that are analyzable as analogies or combinations of analogies, the mapping rules tend to be less regular than those for analogy (Gentner, 1983, p. 162).

The view that interpretation of metaphors involves analogical comparison between domains has been challenged by Glucksberg and colleagues (e.g. Glucksberg & Haught, 2006; Glucksberg & Keyser, 1990; Glucksberg, McGlone, & Manfredi, 1997). Glucksberg and Keyser (1990, p. 6) claim that there is a “fundamental flaw” in Ortony’s (1979) salience imbalance account, since in order for a comparison statement to be of informative value to the listener, i.e. saying something new about the target, it has to have low salience attributes in the target and high salience attributes in the source, regardless of its metaphoricity. Therefore, “salience imbalance cannot distinguish between literal and metaphorical comparisons” (p. 6).

In addition, comparison approaches could not explain expressions that were not merely radically asymmetric, but indeed irreversible. For instance, while ‘alcohol is (like) a crutch’ makes tragic sense, ‘a crutch is like alcohol’ resists interpretation (and ‘a crutch is alcohol’ is Yoda-style poetry). This is difficult to cater for with weighted salience.

Instead of comparison approaches to metaphor, Glucksberg and Keysar (1990) put forward the alternative account of metaphor interpretation in terms of class-inclusion assertions. Literal sentences of the form ‘A is a B’ are often interpreted in terms of class inclusion in a type/token relationship, e.g. ‘a robin is a bird’, while literal sentences of the form ‘A is like B’ typically are comparisons by similarity, as in ‘a robin is like a bird’.

Making a comparison of the terms of a class inclusion sentence, and vice versa, often yields anomalies; for instance, ‘a robin is like a bird*’ or ‘a robin is a sparrow*’ do not make much sense. Similarly, changing places of A and B in ‘A is like B’ gives another anomaly, e.g. ‘a bird is a robin*’, due to the type/token relationship being intrinsically hierarchical and asymmetric. What Glucksberg and Keysar propose is that, much as the literal ‘a robin is a bird’, where the entity ‘robin’ is classified as belonging to the category of ‘birds’, the metaphor ‘my job is a jail’ is interpreted along the lines of ‘my job’ belonging to a category, represented by ‘a jail’. Now, a jail can be seen as belonging to many different categories, such
as ‘legal sentences’ or ‘buildings’. In this particular case, ‘jail’ is seen as representing a category of unpleasant, confining and punishing situations. The word ‘jail’ comes to serve a dual function: it is both the name of the new metaphorical, superordinate category and the name of the more literal prototypical instantiation. In the metaphorical reading, ‘my job’ belongs to the superordinate category. Bowdle and Gentner (2005) note that one weakness of this class inclusion account is that it cannot explain which one of a potentially infinite number of existing and ad hoc categories the source is seen as belonging to in metaphoric interpretation. Glucksberg, et al. (1997) proposed a refined account, where the appropriate category is selected by means of interactive property attribution, where the target and source make interactive but different contributions to the interpretation of a metaphor. As in the original class inclusion model (Glucksberg & Keysar, 1990), the source provides the properties to be attributed to the target, but additionally, in the interactive property attribution model, the target provides constraints on what properties are plausible, lending support to Black’s (1962) interaction view on metaphor.

Due to the relative merits and shortcomings of both comparison and class inclusion approaches to metaphor, Bowdle and Gentner (2005) propose an integrated view in terms of a career of metaphor, relying on Gentner’s (1983) structure-mapping theory. In this view, interpretation of novel metaphors is always done as comparison by means of structure mapping. Then, as metaphoric use of a source gradually becomes conventionalised, a metaphor can be interpreted either by means of comparison or categorisation. The career of metaphor may be extended to dead metaphors, where the connection between the source and target is forgotten (e.g. the metaphorical connection of a biological cell to the humble abodes of mediaeval monks) and ultimately the original source meaning is abandoned (as in the original physical blockbuster type of bombs). Glucksberg and Haught (2006, p. 363) admit that “[t]here is now general agreement that metaphors can be processed either as comparisons or as categorizations”, but argue that it is not a metaphor’s novelty that decides whether it will be interpreted by means of comparison or class inclusion. Instead, according to their quality-of-metaphor hypothesis, a metaphor’s quality or aptness decides the matter.

Really good metaphors work best as categorizations, and sometimes work only as categorization assertions. /.../ Comparisons are resorted to when a categorization does not make much sense; categorizations are used when the metaphor is apt, whether novel or conventional (Glucksberg & Haught, 2006, p. 375).

Glucksberg and Haught (2006) use the approach of reformulating sentences slightly and investigating the effect on our interpretation of their meaning and degree of figurativeness. One example is the insertion of an adjective that is applicable to either the source or the target, or both, in order to establish the differential interpretation of the metaphor ‘A is B’ and its corresponding simile ‘A is like B’. In some cases this gives very different interpretations, as in ‘my lawyer was an old shark’ vs. ‘my lawyer was like an old shark’, where the simile tends to invoke a literal, less awe-inspiring spawned fish specimen.

Wolff and Gentner (2011) set about to resolve the paradox that metaphors can be seen as, on the one hand, directional in projecting information from one domain to another, and, on the other, emergent and symmetric in seeing an otherwise not recognised similarity between two domains. They argue that both views can be encompassed in a two-stage process: first, there is a symmetric matching process in which the two domains are structurally aligned; next, there is an asymmetrical projection of inferences from one domain to the other. They provide support for their view in psychological experiments, where they asked participants to judge whether a sentence of the format “some Xs are Ys” was “comprehensible” or “incomprehensible”, across four types of expressions: literal expressions, e.g. “some birds are robins”; forward metaphors, e.g. “some jobs are jails”; reverse metaphors, where the terms in forward metaphors are swapped, e.g. “some jails are
jobs”; and, scrambled metaphors, where source and target domains were matched from different expressions, e.g. “some pianos are trees”. They found that participants judged the literal expressions significantly more comprehensible than the forward and reverse metaphors, which both were more comprehensible than the scrambled metaphors. However, the participants judged the forward metaphors more comprehensible than the reverse metaphors only when given a response time of 1200 milliseconds or longer; with short response times, 500-600 milliseconds, the ratings of comprehensibility were identical, supporting a symmetry in the initial cognitive processing.

Recently, researchers in neuroscience have started to investigate what resources are recruited in interpretation of metaphors by means of for example fMRI. For instance, Prat, Mason and Just (2012) found that easy tasks of metaphor interpretation activated regions that are highly overlapping with ones previously associated with analogical reasoning, lending support to the view of shared cognitive processing of analogy and metaphor.

Overall, and as a parallel to the debate of the role of relational structure in analogical reasoning accounted for in section 3.1.1, Dedre Gentner has represented one of the camps also in this debate regarding the process of metaphor interpretation. Once again, the two camps have been spurred to sharpen their arguments, with the outcome that refined notions and theories have been developed; I see this as a good example of fruitful research debate.

### 4.1.3. Cognitive linguistics and conceptual metaphor

Lakoff and Johnson (1980) initiated the field of cognitive linguistics with their influential book *Metaphors we live by*. As the name indicates, adherents of cognitive linguistics adopt a cognitive perspective on linguistics, based on findings in cognitive science on how our mind works. Particularly, they assume that we interpret language – not only semantics, but also syntax – in terms of underlying concepts, and that language is not separate from the rest of cognition. One claim of cognitive linguistics – in contrast to the position of e.g. Chomsky (1965) and Pinker (1995) – is that the ability to use language is not innate and due to a particular language processing module in our brains, but embodied and experiential. Lakoff (1987) further criticises the ‘Objectivist view’ that human language can be interpreted out of its human context in a formal manner, such as predicate logic or ‘mentalese’, a hypothetical language of thought (Pinker, 1995). Johnson (1987, p. 175) argues that “a non-Objectivist theory of meaning is a semantics of understanding”, adopting a more inclusive interpretation of ‘meaning’ as experiencing meaningfulness in life, as compared to traditional semantics with its focus on meaning of individual words or sentences.

One central phenomenon recognised within cognitive linguistics is our tendency to conceptualise some conceptual domains – typically of an abstract or novel character – by reference to another, more familiar and concrete domain, by use of conceptual metaphor (Lakoff & Johnson, 1980, 1999). Such conceptual metaphors provide an underlying, coherent conceptual structure, which we use in generation and interpretation of expressed language. Lakoff (1993, p. 208) even states: “The metaphor is not just a matter of language, but of thought and reason. The language is secondary. The mapping is primary…” One example of a conceptual metaphor is ‘Argument Is War’, a general metaphor structure that shapes the way we conceptualise argumentation and makes us able to form sentences, such as ‘He won that argument’ or ‘I attacked every weak point in his argument’. One specific character of conceptual metaphor, as opposed to many other views of metaphor, is that it is a predominantly unconscious phenomenon; we use conceptual metaphors without being aware of them, and their expression often does not come across as particularly figurative. In addition, Lakoff and Johnson (1980) emphasise the ubiquity of the phenomenon. It is an integral part of language, and there are many abstract realms of thought that would be very difficult to express in language without recourse to conceptual metaphor. However, they do
not go so far as to see all language as an inherently metaphorical reflection of the world as in the romantic view of metaphor, but make a difference between direct, embodied experience and more abstract phenomena (Saeed, 1997).

In the light of cognitive linguistics’ assumption of language as a conceptual phenomenon and the account of conceptual metaphor as a mapping between coherent conceptual domains, it is not far-fetched to argue that conceptual metaphor is processed by means of analogical reasoning. Indeed, in their analysis of how metaphor relates to analogy, Gentner, Bowdle, Wolff and Boronat (2001) investigated how we interpret conceptual metaphors, with a particular emphasis on coherence. As detailed in section 3.3.1, Black (1962) introduced the term archetype for a systematic repertoire of ideas that can be extended by analogy to another domain, and he offers the notion metaphor-theme (Black, 1993) for how an abstract connection between two domains may be reused and modified in communication. Similarly, Boyd (1993) discusses how the metaphor of seeing the mind as a computer has sparked off the entire field of cognitive sciences. The other way around, computer architecture has borrowed from descriptions of the mind, e.g. the use of ‘memory’. Such ideas about extended metaphoric systems stand in contrast with localist (Gentner, et al., 2001) approaches to metaphor interpretation, where a sentence or limited set of sentences are analysed in isolation, such as Ortony’s (1979) and Glucksberg’s (e.g. Glucksberg & Keysar, 1990), regardless of their position on the issue of comparison vs. categorisation. As Murphy (1996) points out, one challenge to the use of metaphor systems is that we may view one target domain in the light of several different, possibly inconsistent, source domains. For instance, in addition to ‘Argument Is War’, we can just as well see argument as a journey, a building, or even dancing. Since conceptual metaphor is a largely subconscious phenomenon, it is notoriously difficult to put it to the test empirically. For instance, McGlone (1996, p. 551) found that when subjects were asked to paraphrase ‘Our marriage was a rollercoaster ride’, they came up with salient attributes, such as “exciting” or “there were good days and bad days”, but none of them recognised the alleged underlying ‘Love Is A Journey’. However, this does not rule out that the conceptual metaphor is productive unconsciously. Gentner, et al. (2001) report on a set of studies where the indirect method of measuring reading times was employed. Subjects read short accounts of an event, e.g. about preparing for and participating in a debate, which either were based on one conceptual metaphor consistently, e.g. ‘A Debate Is A Race’, or inconsistently, in which case the entire account reflected one conceptual metaphor, e.g. ‘A Debate Is A War’, apart from the last sentence, which was kept from the race scenario, yielding a mixed metaphor reading. Recruitment of an extended metaphoric system would be consistent with longer reading time of the last sentence, while a localist theory would not predict such an effect. Gentner, et al. found that subjects read the consistent ending sentences significantly faster than the inconsistent ones, but only when they were novel metaphors based on the underlying conceptual metaphors; there was no such effect for highly conventional metaphors.

4.1.4. Cognitive linguistics and polysemy

Apart from metaphors, per se, cognitive linguistics suggests mechanisms for how words come to get new senses, related to previously existing ones, i.e. the phenomenon of polysemy, and historical semantic change over time. Johnson (1987) introduces the concept of image schemata, preconceptual patterns that give meaning to our interaction with the world, through perception, movement, manipulation of objects, etc. Examples of such generic patterns are related sets of CONTAINER, BALANCE and PATH image schemata. For instance, a CONTAINER image schema has the structure of a bounded region in space and a generic object, a ‘trajector’. We interpret the position and motion of the trajector, relative to the bounded space. This image schema is embodied in our experience of interaction with physical objects, such as
seeing that a ball is in a box or by manipulation putting the ball in the box. The experiential correlation of the spatial phenomenon and coming to use the word ‘in’ establishes a basic meaning of the word in relation to the image schema. Now, through metaphorical extension or projection, an image schema can be applied to other phenomena than those we have hitherto interpreted with it. In line with the ideas of Black (1962, 1993) and Boyd (1993), this is a truly creative and constitutive act; we come to see the world in new ways. Johnson (1987) further argues that the direction is not haphazard, but typically projects from a more concrete, physical domain to a more abstract, mental, conceptual domain. For instance, physical balance may be projected to the more abstract issues of a balanced composition of colours, a balanced temper or balance between two mathematical expressions in an equation, resulting in new, related meanings of terms. Johnson (1987) suggests that experiential correlation may offer a mechanism for such projections. For instance, the conceptual metaphor Purposes Are Physical Goals, expressed in ‘we have reached and surpassed our sales targets’, may have arisen from the fact that our early wishes often concerned getting to a particular place. Similarly the conceptual metaphor Up Is More, as in ‘sales are up’, is given the rationale that we have experienced piling up of some stuff. However, the extent to which such experiential correlation is required in coming to see one domain in the light of another domain is uncertain. As we have seen, Gentner (1983) has argued forcefully for the role of structural similarity and consistency in mapping of domains, and it may well be that some mappings are innate to us or brought by culture through language.

Lakoff (1987) takes another tack on the issue of polysemy, introducing idealised cognitive models (ICMs) in his account of categorisation, recognising effects of prototypicality (Rosch, 1973) related to above. ICMs are relatively stable mental representations of phenomena in the world. Typicality effects may arise due to the use of a word in a context that is distant from the ICM against which it usually is interpreted. Take for instance ‘bachelor’, which typically is used against the background of an ICM, representing the institution of marriage in a monogamous society. The Pope, who fits the nominal description ‘unmarried man’ does not come across as your typical bachelor, because he is not eligible to marriage and generally understood with respect to another ICM, that of the Catholic church. Some concepts may relate to a cluster of ICMs. For example ‘mother’ can be interpreted against: a ‘birth model’, a mother physically giving birth to a child; a ‘genetic model’, bringing over genetic material to one’s offspring; a ‘nurturance model’, bringing up and looking after the child, etc. The degree to which a particular candidate fits a category can be probed against the set of identified ICMs. Strong fit within many ICMs would imply a typical, even stereotypical case, here (at least in Lakoff’s American setting) the ‘housewife-mother’. If the candidate does not fit all ICMs, peripheral sub-categories may appear by qualifying their relation to the prototypical mother, such as the ‘surrogate mother’, the ‘working mother’, etc. Overall, this yields a semantic network, where new senses of a word radiate out from a central basic sense, leading to polysemy in a radial pattern. In adopting Johnson’s (1987) notion of image schemas, Lakoff (1987) further sees image schema transformation within a domain as another mechanism behind polysemy, apart from image schema projection across domains, as discussed above. The preposition ‘over’, for example can be analysed using a PATH image schema. However, when going from a more central ‘above’ interpretation of ‘over’, as in ‘John walks over the hill’, to meaning ‘across’ as in ‘John lives over the hill’ (or ‘John is over the hill’, metaphorically referring to his waning prize-fighter career), we have shifted from focusing on the entire path to an endpoint focus, thus yielding another sense of ‘over’. Finally, Lakoff and Johnson (1980) also recognise the power of metonymy, where one entity stands for another, in producing polysemy. Examples include ‘America’ standing for ‘the USA’ (the whole standing for a part), and ‘Washington’ standing for ‘the government’ (a location standing for an institution there). Following this
Analogical reasoning approach, Lakoff (1987) presents a large semantic network for the word ‘over’, which however has been criticised for being too fine grained and resulting in hundreds of different senses of the word. As a way to restrict the productiveness of distinct senses, Tyler and Evans (2001) have developed the principled polysemy approach, where the identification of different meanings of a word is not a sufficient criterion for disambiguating different senses, but that unique patterns of use across contexts and/or grammatical particularities are required as well.

4.1.5. Metaphor in discourse

As will be shown in section 4.3.2, Amin (2012) characterises cognitive theories of scientific modelling as having developed in two phases, where the first phase emerged in explicit contrast to language-based propositional accounts. Amin has put forward the idea of a parallel development of the interplay between language and cognition regarding metaphor in terms of three generations. The first generation represents a traditional view adopted in linguistics and philosophy of language, where metaphor is a purely linguistic phenomenon, an example of figurative language. Metaphor is viewed merely as a dispensable ornamentation, and incompatible with clear thought, a view which we have seen was criticised by Black (1962). A second generation of theories emphasise the cognitive perspective of metaphor, a view in which language is seen as secondary to cognition. Arguably, this perspective has dominated the theories of conceptual metaphor (Lakoff & Johnson, 1980) and the debate regarding interpretation of metaphor in terms of comparison or categorisation, accounted for in section 4.1.2. In a third generation, the specific character of language is again recognised, along with pragmatics and language in use in authentic communication in a social setting. It is not a matter of going back to the first generation, where the cognitive dimension was ignored, but language and thought are seen as partially overlapping fields in close interaction, as recognised by Vygotsky (1962). Glucksberg and Haught’s (2006) use of slight phrase alterations and Gentner’s (2003) view of a mutual relationship between analogical reasoning and relational language point in this direction. As a representative of this third generation, Cameron (2003, p. 19) comments on the view of metaphor in cognitive linguistics:

Language and thought needed to be separated in order to develop the cognitive theory and to highlight its departures from ‘traditional metaphor’, but they are not perhaps as separable as some of the programmatic statements and claims suggest.

4.2. Analogical reasoning and scientific modelling

We continue the investigation of connections between the different perspectives in the theoretical framework by looking at how analogical reasoning relates to scientific modelling. As indicated in Figure 5, the processes of analogical reasoning and scientific modelling share a structure. We approach a target domain and the world, respectively, by comparing them to another entity, a source domain or a model. Indeed, one common perspective is to use the terms interchangeably or as an integrated unit, ‘models and analogies’, when applied to the practice of science.

---

9 Private communication.
4.2.1. Recruitment of analogical reasoning in scientific modelling

Bailer-Jones (2001, p. 110) thinks that it might be fruitful to put scientific models side by side with analogies and metaphors, but offers a word of caution:

To draw inferences about the use of metaphor and analogy in scientific modeling, one needs to tread carefully when assessing whether findings from neighboring disciplines can be integrated – while integrating them is likely to be an important stepping stone in the analysis of scientific models.

Leaving metaphors out for the time being, one broad difference between scientific modelling and analogy is that scientific modelling is exclusive to the aim of saying something about the physical world, while analogies are not restricted to this aim. As pointed out by Sibley (2009, p. 259): “The only distinction between the definitions [of analogies and models] is the implication that scientific models are accepted by a community of experts whereas analogies may be less rigorously evaluated.” The other way around, it is not a clear-cut issue to what degree analogical reasoning is recruited in scientific modelling.

As we have seen, according to Hesse (1966), models are an indispensable feature of scientific theories. In her argumentation for this point, she distinguishes between two different types of models and does so by relating them to the notion of analogy. She gives the example of comparing gases to a set of billiard balls, which was recruited in devising the kinetic theory of gases. The gas molecules are seen to be analogous to billiard balls. However, only such properties of the billiard balls that we want to ascribe to the molecules constitute the positive analogies. Apart from them, there are properties that we do not want to ascribe to the molecules, the negative analogies, and most interestingly, such properties “about which we do not yet know whether they are positive or negative analogies” (p. 8). Hesse calls such undecided cases neutral analogies and points out their potential role in formulating hypotheses and making predictions, and therefore ascribe them utmost importance to the practice of science. Building on Campbell (1920), the first type of model, model₁, refers to such imagined cases which include the positive and neutral analogies, but not the negative analogy, while the second type, model₂, refers to models that include the positive and neutral analogies, but also the negative analogy. It is the models of the first kind that, in Hesse’s view, are a necessary part of scientific theory, but not the ones of second kind, which may be represented as physical scale models.

Bailer-Jones (2001) identifies two potential functions of analogies in science and particularly in relation to scientific models. First: “Analogies can exist as formal relationships between phenomena or, rather, between the theoretical treatment of phenomena” (p. 111). This corresponds to the use described by Gentner (1983) in comparing the atom to a solar system or Hesse’s (1966) comparison between light and sound. However, in addition, Hesse (1953, p. 201, italics in the original) proposed a role for analogy also in the relation between a model and the world: “most physicists do not regard models as literal descriptions of nature, but as standing in a relation of analogy to nature”.

In his set-theoretical approach to scientific theories, the semantic view of theories, to which we will come back in section 4.3, Suppes (1962) shows that the connection between empirical measurements and scientific theories is more complex than is often recognised. He claims that “exact analysis of the relation between empirical theories and relevant data calls for a hierarchy of models of different logical type” (p. 253) and that the relation between these models has to be analysed at each level in the hierarchy. In addition, in establishing such relations: “The central idea... is to restrict models of the data to those aspects of the experiment which have a parametric analogue in the theory” (p. 258). In other words, in Suppes’ view, the connections between the sequential models of a phenomenon are established by means of analogy. Suárez (2003) argues that scientific representation – which
he sees as central to modelling – cannot rely on either similarity or isomorphism, since representation is an inherently directional and non-symmetrical phenomenon. However, as we saw above, Tversky (1977) and Ortony (1979) found that similarity and analogy are not necessarily symmetric or unidirectional from a psychological perspective, so I do not see this argument as the final blow for a structural approach.

All in all, Bailer-Jones (2001, p. 113) argues that analogy may be recruited for both functions, connecting models to other models or to the world, but that it is not a requirement, justifying her view with authentic science cases from astrophysics, and concludes that:

...few existing models in science have not developed beyond the boundaries of the analogy from which they originated, and others may simply not have their origin in an analogy at all. /.../ Moreover, an analogy is a relationship between things or processes while a model is a type of description about some thing or process. If anything, a model could be an analogue, but this is not the issue because the way to evaluate a model is not to judge whether it is analogous to something, but whether it, as it stands (analogous or not), provides access to a phenomenon in that it interprets the available empirical data about the phenomenon.

Rivadulla (2006) gives the example of the notion of black holes, which emerged as a solution to equations within the mathematical formulation of the theory of general relativity, without reference to any other analogous model or previously known phenomenon in the world.

In my view, it is not the asymmetry in the relation between the world and representations of it, as such, that excludes the possibility that modelling or representation is based purely on analogies. One reason for why there is more to it than analogy is the radical ontological difference between the world and representations of it. Indeed, as we shall see, Suppes (1962) does not even focus on the last crucial step of attaching models to the world, the act of representation, but rather on connecting different models to each other. More fundamentally, representation does not have to rely on structural similarity; from Peirce’s (1985) semiotic perspective, not all signs are icons, since a symbol or an index may come to stand for or represent something without being similar to it in any respect.

![Figure 6. Relations between representations of phenomena.](image)

Having come thus far, we may pull the strings together and offer an overview in Figure 6 of how analogical reasoning relates to scientific modelling. At the model level, horizontal connections between theoretical models of different phenomena may be of an analogical kind. This is not controversial. More different voices are heard if we go to the vertical connection between models and the modelled phenomena. As we saw, Suppes (1962) argues for a hierarchy of models, ranging from those of a more theoretical kind and data models closer to empirical data, and that these models are connected by analogy, but he does not take a direct interest in the nature of the connection between data models and the world. However, I would
argue, with Vessey (1991), that different models of a phenomenon, or at least different representations, do not necessarily stand in an analogical relationship to each other. Instead, as put forward by Fredlund, et al. (2012) in their adoption of disciplinary affordances, different representations of a phenomenon provide access to different areas of disciplinary knowledge, and we may not be able to connect the representations structurally or translate between them. We will come back to this point in the reanalysis of the animation of an Otto engine in section 6.2 below. As for the top horizontal connection between different aspects of the world at the phenomenal level, one possibility is that phenomena share some surface similarity, which, as Vosniadou (1989) argues, may be a prompt for searching for deeper structural commonalities. Another specific perspective on this connection between phenomena is found in biology and medicine, where one organism, e.g. a mouse, may serve as a model for another, e.g. us humans. In addition, of importance for science are generalisation and unification of theories, so that one theory may come to account for what previously was regarded as different phenomena in different domains.

4.3. Scientific modelling and semantics

The third relational investigation deals with connections between scientific modelling and semantics. A common ground between these practices is maybe easier to establish by placing two fields of philosophy side by side: philosophy of science takes an interest in how we represent physical phenomena in the world by theories and models, while philosophy of language studies how we represent, refer and give meaning to the world by means of language and other symbolic systems (see the parallel structure in Figure 5).

This section focuses on scientific modelling and discusses different ways in which language and representations have influenced the way the view of modelling has evolved. Once again a word of caution: Hacking (1983) is sceptical regarding the usefulness of philosophy of language in developing our understanding of natural science. When putting forward the role of intervention in the world in doing science and the importance of studying authentic science practices rather than engaging in armchair speculations, he argues that a philosophy based on reflections on language – such as Putnam’s (1973) work on meaning and reference of natural kinds – cannot “teach anything positive about natural science” (Hacking, 1983, p. 92). Nevertheless, I hope to be able to show some interesting areas where the two fields may be relevant to each other, and how the role of language is yet regaining some of its lost ground in recent studies of scientific modelling.

4.3.1. Formal approaches to scientific modelling

Before the 1960s, central to the so called received view on scientific theories within the dominant logical positivist philosophy of science, represented by e.g. Duhem (1991/1914), was the “analysis of theories as empirically interpreted deductive axiomatic systems” (Craver, 2001, p. 55). In the received view, theories are seen as linguistic structures, with a logical vocabulary formalised in first-order predicate calculus and an ‘extralogical’ vocabulary providing its descriptive terms, where deductive and inductive relations between the
descriptive terms are detailed. The connection between the theory and observed phenomena is established by correspondence rules between predicates in an observational vocabulary and predicates in a theoretical vocabulary and universal laws of nature are seen as the main means to bring explanatory power to the theories. In addition, theory change is typically characterised as a succession of theories by means of reduction of one theory into another or replacement (Craver, 2001). However, the received view has been heavily criticized. With a focus on its ability to characterize ‘theories in the wild’, i.e. the construction and use of theories in authentic science practice, Craver points out a number of its weaknesses:

- It does not account well for the multiple, partial and fragmented character of scientific theories, or the typically gradual, piecemeal character of theory change.
- A focus on scientific laws cannot account for causal patterns of explanation and typically are just plain false even within the paradigmatic case of physics (Cartwright, 1983). In addition, there are many fields of natural science, most notably biology, where universal laws are altogether insignificant.
- The restriction to first-order predicate calculus does not square well with the often set-theoretical mathematical formalism used to express many science theories.

An alternative approach, known as the semantic view of scientific theories, touched upon above, was initiated by Patrick Suppes in the early 1960s. Although admitting that the term model is used in a plethora of manners in science, Suppes (1960) suggests that models can typically be reformulated in set-theoretical terms and that the notion of models could replace the predicate calculus approach to scientific theories of the received view:

I claim that the concept of a model in the sense of Tarski [a model as a non-linguistic entity in which a theory is satisfied] may be used without distortion and as a fundamental concept in all of the disciplines from which the above quotations are drawn. In this sense I would assert that the meaning of the concept of a model is the same in mathematics and the empirical sciences. The difference to be found in these disciplines is to be found in their use of the concepts. /…/ Roughly speaking, a possible realization of a theory is a set-theoretical entity of the appropriate logical type. /…/ The important distinction that we shall need is that a theory is a linguistic entity consisting of a set of sentences and models are non-linguistic entities in which the theory is satisfied (pp. 289-290)

As we see, the main focus of the semantic conception approach at its origin was to come to terms with the awkward formalism of predicate calculus and replace it with another formalism; models as interpreted within set theory. In my view, such a perspective disregards the radically different ways in which we interpret models that are structurally or logically similar, but represented by means of different symbol systems (e.g. Johnson-Laird, 1983; Zhang, 1997). Indeed, it is somewhat opaque in what respect the approach is labelled ‘semantic’, apart from the connection between theories as linguistic entities and models as corresponding non-linguistic entities. Still, Giere (1999b) frames his work within the semantic view, or in his words the model-theoretic or later preferred model-based view of theories, and adheres to the basic premise that the structure of theories consists of a family of models, but seeing the relation between models as one of similarity rather than isomorphism. He has come to develop the theory in a more cognitive and ‘naturalistic’ direction, making more room for the individual scientist’s intention with a model. In addition, Giere acknowledges the use of several forms of representation of the models, e.g. scale models, graphs, by use of mathematical formalism of different types, etc., and is thereby sceptical towards the usefulness of formal model theory in understanding science theories.

In contrast to the semantic view, Morrison and Morgan (1999) argue that models should not be regarded as the constituents of scientific theories, but rather as autonomous agents and
mediators between theories and represented phenomena. In addition, regarding the nature of the connection between models and represented physical systems, they recognise that there may be several models of a phenomenon, possibly inconsistent with each other, but useful in different situations: “We do not assess each model based on its ability to accurately mirror the system, rather the legitimacy of each different representation is a function of the model’s performance in specific contexts” (Morrison & Morgan, 1999, p. 28). In a similar manner, Cartwright (1999, pp. 184-185) disapproves of the objectivist traditions accounted for above:

I subscribe neither to the ‘received’ syntactic view of theories nor to this version of the semantic account. For both are cases of the ‘vending machine’ view. The theory is a vending machine: you feed it input in certain prescribed forms for the desired input; it gurgitates for a while; then it drops out the sought-for-representation… /…/ For the whole point of view of the tradition that generates these two views is the elimination of creativity – or whim – in the use of theory to treat the world.

Cartwright sees a particularly difficult problem for objectivist traditions in accounting for the opportunistic way in which scientists recruit ideas from diverse theories when explaining phenomena: “knowledge must be collected from where we find it, well outside the boundaries of what any single theory says, no matter how fundamental and universal we take that theory to be” (p. 181). She gives the example of how the theories of electrodynamics, quantum mechanics and thermodynamics contributed to the Ginzburg-Landau model of superconductivity in a cooperative fashion, ringing of the title of the book, The dappled world, reminding of Nersessian’s (2008a) account of Maxwell drawing analogies from several scientific fields to electromagnetism, related to above. Similarly, in her ‘theories in the wild’ perspective of advances in scientific research, Bailer-Jones (2003) argues against the possibility, adopted generally within the semantic view – although Giere (1999b) is an exception – that models as used in science practices can be translated ‘without distortion’ to a set-theoretic formulation: “…any concept of models, formal or not, that denies that models employ external representational tools and denies that these entail propositions about empirical phenomena disagrees with the assumptions of my approach to an extent that they are incompatible” (p. 63). Black (1962) is emphatic in his rejection of Suppes’ proposal, going against his classification of logically different types of models and reducing them to set theory. He dismisses it laconically as a “different conception” (p. 262).

Frigg (2006) has launched an eloquent attack on the semantic view of theories, framing it as failing in giving satisfactory answers to three crucial problems: The ‘ontological puzzle’ of what kind of objects models are; the ‘enigma of representation’, i.e. in virtue of what a model is a representation of something else; and the ‘problem of style’, accounting for the plethora of different kinds of models. Homing in on the enigma of representation, Frigg agrees with the view of Suárez (2003) accounted for above that representation cannot be based only on isomorphism between the model and what it represents. He rejects the move to take recourse to the intention of a user of a model, for instance as suggested by Giere (1999b) above. Although admitting that the role of the user is essential in representation, it “allows that everything can represent just about everything else by a mere act of fiat…” (pp. 54-55). Next, isomorphism is a relation between structures, and before this relation can be established we have to abstract a structure out of the world of phenomena. Frigg claims, with Cartwright (1999), that such abstraction cannot be done without fitting abstract concepts to corresponding concrete concepts. Crucially, the constituent entities and relations in the structure are the result of such abstraction. Siding with Suárez (2003), Frigg claims that different approaches to relaxation of isomorphism fare no better; neither does Giere’s use of similarity relations: “The claim that M is similar to T remains empty until relevant respects and degrees of similarity have been specified” (p. 61). In conclusion, “structural claims rest on more concrete descriptions of the target system. For this reason, descriptions are an integral part of any
workable conception of scientific representation and we cannot omit them from our analysis” (p. 62). In the respect, Frigg’s approach to the issue of representation is reminiscent of the ‘causal descriptivist’ view on the role of descriptions in reference (e.g. Kroon, 1985). Overall, Frigg (2006, p. 62) argues that the quest for a philosophy of scientific theories without relying on language-based propositions was taken too far in the semantic view:

In the wake of the anti-linguistic turn that replaced the syntactic view with the semantic view of theories questions concerning the use of language in science have been discredited as misguided and obsolete. This was too hasty a move. There is no doubt that the positivist analysis of theories is beset with serious problems and that certain non-linguistic elements such as structures do play an important role in scientific representation; but from this it does not follow that language per se is irrelevant to an analysis of scientific theories or models. Scientific representation involves an intricate mixture of linguistic and non-linguistic elements and what we have to come to understand is what this mixture is like and how the different parts integrate.

Here, Frigg’s line of reasoning has a parallel with the role of language in the development of theories of metaphor, as argued by Cameron (2003) and Amin’s (2012) point that language cannot be omitted from a cognitive view on scientific modelling in section 4.3.2.

Kralemann and Lattmann (2012) point out a parallelism between scientific modelling and Peirce’s (1985) theory of signs, and argue that a semiotic analysis may shed light both on the ontological puzzle and the enigma of representation (Frigg, 2006). They suggest that scientific models essentially are icons, i.e. representing by virtue of similarity, but in a way where “similarity is understood in a more abstract manner” (Kralemann & Lattmann, 2012, p. 4) than for images, e.g. photographs and scale models. Instead scientific models are classified as diagrams, icons that represent by similarity in the relations between the constituent parts, with mathematical equations as an example. Icons are distinguished by the fact that inferences to the represented objects can be made from their inner structure, in terms of a “primary interdependence between the structure of the sign and the structure of the object” (p. 12). In particular, they argue that this semiotic approach may be applied in conjunction with the semantic view of theories. Not surprisingly, I find the parallelism between scientific modelling and semiotics fruitful, since it is the subject of study in this section, but I think that the abstract interpretation of ‘similarly’ is maybe stretched a bit too far and we will come back to the issue whether all models are icons in the reanalysis of article II in section 6.2.

Knuuttila (2011) argues that the program of characterising scientific modelling in terms of representation so far has had a too narrow focus, with the relationship between a model and a real target system as the basic unit of analysis. In contrast, Knuuttila offers the view of scientific models as epistemic tools, concrete artefacts that can be manipulated. Depending on the representational means (e.g. diagrams, natural language or scale models), they afford but also limit our scientific reasoning and thereby also the activity of modelling. Knuuttila and Boon (2011) further claim that the starting point of modelling often is a scientific question of a general character, rather than a clearly delimited target system. Consequently, the mapping of a model to a particular phenomenon may occur at a late stage of modelling, while prior stages may have been guided by other criteria than the representational power, such as mathematical solvability. In this perspective, models have a certain degree of autonomy, similar to Morrison and Morgan’s (1999) view of models as autonomous from theory and phenomena, and mediating between them. Interestingly, Knuuttila and Boon use the case of Sadi Carnot’s development of a model of an ideal heat engine in arguing for their point. Carnot was intrigued by the problem whether there was a theoretical limit to the maximum efficiency of a heat engine and, if this was the case, what that limit would be. They argue that Carnot’s immediate interest was “not primarily that of representing some real target-system more or less accurately, but rather producing a hypothetical device that meets some specific epistemic aim” (p. 319). Based on his experience of physical steam engines, Carnot
considered a body of gas constrained by a cylinder and a piston. The body of gas can be placed in contact with either one of two heat reservoirs at temperatures $T_1$ and $T_2$ ($T_1 > T_2$) or isolated from them. For maximum efficiency, Carnot argued, change in temperature of the body of gas should only occur due to change of volume, implying isothermal expansion and compression when in contact with either heat reservoir, and adiabatic compression and expansion when isolated from them. Consequently, he considered a three-stage process where heat is transferred from the warmer reservoir via the body of gas to the colder reservoir. Now, crucially from the point of view of Knuuttila and Boon (2011), Carnot imagined what would be the consequence if the process were run in reverse, i.e. if heat were to be taken from the cold reservoir and given away to the warm reservoir. This was a completely hypothetical process at Carnot’s time, corresponding to a modern heat pump or refrigerator, which was not yet invented. In other words, Carnot did not intend directly to represent nature, but to construct a thinking device. Considering in combination the three-stage process run forward and in reverse, Carnot argued that all generation of ‘motive power’, i.e. the modern work, has to be accompanied with a heat flow from higher to lower temperature. Only after these theoretical workings of the model were in place, Carnot started to match it with physical engines and findings from empirical experiments.

Emch and Liu (2002) examine the role of scientific models in theory building, in their account of the philosophical foundations of thermodynamics and statistical mechanics, in reference to the semantic view of theories. In their view, the ambition with the semantic view to bring together the views of physicists and logicians on models has not been realised:

A gap still exists in the two conceptions of models from the two communities – the philosophers’ and the scientists’; and while the former has been consolidating and entrenching its results in explaining the formal theory of the semantic (or structural or architectonic) view (or approach), the latter has never relented in its efforts to evoke models for explaining the phenomena at hand, which often defy classification (pp. 1-2).

In addition, in line with Frigg (2006), they argue: “In combating the syntactic view, semanticists went overboard in denying the significant role language (especially a mathematical language) plays in scientific theories” (Emch & Liu, 2002, p. 21).

French and Ladyman (1999, p. 103) are more enthusiastic about the semantic view:

It is, perhaps, an exaggeration, albeit an excusable one, to claim, as Suppe does, that ‘The Semantic Conception of Theories today probably is the philosophical analysis of the nature of theories most widely held among philosophers of science’ (Suppe, 1989, p. 3). Nevertheless, the semantic approach, as adopted by Suppes, van Fraassen, Giere and Suppe himself, does have the distinction of being one of the very few – perhaps the only – global analyses of science in these philosophically fractured, post-Kuhnian times.

Indeed, French and Ladyman (1999) claim that models used in science practice, e.g. visual representations or scale models, typically stand in a ‘partially isomorphic’ relationship to corresponding set-theoretical models. In addition, the semantic conception does not put forward an isomorphic relationship between models and reality. It establishes sequential mappings between models of a more theoretical character and models of data (Suppes, 1962), but stops short of das Ding an sich. The relationship between models and phenomena is acknowledged as a very complicated one, but out of scope for the investigation.

In a science education context, Adúriz-Bravo (2012) proposes that the semantic view of theories may be used as a theoretical foundation of model-based approaches to science teaching, particularly the cognitive approach adopted by Giere (1988). Adúriz-Bravo argues that such a philosophical stance holds the hope that: “A shift of focus from syntax to semantics would imply less attention to formal aspects and more attention to meaningfulness in science education.” Koponen (2007) finds the endeavour of putting model-based reasoning in science education on a solid philosophical ground laudable, but in contrast to Adúriz-Bravo, he sees
some problems with applying Giere’s (1988) interpretation of the semantic conception of theories in science teaching, and proposes some adjustments. First, Giere adopts a scientific realist position within the semantic framework, which Koponen finds more radical than necessary in a science education context. Instead, Koponen supports van Fraassen’s (1980) requirement of theories to be empirically adequate. In addition, Koponen argues that the connection between the models and the phenomena is not covered sufficiently and in particular, he finds Giere’s recourse to ‘similarity’ as the principle with which theories are connected to phenomena too vague. Koponen brings forward Suppes’ (1962) account of matching in a hierarchy of models ranging from theoretical models to data models, but also emphasises, with Cartwright (1999), the bi-directionality of the relation between models and phenomena in a process of matchmaking. The models do not just represent phenomena in a one way fashion; isolated laboratory phenomena are made to fit the models in experiment design. Here, Koponen also connects to Hacking (1983), who argues that science is not just a matter of representing the world, but also intervening in it, by way of constructing artefacts and measurement equipment, thereby creating phenomena that would otherwise not exist.

The structural focus of the semantic view of theories may be compared to Gentner’s (1983) similarly structural approach to analogies, which has been widely adopted in science education. In my view, the criticism of Gentner’s structure-mapping theory in that it does not sufficiently account for the embodied, perceptual character of cognitive processing or situational factors can be used against the semantic view of theories as well. However, where analogical reasoning in general typically involves comparisons of knowledge domains at the same, cognitive level, in scientific modelling two radically different entities – an aspect of the world and a mental representation of this aspect – are brought together. In addition to the embodied character of cognition as built on our neural physiology, required also for analogical reasoning, scientific modelling involves this second additional connection to the world in terms of representation of physical phenomena, which puts additional constraints to a purely structural approach to scientific modelling.

In conclusion, while the semantic conception of theories may provide a useful analysis of how scientific theories relate to models in its formal set-theoretical approach and how such models, closer to or further away from empirical data, relate to each other, it does not aspire to account for how models relate to phenomena and particularly, how scientists use models in practice in the sense put forward by Hesse (1966) or Black (1962). I therefore question its usefulness in model-based approaches to science education research, and believe that studies of scientists’ authentic practices in history and present times à la Nersessian (1999) may be more productive starting points.

4.3.2. Non-formal approaches to scientific modelling

If the received and semantic views deal with establishing scientific theory on a formal footing by means of models, the cognitive basis of scientific modelling has provided another view. In a review of research in this field, Amin (2012) identifies two phases, roughly in parallel with the three generations of interplay between metaphor and cognition related to above.

In a first phase, researchers “found in cognitive science the resources to develop naturalized accounts of scientific theories and scientific reasoning in terms of nonpropositional representations and processes” (Amin, 2012, p. 143). As an example of work in this phase, J. Clement (1994) analyses the problem-solving data related to above on the dependence of the diameter on stretching of springs from the point of view of recruitment of intuitive, non-formal and often imagistic resources. In contrast to Chi, Feltovich and Glaser (1981), who argue that experts use abstract principles in physics problem solving while novices use concrete representations, Clement shows that also experts – including a Nobel physics laureate – use such concrete representations and non-formal intuitions in their
problem-solving exercises. Based on observational phenomena, such as the participants’ own report of use of imagery and intuition, and gestures occurring with making predictions of the physical phenomena, Clement hypothesises about the underlying cognitive structure. In his view, the participants use “dynamic imagery in conjunction with perceptual motor schemas /.../ in ‘running an imagistic simulation’ of an event on the basis of a physical intuition” (p. 204). In other words, the participants envision in a ‘what-if’-fashion what would happen if they were to manipulate the physical system in a particular way, based on previous experiences. Similar to Nersessian’s (2008b) view of mental modelling and relating to diSessa’s (1983) notion of a phenomenological primitive (p-prim), the intuitions are schematic and of ‘modest generality’ across phenomena. Nersessian (1999) sides with the critique of the propositional approach to scientific theories within logical positivism, and argues that non-formal ways of reasoning, such as visual representations, analogical reasoning and thought experiment, play an important role, particularly in the early discovery phase of scientific inquiry. Siding with Clement (1994), she argues that we recruit mental models in the form of simulative reasoning during scientific problem-solving. Here the cognitive perspective of mental modelling comes to meet the practice of scientific modelling, captured in the notion of model-based reasoning.

In a second phase of research on the cognitive basis of scientific modelling, Amin (2012) identifies influences from research on distributed cognition, which emphasises the role of external representations in cognitive processes. As we have seen, proponents of distributed cognition include Hutchins (1995) with examples from navigation as ‘cognition in the wild’, Zhang and Norman (1994) with their studies from psychology laboratory settings, and Clark (2008). These perspectives are now applied to the practice of scientific modelling. Ochs, Gonzales and Jacoby (1996) carried out an ethnographic study where they followed the activities in a physics laboratory focusing on solid states physics. They point out the communicative practice of the researchers identifying with the studied physical phenomena. This is shown in the title quote: “When I come down I’m in the domain state” (p. 329), referring to how one of the researchers follows a whiteboard graph of how an antiferromagnet changes state as part of a certain process. Here, the researcher, the phenomenon, the graphical representation, and potentially also the listener come to conflate in an intriguing way as they try to construct a shared understanding of the phenomenon. According to Ochs, et al.:

These utterances thus seem to have a semantically schizoid, illogical character which blurs the boundaries between the animate subject (physicist) and the inanimate object (physical entity/system) (p. 340).

This communicative practice contrasts both with the distanced way of reporting research findings in writing, typically in the passive voice, with a focus on the described phenomenon, and with the subjective, oral communication where the researchers interact with each other.

As seen above, Nersessian (1999) brings forward visual modelling as involved in model-based reasoning. However, she goes one step further is showing a connection between internal visualisation and external visual representations by giving the example of Maxwell, who used external visual representations in communicating his ideas on electromagnetism, providing instructions for how the reader could ‘animate’ them in order to bring across the dynamics of the phenomena. In line with his view that scientific models can take many different forms, Giere (1999a) presents the case of how gradually more refined diagrams of evidence of geomagnetic reversals were used to convince the community of geophysicists in the 1960s of the theory of plate tectonics. These visual representations show the symmetry along deep ocean ridges with regards to irregular geomagnetic structures, pointing to historical reversals of the magnetic orientation of the earth and, most importantly, an outward movement of the ocean floor. In opposition to the received view of scientific theories, Giere modestly claims that it is at least not implausible that matching of such diagrams from sets of measurements...
plays a role in scientific theory. Building on Hutchins (1995), Giere (2002) has further investigated the role of distributed cognition in scientific research, noticing the importance of interaction with technological artefacts in ‘big science’ such as cyclotron experiments and in accounting for the use of visual representations.

Having accounted for these two phases of research on the cognitive basis of scientific modelling, Amin (2012) argues that language has been largely neglected as a tool for the researcher in analysing the practice of scientific modelling and for the investigated scientists in carrying out the modelling. In this light, Amin sees language as ‘the missing piece’ in cognitive theories on scientific modelling going forward. As pointed out earlier, this reluctance to bring in language possibly emanates from the fact that the cognitive approaches appeared in opposition to logical propositional approaches, based on language only. In fact, “Clark’s work can be seen as a comprehensive attempt to assess and synthesize research in cognitive science that has rejected the a priori assumption that human cognition is best modeled in terms of internal propositional representations and formal operations” (Amin, 2012, p. 159). Nevertheless, Clark (2008) acknowledges that language plays an important role in assigning labels to perceptual patterns which become objects of attention, available for higher-order reasoning, but also in encoding ways of attending to and categorising aspects of the world in a culture and thereby allowing for greater levels of expertise and more effective allocation of cognitive resources than what would be possible without language:

Words and linguistic strings are among the most powerful and basic tools we use to discipline and stabilize dynamic processes of reason and recall. The shift is thus from seeing words and sentences as items apt only for translation into an inner code to seeing them as inputs (whether externally or internally generated) that drive, sculpt, and discipline the internal representational engine (Clark, 2008, pp. 53-54).

Amin (2012) further proposes that advances in cognitive linguistics may be exploited in studying the role of language in directing the attention and coordinating our cognitive resources as a tool for scientific modelling. He notes a parallel between how dynamic processes or animations may be ‘frozen’ in time in static images and how descriptions of processes in terms of verbs may be ‘frozen’ by use of nouns. This process of nominalisation, a type of grammatical metaphor is common in scientific language (Halliday, 1993). Compare for instance the everyday “the crack in the hot glass grew quickly” to the corresponding example from science discourse: “glass crack growth rate depends on temperature”, where the expression ‘glass crack growth rate’ has been grouped together and frozen into a nominal group, and the adjective ‘hot’ is replaced by the quantitative noun ‘temperature’. In addition, conceptual metaphor (Lakoff & Johnson, 1980, 1999) may be recruited for analogical mapping between concrete embodied experiences and more abstract construals, in this case in scientific modelling (e.g. Amin, 2009; Brookes & Etkina, 2007).

4.3.3. The relation between models and representations

I want to finish this section on the connections between scientific modelling on the one hand and semantics and semiotics on the other by bringing up the general question of how models relate to representations. Is it the case that one overarching model manifests itself in several different representations? For instance, Justi and Gilbert (2002) present a typology of different modes of representation of scientific models. Or are different models to be identified with each instantiation of a representation? This would apply to scale models, which are models of the depicted phenomena, but not representations of ideal models. In contrast, in the view of Knuuttila (2011), according to whom representing some phenomenon in the world is only one of the roles scientific models play, each instantiation of a model in some medium is a representation primarily of the model itself. I cannot give an exhaustive account of the model-representation relationship, but merely want to point to the different possible interpretations.
4.4. Overview of the theoretical framework

We have now come to a point where we can summarise some characteristics of the perspectives in the theoretical framework and how they are related:

- As shown in Figure 5, the three perspectives in the theoretical framework – analogical reasoning, scientific modelling and semantics – share a common structure of mapping two different entities.
- In all three perspectives, the question of by virtue of what such mappings can be made has led to investigation and debate within fields such as philosophy of language, philosophy of science, and cognitive psychology.
- Analogical reasoning may be seen as a general cognitive capability that can be recruited for specific tasks, such as in transfer of understanding from one domain to another, and carrying out mapping in semantics and in scientific modelling.
- A broad dividing line may be seen between approaches that emphasise primarily cognition and/or the relational structure of conceptual entities, and approaches that incorporate issues such as pragmatics and context, embodiment, and the study of authentic historical or contemporary scientific development.
- In the area of analogical reasoning, Gentner’s (1983) structural, cognitive approach has been very fruitful, both in providing valuable explanations for central psychological phenomena, but also serving the function of a touch-stone in spurring research with other approaches to analogy and metaphors.
- In semantics, a realisation that both a causal connection to the world and descriptions of what aspect of the phenomenon is intended are required to establish the meaning of words makes causal descriptivism (Kroon, 1985) an attractive route to establishing the mapping between the world and representations of it.
- There does not seem to be a consensus view on how models of a phenomenon relate to each other and to the represented phenomenon within the philosophy of science. The semantic view of theories (Suppes, 1960) is a contender for providing a comprehensive account, but also has many adversaries. A cognitive focus on authentic historical cases (e.g. Nersessian, 1999) may be a more productive approach to scientific modelling, at least within science education.
- As for semantics, Gentner (2003) puts forward that analogical reasoning in conjunction with relational language is part of what makes humans “so smart” in comparison to other animal species. Gentner (e.g. Gentner, et al., 2001) has also investigated the supposed analogical basis of metaphor, even though this approach has been challenged (e.g. Glucksberg & Keysar, 1990). Lakoff and Johnson’s (1980) notion of conceptual metaphor emphasises the systematic role of analogical reasoning in approaching abstract concepts by comparison to embodied experiences, and its role in metaphorical extension of word meaning.
- As depicted in Figure 6, in scientific modelling, analogy may be recruited in connecting models of different phenomena to each other and in connecting models to the phenomena they represent (M. B. Hesse, 1966), although not in all cases (Bailer-Jones, 2001). Similarly, analogy may be involved in mapping different models of the same phenomenon (Suppes, 1962), but not always (Craver, 2001).
- Finally, careful attention to the particularities of authentic language in dialogue seems to be a white spot in educational research on the use of metaphors (Cameron, 2003) and scientific modelling (Amin, 2012), and is an attractive area for future research.
5. Methodological framework

After having detailed the theoretical framework, I will now characterise the individual articles in the dissertation from a methodological perspective and lay the ground for a reanalysis of selected data from the underlying studies. As we saw in Figure 1, the individual articles cover different stages in the process of didactical transposition (Chevallard, 1989). In addition, they also relate to different aspects of the theoretical framework, outlined in Figure 2, and different methodological approaches were used in the conduction of the studies. In Table 1, the articles are characterised along a set of dimensions which are expanded upon in the following.

Table 1. Characterisation of the four articles included in the dissertation.

<table>
<thead>
<tr>
<th></th>
<th>Article I – senses of entropy</th>
<th>Article II – Otto engine animations</th>
<th>Article III – teacher students’ analogies</th>
<th>Article IV – first-graders’ analogies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role in didactical transposition</td>
<td>Science content</td>
<td>Student understanding</td>
<td>Teaching approaches</td>
<td>Teaching approaches</td>
</tr>
<tr>
<td>Science content</td>
<td>Entropy</td>
<td>Thermal processes, ideal Otto cycle</td>
<td>Thermal processes, gas expansions</td>
<td>Heat transfer and mixing</td>
</tr>
<tr>
<td>Theoretical framework focus</td>
<td>Semantics</td>
<td>Scientific modelling</td>
<td>Analogical reasoning</td>
<td>Analogical reasoning</td>
</tr>
<tr>
<td>Participants</td>
<td>N/A</td>
<td>Wide range (N=19)</td>
<td>Physics preservice teachers (N=8)</td>
<td>First-graders (N=25)</td>
</tr>
<tr>
<td>Research design</td>
<td>Theoretical investigation</td>
<td>Flexible, incremental</td>
<td>Inductive/deductive</td>
<td>Inductive/deductive</td>
</tr>
<tr>
<td>Type of data analysed, data collection approach</td>
<td>Written accounts of entropy and text corpora</td>
<td>Think-aloud protocols and interviews, video-recording</td>
<td>Collective generation of analogies, video-recording</td>
<td>Generation of analogies and drawings, video-recording</td>
</tr>
<tr>
<td>Data analysis</td>
<td>Tools from linguistics and philosophy of language</td>
<td>Bottom-up categorisation of the participants’ statements</td>
<td>Assessment of ownership of provided vs. self-generated analogies</td>
<td>Assessment of ability to generate analogies of abstract phenomena</td>
</tr>
<tr>
<td>Distributed cognition and scaffolding</td>
<td>N/A</td>
<td>Interpretation of computer animations</td>
<td>Collaborative generation of analogies</td>
<td>Interaction with phenomena and drawings</td>
</tr>
</tbody>
</table>

5.1. Circumstances of the articles and my contribution

The research presented in the four articles has been carried out in collaboration with other researchers: my supervisor Helge Strömdahl and fellow graduate students Johanna Andersson and Fredrik Jeppsson. I am convinced that such collaboration has contributed to improved quality of the research and that I would not have been able to write these articles alone. In addition, it is more fun and rewarding to work with other people than to work all by yourself! Nevertheless, as a doctorate student, I have to be able to point out my own contributions to the different parts of the dissertation.

Article I emanated from a course in linguistics, given by Richard Hirsch, with a focus on cognitive linguistics and metaphors, which Fredrik Jeppsson and I took in 2009, and many of the directions in this field were provided by Richard, as the suggestion to include a text corpus analysis. The idea to focus on entropy was partly inspired by Risto Leinonen as we learnt that he has taken an interest in the concept as part of his doctorate studies. Helge Strömdahl
contributed with the two-dimensional semiotic/semantic analysing schema (2-D SAS), primarily as a structure for the educational implications, and in later stages of finalising and revising the manuscript. The bulk of the analysis of the different senses of entropy, their relationships and the educational implications were carried out by Fredrik Jeppsson and me in tandem with equivalent contributions, while I took a greater responsibility for the literature review with regards to interpretations of entropy in physical sciences and information theory.

The main story in article II, with its focus on views of scientific models, emerged after quite some time of incremental data collection and analysis. I came up with the idea of interviewing natural science students and teachers regarding combustion engines as a ‘science, technology and society’ (STS) approach to thermodynamics at an early stage of my doctorate studies and bounced the idea of using computer animations with Helge Strömdahl in supervision. Alternative tacks on the subject at different stages included conceptual understanding of thermodynamics and a cognitive focus on interpretation of dynamic representations, and Helge was influential in formulating the main message of the article. I carried out the data collection at three different occasions – one of which together with Helge – and performed the transcription, the bulk of the analysis and writing of the manuscript. Apart from the supervisory suggestions on the manuscript from Helge along the way, important input was provided by Konrad Schönborn as part of course assignments, particularly with regards to the literature on visual representations.

The interest in self-generated analogies was sparked off by reading Blanchette and Dunbar’s (2000) article on the matter. For article III, Fredrik Jeppsson and I again carried out much of the work in close cooperation, including designing the study, coding the transcripts in MAXQDA and analysing the data. In the data collection and transcription, we took responsibility for one group of participants each. I performed most of the literature review.

Article IV started with the idea of adapting the approach of using self-generated analogies to much younger participants than previously studied, as part of a larger initiative, led by Shu-Nu Chang Rundgren, to investigate science teaching for elementary school. Shu-Nu, Fredrik Jeppsson and I laid out the foundation of the design, Fredrik and I carried out the data collection and transcription, where I took the responsibility for the general introduction to analogies and the mixing marbles exercises, while Fredrik ran the heat transfer exercises. Fredrik and I did most of the analysis, supported by Shu-Nu and Johanna Andersson, who joined the project half-way through, providing hands-on experience in elementary science teaching and with a particular focus on the children’s drawings. Also in this study, I took a primary responsibility for the literature review with regards to children’s conceptions of heat and mixing and their analogical reasoning, where Goswami’s (1992) book should be pointed out as particularly influential.

5.2. Paradigm and research design
As we have seen in the theoretical framework, while the cognitive tradition can be criticized for studying the mind of the individual in isolation, the socio-cultural tradition has been criticised for ignoring cognition or deferring it to an entirely social phenomenon, enacted in the communicative ‘cloud’ of the participants in a community. For an overall discussion of the merits and shortcomings of cognitive vs. social perspectives on educational science, see for example the debate in Educational Researcher (Anderson, Reder, & Simon, 1996, 1997; Greeno, 1997). Against this background, I generally sympathise with the perspective of distributed cognition as related to above (e.g. Clark, 2008; Hutchins, 1995; Zhang & Norman, 1994), in that the individual’s cognition cannot be ignored, but has to be analysed in interplay with other people and external representations. Indeed, Vosniadou (2007) suggests that distributed cognition may be a productive approach to closing the ‘cognitive-situative divide’ in conceptual change research. As a parallel, the notions of scaffolding, introduced by Wood,
et al. (1976), and Vygotsky’s (1962) *zone of proximal development* (ZPD) give emphasis to the difference between what an individual can do and learn in isolation and when supported by others and by representational tools. Interestingly, however, in contrast to the distributed cognition view, the scaffolding metaphor and ZPD both imply the goal that – like the building – the child once will be able to stand by itself, through a process of internalisation.

The three empirical studies reported upon in articles II, III and IV, all adopt a primarily cognitive perspective, manifested in an interest in the mind of the individual. However, in trying to make the participants perform at the top of their abilities, we have striven towards providing appropriate scaffolding. In article II, the participants’ task was to interpret two computer animations of an Otto engine. The animations can be seen as a kind of scaffolding and *shared objects of attention* (Schultz, Säljö, & Wyndhamn, 2001) between the interviewer and interviewee: imagine, for instance, how interviews on the functioning of combustion engines and the abstract issue of scientific models would have unfolded without them. In article III, the main form of scaffolding is the collaboration with peers in coming up with analogies and justifying them before the researchers and fellow students. Finally, in article IV, in adapting the approach to the considerably younger group of participants – 7- to 8-year-olds – the scaffolding was of a more tangible kind: interaction with physical phenomena, combined with making their own drawings of things that ‘work in the same way’ and explaining them to the researchers. In addition, particularly in the two last studies, we have had the ambition to establish naturalistic or authentic teaching settings – not in the sense of mimicking in school what happens in society outside its walls – but in making the intervention and data collection resemble regular school practice, albeit hopefully of a novel kind. We wanted to mitigate the pitfalls of laboratory psychology, somewhat sarcastically described by Bronfenbrenner (1977, p. 513, italics in the original) as ‘the science of the strange behavior of children in strange situations with strange adults for the briefest possible periods of time’. In Dunbar and Blanchette’s (2001) words, *in vivo* cognitive studies in authentic settings should be mixed with controlled *in vitro* studies in the psychology laboratory for effective investigation of cognitive phenomena. As mentioned, however, due to research considerations, differing somewhat from those of regular teaching, an authentic ideal has been counter-balanced by, for instance, an interest in how far the participants can reach with different levels of scaffolding. We as researchers have therefore given away less information to the participants, at least initially, than what the typical teacher would be expected to do. All in all, for good and for bad, we have allowed ourselves some opportunistic cherry-picking from different research traditions.

Lincoln and Guba (1985) propose naturalistic inquiry, studying phenomena in their natural contexts, as an alternative to the previously dominating positivistic paradigm of research striving to control phenomena by isolating them. Apart from the natural setting, naturalistic inquiry is characterised by, among other things: human interaction as the main research instrument due to our unique responsiveness and adaptability; predominantly qualitative methods of inquiry; emergent (as opposed to *a priori* fixed) design; purposive (as opposed to random) sampling of informants, typically aiming towards maximum variation; inductive data analysis in line with grounded theory, i.e. Glaser and Strauss’s (1967) idea to analyse empirical data without being informed by a theoretical framework beforehand, and gradually formulate a ‘substantive theory’. To what degree do the three empirical articles II, III and IV share the characteristics of natural inquiry? As mentioned, articles III and IV have in common an ambition of authenticity. However, the studies were short interventions in the everyday life of the participants, rather than the prolonged engagement which characterises typical ethnographic research. The use of qualitative methods and the human being as the principle research instrument was adopted in all three studies. Article II evolved in an emergent fashion where the transcripts were analysed in a bottom-up fashion without relying
of a preconceived theory of the studied phenomenon, and with an approach to sampling of informants very much in line with Lincoln and Guba’s (1985) ideas. We started with a convenience sample of easily accessible informants, but then followed the approach of serial selection of informants where analysis of data from one sample informs the decision of the next sample, continually adjusting and refining the focus until redundancy has been reached. In contrast, articles III and IV used a more fixed, deductive approach, both primarily relying on Gentner’s (1983) structure-mapping theory in analysing the generated comparisons, even though the theme of ownership in article III evolved in a more emergent bottom-up fashion when analysing the transcripts. We had decided on the set of informants beforehand for articles III and IV, as they report on short interventions. Interestingly, however, if seen as two parts of a wider investigation of self-generated analogies for thermal phenomena, article IV may be regarded as a sequel to article III, involving the particular case of first-graders; if first-graders were capable to come up with own analogies for these phenomena and the associated rather abstract concepts, it would strengthen the case that analogies is an ‘untapped resource’ in teaching (May, et al., 2006) – and indeed some of them did express such ability!

5.3. Quality in qualitative research

Larsson (1998) sees qualitative research as boiling down to ‘characterising something’, bringing forward its inherent quality. Against this background, he proposes a set of quality criteria for the practice of qualitative research. Regarding the quality of results, one distinguishing idea of qualitative research, primarily taken from an ethnographic tradition, is to strive for establishing richness in meanings of the encountered phenomena and situations. Trying to establish the richness of meanings of the term ‘entropy’ is the very objective of article I, although in a more restricted, semantic sense of the word ‘meaning’ than typically adopted in qualitative research.\(^{10}\) At the same time, we restrict ourselves to the approach of ‘principled polysemy’ (Evans, 2005; Tyler & Evans, 2001) in balancing such richness in meanings against a discrete set of stable senses of the word that could be listed in a dictionary. In article II, the aim of the study can be framed in terms of establishing the meaning that the participants assign to the animations they interact with. Interestingly, what characterise the vehicle mechanics teachers is that, against a background of their expert insight into the functioning of combustion engines, they see an idealised interactive animation as lacking meaning in relation to the phenomenon. In contrast, at the level of the researchers’ interpretation of the situation, the stance that these participants adopt is highly meaningful. This perspective of participants in a study exploring the richness of meanings in a natural phenomenon by means of external representations, and of the researchers exploring the richness of meanings in the participants’ interaction can also be adopted for articles III and IV. Here, the participants used their spoken language, gestures, drawings, etc., in conveying their ideas of the natural phenomena in focus, and were encouraged to come up with as many ways to see the phenomena as possible. Their degree of success in this endeavour expresses their ability for analogical reasoning, seeing things in terms of other things, and for creating multiple representations of a phenomenon.

Good structure is another quality criterion, which at first sight may be seen to have to be traded off against the richness in meanings. However, Larsson (1998) means that any tension between structure and meaning ideally dissolves if the account harbours exactness of the interpretation. In addition, good structure in the design and communication of research is manifested in consistency, established at two levels: first, that there is a ‘harmony’ running

\(^{10}\) Note, however, that it is the qualitative sense of ‘meaningfulness’ that Johnson (1987) has in mind in his approach to a cognitive semantics.
through the research question, data collection approaches and presentation of results; second, that there is a consistency in the findings, i.e. that the constituent construals fit together in the story as a whole. While it is hard to assess the exactness of interpretations of ones texts as an author, some words on the consistency could be worth mentioning. One common stance is that the research question should drive data collection, analysis, and other subsequent stages of research. In real life, this often does not occur, but the research question is formulated, or at least heavily revised, at late stages of writing up ones findings. Larsson points out that in the ethnographic tradition the researcher is advised to enter a field of study with open eyes or a general interest in mind rather than with a pre-formulated research question. This kind of stance was adopted in article II, where the focus on idealised models emerged in the analysis of the data from the second data collection session. Obviously, the resulting article should be well structured and consistent, but I side with Larsson in that there may be many ways to reach there.

Further, research should be judged with regards to theoretical contributions to its field. Here, Larsson (1998) proposes the heuristics criterion: that the given account makes it possible to see things in a new way:

A successful interpretation thus results in a possibility to regard reality differently or that new categories of thought are created. These may be concepts, mechanisms, or the fact that a phenomenon may be connected to a context which makes something until then incomprehensible more reasonable – you can understand it.

Possible theoretical contributions of the research reported on in this dissertation are discussed in the following section on ensuring trustworthiness and brought up again in the final implications section.

5.3.1. Approaches to ensuring trustworthiness

Validity and reliability, two standard quality criteria of quantitative research, are difficult to apply straight off to small-scale flexible, qualitative research. Lincoln and Guba (1985) propose a widening of the scope and discuss quality in terms of ensuring trustworthiness: how can the reader and research community at large come to be convinced of the plausibility and relevance of a study’s results and conclusions? In striving towards establishing trustworthiness, our overall ambition has been to provide a high degree of transparency, both in the descriptions of the methods used and in the analysis of the results. When it comes to the method, transparency would be manifested as reproducibility. In contrast to psychology laboratory set-ups, many factors vary in an uncontrolled way in the ‘wild’ approaches adopted in articles III and IV, and different results sure would be obtained if the studies were to be transferred to other contexts or groups of participants. Still, the objective of the method descriptions has been for a reader and fellow researcher to be able to understand how the data collection and analysis was conducted at a level of detail sufficient for making a comparative study with an identical set-up. In presenting the results, we have used rather extensive transcript excerpts of the communication taking place, complemented with framing of the contexts and our analyses of what is happening in the event. This gives the opportunity to the reader to make his or her own assessment of the claims and scrutinise the justification of any categorisations, which in our view is superior to just presenting the results of such a classification as relative frequencies or with analytical statistics. Such transparency may be interpreted in terms of reliability:

In a more general sense, researchers using flexible designs do need to concern themselves seriously with the reliability of their methods and research practices. This involves not only being thorough, careful and honest in carrying out the research, but also being able to show others that you have been (Robson, 2002, p. 176).
One common approach to ensure reliability of coding in a scheme of categorisations is for two or more researchers to perform the coding independently in parallel, subsequently compare their scores and calculate an inter-rater reliability, for instance, if applicable, measured with ‘Cohen’s kappa’. This is a route that we considered in article III, where we made a categorisation of the participants’ comparisons. However, in line with the overarching ambition of qualitative research of trying to achieve richness in the understanding of the phenomena, we chose to carry out the classification and further analysis of the transcripts together, sitting side by side.

As opposed to Lincoln and Guba (1985), Maxwell (1992) means that the notion of validity may also be applied to qualitative research, although in ways different from those of quantitative research, and in relation to the more fundamental issue of coming to understand the studied situations. Interestingly, from the point of view of the subject of this dissertation, he frames validity in terms of discrimination between accounts, or representations, of a phenomenon:

…the ways in which researchers make these discriminations do not pertain entirely to the internal coherence, elegance, or plausibility of the account itself, but often refer to the relationship between the account and something external to it – that is, the phenomena that the account is about. Validity, in a broad sense, pertains to this relationship between an account and something outside of that account, whether this something is construed as objective reality, the constructions of actors, or a variety of other possible interpretations (Maxwell, 1992, p. 283).

In an attempt to make explicit some of qualitative researchers’ implicit considerations, Maxwell (1992) proposes five categories of validity: descriptive validity; interpretive validity, theoretical validity; generalizability; and, evaluative validity. These are discussed in turn below in relation to the studies of this dissertation, and where applicable to Larsson’s (1998) roughly parallel set of criteria.

Descriptive validity relates to the factual accuracy of accounts, “matters, on which, in principle, intersubjective agreement could easily be achieved, given the appropriate data” (Maxwell, 1992, p. 286). Maxwell gives the example that if a child throws an eraser in a classroom, describing the act as ‘throwing’ is rather uncontroversial, provided the act has been properly registered. In our empirical studies, while relying on video and audio recording, the basic object of data analysis has been transcriptions of the oral exchanges among the participants and with the researchers. Gestures have sometimes been recorded, typically in brackets as interpretation of the statements when transcribing, but they have not been attended to systematically. Other dimensions, such as prosody or pitch, have been largely omitted, which hinders reanalysis in these respects. The quality of the audio recordings in articles II and III was rather high, so we have good confidence in these verbatim transcripts. The situation was more troublesome for article IV, where many children typically talked and yelled at the same time, and the children that described their drawings often did so shyly in a low voice. Here, only selected episodes were transcribed and with less confidence.

Interpretive validity relates to what the situations in which they interacted mean to the participants themselves. In contrast to the matter-of-fact descriptions above, this always requires the researcher to make inferences from the participants’ statements (Maxwell, 1992). Lincoln and Guba (1985) put forward member checking, also known as respondent validation, where you as a researcher go back to the participants of a study at different stages of analysis to see if you have understood them correctly, as a key approach to this matter, in order to strengthen the credibility or your results. Member checking can be done informally throughout a study, but also formally after interpretations have been formulated. Robson (2002) is more sceptical towards member checking and sees challenges, such as the
respondents being biased towards agreeing with the researchers’ interpretations in order to please them. In carrying out the research included in this dissertation, we have not used member checking systematically; not due to suspected bias, but rather because we did not expect useful outcomes. If confronting participants in hindsight with what they may have thought, we would expect something along the lines of: ‘Yes, you are probably right. But I had not thought about it much.’ Particularly when it comes to the study of children’s behaviour and implied ways of thinking, there is often a large gap between what they can do and what they can formulate and reflect upon. Worse, can they reflect upon what they are not yet capable of doing, such as, for instance, in article IV, conceptualising heat as some kind of substance contained in warm objects? Nevertheless, we carried out interviews with a few of the children after the analogy-generation and drawing exercises, but rather in order to establish a calmer setting for talking together, with the children’s drawings as a shared object of attention (Schoultz, et al., 2001). Instead of formal member checking, we typically aimed for making sure at the time of interview or observation that we had interpreted the participants correctly, by asking whenever we were not sure what they meant or how they had interpreted a situation. This may be seen as a kind of informal member checking, going beyond bare ‘talk aloud’ approaches, with minimal researcher intervention. In our analyses, we have read and interpreted the transcribed statements and tried to put ourselves in the shoes of the participants. In addition, we discussed our results from article IV with the involved teachers and their colleagues some time after the study, and we had the opportunity to discuss our interpretations in article III with some of the participants in subsequent courses in their teacher training program. Overall, the issue of interpretive validity, and particularly member checking, is probably a point where cognitive and ethnographic approaches are particularly difficult to reconcile. Cognitive approaches assume a cognitive theory, maybe accessible only to the researcher but not to the involved studied participants, while the ethnographer wants to understand the phenomena as interpreted within its culture. In retrospect, I find our sparse use of member checks problematic, not from the point of view of validity, but from an ethical one. I sympathise with Lincoln and Guba (1985), in that our participants could have been more involved in the interpretations, since:

It protects them from being excluded in the formulation of knowledge that purports to be about them and so from being managed and manipulated, both in the acquisition and in the application of the knowledge, in ways that they do not understand and so cannot assent to or dissent from (Heron, 1981, pp. 34-35).

*Theoretical validity*, in turn, relates to giving a plausible *explanation* of a situation or phenomenon. In Maxwell’s (1992) view, what characterises this category is that, as opposed to the preceding two categories, different theoretical stances might give radically different accounts; the theoretical level of understanding is inherently controversial. Theories are built up with concepts and relationships between these concepts, which correspond to two aspects of theoretical validity: “the validity of the concepts themselves as they are applied to the phenomena, and the validity of the postulated relationships among the concepts” (p. 291). In a similar vein, but emphasising even further the comparison to previous research or alternative perspectives, Larsson (1998) talks about the *discourse criterion* of validity: that the arguments put forward should be able to stand against other arguments that have been or could be made with regards to the studied phenomenon.

Discussing the articles one by one with regards to theoretical validity, article I takes a semantic view on scientific terms, in this case ‘entropy’, and thereby makes use of central concepts, such as assuming that a word may have several, distinct ‘senses’, corresponding to ‘referents’, and that these senses are connected to each other in a radial fashion. As we have seen, the nature of these concepts is far from generally agreed upon. For instance, the line between pragmatics and semantics is blurred and delimitation of objective senses of words
may be a vain project. Nevertheless, I think that a semantic approach to understanding contributes with an interesting complementary view to the purely cognitive notion of ‘conceptions’. In the analysis of the term ‘entropy’ with the 2-D SAS approach (Strömdahl, 2012), the schema was not merely applied straight off, but also tested and, to some degrees, developed or refined. In particular, as part of the investigation, it became clear that a term may have several senses, all within the same category of scientific modelling, but relating to different scientific fields. In the case of ‘entropy’, this corresponds to the interpretations of the term in classical thermodynamics, statistical mechanics and information theory.

Article II focuses on the dividing line of accepting idealisations and simplifications in models in theoretical scientific traditions, but not in a practical vehicle mechanics tradition. As we have seen, scientific modelling is a large field of study in its own right, but a sceptical view of outsiders, with an in-depth knowledge of the involved physical phenomena, is rather novel and therefore potentially controversial. The choice to discriminate between theoretical and practical traditions of knowledge is rather commonsensical and could probably be replaced with a more established typology of knowledge, such as the classical Greek triple: episteme, techne and phronesis, although they were found frustratingly difficult to delimit and apply to the modern world.

Articles III and IV rely on Gentner’s (1983) structural approach to analogy, which is widely applied in science education and cognitive psychology, but as we have seen in the theoretical background above, it has also been criticised for its failure to account for semantic and pragmatic aspects of analogies. In addition, as pointed out by Goswami (1992), even when accepting a structural approach, analysis of individual comparisons is not always a clear-cut issue, but depending on the constituent entities you choose. For instance, as an example from article IV, comparing the heating of a frying pan to the process of a child becoming warmer may be interpreted as corresponding entities shifting attributes, i.e. a literal similarity, or in terms of going from one state to another, connected by a causal relation, an analogy? In article III, we frame the different treatment of self-generated vs. teacher-generated analogies as a matter of the participants assuming more ‘ownership’ of their own analogies. By following Savery (1996) and Enghag and Niedderer (2008) in operationalising ownership as a set of behavioural indicators, rather than alluding to participants’ psychological states. In this way, we avoided going into full-fledged motivation theory, which might have provided a more rigorous analysis, but it was considered too advanced for bringing across the point we found most intriguing in the study, i.e. that the students spent considerably more time developing their own analogies than elaborating on those previously provided by the lecturer.

Having accounted for the three types of validity above, Maxwell (1992) goes on to discuss the issue of generalizability, referring to “the extent to which one can extend the account of a particular situation or population to other persons, times, or settings than those directly studied” (p. 293). Generalizability is notoriously difficult to accomplish in small-scale qualitative research (indeed, the uniqueness of a particular case may be its main selling point!) and, as mentioned, qualitative researchers are typically more concerned with richness of the accounts of the actually studied phenomena. However, Maxwell suggests distinguishing between internal and external generalizability, where the former relates to persons or events within the studied community but who were not directly observed, and the latter to the case of going outside the studied community. In contrast, Larsson (1998) connects generalizability to the heuristics criterion mentioned above, coming to see things in a new way:

The generalisation from the case happens when those who have been persuaded by the interpretation keep it in mind as they reflect on other cases and thus may discover the relevance of the description that the qualitative analysis resulted in, if they find it useful in the particular case.
With its theoretical analyses, article I may be seen as contributing to a broadening of terms to which both 2-D SAS (Strömdahl, 2012) and ‘principled polysemy’ (Evans, 2005; Tyler & Evans, 2001) apply, hence a case of generalization. Regarding the 2-D SAS analysis, ‘entropy’ was different from the previously analysed ‘heat’ and ‘temperature’, in that it is only rarely used in non-scientific discourse and thereby largely lacking a ‘non-formal sense’. In addition, it is connected to empirical measurement only in an indirect way by being derived from other physical quantities and, as we have seen, has many different senses within different scientific disciplines. When it comes to ‘principled polysemy’, article I is the first study that used the approach to analyse terms, as used in scientific discourse; Evans’ (2005) study of polysemy of the word ‘time’ related only to its senses in everyday language, but not to subtleties regarding time as a physical quantity, such as differences in relativistic vs. Newtonian interpretations. In article II, as mentioned, we followed the approach of adding new participants in an incremental fashion. It was not done by a process of random sampling, but purposefully in order to span a range of different levels of experience of combustion engines in theoretical and practical traditions of knowledge, and expose the variation in their ways of reasoning. Although not being able to categorically state, for instance, that all vehicle mechanics teachers are sceptical about idealised models of Otto engines in secondary teaching, we can point to the existence of at least two such teachers! From a heuristic point of view, the study has succeeded if it makes it possible to see that idealisations made within science can be too radical and therefore not useful from the point of view of practical traditions of knowledge. Moving to article III, the wealth of the analogies generated and the commitment to the exercise shown by the participants is, in our view, quite remarkable. Could this be transferred to another setting? We do not know, and it is a matter of empirical investigation, but my hunch is that it could, due to the ambition in the design to unleash the creativity of the participants. At the same time, a tendency to get stuck in idiosyncratic explanations is also likely to be reproduced. Finally, we were initially cautious about our claims in article IV with regards to the scope of the applicability and homed in on ‘8-year-olds’, ‘two abstract thermal phenomena’, and the like, but one of the reviewers encouraged us to go for the more inclusive: “Young children’s analogical reasoning in science domains”.

With his fifth and last category, *evaluative validity*, Maxwell (1992) takes a more normative tack, e.g. whether or not it was appropriate for young Dennis to throw that eraser. From a normative perspective, regarding the approaches used in this dissertation, I would just like to express my sympathy for emphasising the phenomenon of polysemy, that different words can have several related senses, rather than trying to chivvy out students’ misconceptions about a phenomenon. Likewise, I am attracted to the genuinely constructivist approach underlying self-generating analogies: to build upon what the students already know.

Larsson (1998) ends his list of validity criteria with the *pragmatic criterion*: that the findings and conclusions of research have consequences for the studied practice, corresponding to the sceptical “so what?” question. In this regard, the approach of performing a linguistics analysis of the scientific term ‘entropy’ in article I was at an early stage compared, in a not very flattering way, to Hermann Hesse’s (1987/1943) novel *The glass bead game*: an amusing intellectual pastime, but with little relevance to the surrounding world, and, in particular, to the practice of science teaching and learning. I acknowledge the tension between theoretical investigations and practical relevance, but I hope that we have been able to recover some of the relevance with the rather in-depth analysis of the educational implications, particularly of the recognition that the macroscopic and microscopic senses of entropy are hard to relate to each other. The empirical studies underlying articles II, III and IV are easier to relate to the educational practice. If article II could make teachers reflect on the use of models or representations across the theoretical/practical divide, it would potentially have an influence on the teaching practice. The studies resulting in articles III and IV are put
in a naturalistic setting with the double ambition of collecting data for research purposes and sketching an example of a teaching approach, and are therefore more directly applicable.

Finally, having undergone the process of peer review by members of the research community has strengthened the quality of the argument in the manuscripts through cycles of iterative revision, and has added to the trustworthiness of the studies by appealing to the authority of the journals.

5.4. Reanalysis of selected data in the articles

As mentioned, the ambition in the cover story is not merely to present an overview of the results from the individual studies, but to reanalyse strategically selected examples of the data collection instruments and gathered empirical data from the point of view of the three perspectives of the research framework, i.e. analogical reasoning, semantics, and scientific modelling. Chapter 6 serves the two purposes of: 1) showing how the perspectives relate to one another by means of example cases, and; 2) in line with the overall ambition of qualitative research, providing a richer account of the examples through triangulation in adopting different perspectives on the same data. This emphasis of the richness can be contrasted with a common purpose of triangulation in establishing reliability and robustness, i.e. showing that you can reproduce roughly the same results through different approaches.

The reanalysis of article I consists of adding a new layer of interpretation, an exercise made possible through reading and conducting additional research in other related academic traditions after the study was published. A more in-depth study of cognitive linguistics in Amin, et al. (2012) and Jeppsson, et al. (2012) led to a new interpretation of entropy as a substance in the data also in article I, and Jeppsson, et al. (2011) contributed to thoughts on metaphors involving entropy. Further investigation of the issue of reference in the thermal domain (Strömdahl, et al., in progress) within the philosophy of science and the philosophy of language has given further reflections on the reference of entropy.

Next, the reanalysis of article II focuses primarily on the structure of one of the two computer animations of ideal Otto engines that was used in the interviews, and therefore not on the respondents’ interpretations of the animations. This animation serves as a good example for a discussion of the degree to which structural similarity is involved in connecting different representations of a phenomenon and connecting them to the represented phenomenon, addressing research question R1.

As for article III, the reanalysis deals with the connection between the teacher student’s analogies and scientific models, thereby addressing research question R2. Both models and analogies involve making mappings between domains, but one broad dividing line is that scientific models have some representational pretention in relation to the world, while analogies are not required to have such connections. Here, two analogies generated by the teacher students, which differ with regards to the potential faithfulness of representation, are brought to the fore, and discussed from the point of view of science learning.

Research question R2 is kept in mind also in the reanalysis of the data from article IV. Here, we reconsider the two collages created by first-grader Lisa, representing different aspects of the phenomena of ‘heat’ and ‘mixing’. These drawings epitomise the dissertation in constituting multiple self-generated analogies and models of phenomena, and showing that words may have several related meanings: polysemy.
6. Summary, reanalysis and discussion of the individual studies

We now turn to a reanalysis of selected parts of the empirical data from the four articles, from the points of view of the three perspectives in the theoretical framework, i.e. analogical reasoning, semantics, and scientific modelling. The original focus in the articles with regards to the theoretical framework is indicated in the schema to the right. The sections begin with a brief summary of the general design, research questions, and main findings of the studies, and then turn to complementing the analysis from the remaining two perspectives of the framework, here in dotted lines. As presented above, this reanalysis was informed by the following research questions:

R1. To what degree is analogy involved in connecting different representations of a phenomenon to each other and to the represented phenomenon?
R2. How do students’ self-generated analogies relate to scientific modelling?

6.1. Article I – Connections between different senses of entropy

In article I, we developed a semantic network for the word ‘entropy’, shown in Figure 7, based on five distinct senses of the word and how they are related to each other. We followed the approach of ‘principled polysemy’ (Evans, 2005; Tyler & Evans, 2001) in establishing a set of stable senses of the word, adapted in the way that different kinds of referents would discriminate between two senses of the word.

The study was conducted in order to answer the following research questions:

- What are the distinct senses of the word entropy and how are these senses of entropy related logically and historically?
- What are the educational implications of the answer to the question above regarding teaching and learning the scientific senses of entropy?

In the development of the semantic network in Figure 7, Boltzmann’s and Gibbs’ statistical interpretations were seen as the central, sanctioning Statistical Sense of the word entropy, due to its generative character, in spite of the historical precedence of the macroscopic Thermodynamic Sense which Clausius had in mind when coining the word in 1865.11 Even though the Thermodynamic Sense and Statistical Sense relate to a largely overlapping domain of thermal phenomena and the same physical quantity, we identified them as distinctly

11 It should be noted that by 1865, Clausius fully embraced a microscopic, kinetic theory of gases, to which he also had contributed himself (Müller, 2007). Qualitatively, he related entropy to the “disgregation of a body… depending on the arrangement of the particles of the body” (Clausius, 1867, p. 356), reminding of the proposed view of entropy as the spreading of a system’s energy (Leff, 1996). However, a quantitative microscopic interpretation of entropy was introduced first by Boltzmann.
different senses, due to their interpretation in different theories, with radically different models; in our view, they have different referents. Gibbs’ formulation of entropy as a function of probabilities of microstates was borrowed by Shannon in his development of information theory, but leaving behind its connection to physical systems. Here, the two senses share the mathematical formalism, but refer to very different entities: an aspect of a microscopic model of a thermodynamic system and information contained in a message, respectively. The Disorder Sense, the adoption of the metaphor ‘entropy is disorder’ for instructional purposes or outside natural science, as in art theory, can, in our view, only be conceived of in a microscopic interpretation; there has to be some constituent entities that are disordered in relation to each other. However, its negative connotations cannot have been taken from the probabilistic, neutral view of statistical mechanics, but is likely to derive from the ‘heat death’ implications of the second law of thermodynamics, thus motivating the complementary horizontal arrow coming from the Thermodynamic Sense in Figure 7.

From a teaching perspective, the facts that there are many ways of approaching the concept of entropy in different scientific domains and that it is difficult to see how these domains are connected provide formidable challenges to students. Baierlein (1994) offers an algebraic approach to justifying the macroscopic-microscopic connection in physics by deriving the macroscopic relation $\Delta S = Q/T$ for the case of isothermal expansion of an ideal gas, assuming the number of microstates of a particle to be proportional to the volume. However, it is still difficult to see how a macroscopic treatment, using ideal heat engines as referents, relates conceptually to a microscopic approach using energy distributions across particles.

Passing on to the reanalysis of the matter treated in the article, one fundamental question is what the relation is between the Statistical Sense and Thermodynamic Sense of entropy. One line of reasoning holds that thermodynamics can be reduced to statistical mechanics, meaning for example that thermodynamics can be derived from or explained by statistical mechanics or that the latter applies to a wider range of phenomena, i.e. is a more general theory. Sklar (1993) fits this endeavour in the larger undertaking of reducing theories of observational macroscopic phenomena to theories of their microscopic constituents, but is sceptical about the prospects of success in reducing thermodynamics to statistical mechanics in totality. Problems emerge when we look at the details of the two theories and try to
establish correspondences between individual concepts. Some concepts are rather unproblematic. For instance the volume of a system as interpreted in thermodynamics maps neatly onto the volume in statistical mechanics. However, already seemingly clear-cut concepts, such as ‘heat’ and ‘pressure’ offer subtleties and with ‘temperature’, matters become even more intricate. As we have seen, temperature is fundamentally the inverse of the derivative of the entropy with regards to the internal energy for isolated systems, while the commonly put forward identity of temperature with the average kinetic energy per particle is limited in its application. In addition, statistical mechanics offers interesting generalisations beyond thermodynamics, such as the potentially negative temperature of nuclear spins of the atoms in a crystal. The most challenging issues, however, are introduced with entropy:

The concept of entropy is the most purely thermodynamic concept of all. Its primary meaning is fixed entirely by the role entropy plays in the characterization of systems that follows out the method introduced by Clausius of proving entropy’s existence as a state function from the basic consequences of the Second Law. Given the abstractness of entropy and its place high up in the theory and unrelated to immediate sensory qualities or primitive measurements (as temperature is related to these), it is not surprising that in seeking its statistical mechanical correlate of thermodynamic entropy we have the least guidance from the surrounding embedding theory. [...] The additional levels of complexity in the reducing theory may leave some openness in what to choose as the surrogate for entropy. This is exactly what we find – that is, a wide variety of ‘entropies’ to correlate with the thermodynamic concept, each functioning well for the specific purposes for which it was introduced (Sklar, 1993, p. 354).

As Grad (1961) emphasises, one has to settle upon one out of many alternative Statistical Senses of entropy given the particular circumstances, constrained by characteristics of entropy, such as its additivity and maximum value at equilibrium. Different entropies are formulated when following Boltzmann’s approach of partitioning a phase-space of a system in discrete boxes or Gibbs’ approach with probabilities of states across ensembles. In article I, we handle this plethora of different statistical mechanics treatments of entropy in terms of different statistical sub-senses of the word and when establishing a connection to the Thermodynamics Sense, each case would have to be discussed in turn. Further extensions of entropy are conceivable in both macroscopic and microscopic interpretations, such as the ambition to account for non-equilibrium phenomena and the entropy of black holes. Bartels (1995) points out the connection between black hole entropy and traditional interpretations in thermodynamics and statistical mechanics as particularly interesting, due to its being at the limits of our current physics theories at the time being; we still wait for an integrated account of quantum mechanics and general relativity, and cannot ascertain how black hole entropy relates to the other entropies.

Shannon’s (2001/1948) introduction of the Information Sense of entropy bears all the hallmarks of a *bona fide* analogy in relation to the Statistical Sense. In line with Gentner’s (1983) structure-mapping theory, two superficially very different domains – thermodynamic systems and communication systems – are compared, and found to share a structural similarity. Indeed, from Bartel’s (1995) philosophy of science point of view, the crucial issue is whether different entropies refer to the same thing in the world and he therefore immediately categorises entropy in the Information Sense as an altogether different beast than the physical quantity underlying the Thermodynamic Sense and Statistical Sense; the relation is one of analogy, not identity. Shannon declares:

Quantities of the form \( H = -k \sum p \ln p \) (the constant \( K \) merely amounts to a choice of a unit of measure) play a central role in information theory as measures of information, choice and uncertainty. The form of \( H \) will be recognized as that of entropy as defined in certain formulations of statistical mechanics… (Shannon, 2001/1948, p. 12)
In other words, structurally indistinguishable relations – \( H \) and \( S \) as functions of \( p_i \) – are given radically different semantic interpretations, hence a case of metaphorical extension of the word entropy (Johnson, 1987). However, in contrast to Johnson’s view of coinciding experiences as a basis for metaphorical extension, Shannon probably became aware of the mathematical structural similarity of the problems at hand. Another thing that differs between the framework of conceptual metaphor and Shannon’s adoption of the term entropy in information theory is that conceptual metaphor deals with unconscious mapping between two domains as a part everyday language, while Shannon adopted entropy in a very thought-out and technical manner. In this new developing domain of information theory, entropy took on the new Information Sense, separated from any connections to physics. As we point out in Jeppsson, et al. (2011), the statement ‘entropy is information’ has, in fact, come to be a valid identity or category inclusion statement within information theory, where it has lost its metaphorical interpretation. The ‘career of metaphor’ (Bowdle & Gentner, 2005) has been a rather quick one in this case. Interestingly, having come to see the structural similarity between thermal and informational phenomena sparked off an investigation whether there are more fundamental connections, a kinship also in the semantics. Jaynes (1957) shows how statistical mechanics may be formulated with an information theoretical approach and, the other way around, information processing in reality has to be carried out in compliance with the laws of physics, and leads to increased entropy. For instance, Bennett (1982) shows that Maxwell’s demon – an imagined device placed between two gas cylinders, letting through only molecules coming from the right but blocking those from the left, which would lead to decreased entropy in the system – in fact is not a threat to the second law of thermodynamics. However, it is not, as had been previously argued, the act of discriminating between particles from the right and left, i.e. measurement, that is necessarily irreversible, leading to an entropy increase. Instead, it is the act of erasing or ‘forgetting’ the demon’s previous measurements, preparing it for new measurements, that introduces the irreversible step and thereby leading to an entropy increase counterbalancing the entropy taken out of the system by the demon. Such two-way connections motivate the double-headed arrow between the Statistical Sense and Information Sense in Figure 7. This reciprocal relationship between the Statistical Sense and the Information Sense of entropy fits with Black’s (1962) interaction view of metaphor or the mutual alignment variant of analogy (Kurtz, et al., 2001), where the understanding of both domains is deepened by establishing their connection.

Apart from the four senses previously recognised in the history of the development of ‘entropy’ as a word, we also described a more novel Homogeneity Sense, an interpretation of entropy as the qualitative state of homogeneity, which a system either is in or not. We provided translated example quotes from a Swedish text corpus (Språkbanken, 2009), such as:

[A] three ton bitumen cube that is cubic at first, but after a while slowly settles into a flat lump of asphalt. It creates a powerful image of how the form of matter is smoothed out into entropy. (1)

He seems to suggest that it is only a matter of technology to avoid entropy. Well, of course you could say that whether an industry is polluting or not can be decided by seeing if the smoke goes out of or into the chimney. (2)

12 Ironically, from the point of view of this dissertation, one of the crucial steps Shannon (2001/1948, p. 3) takes in his seminal paper is to realise that information should be treated in isolation from semantics: “The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.”
Human technology will certainly not be totally free from entropy, but the degree of entropy can vary significantly between different technologies. (3)

We suggested that the Homogeneity Sense has derived from the Disorder Sense through a process of image schema transformation (Lakoff, 1987), where there has been a shift from the dynamic process of a system getting disordered to an end state of disorder, much as ‘over’ has taken on a new sense in ‘John lives over the hill’ compared to ‘John walks over the hill’. At our initial analysis, we were puzzled of the Homogeneity Sense and generally quite forgiving of non-scientific uses of the Homogeneity Sense, as (1) above, but critical towards its application in scientific or technological writing, as in (2) and (3). After having performed conceptual metaphor analyses of language involving entropy in textbooks (Amin, et al., 2012) and spoken language during problem solving (Jeppsson, et al., 2012) an alternative interpretation of (1) - (3) becomes possible. While energy often is spoken about metaphorically as a substance-like entity that resides in objects, entropy is conceptualised as a value on a scale, particularly in conjunction with energy as in: ‘as more energy entered the system, the entropy increased’. However, although more unusual, it is not impossible to conceive of entropy as substance-like, as in The Feynman Lectures on Physics:

Any heat $dQ$ that has been absorbed from the substance at temperature $T$ has now been converted by a reversible machine, and a certain amount of entropy $dS$ has been delivered at the unit temperature as follows: $dS = dQ/T$ (Feynman, Leighton, & Sands, 1963, pp. 44-10).

In fact, as mentioned, the Karlsruhe Physics group (e.g. Herrmann, 2000) has suggested to regard entropy as the central extensive substance-like quantity of thermal physics. In this vein, “smoothed out into entropy”, “avoid entropy” and “free from entropy” may be interpreted in terms of generation of an amount of entropy, or trying to avoid it, rather than reaching the qualitative state of homogeneity.

When it comes to the teaching of thermodynamics, the adoption of the Disorder Sense of entropy is the most controversial, and Lambert (2002) has launched a quest for its removal in science teaching, particularly at the university level, which has been largely successful when it comes to textbooks, but is probably difficult to complete in terms of the minds of teachers and teaching in practice. In this respect, the finding by Amaral and Mortimer (2004) that the empirical zone of the conceptual profile, including the idea of entropy as disorder, serves as a way for students to connect their intuitive ideas with a more sanctioned, received view of the second law of thermodynamics using free energy and distribution across particles is very interesting indeed. Even if the idea of entropy as disorder is not in line with current physics or chemistry, it may serve as a communicative devise or stepping stone towards a more advanced understanding.

While Article I and our subsequent study on metaphors for entropy (Jeppsson, et al., 2011) bring up metaphors that may lead to a scientific understanding of entropy, e.g. ‘entropy is information’, Zencey (1991) tells the complementary story of how entropy itself has come to be used metaphorically in society, where the entropy of the bitumen cube applied in the domain of art mentioned above may serve as an example. Zencey shows how ‘entropism’ has become an all-encompassing root metaphor for processes in nature and society, applied, for example, to economics, the philosophy of history and dystopic fiction. He traces the enticement of the entropy metaphor to a number of different factors, including: its pessimistic connotations to increasing disorder; our energy-conscious age, prone to borrow from the field of thermodynamics; an appeal to the authority of science; and, the dialectical tension between nature’s tendency for degradation and man’s struggle against it.
6.2. Article II – Representations of the Otto engine

Article II provides an analysis of two computer animations representing Otto engines, and how 19 respondents interpret these animations. The first ‘detailed animation’ shows the functioning of a cylinder of an Otto engine as a short cyclical drawn film clip, where the parts of the engine are depicted in a fairly realistic fashion. In the second, ‘interactive animation’, a cylinder is depicted in a more schematic stick-figure fashion, together with a PV graph and numerical readings of involved physical quantities throughout an ideal Otto cycle.

The respondents were selected to represent a wide range of experience of combustion engines in theoretical and practical traditions of knowledge, a notion introduced by Molander (2002). Upper secondary school students and teachers specialising in natural sciences represented a theoretical tradition of knowledge, while students and teachers specialising in vehicle mechanics represented a practical tradition of knowledge, and PhD and MSc students in vehicle system engineering were seen as having an in-depth insight into theoretical traditions of knowledge, but also some knowledge of practical traditions.

The study addressed the following research questions:

- What aspects of modelling are focused on in a theoretical or practical tradition of knowledge, respectively?
- How is the interpretation of two computer animations depicting models of a technical artefact and its function influenced by the level of experience in a theoretical or practical tradition of knowledge?

The most intriguing finding in the study was that the rather crude stick-figure drawing and理想isations made by use of an ideal Otto cycle in the ‘interactive animation’ were not seen as conducive to learning about the functioning of engines by the vehicle teachers representing a practical tradition of knowledge. The vehicle teachers and students quite liked the ‘detailed animation’, but would have preferred an even more realistic slow-motion video film of the functioning of an engine and hands-on readings of the pressure of a real, physical engine. This may be contrasted with the theoretical traditions of knowledge, where such simplifications and idealisations were seen as an integral part of modelling in science and engineering, facilitating a quantitative, algebraic treatment of the phenomena.

In the following, we revisit the ‘interactive animation’ with particular regards to the way the different synchronised parts of the animation relate to each other and the underlying phenomenon they portray. The end state of the animation is shown in Figure 8. Overall, the three representations of an engine, a stick-figure drawing, a PV diagram and numerical measures, may be seen as an example of a cascade of inscriptions (Roth, et al., 1997). Provided sufficient knowledge of the function of engines, we can establish correspondences between parts in the stick-figure drawing to the left and parts in a physical engine. In this respect, the drawing is an example of an icon, similar to what it represents (Peirce, 1985).
Having established the representational relation between the image and a cylinder in an engine as one of rough superficial similarity, we come to the philosophical issues of reference and representation. In virtue of what does the drawing represent or refer to the engine? Following the causal theory of reference (Kripke, 1980; Putnam, 1973), we can track the reference to a moment of creation and the designer’s intention: this drawing is altogether meant to refer to an Otto engine. Due to the concrete character of the drawing, with some knowledge of engines, an interpreter could be able to pick out the details of the drawing and see that it refers to an engine. If you do not have such experience of engines and representations of them, you may need time and help in seeing the critical components. For instance, it took time for the natural science students to discern the red capital Ts turned upside-down and understand that they represent valves, which are crucial in understanding in what phase of the cycle the engine is. However, once you have discerned them and grasped their role in the whole, such physical objects are rather unproblematic. But this is not all the designer intended with the drawing. For instance, changes between the four phases or strokes of a cycle of the engine are represented by opening or closing of the valves, letting gas into or out of the gas chamber, which at the moment of change turns colour. In this regard, how does the drawing refer to or represent, for instance, the intake stroke, the phase in which a mixture of air and fuel is taken into the cylinder? The intake stroke is an example of a theoretical term, and I am in favour of causal descriptivism (e.g. Kroon, 1985) in that you need descriptive statements in addition to the causal link to the phenomenon in order to delimit its meaning. You cannot simply point to the drawing at a particular moment and say that it now represents or refers to the intake stroke. The same goes for most involved physical quantities, such as pressure, temperature or entropy, even though the volume may be less problematic.

Moving over to the top right part in Figure 8, there is a PV graph of an ideal Otto cycle, where the pressure in the gas chamber is drawn against the volume, in synchronisation with
the motion in the drawing to the left as the cylinder goes through its four strokes. What is the relationship between, on the one hand, the graph and, on the other, the stick-figure and a real Otto engine? First, I would like to point out something that to me is rather obvious and trivial: The representation does not look like the represented phenomenon in any way. It is far from similar to it, let alone isomorphic. But, on the other hand, it is not an entirely arbitrary representation either. Being an graphical representation of a scientific model, Kralemann and Lattmann (2012) would classify it as an icon, due to the fact that its inner structure has implications regarding the functioning of the engine it represents. However, I think that the fact that the structure of the graph says something about the structure of the stick-figure drawing or a physical engine is not sufficient to claim structural similarity and iconicity, since none of the constituents or relations between them can be mapped across the representations. From a historical point of view, the use of PV graphs can be tracked down to Southern’s invention of the engine indicator, a measurement device to plot the pressure against the volume of a running steam engine, in the late 18th century (Walter, 2008). The phenomenon-representation connection was established by having a contraption draw a graph of its functioning, mechanically, in real-time. In my view, in Peirce’s (1985) classification, this causal, pragmatic representation has a strong index aspect to it. Later, PV graphs came to be used in representing also idealised engine cycles, such as in Clapeyron’s popularisation of the ideal Carnot cycle (Müller, 2007), which inspired engineers such as Diesel to develop ever more efficient physical engines (Bryant, 1976). Once again, there is a moment of baptism and a conveyance of ideas in culture along the line of the causal theory of reference (Kripke, 1980; Putnam, 1973). The way PV graphs have come to represent ideal or physical engines also illustrate the role of intervention in nature, here by means of measurement, as pointed out by Hacking (1983), and diSessa’s (2004) idea of scientists as designers of representation. From a cognitive linguistics perspective, the PV graph is based on the underlying conceptual metaphors States Are Locations (Lakoff & Johnson, 1980) and Change Of State Is Movement Along A Path (Jeppsson, et al., 2012). Particular lines and curves in the graph can be connected to distinct processes in the cycle, such as the downward slope at the top which is drawn as the gas expands during the power stroke, actually contributing to a structural, iconic interpretation. In addition, the synchronisation of the two representations is probably helpful in recognising the connection between the drawing to the left and the graph, reminding of Johnson’s (1987) emphasis on the connection between coinciding experiences. However, increasing idealisation may be an obstacle to seeing how the representation refers to the studied phenomenon; no combustion engine could ever produce a ‘real’ PV graph or indicator diagram looking like the one here, as kindly pointed out by the vehicle mechanics teachers.

Finally, we can compare the PV graph to the set of fictitious physical quantity readings underneath. What is the relationship between the graph and the consecutive set of number sextets? It is tempting to talk about isomorphism, at least between a number pair of pressure and volume readings and a point in the graph. However, I would claim, with Vessey (1991), that the translation between table and graph formats here implies crossing the divide between discrete and continuous representations, making an exhaustive matching of elements impossible, and we can therefore reject isomorphism. Both representations are technically probably generated from the same digital set of data, but mentally, as pointed out by Vessey, it is hard to translate between these modes of representation. Translation requires knowledge of the procedures of sampling, when going from a continuous line to discrete point measures, and extrapolation in the other direction. These are standardised and conventionalised methods in science, but not relying on structural similarity.

All in all, how does the semantic view of theories fare with respect to the ideas that models are connected to each other and represent phenomena exclusively by virtue of similarity or isomorphism? I would side with Suárez (2003) in answering: Not very well.
**6.3. Article III – Teacher students’ self-generated analogies**

The study reported in article III aimed at exploring how the approach of *self-generated analogies* (Blanchette & Dunbar, 2000; Wong, 1993a, 1993b) might be applied in the teaching of thermal phenomena at university. In addition, we tried to establish a context close to regular teaching, by letting the participants interact in group work. Accordingly, two groups of four preservice physics teacher students in each were asked to generate analogies for two thermodynamic processes, reversible adiabatic expansion of an ideal gas, and free adiabatic expansion of an ideal gas.

The study was guided by the following research questions:

- *How do students perform analogical reasoning in a group exercise when they are asked to come up with analogies for thermal phenomena?*
- *How is ownership of learning manifested as students come up with analogies?*

The participants generated many analogies, some of which were elaborated to a considerable depth. Interestingly, all analogies were based on microscopic modelling, where entities such as balls or people corresponded to particles in a thermodynamic system, but no analogies were based on similarities in algebraic structure to other scientific models, such as Carnot’s waterfall analogy. This preference of microscopic approaches lends support to Reif’s (1999) view of macroscopic thermodynamics as particularly difficult to visualize. In addition, the students developed the self-generated analogies to a greater depth than analogies they had been presented to in the preceding teaching of thermal physics. However, in contrast to Blanchette and Dunbar (2000), who attributed such differences in the interpretation of teacher-generated and self-generated analogies to conceptual challenges in grasping the teacher-generated analogies, we related these differences to the notion of ownership (Enghag & Niedderer, 2008; Savery, 1996). The participants recognised that their own analogies were not perfect and it was up to them to scrutinise how far the correspondences could be taken and at what point the analogies would break down. The other way around, it is easy to take for granted that the teacher has chosen to present only such analogies that are highly appropriate and forget that they also will break down when stretched too far.

When it comes to the particular analogies, two of them are now revisited from the point of view of how models relate to analogies. One of the two groups came up with the analogy between a thermodynamic system and a jar of angry bees and both groups made use of different variants of ideal, macroscopic balls colliding with each other and with the surrounding walls, reminding of Maxwell’s billiard ball models of gases, related to above.

Overall, I side with Sibley (2009) in that scientific models are restricted to representations that in some way have been acknowledged by the research community, in being useful for predictions, explanations, etc., while analogies not necessarily carry such pretentions. Scientific models may encompass considerable simplifications; we may be well aware of aspects that differ between the natural phenomena and models of them. However, crucially, we may play pretend: What if nature is constituted by a certain set of entities and
behaves in a certain manner? In the case of the teacher students, this act of imagination was facilitated by a tactical choice in the setting: comparison of particles in a thermodynamic system interacting only at collisions with balls that bounce on walls in free space, where the influence of gravity and air resistance can be neglected. For good and for bad, we may even conflate the conceptions of the phenomenon and its representation: going beyond pretending or imagining the phenomenon as if it behaved like the model, what if the phenomenon actually works like the model? Students easily forget about the idealisations that necessarily have been made in developing the model and its tentative character when framed in the scientific endeavour as a whole. However, as we saw in section 3.3.1, they are in good company, as Maxwell gradually went from the as if mode of thinking to committing to the actual existence of electrical fields (Black, 1962). In Hesse’s (1966) words, ascertainment of the correspondence of neutral analogies is one of the main uses of scientific models.

The case of the analogy with the angry bees is a quite different story. The students are not tempted to believe that a thermodynamic system actually works like a can of bees. They cannot further rely on the authority of their teachers or scientists in that the match between particles and bees is a particularly good one. They are out on their own and it is up to them to ascertain the merits and shortcomings of the analogy, a case where assuming ownership for one’s learning and for the particular generated analogies becomes particularly important. While many correspondences between a gas and a billiard ball model of the gas are unproblematic and taken for granted, this is not the case for the angry bees. As pointed out by Heywood and Parker (1997), however, this is not necessarily a bad thing from a learning perspective; with a strained or even silly comparison, the students are forced in a creative way to reflect upon in which respects the analogy holds and where it breaks down. In my view, the gradual development of the analogies by taking into account new aspects one by one – such as the size of individual bees and their ability to rotate with respect to the three dimensions of space – and investigating what they may correspond to across the domains was much facilitated by the design choice to have the participants work together in groups.

6.4. Article IV – First-graders’ self-generated analogies

The study eventually resulting in article IV was initiated based on the perceived usefulness of asking university students to generate their own analogies reported in article III. There was also an increasing focus on teaching children at younger ages at our institution, which encouraged us to endeavour into this new domain of study: would it be possible to adapt a similar approach to first-graders? Compared to the university students of article III, we were now confronted with several additional challenges.

First, admittedly, it may be difficult to bring about analogical reasoning among young children. Indeed, Piaget, et al. (1977/2001) claim that full-fledged analogical reasoning typically is inaccessible to first-graders, although this has been challenged by subsequent research (Goswami, 1992). Second, these children’s understanding of thermal phenomena is very limited compared to that of university physics students. In addition, the collaborative group work approach would be difficult to use among such young children.
The participants were one class of first-graders and the study design was developed in collaboration with their two teachers. The concept of ‘analogy’, which was taken to be unknown to the children, was introduced by showing the children images of a car, a bicycle and a walking girl, and asking them in full class in what respects these items ‘work in the same way’. Next, they were exposed to two different experiments in smaller groups: marbles getting increasingly mixed as they are rocked back and forth on a see-saw, and: a frying pan being heated when placed on a hot-plate. The children were asked to predict and then carefully observe what happened at different stages of the experiments and explain any discrepancies in a POE fashion (White & Gunstone, 1992). In a final stage, the children were asked to represent the phenomena by means of drawing and to come up with analogies for them, once again in terms of things that work in the same way.

The study was guided by an overarching research question:

- **How does children’s analogical reasoning differ between teacher-generated and student-generated analogies?**

In addition, the following complementary research questions helped provide explanations for the children’s differing analogical reasoning in different circumstances:

- **How is children’s ability to perform analogical reasoning affected by the conceptual difficulty of the target domain?**
- **How do different kinds and degrees of scaffolding influence children’s analogical reasoning?**

Coming to the results, the children as a collective had no problems identifying functional ways in which the car was similar to the bicycle. For instance, the car’s steering wheel was found to correspond to the bike’s handlebars, since they make it possible to steer and turn the vehicles, regardless of their superficial physical dissimilarity, hence a *bona fide* analogy! This ability for analogical reasoning was extended also to the walking girl, a conceptual domain more distanced from the two vehicles, but the children retained their structural focus. For instance, one child realised that the steering wheel of the car could be seen as corresponding to the legs of the girl, once again because they make it possible to turn, after a moment of hesitation based on the immediate association between the steering wheel and the hands that hold it. Intriguingly, some – but far from all – of the children were also able to generate their own comparisons with a predominantly structural or functional focus, for the two physical phenomena they had encountered.

Here, I would like to revisit two drawings made by the first-grader Lisa representing ‘other things that work in the same way’ as the two encountered phenomena of heat transfer from a hot-plate to a frying pan (Figure 9) and mixing of marbles as they are rocked back and forth on a board (Figure 10). Overall, you can see a personal aesthetic or style in the drawings; they are both collages of rough sketches portraying different aspects of the phenomena, in the first case with neat little captions of the distinct ideas.

On the left of the drawing of heat transfer (Figure 9), Lisa has drawn the sun as it heats the water, in which people are bathing. She has used the common convention of drawing straight yellow rays of light radiating out from the sun, but also, more ingeniously, spirals over the water, presumably representing the heat of the water. Is this an analogy? In article IV, our answer is yes. Even though we did not ask her explicitly of how the drawing relates to her experience of the hot-plate experiment, the sun maps to the hot-plate and the frying pan to the water, thereby establishing a relation between a source and recipient of heat in the two moderately close domains.
The other items of the drawing represent aspects of heat that Lisa associates to, in line with the findings of Albert (1978) among other children of similar age, but without a clear relational focus. These include the heat of a radiator and the heat of a shining light bulb (notice the spirals indicating heat on top), the sensation of hotness as you move yourself, the heat of an engine, the lava running down a volcano and, the less common reflection that the dry stem of an apple tends to burn, in contrast to the juicy, fleshy part of the apple. However, what strikes me is the richness and diversity of her conceptions of what is going on. Once she has caught an aspect of the phenomenon, she quickly draws a sketch of it, possibly as a note for her memory in line with Ainsworth’s (2010) view of drawing as a tool for your own learning. Once this is done, she is not contented, but quickly moves over to another, related phenomenon, showing yet another meaning of the word ‘warm’.

Lisa’s other drawing – that of the marbles being rocked back and forth (Figure 10) – is, if possible, even more varied. Rather than sticking to one concept, heat in the first drawing, here she picks out quite different aspects of what is going on. At top left, there is a marble track where two marbles move in opposite directions and collide. In this image, she focuses on the aspect of collision and simplifies the scenario by restricting herself to a one-dimensional trajectory and a previously known domain. At bottom left, she has drawn a game of dodge-ball, where children try to hit each other with soft balls, which she explicitly maps to the event of the marbles bouncing on the short sides of the board, and hitting another child corresponds to the marbles hitting and being deflected from a plastic separator taped to the board, a quite different, but complementary aspect of the marble game.
At the top right, there is a sad girl who has dropped her groceries on the ground, creating “an awful mess”. In a more abstract comparison than the other ones, she relates the marbles getting mixed up to the groceries getting mixed up on the ground, reminding of the previously common, but controversial introduction to entropy in terms of a ‘messy room’ in university textbooks, yet again a quite distinct aspect. All these three representations stand in an analogical relation to the board with marbles. In addition, this way of picking out different aspects of a phenomenon and relating them to other known phenomena reveals an awareness of the gist of scientific modelling: that there can be many models of one phenomenon, each focusing on a particular set of objects and that considering many such models in combination may give a richer understanding than when sticking to one model alone. As we have seen, such ability to handle several complementary conceptions of a phenomenon in mind has been discussed in terms of multiple analogies (Spiro, et al., 1989) or different zones of a conceptual profile (Mortimer, 1995), and it is truly astonishing to see it expressed by an 8-year-old child.

---

13 See e.g. Lambert (2002) and Jeppsson, et al. (2011) for a discussion of the ‘messy room’ analogy.
7. Conclusions and implications

In this final chapter, we revisit the auxiliary and research questions of the dissertation after having developed the theoretical framework and applied it in the reanalysis of the included articles, and end by drawing some conclusions for science education research and practice.

7.1. Revisiting the auxiliary and research questions

A1. How have the fields of analogical reasoning, semantics and scientific modelling been used in the study and development of thermodynamics and thermodynamics education?

The research on analogical reasoning, semantics and scientific modelling, and their application on thermodynamics education have followed intertwined paths in roughly the same direction. After a ‘cognitive turn’, structural approaches and the use of computer simulations of cognitive structures were popular in the study of analogical reasoning. In particular, Gentner’s (1983) structure-mapping theory has been influential in accounting for analogies in science education, but also criticised for neglecting factors, such as the meaning of the modelled objects and the contexts in which analogical reasoning is carried out. In science education, the quest for developing the perfect analogy prior to teaching still seems to be largely in vogue, even though the more genuinely constructivist idea of self-generated analogies picks up interest from an increasing range of educational researchers. With regards to scientific modelling, structural approaches, such as the semantic view of theories, have received less interest in science education, and have been found largely incapable of capturing the creative spirit of the scientific endeavour. Here, Nersessian’s (1999) focus on cognitive aspects in authentic cases from the history of science has been more fruitful. All in all, however, going forward, I agree with Amin (2012) that semantics and language in general is the most underexploited field of metaphors and models in science education. After the cognitive turn, propositional approaches were abandoned and with them the entire field of language, but further progress would require bringing language back into the picture.

A2. How are the fields of analogical reasoning, scientific modelling and semantics related to one another?

As shown in Figure 5, the three different perspectives in the theoretical framework share the structure of mapping two entities, where one typically is of a more concrete kind, anchored in the world, and the other is more abstract or conceptual. Then again, when you look at the closer details, peculiarities and inconsistencies appear, altogether comprising a complex pattern. However, I would once again like to point out some features of parallelism. First, as pointed out by Hesse (1966) analogical reasoning is clearly recruited in many stages of scientific modelling, including coming to see connections between models of different phenomena, but also, potentially, between some different representations (but not all) of one individual phenomenon. Similarly, and emphasised particularly in cognitive linguistics, language as a semantic system is not completely arbitrary, but connected systematically and metaphorically to the world through our experiences. Second, as mentioned above, the way we are acutely but unconsciously sensitive to particular expressions in language is intriguing. How come ‘my lawyer was an old shark’ and ‘my lawyer was like an old shark’ (Glucksberg & Haught, 2006) bring such different thoughts to mind? As we have seen, this is still an unsettled issue in the debate regarding the connection between metaphor and analogy. How such subtleties influence our learning – for instance, does it matter if we say ‘the energy in the box’ or ‘the energy of the box’? – we simply do not know. However, my bet, as expanded upon in Jeppsson, et al. (2012), is that our capacity for flexible metaphorical interpretation is vast and that as teachers, we should not restrict ourselves to using a literal, correct language, but have to allow ourselves, and cannot avoid, exploiting the possibilities in metaphoric
language. Third, we have seen how the phenomena of polysemy, multiple models of a phenomenon and multiple analogies largely tell the same story. Rather than saying that there is only one correct meaning of a word or one correct model or analogy, there may be several, complementary versions, all emphasising different aspects and useful in different circumstances. This perspective characterises learning and conceptual change more as an issue of enrichment and coming to embrace a broader repertoire of ways of reasoning (Caravita & Haldén, 1994), than replacing a misconception with a purported correct view.

R1. To what degree are different representations of a phenomenon connected to each other and to the represented phenomenon by means of analogy?

Hesse (1966) suggests that analogy might be used to connect a model to the phenomenon it represents and Suppes (1962) argues that modelling involves connecting different models of a phenomenon by means of structural similarity, i.e. analogy. In the reanalysis, these issues were investigated primarily with regards to the interactive animation of an Otto engine used in article II. This animation provides examples where the connections between representations of a phenomenon and the same idealised model of it do not rely on isomorphism (or any other structural ‘morphism’). As a parallel to the view of Baier-Jones (2001) and Rivadulla (2006) on relating models of different phenomena, this does not mean that structural relations are not common or important in relating models of the same phenomenon; structural relations just do not tell the entire story. Vessey (1991) points out translations between representations that lack a structural fit as particularly challenging in coming to grasp a particular topic, which here may apply to interpretation of the PV graph in relation to the structurally dissimilar stick-figure image. The other issue, the relationship between a representation and the represented phenomenon, is very intriguing indeed and can be approached within all three perspectives of the theoretical framework. By virtue of what is a scientific model connected to the aspect of the world it represents? How is a word or other symbol connected to its referent?

In contrast to Hacking (1983), I would argue that approaches from the philosophy of language, such as the causal approach to connecting concepts to referents in the world (Kripke, 1980; Putnam, 1973) or causal descriptivism (Kroon, 1985) may give insight into the issue. Similarly, ideas from semiotics, particularly Peirce’s (1985) notions of icons and symbols, signs that do not represent by virtue of resemblance, may provide a broader perspective on the phenomenon-model connection than a strict structural perspective.

R2. How do students’ self-generated analogies relate to scientific modelling in education?

First, again, the richness in the analogies generated by the university students in article III and by 8-year-old Lisa in article IV is remarkable. These participants have clearly embraced the idea that there is not one correct or perfect analogy for the encountered phenomena, an important insight also with regards to the practice of scientific modelling. Second, the analogies do not appear in a vacuum or are due to a sudden recognition of structural similarities, but are related to the participants’ experiences in life and from previous teaching. All analogies in article III correspond to a microscopic approach to thermal phenomena, several of which relate to sanctioned scientific models, such as billiard ball or dog flea models, and I am still bewildered as to how Lisa came to associate the mixing marbles to the dropped groceries spread out on the ground, reminiscing of the messy room disorder analogy for entropy. Further, it is my deep belief that the exercise of generating your own analogies, preferably in a collaborative setting, takes away some of the anxiety in characterising a model completely correctly and emphasises the creative, provisional and pragmatic character of scientific modelling: a powerful take-home message with regards to the nature of science. Still, the students should not be left to themselves in the exploration of their analogies; they should be encouraged to evaluate their analogies and advised not to pursue the development of a particular analogy if it is leading too far away from sanctioned science.
7.2. Theoretical and educational implications of the dissertation

As emphasised by Larsson (1998), research should be assessed against the degree to which it has given a theoretical contribution to its field in terms of coming to see things in new ways. If any such contribution has been achieved with this dissertation, it would be due to: 1) application of novel research designs in a science education setting; and, 2) investigation of how the three perspectives of the theoretical framework relate to one another and may inform science education.

Regarding the first factor, article I adopts a semantic perspective and cognitive linguistics theories in the investigation of the meanings of ‘entropy’, a term that is primarily used among specialists of natural science, information theory, arts, etc. Cognitive linguistics, including conceptual metaphor theory, is an emerging framework for science education research, but still underexplored, particularly with regards to learning and developmental issues. The key contribution of article II is the finding of the divergent views on scientific models among experts of a phenomenon in practical vs. theoretical traditions of knowledge. Even though it has been known for long that people of different trades may view a phenomenon from different perspectives, the views on idealised scientific models among vocational students and teachers have not been recognised before and the potential scepticism of idealisations among colleagues is well worth considering for science teachers. The use of self-generated analogies, adopted in article III, has received increasing recognition in science teaching, even though teacher-generated analogies still dominate the scene. What is particularly novel with the article is the approach of asking the participants to come up with analogies in collaboration, and thereby taking advantage of the scaffolding provided by peer communication. The finding in article IV that some first-graders are capable of coming up with their own analogies for rather abstract scientific concepts is arguably the most remarkable result of the included studies. This was made possible by aiming for authenticity and providing scaffolding in the form of interaction with the investigated artefacts, the natural phenomena and making representations through drawing. As mentioned, it further lends support to the view that we should encourage children to generate their own analogies and recognise and build on any spontaneous analogies they come up with in primary teaching (May, et al., 2006). In these studies, we made more use of scaffolding than what is customary in typical psychology laboratory research, but I would encourage even more scaffolding, if applying our ideas in regular teaching or in future studies of cognition in the wild.

When it comes to the second factor, there is no ambition to move forward the positions in philosophy, semantics or cognitive psychology, per se, but to give an example of how they may be applied in the field of science education. In my view, recognition of polysemy and that there may be several scientific models or analogies for a phenomenon offer attractive alternatives to the view of conceptual change as replacement of a misconception for a supposedly correct scientific concept. This has important implications for science education research, as well as for the teaching practice: We should offer our students new ways of seeing things, and encourage them to connect to their prior knowledge, rather than just telling them that they have been wrong. Wiser and Amin (2001) have shown how recognition of polysemy can be applied in science teaching, contributing to a more refined view of the concept of heat and of the nature of science, and article II may be seen in the light of moving away from identification of deficiencies; there may be many perfectly valid ways of seeing a phenomenon, rather than one single expert understanding. Finally, articles III and IV have a more direct connection to the science teaching practice, as they were a kind of design studies or possible approaches to consider in science teaching. Asking students to generate their own analogies for phenomena may be a powerful tool towards making the students ‘talk science’ (Lemke, 1990), assume ownership for their learning (Enghag & Niedderer, 2008) and develop insight into the creative, pluralistic aspects of science.
8. References


Clausius, R. (1867). *The mechanical theory of heat, with its applications to the steam-engine and to the physical properties of bodies*. London, UK: John van Voorst.


Lee, A. (2010). When the article is the dissertation: pedagogies for a PhD by publication. In C. Aitchison, B. Kamler & A. Lee (Eds.), *Publishing pedagogies for the doctorate and beyond* (pp. 12-29). Abingdon, UK: Routledge.


Studies in Science and Technology Education
ISSN 1652-5051

44. Anna Lundberg (2011) Proportionalitetsbegreppet i den svenska gymnasie-
(Licentiate thesis) Linköping University
45. Sanela Mehanovic (2011) The potential and challenges of the use of dynamic software
in upper secondary Mathematics. Students’ and teachers’ work with integrals in
Linköping University
47. Veronica Flodin (2011) Epistemisk drift - genbegreppets variationer i några av
forskningens och undervisningens texter i biologi. ISBN 978-91-9795-161-6
(Licentiate thesis) Stockholm University
thesis) Karlstad University
49. Mats Lundström (2011) Decision-making in health issues: Teenagers’ use of science
University
attityd till naturvetenskap. ISBN: 978-91-7519-988-7 (Doctoral Dissertation)
Linköping University
51. Pernilla Granlind Enochson (2012) Om organisation och funktion av människo-
kroppens organsystem – analys av elevsvar från Sverige och Sydafrika.
52. Mari Stadig Degerman (2012) Att hantera cellmetabolismens komplexitet –
(Doctoral Dissertation) Linköping University
Linköping University
- Students’ beliefs, knowledge, and motivation. ISBN 978-91-7459-398-3 (Doctoral
Dissertation) Umeå University
Umeå University
och teknik i skola och samhälle – Innehåll, medierna och utbildningens funktion.
57. Thomas Lundblad (2012) Simulerad verklighet i gymnasieskolans fysik: en
designstudie om en augmented reality simulering med socio-naturvetenskapligt
Stockholm University