DEPARTMENT OF PHYSICS
UNIVERSITY OF JYVÄSKYLÄ
RESEARCH REPORT No. 7/2007

# PUPILS' MENTAL MODELS OF A PULLEY IN BALANCE, AND HOW THE MODELS ARE CHANGED BY SUCCESSIVE PULLEY DEMONSTRATIONS

BY
OLAVI HAKKARAINEN

Academic Dissertation
for the Degree of
Doctor of Philosophy

*To be presented, by permission of the*
*Faculty of Mathematics and Science*
*of the University of Jyväskylä,*
*for public examination in Auditorium FYS-1 of the*
*University of Jyväskylä on November 9, 2007*
*at 12 o'clock noon*

Jyväskylä, Finland
November 2007

ISBN 978-951-39-3185-8
URN:ISBN:978-951-39-3185-8

Publisher University of Jyväskylä, Department of Physics

Supervisor Professor emerita Maija Ahtee
University of Jyväskylä
Jyväskylä, Finland

Professor Jukka Maalampi
University of Jyväskylä
Jyväskylä, Finland

Reviewers Dosent Heikki Saari
University of Joensuu
Joensuu, Finland

Dosent Ismo Koponen
University of Helsinki
Helsinki, Finland

Opponent Professor Jouni Viiri
University of Jyväskylä
Jyväskylä, Finland

Olavi Hakkarainen

**Pupils' mental models of a pulley in balance, and how the models are changed by successive pulley demonstrations**

KEYWORDS: teaching of physics, mental models, cognitive conflicts, conceptual change, learning intervention, weight, variation theory.

## Abstract

This study explores changes in the Finnish comprehensive school pupils' (age eleven to fifteen) conceptions of hanging objects in balance under successive demonstrations. The study aims to address the understanding of a young pupil's conceptual change and to capture the factors of the instructional processes which foster or hinder the process of conceptual change. The aim of the study is to promote science teaching in comprehensive schools.

The theoretical framework of the study is based on the constructivistic view of knowledge where teacher's central task is to know pupils' ideas and conclusions, and also to guide and to help pupils to formulate interpretations based on the physics context.

The theoretical part of the study considers historical conceptions of weight, the conceptions pupils have about the concept of weight, the descriptions of different theories about conceptual change, and the learning methods applicable for changing pupils' conceptions towards the scientific concept of weight.

In the empirical part of this dissertation phenomenographic method has been applied to reveal the many different ways pupils experience the same phenomenon. First, pupils' conceptions about a pulley in balance were investigated with a paper-and-pencil test. Pupils' explanation categories were understood as Vosniadou's mental models. It turned out that only about 2% of the fifth graders and about 10% of the ninth graders seem to grasp the scientific Motion model that leads to the concept of gravity.

A teaching episode, including pre-activities and a learning intervention consisting of demonstrations showing the pulley in balance in three different positions, was planned to change pupils' conceptions. Also, the order of the demonstrations was varied in different groups of pupils. According to the pre-activity, in which pupils compared manually the weights of two objects, only 5% of the fifth graders and 50% of the ninth graders used scientific (pressure) model. After the learning intervention about 40% of the fifth graders and 45% of the ninth graders seem to use scientific model. In the ninth grade differences in reaching the scientific Motion model were large, varying from 27% to 55% depending on the order of the demonstrations. Therefore, it is of a great importance in which order the demonstrations are performed.

About six months after these tests a delayed study was carried out. For some of the fifth and ninth grade pupils who passed the teaching episode the same paper-and-pencil test used earlier was repeated. The understanding that

some of the fifth graders seemed to have formed from the weight concept after the teacher-independent learning intervention had almost totally vanished in the delayed study (9%). In the ninth grade about 45% of the pupils achieved general understanding about the pulley in balance by transferring the scientific explanation from the pulley demonstrations to the paper-and-pencil test.

The success of the learning intervention in the case of the fifth and ninth grade pupils can be explained with Marton's theory of variation because the use of successive demonstration is based on simultaneous variation. A particular way of experiencing something is a set of related critical aspects focused at the same time. Teachers can help pupils' learning by using appropriate variations in the successive demonstrations so that the pupils will recognize the cognitive conflict present in their earlier thinking.

Olavi Hakkarainen

**Oppilaiden tasapainossa olevaan väkipyörään liitämät mentaalimallit ja kuinka ne muuttuvat peräkkäisten väkipyörällä tehtävien demonstraatioiden vaikutuksesta**

HAKUSANAT: fysiikan opetus, mentaalimallit, käsitteellinen ristiriita, käsitteellinen muutos, oppimisinterventio, paino, variaatioteoria.

## Tiivistelmä

Tutkimus käsittelee peruskoulun 5.–9. luokan oppilaiden käsityksien muuttumista, kun heille esitetään peräkkäisiä demonstraatioita herkkäliikkeisellä väkipyörällä, johon on ripustettu yhtä painavat mutta erikokoiset kappaleet. Tarkoituksena on saada selville, kuinka oppilaiden käsitykset muuttuvat opetuksen aikana ja mitkä seikat edistävät ja estävät käsitteellistä muutosta. Perimmäisenä tavoitteena on kehittää peruskoulun luonnontieteiden opetusta.

Tämän tutkimuksen teoreettinen perusta on konstruktivistinen näkemys, jonka mukaan opettajan keskeinen tehtävä on tuntea oppilaiden ennakkokäsitykset. Opettajan on myös hallittava opetettava aihe syvällisesti, koska hänen on autettava oppilaita tekemään omia fysiikkaan perustuvia tulkintoja.

Väitöskirjan teoreettinen osa sisältää yhteenvedon siitä, millaisia käsityksiä painosta ja painovoimasta on ollut historian kuluessa, millaisia käsityksiä oppilailla on tällä hetkellä, kuvauksia käsitteellisen muutoksen teorioista ja opetusmenetelmistä, joilla käytännössä käsitteellinen muutos voidaan saavuttaa.

Kokeellisessa osassa sovellettiin fenomenograafista tutkimusmenetelmää kartoitettaessa, miten oppilaat kokivat tarkasteltavan ilmiön. Esitutkimuksessa kerättiin ja analysoitiin paperitestillä oppilaan käsitykset kahden erikokoisten mutta samanpainoisten kappaleiden painoista toisiinsa verrattuna sekä käsin punnittaessa että väkipyörään ripustettuina. Oppilaiden selitykset oli ymmärrettävissä Vosniadoun mentaalimallien avulla. Osoittautui, että vain 2 % viidesluokkalaisista ja noin 10 % yhdeksäsluokkalaisista selitti kappaleiden asennon väkipyörässä "Liike" -mallin mukaan, jota voidaan pitää tieteellisenä mallina, koska sen taustalla on käsitys painovoimasta.

Oppilaiden käsityksien muuttamiseksi suunniteltiin opetusepisodi, joka sisälsi valmisteluosion ja kolme peräkkäistä demonstraatiota sisältävän opetusintervention muuttamalla väkipyörään ripustettujen kappaleiden asentoja toisiinsa nähden. Myös demonstraatioiden järjestystä muutettiin kolmessa oppilasryhmässä. Valmisteluosiossa oppilaat vertasivat käsin punniten kahden kappaleen painoa. Tällöin ainoastaan 5 % viidesluokkalaisista ja 50 % yhdeksäsluokkalaisista käytti tieteellisesti perusteltavissa olevaa paineeseen pohjautuvaa mallia. Opetusintervention jälkeen havaittiin noin 40 % viidesluokkalaisista ja 45 % yhdeksäsluokkalaisista käyttävän kappaleiden

liikkumiseen (tai oikeammin liikkumattomuuteen) pohjautuvaa mallia. Yhdeksännellä luokalla erot ryhmien välillä vaihtelivat 27 prosentista 55 prosenttiin riippuen demonstraatioiden järjestyksestä. Sen vuoksi opettajan tulee myös miettiä, missä järjestyksessä demonstraatiot kannattaa näyttää.

Noin kuusi kuukautta opetusintervention jälkeen osalle oppilaille suoritettiin viivästetty tutkimus. Heille esitettiin sama paperitesti, jota käytettiin esitutkimuksessa. Opetusintervention tulos, että viidesluokkalaiset olisivat siirtyneet käyttämään tieteen mallia, näytti lähes tyystin kadonneen (9 %); sen sijaan yhdeksäsluokkalaiset käyttivät opetusinterventiossa opittua tieteen selitystä myös paperitestissä (noin 45 %).

Martonin variaatioteoria selittää opetusintervention menestyksen, sillä peräkkäisissä demonstraatioissa varioitiin ilmiön (yhtä painavat kappaleet) havaitsemisen kannalta kriittistä piirrettä (tasapaino säilyy). Jotta oppilas kokee ilmiön tietyllä tavalla, hänen tulee kohdata samanaikaisesti useita vertailukelpoisia ilmiöön liittyviä näkökulmia. Opettaja voi auttaa tätä prosessia valitsemalla ja varioimalla perättäisiä demonstraatioita niin, että oppilas havaitsee ristiriidan aikaisemmassa ajattelussaan.

## Preface

Research for Science education is a multidisciplinary field with many different aspects. Many research problems in this field require creative processing because no simple methods are available for solving them. This has been the most challenging and rewarding aspect of making this thesis.

The doctoral work presented in this thesis was carried out between 2000 and 2006 in the area of Jyväskylä and Helsinki. The work was carried out mostly along with teaching duties in the secondary school in Tikkakoski; the time was quite intensive, because simultaneously I gathered data to my research. In the time of Tikkakoski it was also possible to me to be a temporary full-time research student in the Finnish Graduate School. This time was of great importance because it made it possible for to me to classify the data I had earlier gathered. I am very grateful for this to the Finnish Graduate School.

As early as 2003 I got a possibility to introduce my work at the 4th ESERA conference in Noordwijkerhout. I am grateful for these days to Department of Physics in University of Jyväskylä.

I wrote the first drafts for the articles in Tikkakoski. Articles were then polished and rewritten many times in Espoo for the publication. Finally the overview of articles was also done in Espoo. All these experiences during the research work have been invaluable.

I warmly thank both of my supervisors, professor emerita Maija Ahtee and professor Jukka Maalampi. Maija Ahtee has patiently, through her wide knowledge, supported me during my doctoral work. She has always been able to give her time for discussion concerning the problems with my work. Jukka Maalampi has given a lot of advice in practical questions concerning my dissertation. I thank also Professor Jouni Viiri for his hard work as a pre-examiner of my dissertation.

I want to thank my principal Mika Risku for his positive attitude towards my research and Tuula Rantatupa, Tuuli Murtorinne for their assistance with the final polishing of the language. I am also grateful to Chris Johnston, for providing a native English speaker's perspective. I also thank my colleagues Leena Hirvonen, Säde Kiesi, Pia Lindroos, Pauli Nikula, Markku Pasanen, and Jarkko Helin for pleasant and productive collaboration in teaching physics in the secondary school in Tikkakoski.

I want to express gratitude to my Mother for all the support and encouragement. I would like to thank my wife Seija for her understanding and support and patient love. Many thanks to my daughters Sanna and Outi and grandchildren Mathias and Oliver for putting all these things in right perspective.

Looking back on these years which are summarized here the doctoral work has been a very positive experience which has taught me much more than I initially expected. I hope that this work can provide new ideas and questions for future studies.

Espoo, Sebtember 2007
Olavi Hakkarainen

## List of original publications and author contributions

This thesis is based on the articles referred in the text by their Roman numerals. Some unpublished results are also presented.

I Hakkarainen, Olavi and Ahtee, Maija (2005). Pupils' mental models of a pulley in balance. *Journal of Baltic Science Education*, 2005, 2(8), 26-34.

II Hakkarainen, Olavi. Mental models in manual weight comparisons between two objects of different size. *Themes in Education*, Vol. 6, No 2, 151-167.

III Ahtee, Maija and Hakkarainen, Olavi (2005). Importance of the order of demonstrations in changing pupils' conceptions. *NorDiNa Nordic Studies in Science Education*, 1, 31- 42.

IV Hakkarainen, Olavi and Ahtee, Maija. The durability of conceptual change in learning the concept of weight in the case of a pulley in balance. *International Journal of Science and Mathematics Education*, 5, 3, 461-483.

The author of this thesis has performed all data gathering, numerical and statistical operations, drawings and tables in all of these four publications, of course in close association with the supervisor exchanging of views in planning. Classification to categories I have done with the supervisor Maija Ahtee. I also have written the first drafts of the above manuscripts and then refining them with the supervisor for the publication. I suppose that my share is about 80% of all we have done.

**CONTENT**

# THEORETICAL PART

## 1. Introduction

### 1.1 Background for the study

I have worked for almost 30 years as a teacher of mathematics, physics and chemistry in lower secondary school grades 7-9. When teaching I have noticed that students meet with difficulties in understanding the concept of force. In the lessons pupils do experimental tasks in pairs or, if there are not enough experimental tools, in groups of four pupils. When teaching theory I used a lot of demonstrations to illustrate the concepts, therefore I have traced purposeful laboratory practical and demonstrations (e.g. Hakkarainen 1995; Asunta and Hakkarainen 1996) during many years to illustrate the concept of force (Hakkarainen 2000).

The National Board of Education creates and decides about the core curriculum, which is the basis for the local curricula (1994). In the lower secondary school the concept of force is studied in the seventh, in the eighth or in the ninth grade depending on the local curricula. In my school teaching the concept force and mass took place in the eighth grade. The pupils had possibility to meet the concept of force also earlier under environmental and natural studies in pre-school education and in earlier grades in the comprehensive school depending on the local primary school curriculum.

It is well known that pupils have different concepts of force (see for example Viiri 1995, Driver 1983, Gunstone and White 1981). Numerous studies have been done and are continuing to be done. There are many key ideas about the force concept that students have to assimilate in order to understand the concept properly. These are e.g. force as the cause of acceleration, forces like gravity as an action-at-a-distance and contact forces, force as an interaction between bodies, balanced and unbalanced forces, etc. Most of the studies have concentrated on considering students' ideas about force acting on objects in motion (Tao and Gunstone 1999; Mildenhall and Williams 2001; Ioannides and Vosniadou 2002) and also about forces acting on objects at rest (Minstrel 1982). Many of these possibilities are seen through studying the pulley.

In the history of science the pulley was an instrument by which Atwood in the18$^{th}$ century made gravity and accelerative movement visible (Greenslade 1985). Atwood also measured the exact numerical value of the acceleration of the gravity (Atwood's machine). The pulley shows the effect of gravity on the accelerating body but it also demonstrates the effect of the gravitational force when body is at equilibrium. It is well known that pupils use alternative conceptions as to weight and force, and also to explain the function of a pulley. In this work, the aim is to find a way how to help pupils to see how gravitational force works in the pulley system.

Johnson-Laird's theory of mental models (chapter 3.4) gives information how the human mind is constructed and how it works. The theory provides general information about human thinking. Mental models constructed in Vosniadou's theory are an extension of Johnson-Laird's work, and it gives me the possibility to interpret conceptual change in pupils' thinking with the function of pulley.

## 1.2 Overview of the Dissertation

It is well known that pupils arrive at the classroom with preconceptions of the topic to be taught. These preconceptions can in many cases be beneficial to the learning process but sometimes also detrimental. This is the starting point to my study. As a physics teacher I search for the methods, how I get the pupils change their wrong conceptions towards the scientific explanation. This dissertation is an effort of that kind.

The dissertation consists of a theoretical and an empirical part. In the theoretical part I introduce the theories, the concepts, the methods, and the instruments I used in my study. In Chapter 2.1 I treat the concepts of mass and weight and their development in the history of physics. Pupils' conceptions about mass and weight are considered in Chapter 2.2. The most important research and teaching instrument is the pulley and in Chapter 2.3.

Chapter 3 tackles Conceptual Change. According to the constructivist view of learning a pupil is seen as an active builder of his/her knowledge. In order for learning to take place it is necessary for the student to become aware of her/his own knowledge. Learning requires not only enrichment of knowledge and integration of new information, but also reorganisation of existing knowledge called usually a conceptual change.

In the last few decades many theories have been constructed about conceptual change (Posner et al. 1982; Carey 1985; diSessa 1993; Chi, Slotta and de Leeuw 1994). The meaning of theories is two-sided. First if we have a useable theory, the theory guides our thinking of the subject we are studying. However, we must be aware that we are critical when using the theory. We must be not blind to its weak points and should be ready to improve it and continue reconstructing the theory. Second, with the creating process of theory we get more knowledge, when we compare the experiments with the theory and reconstruct the theory suitable to our observations.

During the last two decades cognitive psychology and developmental psychology have given us tools to study the human mind. One of the most influential theories that has been formulated in cognitive psychology is Johnson-Laird's 1983 theory of mental models (Johnson-Laird 1983). The theory seeks to provide a general explanation of human thought. At its core there is the assertion that people represent the world they are interacting with through mental models. Mental models in Vosniadou's theory resemble those to Johnson-Laird's theory. Mental models here are analogical representations that preserve the structure of the thing they present. By observing how pupils perceive physical phenomena, it can be determined which of these models the pupils are currently using and the presumptions and constraints in the learning process can be identified. Chapter 3 contains also a description of these two models.

In this research I have studied what kind of conceptions pupils (age 11-16) have about weight in the context of pullies, and how these conceptions can be changed. There are two main strategies in teaching to promote conceptual change, namely the use of cognitive conflicts and that of analogies. These can be used in different ways in teaching. Some of these methods are dealt with in chapter 4. Marton's variation theory is also introduced there.

In the beginning of the empirical part in the Chapter 5 I introduce, for the first time, the exact research questions. Then there follows a description of the design of the research with the corresponding Learning study, and the methodology which consists of subjects, data collection and data analysis. The publications are introduced in Chapter 6. The original publications are included in the Appendix. Duplication of the original publications is avoided, but, however, discussion and conclusion demands some presentations from the published material.

Results and discussion are presented in Chapter 7. In this dissertation it is studied how pupils' predictions and explanations develop into mental models. There are also the results of how, by using cognitive conflicts, the teacher can guide and promote pupils' discernment of variation, and thus also contribute to learning. The results and discussion in the dissertation differs slightly from those presented in the original articles in the Appendices, because the knowledge and the interpretative skills of the researcher have developed and changed during the research. Finally the mechanism of cognitive conflict in changing pupils' conceptions is proposed.

Chapter 8 introduces the statistical method used in this study to estimate the reliability of the research. Then experimental validity is evaluated. The mental models of this study are then compared with other models from the theories of conceptual change.

In the beginning of Chapter 9 there are introduced pupils' conceptions of weight in the pulley context. There is also a description how pupils improve their ability to resolve a cognitive conflict. In the Chapter 10 I present the conditions where learning takes place according to the variation theory. Then I suggest how the Space of learning could be designed by a teacher in instruction. Finally, I present the conditions under which pupils may discern the critical aspects of a phenomenon they are learning more effectively.

### 1.3 The aim of this study and the main research question

Pupils' conceptions about force have been studied abundantly, in Finland e.g. by Viiri (1995) and Savinainen (2004). It has been found that Aristotelian physics thinking still flourishes among students, and that the interference structure that generates the medieval explanations of motion appears to have striking similarities with the intuitive inference structure of students who provide intuitive explanations of motion. Impetus, or force, was considered by medieval scholars and, similarly, by many present-day students as a property of the moving object. The students must now learn out of their intuitive wrong conceptions and learn in the new Newtonian force concept and the Newtonian three laws in about ten years. There are many difficulties: to distinguish between mass, weight and force, to distinguish between weight, contact force and action at a distance, to distinguish between velocity and acceleration and to distinguish between action and reaction. The students have to learn, to observe and to use these concepts. These concepts must also be observed and used in different situations. Learning and teaching of all this is not an easy task. The purpose of this study is find methods to help this process.

Learning is seen as a constructivistic process, in which the pupil personally constructs knowledge with the help of the teacher. Learning occurs as a conceptual change or series of changes, which take place in pupil's mind at the moment when the pupil sees three pulley demonstrations. The main question in my study is:

*How do pupils' ideas about weight change and develop during the successive demonstrations?*

# 2. Mass, weight and pulley

In this chapter I will begin by presenting a brief historical overview of the development of the concepts of mass and weight. My presentation is mainly based on Jammer's books on historical development of the concepts of mass and force (Jammer 1957, 1961). In the history of science the pulley was an instrument which made Newton's second law visible through Atwood's machine. It serves the same purpose to the pupils. The function of the pulley helps pupils to perceive the force concept as a gravitational phenomenon. After this brief survey of history I shall deal with pupils' conceptions of force and mass. In the last section there is a short introduction of the working of the pulley.

## 2.1. Mass and weight in history

### 2.1.1 Concept of mass

The starting point of students' difficulties of learning the concept of mass is that mass is given in three different meanings: *inertial mass, gravitational mass* and *quantity of matter* (Kurki-Suonio 2005). Pupils first come across with the concept of mass in the meaning of quantity of matter. The mass of a body refers to the amount of matter the body consists of. However, *quantity of matter* is measured by the gravitational interaction by using weighing. The inertial mass is learned later in the laws of Newton. When the ideas of quantity of matter and weight dominate the understanding of mass the whole of the concept is difficult to assimilate. These difficulties are evident also in the history of the mass concept (Kurki-Suonio 2005, Jammer 1961).

Jammer (1961, 7 - 29) gives a historical view of the development of the concept of mass. The word mass as used in physics is derived from Latin *massa,* but philological studies show that the word mass is originated either from the Greek *maza* (barley-cake or ordinary bread) or possibly from the Hebrew *mazza* (unleavened bread).

The rise of commerce and the exchange of goods already in prehistoric times called for ways to measure the quantity of goods, such as grain and metals, because use of counting was impracticable. This led to the idea of the quantity of matter, the historical predecessor of the concept of mass. Two methods were available and both were applied in commerce: the determination of weight and the determination of volume. Weighing with balance scale was a general way to compare amounts of goods from Antiquity to Modern Days (Galili 2001). Just in the 18$^{th}$ century Euler stated explicitly that the matter (mass) of a body is not measured by its volume. In Antiquity there were not measures of the quantity of matter.

In the early Middle Ages the assimilation of Platonic with Judeo-Christian philosophy proved to be of great importance for the further development of the concept of mass. The new impulse came from theological speculations. For the concept of matter and mass, the following three theological topics were important: creation, death, and transubstantiation, corresponding to the problems in natural philosophy of production, annihilation, and transmutation of matter. These problems were related to the principle of conservation of mass. In the 13th century the theologian Aegidius Romanus while considering the Eucharist, suggested that in addition to weight and volume there is a third measure of matter that is independent of determinations of volume or weight. This new measure of matter differed from the earlier quantities in the sense that it had no dimension.

The concept of quantity of matter played an important role in the discovery of inertial mass. According to Jammer (1961, 48 - 58), Buridan in the 14th century used it in the theory of impetus. His idea was the proportionality between impetus and quantity of matter. The heavy objects contain more matter and therefore also more impetus than the lighter ones of the same size. Hence, the heavy body moves, due to its greater impetus, further than the light object under the action of same motive force (same velocity). In other words, in the impetus theory the quantity of matter in the body of a physical object determines the resistance the object exerts against the moving force. Resistance here is still considered as a real force, and it is surprising to see how close to the concept of inertial mass the impetus theory goes.

Galileo was not able to explain why the velocity of free fall is independent of the weight of the falling object, a result he got through thought experiments (Atkinson and Peijnenburg 2004) and through his experiments with metallic spheres on inclined planes. There are a few passages from him in which he suggests the idea of inertial mass, but he never formulated a distinct description of it.

Kepler was the first to articulate the notion of inertial mass. At first he realized that the traditional notation of the circular and uniform motion of celestial bodies had to be given up. He discovered that the assumption of elliptical orbits for planetary motion agrees perfectly well with the observational data. Therefore he tried to find a dynamical explanation of the planetary motion.

In early writings the influence of Neoplatonic thinking in Kepler's reasoning is evident: *"Every celestial sphere, because of its materiality has a natural inability to move from place to place, a natural inertia or rest whereby it remains in every place where it is set by itself"*. Later the process of motion appears as two opposing factors. In planetary motion, he claims: *"the transporting power of the sun and the impotence of the planet or its material inertia strive against each other"*. Inertia according to Kepler is not only the inability of matter to transport itself from place to place, but also an active unwillingness to stop motion from without. This unwillingness or resistance stands in direct proportion to the quantity of matter. *"Inertia or opposition to motion is characteristic of matter; it is stronger, the greater the quantity of matter in a given volume"* (Jammer 1961, 49 - 58).

*Newton's conception of mass*: In Principia 1687, Newton defined the concepts of *mass, momentum* and, *inertial force.* According to Jammer (1961, 59 - 74), he based his definitions on Kepler's conception, and on important conclusions from impact experiments and the dynamic rotational motion, due to Huygens. In Definition 1 he defined the concept of mass as follows: *"The quantity of matter is the measure of the same, arising from its density and bulk conjointly".* There "the same" means equality sign, and "bulk and density conjointly" refer to multiplication. The term "bulk" is used instead of volume; besides the concept of mass as well as density had no unit. Definition 2 concerns momentum: *"The quantity of motion is the measure of the same, arising from the velocity and quantity of matter conjointly."* Definition 3 describes the inertial force of matter: *"a power of resisting, by which every body, as much as in it lies, continues in its present state, whether it is of rest, or of moving uniformly forwards in a right line."* However, in the three laws of motion, which Newton introduces after these definitions, the concept of mass does not appear explicitly.

Newton's definitions of mass and of inertial force became the subject of numerous comments and even of severe criticism. Jammer (1961, 70) gives two explanations to Newton's definition 1 of mass. According to his first explanation Newton simply repeated Kepler's idea. In the second explanation Newton made a difference between the conceptions such as quantity of matter and inertia on one hand and gravity on the other hand. Newton considered such qualities that are found in experiments like extension, impenetrability, and mobility as universal properties of all bodies. The same held for inertia and quantity of matter but not for gravity, which decreases with the distance of the body from the centre of the Earth.

Newton showed that gravity or weight (*the passive gravitational mass*) is proportional to quantity of matter (the inertial mass of body) in series of pendulum experiments. These experiments justify the use of the balance for the determination of mass. When it was further observed that pendulum clock operates in a different manner in a different place in the Earth, Newton seems to have concluded that the quantities of matter or mass and weight are different concepts. More than fifty years later, John Bernoulli (1742) was the first to clearly state that "mass" times the acceleration of free fall is the weight of the body.

To sum it up, Newtonian mechanics distinguishes three different kinds of mass: (1) *inertial mass,* which by Newton's second law of motion is definable through its reaction to a force independent of mass; (2) *"passive gravitational mass"*, the gravitational force is proportional to the mass of the attracted body; (3) *"active gravitational mass"* from the principle of "action and reaction" follows that the force is also proportional to the mass of the central body, recently (Jammer 1961, 122-123).

*Later development of the concept of mass*: Since Lavoisier had (1789) shown that the principle of conservation of mass holds for chemical reactions, it was clear that it was actually mass, not weight that was conserved in physical and chemical phenomena. Later on atomic masses became essential in chemical knowledge. However, as Jammer (1961, 86) emphasizes the formal definition of the concept of mass was not given in the eighteenth century. Usually definitions were synonymous with "quantity of matter" without supplying any other operational interpretations. The only exception was due to Leonard Euler. In his book *Mechanica* (1760) Euler stated that the conception of mass characteristic of an individual physical body was determined as the ratio of force to acceleration.

In the middle of the nineteenth century the increasing influence of the positivistic point of view to natural science led to an intensive criticism of the concept of force. It was claimed that new kinematics possessed logical and methodological priority over dynamics. Therefore the reform of dynamics to kinematical conceptions was regarded, e.g. by Saint-Vienant and Mach, as an important task for theoretical mechanics.

Saint-Vienant (1851) took as his starting-point the principle of conservation of linear momentum in the collision of two bodies. According to him two bodies have equal masses if their velocities after impact are equal. At the end of the century Andrade claimed that the definition of mass has to be given by the formula $m_1 : m_2 = |\Delta v_1| : |\Delta v_2|$.

Mach (1883) considered two bodies A and B having masses $m_A$ and $m_B$ interacting (magnetically, electrically, gravitationally; it doesn't matter how) with each other, but being unaffected by all the other particles in the universe. He also presupposed five statements that included three experimental propositions and two definitions like "*Moving force is the product of the mass-value of a body into the acceleration induced in that body*". After interaction the ratio of the two bodies $m_{A/B} = m_A/m_B$ is inversely proportional to the negative ratio of their accelerations, $a_A$ and $a_B$: $m_{A/B} = - a_B/a_A$. The minus sign comes because the accelerations are oppositely directed. If the mass of either body is set equal to the standard unit mass ($m_B$ = 1 kg), then Mach's operational definition of mass, $m_A = - a_B/a_A$ is obtained (Hecht 2006).

The procedure, in which the definition of mass was produced, has been criticized. Attempts to formalize Newton's mechanics by an explicit definition of mass have never been very successful. The logical structure of Newtonian mechanics seems to defy all attempts to complete such an analysis that involve explicit definitions of the fundamental terms. Therefore, opposite for instance to geometry, in mechanics semantic rules or correlations with experience are considered and defined through operational measurements, and then joined with the theory itself. Mach's definition of mass followed this principle. Mach did not say what "mass" really is but gave a rather advanced definition of the concept relating to the quantitative determination certain operational procedures. His definition is actually equivalent to Newton's third law (Jammer 1961, 120-121).

Savinainen (2004) has considered the development of the concept of mass based mostly also on Jammer (1999, 1997). He finishes his review with Jammer's statement about Mach's second Definition (Jammer 1999, 221):"The product of the mass and the acceleration induced in that body is called the moving force". Mach's presentation is problematic because his definitions of mass and force refer to an inertial reference frame but he does not consider whether the assumption of such a reference frame presupposes the concept of force and hence leads to a vicious circle; knowledge of forces is based on some theory about masses and vice verse (Jammer 1999, 240). All the same, nowadays the tendency is to simply accept the basic veracity of Mach's approach (Hecht 2006).

### 2.1.2 Concept of weight

During it's long historical development the word force has been given many different meanings, most of them considering force as a property of physical objects. Initially, force was associated with motion, because there was no distinction between uniform motion and accelerated motion as both of them were thought to be explained by a force exerted on the body.

*Weight in ancient thought*: According to Galili (2001) the history of weight concept in science started from the notions of heaviness (weight) and lightness (levity). First Greek philosophers interpreted the origin of these qualities. Atomists explained the heaviness by a centrifugal effect within cosmic vortexes, and the lightness as resulting from the flowing out of small atoms. Two theoretical conceptions prevailed in Greek science: Plato's thinking and Aristotelian thinking.

Plato interpreted weight as a tendency of bodies towards their kin. Aristotle introduced weight in his cosmology. Weight proclaimed a tendency of objects to restore the violated order in which fundamental elements were spatially organized along a line from the centre of the universe. He manifested that the permanent seeking of appropriate state of rest is the reason for the natural motion of any object. For example, a stone that is composed of element "soil", which is "heavy", has a tendency to move to the physical position of the element, which is the centre of the Earth. Smoke, mainly composed by air, which is "light", tends to move up to the universe, the natural position of the "air" element. These two kinds of motion are called "natural motion" because they are caused by forces which are internal properties of the objects (Ioannides and Vosniadou 2002; Sequira and Leite 1991).

To explain for any motion opposing the natural tendency "violent motion" of the objects (e.g. heavy objects moving upwards), Aristotle used the idea of an external "mover", which would be required to both overcome the natural tendency of the object to fall, and act as long as the object was in motion. Thus, the velocity of an object in natural motion would depend on the weight of the object while the velocity on an object in violent motion would depend on the "mover". The weight of object is seen as the efficient cause of such motion. Although weight was distinguished from force, weight may interact with it, producing the natural motion of the object and its velocity. In violent unnatural motion, weight resisted the moving force (Galili 2001). An external force in "violent motion" can be exerted on an object only by a living agent in direct contact with it or indirectly through some connection such as rope. Non-living things are barriers that stop or guide motion, but they do not exert forces. In Aristotelian theory in "natural motion" (free fall) an increase in speed can be achieved by an increase of weight as the object gets closer to its natural place to the centre of the Universe (Halloun and Hestenes 1985).

Weight in the science of the Middle Ages preserved the interpretation of weight as a property of the body and not as a force. In Aristotelian physics the speed of natural falling was directly proportional to weight. It was discovered in the Middle Ages that objects accelerate while falling and that therefore the original concepts had to be modified. Then weight was split into two components: the natural (habitual) *still weight* (or pondus) which remained unchanged and the *actual gravity (accidental)*, reflecting the apparent rise in the speed of falling. With time, impetus replaced the actual gravity and served Buridan to describe the accelerating nature of falling. In the Copernican picture of the world the Earth lost its central position of the Universe, and the ancient Platonic idea of "a tendency of bodies towards their kin", was revived to justify the natural downward pull, this time towards every heaven by body (Galili 2001; Sequira and Leite 1991).

*Weight in Galileo's thought*: Galileo followed the Platonic thoughts, but in 1608 he suggested a way to measure the difference between "dead weight", weight at rest, and that in motion. Galileo eventually reached a conception akin to that of Archimedes. Altogether Galileo's contribution was important, since weight no longer was regarded as related to speed, and the proportionality between the strength of the downward pull and amount of matter in the object was affirmed. However, the terms used by Galileo, all carried the idea that heaviness was measured by weighing. In summary, in Galileo's opinion weight was a characteristic of a single object, a one-way downward pull (Galili 2001), and it was his idea of the existence of a vacuum that led him to demonstrate that different objects have equal falling times when the fall occurs in a vacuum, and almost similar falling times when the fall occurs in the air (Sequira and Leite 1991).

*Newton's conception of weight*: After Galileo the cause of gravity was first related to the motion of heavenly bodies (Galili 2001). Jammer (1957, 81-93) suggests that Kepler had already understood the universal character of attraction, usually is attributed solely to Newton. However, Kepler still thought along peripatetic lines assuming that force is proportional to the velocity of moving object. Newton's law of Universal Gravitation introduced the gravitational force, an attractive central force between the heavenly bodies. This gravitational force was generalized to the force of gravity, and identified on the Earth as weight. This meant that weight is characterized by a pair of bodies, and it is no more a universal quantity.

Newton defined weight operationally: *"weight is always known by the quantity of an equal and contrary force just sufficient to hinder the descent of body"*. There was no doubt about whether his gravitational and operational definitions always provide the same weight. To Newton the equation "weighing result = gravitational force" was valid. Newton operated also with particular *relative quantities*, where he used the pair of opposites as true-apparent, and mathematical-vulgar. He applied them to reflect undesirable errors of measurements due to human limitations. So the notion of *true-weight* was reserved for the gravitational force, and *apparent-weight,* was joined by Newton to represent the results of weight in the presence of defective factors, which might deceive laymen. The nowadays-used "apparent weight" does not coincide with that in Newton's perception (Galili 2001).

*Weight after Newton*: The twentieth century brought along a change in understanding the concept of mass. The concept of absolute space-time was reconsidered and it was eventually replaced by the relativistic view of Einstein. The role of the observer was introduced in a completely new sense. Einstein's Principle of Equivalence prevents one from making distinction between gravitational and inertial forces by simple weighing. Also the subsequent progress, in the areas of macrophysics as well as microphysics (astrophysics, cosmology, elementary particles, atoms, molecules, etc.) does not require the weight concept. In everyday life and in introductory physics courses weight has remained as a relevant and useful concept, albeit difficult to learn.

In the following the definitions of the terms gravity, gravitational force and weight are given to prevent possible confusion.

Gravitational force is a force that attracts objects towards each other according to Newton's gravitational law.
The force of gravity on an object is caused by the mass of the Earth.
The weight of an object is the measurement of the force of gravity on that object.

## 2.2 Pupils' conceptions of mass, and weight

Children's alternative conceptions in physics have been studied for three decades in a wide range of physics domains. The studies are documented, e.g. in the bibliographies of McDermott and Redish (1999), Pfundt and Duit (2001) and Duit (2006). Novak originated in the University of Cornell a seminar series on students' misconceptions (e.g. Novak 1987, Novak 1993). Some studies on physics education are summarized in International Handbook of Science Education by Fraser and Tobin (1998), and about physics teaching and learning by Arons (1997). I have listed here some books and articles in which students' misconceptions in domain of physics are documented (e.s. Driver, Guesne and Tiberghien 1985; Osborne and Freyberg 1985; Driver, Squires, Rushworth and Wood-Robinson 1994; Treagust, Duit and Fraser 1996; Schnotz, Vosniadou and Carretero 1999; Magnani, and Nersessian, 2002; Sinatra and Pintrich 2003).

Pupils' misconceptions of motion and weight have received much interest in studies concerning students both in lower and upper secondary school and at the university (e.g. Enderstein and Spango 1996; Galili and Bar 1992; Watts and Zylbersztajn 1981; Viennot 1979), but studies about students' understanding of the development mass and force concepts have been done only by a few science education researchers (e.g. Chi, Slotta, and de Leeuw 1994; Bliss, and Ogborn 1994; Vosniadou 1994; Stavy 1990). Weight is a special kind of force, namely gravitational force. Therefore, if one examines children's understanding of the conservation of matter at the quantitative level, one also has to study children's understanding of force. Piaget and his colleagues concentrated particularly on studying the development of children's cognitive stages, especially with conservation tasks (Piaget 1972; Piaget, and Inhelder 1974). Driver's first works contained research with primary school pupils but later studies were done also with students in upper secondary school and at the university (Driver, Guesne and Tiberghien 1985; Driver, Leach, Millar and Scott 1996). In Finland students' misconceptions of force have been studied at the Technical University by Viiri (1995) and in upper secondary school by Savinainen (2004).

Initially young children cannot tell the difference between mass and weight. A child acquires a simple concept of weight/mass at a very early stage of his development. By picking up, holding and dropping different objects the child experiences weight before he learns to talk (Bliss and Ogborn 1994). The concept of weight based on common sense is learned without any deliberate teaching for example through personal experience as pressing force for example on one's palm without any deliberate teaching. Gradually, with maturation, the conception of weight as quantity of matter develops (Bar et al. 1994).

According to Galili and Bar (1997), children have two different knowledge facets about weight. Younger children emphasize the *active* facets which manifest themselves as a sensation of weight (a pressing force; according to Vosniadou 1994 an internal force) and as the ability of an object to perform certain actions such as breaking other objects when placed or dropped upon them (Smith, Carey and Wiser 1985; Piaget 1972). The *passive* nature of gravity, exemplified by things such as resistance to vertical or lateral movement, the connection between weight and amount of substance, or the understanding of density, is recognized later (Watts 1982; Nose, Torosantucci and Vincentini 1988). Conceptual distinction between true and apparent weights is not held.

Palmer (2001) asked 112 students from grades 6 and 10 whether gravity acted upon a series of moving or non-moving objects in some everyday situations. One result was that majority of students considered that gravity did not act on the object which was buried in the soil, and the reasons were that gravity pulls things down to the ground but not below the ground, and/or gravity is associated with air but there is no air below the ground. There were also opinions that gravity pushed up on rising objects. This idea conflicts with some earlier studies (e.g. Berg and Brouwer 1991, Smith and Peacok 1992) where it was found that pupils believed there is an upwards force on a rising ball attributed to the force of the person or the force of the throw.

Children have misconceptions about mass/weight concerning the shape of an object. When the shape changes, then also the weight/mass of the object changes; the amount of matter depends on the form of a cup inside which the matter is (Piaget 1972; Piaget and Inhelder 1974). When children grow older they succeed in doing classic plasticine tasks; about six-to-seven-year-old children use in these tasks explanations of identity, reversibility, compensation, and additivity. They also start to apply the logical ability to deal with certain weight conservation tasks.

In weight conservation tasks which involve changes of state, children's responses are heterogeneous. The majority of pupils, who did not answer the melting of ice task correctly, believed that ice is heavier than water of the same amount or that water has no weight. Few of the younger ones believed that same quantity of water is heavier than ice. It was the same with the evaporation of acetone. The younger pupils in the sample tended to believe that gas has no weight and the older ones believed that a liquid is heavier than a gas of the same quantity. The development of conservation of weight in the processes of dissolving sugar in water and expansion of water by heat is very similar to the development of conservation in the melting of ice task (Stavy 1990). Galili and Bar (1997) made same conclusions.

It seems those pupils' abilities to understand the conservation of weight depends on the nature of transformations. First the changes in plastic deformations are understood, then the changes in the process of change of state from solid to liquid, and finally in the process of change from liquid to gas (Stavy 1990).

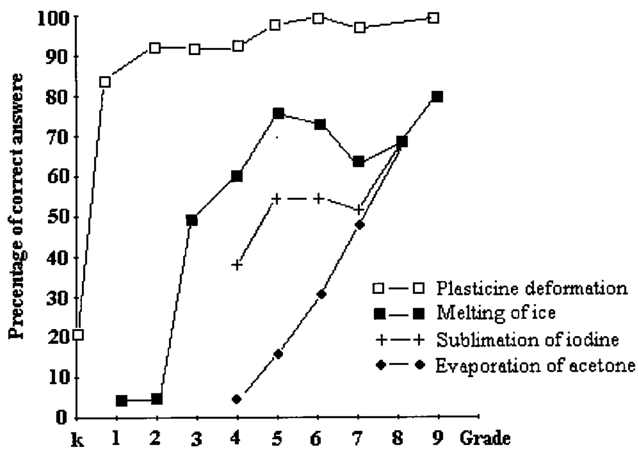


FIGURE 2.1. Percentage of correct answers to different tasks of conservation of weight. The answers are from nursery school to lower secondary school (from Stavy 1990).

There is also the effect of context on weight conservation (Stavy 1990). In two essentially identical tasks pupils' responses differed. Such as the evaporation of colourless and thus invisible acetone, and the evaporation of colourful visible iodine were compared. More pupils in the fourth to seventh grades conserved the weight of iodine than did with regards to acetone, see figure 2.1.

In summary according to Gallili and Bar (1997) it appeared that children's knowledge of weight is mainly characterized by two schemes: first weight is the pressing force featuring particular objects and sensed heaviness related to a muscular effort, and secondly weight is the amount of matter. The point here is that children do not invent the force of attraction of the Earth; the Newtonian reciprocal nature of weight. The weight of an object depends on its position. Usually objects lying near the ground are regarded as heavier than the ones higher up. If an object is moving upwards it contains extra force coming from the thrower. For many pupils in lower secondary school the more undifferentiated concept of heaviness is closer to their understanding than mass or weight (Hewson and Hewson 2003). Some objects have mass or weight (brick), while others do not (pin, hair).

*Development in understanding of weight and force concepts*: The development of pupils' force concept has been studied by Vosniadou and Ioannides (1998) and Vosniadou (1994). Ioannides and Vosniadou (2002) studied also the meaning of the word 'force'. Their results of developmental changes in the meaning of force showed that most children made use of a small number of relatively well-defined and internally consistent interpretations of force. The meaning of force varied significantly with age. The observed meanings can be grouped into two categories: those that appear to be based on everyday experience ("initial theories") and show no influences of Newtonian theory and those that have been influenced by instruction or by scientific theory ("synthetic theories" Vosniadou 1994).

In the study of Ioannides and Vosniadou (2002) there were two kinds of initial meanings of force, not influenced by instruction: younger children thought that force is an internal property of big and heavy objects related to their weight (internal force meaning), while older children thought that force is an acquired property of inanimate object that move as the result of an agent pushing or pulling them (acquired force meaning). The acquired force meaning was well established by the age of twelve. It appears that the children who adopted this meaning of force had differentiated force from weight.

There were also some combinations of the above-mentioned internal and acquired meanings that were called *hybrid* meanings showing a transition from the internal force to the acquired force (Ioannides and Vosniadou 2002). The *hybrid* meanings of force are created spontaneously not influenced by instruction. At about the age of fifteen, as a result of instruction, various synthetic meanings of force have been created through assimilation of the notions of gravitational force and the force push/pull. This causes the exiting explanatory framework increasing fragmentations.

When pupils' ideas of mass and weight, given in the previous section, are compared with the history of science ideas, it is amazing to see how near each other they are. For example the development of the weight concept from the Aristotelian to Galileo's conception is seen in pupils' thinking. The same trend is seen in the concept of mass and the conservation of mass. It is only Newtonian physics that is not seen in pupils' understanding without intensive school instruction (Ioannides and Vosniadou 2002; Galili 2001).

But there is not, however, an agreement among researchers whether the pupils' reasoning is along the same lines as in the history of science. For example according to Vosniadou and Ioannides (Vosniadou 1994; Vosniadou and Ioannides 1998) the naive framework theory of physics constrains the process of acquiring knowledge about the physical world in ways analogous to those that research programs and paradigms have been thought to constrain the development of scientific theories (Kuhn 1970). Chi, Slotta and deLeew (1994) noticed that the medieval researchers connected the most difficult preconceptions (like force, light, and fire) to material category in the same way as present-day pupils do. Today's students just like people in the Middle Ages use reasoning primitives like "More causes lead to more effect" when explaining their every-day experiences (see diSessa's phenomenological primitives chapter 3.2). Viiri (1995) underlines that even if students' ideas are coherent with theories in science history they differ as far as their philosophical background is concerned.

## 2. 3 Pulley in balance

In this work I have used a small revolving pulley (fig. 2. 3). It is familiar to pupils, because they have used it in school. The apparatus is composed of a light and mobile wheel, which can revolve freely around its fixed horizontal axis. Two bodies A and B are hanging on pulley wheel at the end of a thin, flexible, inelastic and light string. These assumptions mean that the mass of the string is small, the string does not slide, and that the pulley rotates around its axe without friction.

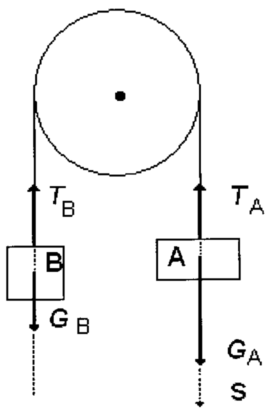


FIGURE 2. 3. A pulley.

The only forces acting in the bodies are $G_A$, $G_B$, and the tensions of the string $T_A$ and $T_B$. According to Newton's second law, the equations of motion for each body are as follows:

$$m_A a_A = G_A + T_A \quad \text{and} \quad m_B a_B = G_B + T_B,$$

where $m_A$ and $m_B$ are the masses, and $a_A$ and $a_B$ accelerations of the bodies, respectively. When we consider the motion of the string we get an expression in the acceleration of the string

$$a = \frac{m_A - m_B}{m_A + m_B} g \quad ,$$

where $g$ is gravitational acceleration. When $m_A = m_B$, then the bodies may stay at balance, in whatever position the bodies are placed, because in practice the slight difference in the masses due to the different lengths of the string on the opposite sites of the pulley is compensated by the friction in the axes of the pulley. If $m_A < m_B$ or $m_A > m_B$ we can see accelerated motion (Kurki-Suonio and Kurki-Suonio 1995, 143-144, and 323-326).

At present we know that the pupils bring to the classroom various everyday conceptions when they start learning science. There are a few studies done with a pulley to find these misconceptions. Gunstone and White (1981) used a bicycle wheel as a pulley. They found with that 27% of first year university physics students reasoned the lower hanging block of wood to be heavier than the hanging bucket of sand (of the same mass). Some of these students drew even inappropriate analogies to seesaws or beam balances. In Mohapatra's and Bhattacharyaa's (1989) pencil-and paper test concerning a pulley in balance, about 60% of the ninth graders stated that downward force on the lower hanging body was more than on the higher hanging body even if it was mentioned in the question that the two bodies were of equal mass. The researchers concluded that the pupils applied the image of a physical balance to the case of the pulley. According to McDermott, Scaffer and Somers (1994) many students had serious difficulties with the acceleration, the internal and external forces, and the role of the string. Most students understanding in pulley context did not develop spontaneously.

These pulley studies give information about pupils' misconceptions, but they do not tell us how to change these conceptions. In this study I will map out pupils' misconceptions in pulley context and then I try to change these meanings to find out how pupils' thinking works when they try to figure out a proper explanation and then to find factors that promote or hinder learning. My aim is also to find out if it is possible to guide pupils' thinking and learning towards a more dynamic and versatile way with a series of demonstrations.

## 3. On conceptual change theories

Science researchers are divided into two categories: first educational researchers found that students bring alternative conceptions about scientific phenomena to the learning experience. This is seen in studies on pupils' behaviour in teaching-learning context. These existing conceptions can serve as scaffolds or as barriers when learning new conceptions. Later cognitive and developmental researchers produced theories how children's understanding develops in science. Science education research and cognitive developmental research have greatly contributed to our understanding of learning. To change alternative conceptions is not enough simply to inform students of scientific conceptions. Learning requires not only enrichment of knowledge and integration of new information but also reorganisation of existing knowledge called as conceptual change. Finally there is a section on mental models.

### 3.1 Conceptual change in science educators' perspective

Initially conceptual change has been considered as developmental research based on the Piagetian views, but studies are strongly connected with instructional processes, aiming to reveal the nature of conceptual change (Schnotz et al. 1999). Turning to the history of science, researcher likened the change in individuals' conceptions to that experienced by a scientific community during a paradigm shift (Kuhn 1970).

Posner et al. (1982) believed that there were analogous patterns of conceptual change in learning. Assimilation is a weak form of the change in which existing concepts are used to deal with new phenomena. However, in accommodation the student has to replace or reorganize his central concepts. Inquiry and learning occur against the background of the learner's current concepts. Posner et al. identified four fundamental conditions that need to be fulfilled before conceptual change can happen. First, individuals must become dissatisfied with their existing conceptions. Second, individuals must find the new conceptions intelligible. Third, students must regard the new conceptions as plausible. Finally, students must find that the new conceptions are fruitful; the new conceptions must lead to new insights, and the accommodation of the new conception may follow. An intelligible conception is sensible if it is non-contradictory and its meaning is understood by the student; plausible means that in addition to the fact that the student knows what the conception means, he or she finds the conception believable; and, the conception is fruitful if it helps the learner solve other problems or suggests new research directions. Posner et al. (1982) insist that a plausible conception must be intelligible and a fruitful conception must be intelligible and plausible. Resulting conceptual changes may be permanent, temporary and or too tenuous to detect.

There are some other conditions for conceptual change. Posner et al. (1982) claimed that an individual's current concepts, his conceptual ecology, will influence the selection of new central concept. Strike and Posner (1982) expanded the conceptual ecology metaphor to include anomalies, analogies and metaphors, exemplars and images, past experiences, epistemological commitments, metaphysical beliefs, knowledge in other fields and competing concepts. Hewson and Thorley (1989) used term "status" to denote the extension of conceptual change. If the status of concept is high the pupils may adopt that concept easier than a low one. Hewson and Hewson (1992) asserted that a person's conceptual ecology plays a critical role in determining the status of the person's conception since it influences his or her judgement as to whether the conditions for conceptual change have been met.

One of the common instructional strategies to foster conceptual change is to confront students with discrepant events that contradict their existing conceptions. This is intended to invoke a conceptual conflict that induces students to reflect on their conceptions as they try to resolve the conflict. Dreyfus, Jungwirth and Eliovitch (1990) have pointed out that sometimes the contradictory information can be threatening to students who do not have enough knowledge for solving the conflict. Niaz (1995) noticed that some students ignored the conceptual conflict protecting their own conceptions, whereas Trumper (1997) observed that student reacted to conceptual conflicts in ways that did not lead to conceptual change: they failed to recognize the conflicts, they recognized the conflicts but avoided it by passively relying on others, and they resolved the conflict partly using alternative conceptions. Similar outcome has been obtained by Demastes, Good and Peeples (1996) who classified changes into four categories: a) cascade of changes (a) series changes triggered by one change by one conception, (b) wholesale changes (alternative conceptions discarded in favour of scientific conceptions, (c) incremental changes (alternative conceptions changing incrementally into scientific conceptions), and (d) dual constructions (students holding two logically incompatible conceptions).

Initially, accommodation in the conceptual change was considered as a prompt change, but there are also other opinions. Fensham, Gunstone and White (1994) think that conceptual change is rarely prompt, but more often "an accretion of information and instances that the learner uses to sort out contexts in which it is profitable to use one form of explanation or another". They called this "conceptual addition". Linder (1993) offers the idea of "conceptual fitting" and suggested that the learner has a range of conceptions which are invoked according to specific contexts. Maloney and Siegler (1993) extended this view and proposed the notion of conceptual competition and suggested that different competing conceptions coexist in the learner and that after a long period of learning one of these achieves dominance. Dykstra, Boyle and Monarch (1992) asserted that conceptual change is a progressive process of refinement of students' conceptions, and they propose taxonomy of conceptual change process, which consists of differentiation, class extension, and reconceptualization. Correspondingly, Niedderer and Goldberg (1994) described conceptual change as a process of changes from the learner's prior conceptions to some intermediate conceptions and then to scientific conceptions.

Tao and Gunstone (1999) noticed that, students oscillated like Newton (e.g. Steinberg, Brown and Clement 1990) between alternative and scientific conceptions during instruction. The students change from one context to another before a conceptual change independent of context appears, thus conceptual change is also dependent on context and unstable. Students need to reflect on instructions, which lead to change, and reconstruct their conceptions. They need to be able to perceive commonalities and accept the generality of scientific conceptions across contexts.

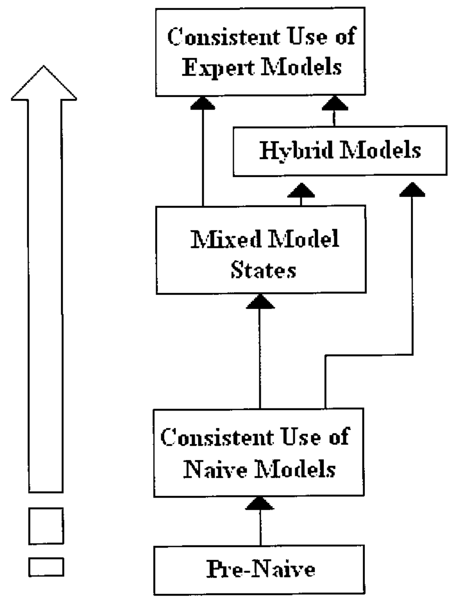


FIGURE 3.1. Stages of conceptual development (Bao and Redish, 2003).

Bao and Redish (2003) proposed empirical justified model for students' conceptual process of physics (see Figure 3.1). The model consists of five typical stages. First is the pre-naive stage. The students develop or possess a logically inconsistent model of the concept and they fail to provide any sound reasoning for the physical phenomenon. Many students can't provide any kind of reasoning. They fail to recognize the relevant variables in the context and do not find causal relations in their reasoning; they may explain the result even with the cause. During instruction, it is widely noticed that students can create new alternative conceptual understandings that are different from both their previous knowledge and scientific knowledge. These observations correspond to the third stage mixed model states. These reconstructions are made-up models in different context settings and they are only partly consistent with scientific models. Also, starting from the third stage, students may develop hybrid models as a result of instruction. Finally, students reach the consistent use expert's models.

## 3.2 Conceptual change in cognitive and developmental researchers' perspectives

There are two categories in conceptual change theories. First, many developmental psychologists describe concepts as deeply rooted in *naive theories* (e.g. Carey 1985 and Vosniadou 1994). *Theory* in these connections means that there is a coherent body of knowledge involving causal and explanatory understanding. Second, some researchers use *novice-expert paradigm* (like diSessa 1993 and Chi, Slotta and de Leeuw 1994) according to the fact that a body of knowledge involves both a collection of facts and some procedures to operate on them.

*Change in domain-specific theory*: Cognitive developmental research has far confirmed the hypothesis that the knowledge acquisition process starts at birth. Infants start rapidly to construct certain fundamental understandings of the physical and social world. Carey's work (1985) is seen as pioneering in the field of researching young children's conceptual change. It is based on the statements that conceptual change is essentially a process of theory change caused by increasing domain-specific knowledge. Children change their theory from few-theory-like conceptual structures, which lead to new theories, through restructuring. Generally, the learning process is situated in the expert-novice paradigm and conceptual change described as taking place through weak restructuring and strong restructuring. The weak restructuring is seen as regular learning as the change is in the relationships between concepts. Radical conceptual change in her account is the change in the ontology or essence of the concepts themselves. According to this perspective, initial conceptual structures undergo radical changes when the child develops and reaches adulthood.

*The change in ontological categories*: Chi and Slotta (1993); Chi, Slotta and de Leeuw (1994); Chi (2005) tie the meaning of concept to the ontological category to which a concept is assigned under novice-expert paradigm. Although this statement is simply a definition of the concept of ontology, what is more important is the character of these ontological categories. In the theory of Chi (1994) there were three primary ontological categories ("trees"): Matter, Processes and Mental States. There was also a hierarchy of subcategories embedded within each of these major categories or "trees". Processes category can be divided into Events, Procedures, or Constrains-Based Interactions; Matter category can be divided into Natural Kinds and Artefacts subcategory. Many physics concepts like heat, light and force belong to the Constraint-Based Interactions, which is a subcategory of Processes. Students tend to consider these concepts belonging to the wrong Matter category.

The strongest conceptual change in Chi's theory is the changes from one primary ontological category to another. The first science concepts in the classroom or elsewhere encountering on, that the learner classifies these concepts to the Matter category. If the concepts' actual ontology differs from the student's initial classification, the process of conceptual change may be very difficult. When the new knowledge demands radical changes in the student's thinking, the pupil must perhaps first to construct up the new ontology category Constrains-Based Interactions and after that the suitable subcategory. That is why the ontological changes occur first and after that the epistemological changes. When a new knowledge only enriches the structure of knowledge is easy for the student to join it in his old structure of knowledge compared to the case that requires changes in the structure.

*The change in framework theory*: Vosniadou (1994) presented a theory, in which conceptual ideas such as framework theories, specific theories and mental models are examples of the broadest theoretical constructions that children construct (see Figure 3.2). Students may construct mental models under the control of a naïve framework theory, which consists of certain specific fundamental assumptions. In addition, Vosniadou suggests that students have specific theories for describing correlations within groups of phenomena related to physical concepts. These have developed in a child subject into the constraints posed by the framework theory. Mental models, on the other hand, are created when the child faces a problem and solves it, and are more compact and concrete than specific theories.

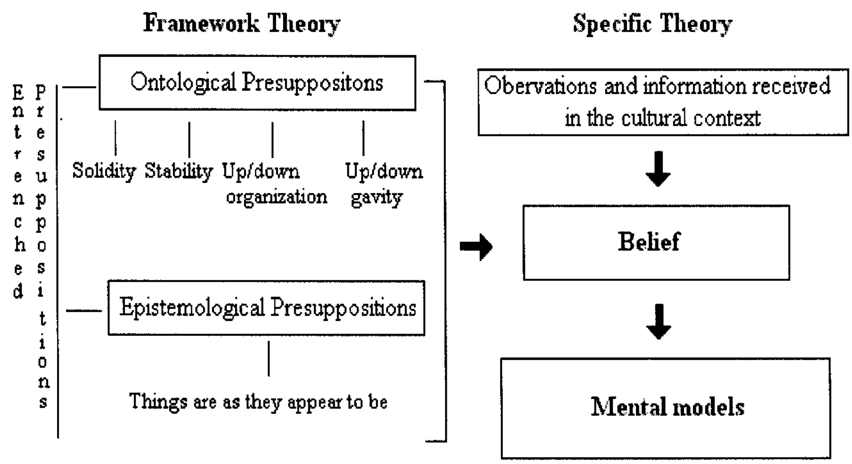


FIGURE 3.2. Children's hypothetical conceptual structure underlying initial mental models (Vosniadou 1994).

Vosniadou states that in the context of the present theoretical framework, the simplest form of conceptual change is the *enrichment* of an existing conceptual structure. *Enrichment* is the simple addition of new information to an existing theoretical framework through the mechanism of accretion. It is assumed that this is a relatively easy form of conceptual change. *Revision* is required when the information to be acquired is inconsistent with existing beliefs or presuppositions or with the relational structure of a theory. It is argued that the revision of a specific theory is easier than the revision of a framework theory. When the beliefs of a specific theory are constrained by a framework theory, conceptual change can be very difficult to achieve. The change in a framework theory is difficult because the presuppositions of the framework theory represent relatively coherent systems of explanation, based on everyday experience and tied to years of confirmation.

Developmental, cross-sectional research has shown that children's initial representations (*initial mental models*) of physical world undergo important conceptual changes with development and learning. Some of these changes cannot be directly related to the teaching of science whereas others can. Initial conceptual structures can change as a result of children's enriched observations in the cultural context, or of other kinds of cultural learning rather than as a result of specific science instruction. *Spontaneous changes* are this kind of changes. Other kinds of changes are products of science instruction. *Synthetic mental models* have been explained as representing attempts to synthesize the currently accepted *scientific* explanation with the aspects of pupil's own initial concept. Changes like these are *instructionally based* changes (Vosniadou and Ioannides, 1998).

There are also some combinations of internal and acquired meanings called *hybrid* meanings showing a transition from the internal force to the acquired force. The hybrid meanings of force are created spontaneously not influenced by instruction (Ioannides and Vosniadou 2002).

Vosniadou´s theory (1994) makes a distinction between *framework theories* and *specific theories*. In physics, the framework theory includes basic ontological and epistemological presuppositions that define the physical object (for example Spelke's constraints, Spelke 1991). Specific theories are embedded within infants' naive framework theory and are constrained by them. Therefore, teachers should take into account the order of instructional materials when they design the teaching. The teaching is effective only if it takes into account presuppositions in children's naive framework theory. If teaching is focused on misconceptions in specific theory it leads only to formation of new misconceptions because the learner tries to explain his misconception in the wrong explanatory framework, under the wrong ontological and epistemological presuppositions and it is possible that it causes new synthetic model. For example, some children possess the misconception that the Earth is a hollow sphere inside which people live on a flat ground. It is not enough for a teacher to claim that the Earth is not a hollow sphere because children do not understand the up/down concept on the surface of a sphere. Children need teaching about gravity and why the large ground seems flat.

*The change in constructions of causal explanations*: In diSessa's theory (1993) the knowledge acquisition process follows novice-expert paradigm. An expert's scientific knowledge is tied to physical laws and principles. Initial knowledge consists of pieces, which diSessa calls p-prims (phenomenological primitives or facets). Sometimes diSessa's p-prims are general in character and sometimes they refer to specific physical situations.

Bao and Redish (2002) used the term *reasoning primitive* to describe the abstract primitive understanding of diSessa's p-prims and the term *facet* to refer to the mapping of a *reasoning primitive* into a physical situation. In this system, misconceptions of scientific concepts are explained as being triggered by particular *reasoning primitive*, which is true in many situations, but in some physical context they can form misconceptions. For example, in the impetus theory misconceptions are traced to a class of facets, such as "force is a mover". These facets are associated with *reasoning primitives* "a continued cause is needed to maintain a continued effect", which itself is not a misconception. In gravity context, the *reasoning primitive* "more cause leads to more effect" may produce the *facet* "a big body falls faster than a light one", which is correct in many everyday situations, but not in physics. diSessa's reasoning primitives seem in many cases to be close to Spelke's five constrains of behaviour of physical objects.

According to diSessa, conceptual change occurs through the reorganization of p-prims or through an increase in internal coherence and in systematisation of the collections of p-prims that serve as explanations. The most important kind of conceptual change is described to be a change in the function of p-prims, which cease to be self-explanatory and become tied to more complex knowledge structures, such as physics law and principles. These pieces of knowledge are very near scientific concepts. An important difference between novices and experts is that novices, unlike the experts, are not aware of the hypothetical status of the presuppositions and beliefs which constrain the way they interpret new information.

*Socio-cultural view*: Ivarsson, Schoultz and Säljö, (2002) take a more radical position. According to them, a conceptual change results from changes in the way students use intellectual tools in various contexts, and the actual change occurs as the societal level. Scoultz, Säljo and Wyndhamn (2001) have shown that human reasoning is a tool dependent on nature. According to them when children's reasoning is supported by a cultural artefact (like the globe they seem to be familiar with highly sophisticated modes of reasoning. This is considered a rather strong argument for a sociocultural interpretation of mind.

### 3.3 Overview of the theories of conceptual change

The theories of the conceptual change (Carey 1985; Chi 1992; diSessa 1993; Vosniadou 1994) describe and analyse the relation between a learner's earlier thinking and new knowledge in order to find reasons for learning difficulties and acquired misconceptions. According to these theories difficulties in learning depend not only on the classroom atmosphere but also on the learner's earlier structure of knowledge, which often do not harmonize with new and studied knowledge. They have a focus on the changes in learners' thinking compared with the naive thinking and they are interested in both the quality of changes and degree of difficulties.

The theories of conceptual change distinguish two kinds of changes in thinking during a learning process. When new knowledge only enriches then old knowledge is easy for the student to change compared to knowledge that requires more radical thinking. The definition of the conceptual change of Vosniadou et al. (2001, 383) is determined only as a radical change. The characteristics of radical conceptual changes are both subject specific nature and the radical reorganisation of earlier knowledge. These conceptual change theories state that learning difficulties concerning scientific concepts do not only depend on the complexity and the abstraction of new knowledge but it can also be based on earlier knowledge and its character. The theories are not interested in learning process, but they have more focus on changes in the structure of knowledge. Learning is a long process during which initial structures based on learner's interpretations of everyday experiences are continuously enriched and restructured.

The knowledge acquisition seems to be a gradual process characterised by erroneous and inadequate conceptions, which according to Vosniadou (1994) are explained as attempts on the part of the learners to assimilate the scientific information to their initial conceptual structures without violating the presuppositions of the framework theory.

In an instructional context, prior knowledge is seen crucial in teaching pupils science concepts and prior knowledge may be a starting point for learning, but not only the cognitive factors affect conceptual changes. The critics are mainly levelled at the rational nature of these theories, because they neglect the social and affective aspects (Duit and Treagust 2003). Dykstra (1992) claims that group factors like a "town meeting" atmosphere can promote concept learning. There are also other factors, including motivation, social attitudes and instructional contexts which influence the conceptual change process (Pintrich, Marx and Boyle 1993). Social and affective perspective of conceptual change is dealing with motivational beliefs about learning and how the pupils see themselves as learners (Tyson et al. 1997). Moreover, pupils' self-efficacy is linked with their cognitive engagement. There seems to be a correlative relationship between metacognition and self-efficacy judgements. Metacognition, in turn, supports conceptual change making verbalisation and modelling easier for pupils (Pintrich 1999). According to Sinatra and Pintrich (2003) the learners are seen intentional if they pursuit understanding over and above the requirements of school tasks. In order to achieve this kind of learning, pupils must be purposeful and methodical and able to monitor and regulate their learning in a metacognitive manner. It is also assumed that intentional learning is something that develops and can be cultivated by instruction. Vosniadou (2003) remarks that conceptual change can happen without intentional learning but intentional learning may greatly facilitate conceptual change because many students who are not intentional learners nevertheless achieve minor reorganizations of prior knowledge in school settings. She also claims that particular attention should be paid to issues that have to do with metacognitive control and metaconceptual awareness.

Mayer (2002) has compared four views of conceptual change: misconception repair (Chi, Slotta and de Leeuw1994), knowledge-in-pieces (diSessa 1993), synthetic meaning (Vosniadou 1994), and sociocultural view (Ivarsson, Schoultz and Säljö 2002). Mayer concludes that the two most challenging ingredients for research on conceptual change are mechanisms embedded in detailed *theories* and *methodologies* that generate relevant data for testing the theories. He also points out that intervention studies offer an important component in any program of research on conceptual change because the ultimate challenge is to use the understanding of conceptual change to improve science education.

### 3.4 Mental models

Mental models are representations of reality that people create to understand specific phenomena. In this study I am looking for the ways how pupils understand the weight concept i.e. I try to map pupils' mental models related to the objects hanging in the pulley in equilibrium. The main aim is to find how to change the faulty models with the scientifically correct ones through teaching. Two books on mental models were published in 1983, one by Johnson-Laird, which could be called the theoretical approach to mental models, the second by Genter and Stevens, consisting of a set of papers that could be called the instructional approach (Greca and Moreira, 2000).

*Jonhson-Laird's theory of mental models*: Johnson-Laird (1983) proposes three types of mental representations:

1. *propositional representations*, which are strings of symbols that correspond to natural language,
2. *mental models*, which are structural analogies of the world,
3. *mental images*, which are the perceptual correlates of models from a particular point of view.

Human beings understand the world by constructing mental models in their minds. The role of mental model is to account for the individuals' reasoning both when they try to understand discourse and when they try to explain and predict the physical world behaviour. Propositions are linked with each other by a particular syntax, which can be verbally expressed and whose truth value depends on their interpretation according to a mental model. The semantic mental language maps propositional representations onto mental models. Mental images correspond to a view or visualization of the model from a given perspective, as a result either of perception or imagination; they represent the perceptible features of the corresponding real-world objects in individual human reasoning. Models may be basically propositional or may rely mainly on images, or both. Despite the fact that mental models are incomplete, unscientific, quite unstable, frequently deficient in a number of ways, and perhaps including contradictory, erroneous and unnecessary concepts, it is important to emphasize nevertheless that mental models are working models of situations and events in or of the world. It is possible through their mental manipulation to increase understanding, and explain phenomena and they may act to result predictions. (Sasse 1997; Creca and Moreira 2002; 2000; 1997).

Vosniadou (1994; 2002) used and developed mental models further. She argues that the ability to form mental models is a basic characteristic of the human cognitive systems and that the use of models by children is the foundation of more elaborate and intentional use of models by scientists. According to her the mental models constructed by children and laymen have predictive and explanatory power and they can be used as mediating mechanisms for revision of existing theories mechanisms, and for the revision the construction of new ones. Nevertheless, it is possible that some mental models, or parts of them, which have proven useful in the past, are stored as separate structures and retrieved from long-term memory when needed.

*Conceptual models in instruction*: Generally a *conceptual model* is a representation created by researchers, teachers, engineers. When the statements of the theory are related to a simplified and idealized physical system or phenomenon, the resulting description is a *physical model* or *scientific model*. The aim of these models is to help pupils to understand the teaching or states of affairs in the world. They are the objects of learning, and therefore external to the pupils. Physical models are complete representations with scientifically accepted knowledge, and therefore they are shared by a given community, and have their coherence with the scientific knowledge of that community. These external presentations can materialize as mathematical formulations, analogies, or as material artifacts.

In teaching it is usual to assume that pupils have acquired or constructed the conceptual model that has been presented to them. The situation is usually more complicated. Pupils create their own conceptual models, *working models*, which differ from those of *physical models*. Firstly, because pupils do not have the necessary knowledge of the situation in which they can interpret them as conceptual models. One is able to see the details, but the wholeness will be meaningless. Secondly, because pupils often do not understand that a conceptual model is a simplified and idealized representation of phenomena or situations, without being told the actual phenomenon or situation (Greca and Moreira 2000). It is also usual that pupils have already knowledge about the subject of instruction, then it is possible they connect the new information to the ideas they already have (*initial model*), and a new conceptual working model (called *hybrid model*, Vosniadou 1994) is formed.

*Mapping mental models*: In instruction, *modelling is understood as a facilitating process for the construction of adequate mental models*, which *will help to understand physical models* (see e.g. Saari and Viiri 2003). There are in literature a few people who map pupils' mental models for studying how pupils understand the surrounding world. There is, however, no general agreement about the way of doing the mapping; Vosniadou 1994; 2001; Ioannides and Vosniadou 2002; Coll and Treagust 2003; Schoultz, Säljö and Wyndhamn 2001.

According to Vosniadou an important characteristic of mental models is that they can be explored extensively in order to generate predictions and explanations. The method of getting data is very multi-faceted. Mental models were obtained by collecting data from children's answers to verbal questions with drawings and models made out of play-dough. One important aspect of the methodology is the use of generative questions. Generative questions cannot be answered on the basis of stored information but require a genuine solution to a new problem. It is assumed that in order to answer a generative question the pupils must create a mental representation or mental model and explore it in order to derive a relevant answer from it.

Finding mental models Schoultz, Säljö and Wyndhamn (2001) interviewed pupils in a different way than Vosniadou. They modified a usual interview situation by using a globe as a resource for thinking for the child in the interview situation. The outcome of study is radically different from that reported in earlier literature. All children immediately accepted the globe to be a model of the earth.

Coll and Treagust (2003) tested the understanding of metallic bonding in three academic levels: secondary school pupils (aged 12), undergraduate and postgraduate students. They used semi-structured interviews including the use of Interviews-About-Events focus cards depicting metallic properties and cards containing depictions of models from curriculum material.

In my study, I was interested in pupils' explanation models before and after every three successive demonstrations. I wanted to see how the original mental model changed. In my teaching I avoided giving instruction that might have affected pupils' own understanding. The instruction was given therefore with demonstrations. The demonstrations themselves were the instruction. The mapping process was fast, not like in instruction, slow and time-consuming modelling.

# 4. Learning and teaching physics

This study is dealing with the knowledge construction process of pupils aged 11-16. In order to produce knowledge construction the main aim of physics education and teaching is to get together the new scientific phenomena, the learning concepts and their interrelationships. The teacher plans experiments including the phenomenon and produces explaining crucial concepts. Finally the explanation is tested by a new experiment, which is closely related to primary experiment. The pupil's task is to do findings on the experiment and to create the explanations on the basis of their own ideas. According to the constructivist view of learning, the teacher's central task is to know the pupils' ideas and pupils own conclusions, and also to guide and to help pupils to do interpretations according to the physical context.

## 4.1 General instructional process

Instruction and its core conditions as general are discussed in my work very briefly. Havu (2000) has treated the relationships of the concepts included in teaching. These are described in figure 4.1. As a starting point she has taken Jackson's model (1968), in which teaching is divided into three phases: pre-interactive, interactive and post-interactive phase. The first two ones form the instructional phase and reflection takes place in the post-interactive phase.

Jackson uses term teaching as an activity that occurs during the interaction between the teacher and children. Kansanen (1999) called the broader understanding as the instructional process, consisting of planning, aims and reflection. He sees Jackson's three phases as a cyclic and interactive process between its components.

Shulman (1986) considers the instruction as a method and actions occur during interaction phase. Instruction is part of teaching, which includes pedagogical solutions for example organizing and managing the group, presenting explanations and descriptions and checking work; and interacting with the pupils through questions, probes and reactions (Shulman 1986, 17; 1999).

Learning is something that happens in pupils' mind, depending on various personal and contextual factors (Kansanen, 1999, 85). Niededder and Schecker (1992, 75) define the concept of learning in the field of science education as a change in the elements of the cognitive processes, which results from developmental processes of the cognitive system interacting with external situations. Learning is something internal and studying is external. Learning is seen as a result of the studying process, and it may occur at several levels e.g. at a behavioral, cognitive or social level. Thus it is very rarely possible to study learning; much more often we should study studying (Malinen 1997). For that reason it is possible that studying instruction, studying and learning are reciprocal and together named as the instructional process. The instructional process is a phase in which the students and the teacher are in interaction aiming at improving the pupil's skills (Havu 2000).

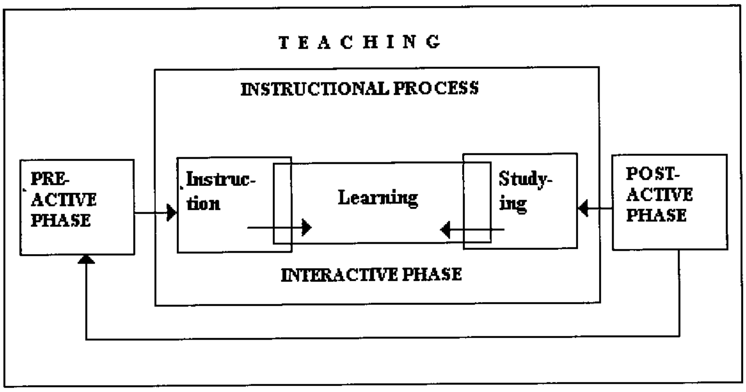


FIGURE 4.1. The understanding of the relationships between the concepts included in teaching (Havu 2000).

In this study I have not taken into account many aspects including the teaching-learning process. In the study my focus was mainly on the pupil's cognitive processes. The instruction came to my study via the demonstrations planning by the teacher, the studying consisted of the pupils seeing demonstrations in their own way and making their own interpretations without adding teachers' manipulation, and the learning consisted of finding out whether the interpretations were the same or changed during and after successive demonstrations. Particularly, I was interested in finding out how the pupils noticed cognitive conflicts and in the importance of variation in demonstrations.

## 4.2 Teaching approaches to promote conceptual change

Scott, Asoko and Driver (1991) have identified two main groupings of strategies in teaching to promote conceptual change. The first grouping is that of strategies which are based upon cognitive conflict and resolution of conflicting perspectives. Strategies like that which emphasise *cognitive conflict* and the resolution of that conflict by the learner may be seen to derive from a Piagetian view of learning, in which the learner's active part in reorganising their knowledge is central. The second grouping is that of strategies which build on existing ideas and extend them through e.g. *metaphor or analogy* to a new domain. It is *Variation theory* (Marton and Booth 1997; Marton and Tsui 2004) which gives guidance to teacher how to help students to learn by discerning the critical aspects in a phenomenon.

### 4.2.1 Cognitive conflicts

The strategy to foster conceptual change is to confront students with discrepant events (anomalous data) that contradict their existing conceptions. Kuhn (1972) noted the importance of anomalies in science, arguing that persistent anomalies play a key role in scientific revolutions. The use of anomalous data in the classroom has been guided by assumption that in many cases science students are like scientists (see for example Posner et al. 1982; Driver 1983; Dreyfus et al. 1990; Viennot 1979; 2001). There are four crucial elements in this assumption. First, like scientists the science students possess beliefs of how the physical world operates. Second, students notice if new data are incompatible with their prior beliefs. Third, they recognize that these anomalies pose a threat to their current theories. Finally, like scientists students and children will sometimes choose to adopt an alternative theory in response to data that are anomalous for their prior theory. The anomalous data produce cognitive conflict or cognitive dissonance, and students will resolve the cognitive conflict by bringing their personal conceptions closer in line with scientists' conceptions. In brief, an unexpected result or observation which does not fit with preconceptions creates cognitive conflict.

Teaching strategies that are based on a cognitive conflict and its resolution can be divided in two groups: *discrepant events* and *conflict between ideas.* Strategies using *discrepant events* involve situations where the student's existing ideas are made explicit to the student through an exposing discrepant event and then challenged in order to create a state of cognitive conflict through pupil's explanation and scientific explanation of that event. Finally the teacher encourages and guides the pupil to accommodate and to develop a new conceptual model consistent with the scientific view (Nussbaum and Novick 1982).

In the second strategy the aim is to pay attention to two types of physical reality. A pupil's cognitive structure relates to a certain physical reality, which is different than the actual physical reality. A conflict is then produced between these *two different ideas,* cognitive structures related to the same reality (Stavy and Berkovitz 1980, 679).

It has been developed several teaching approaches, which are suitable for students to resolve differences between ideas from a range area of other students, the teachers, and science text.

"Generative Learning Model of Teaching" (Cosrove and Osborne 1985) consists of four phases. In the first *Preliminary phase* the teacher needs to understand the scientist's view, the pupils' view and his own view about the learning subject. In *Focus phase* the learner is engaged in clarification of his own views of the concept within a real everyday situation. In the third *Challenge phase* learners debate their current views with each other and the teacher introduces the science view if necessary. The last *Application phase* gives opportunities for application of new ideas across a range of contexts. The researchers stressed that learners did not pay attention to alternative ideas until they could be rendered intelligible and plausible by experimentation, demonstration or reference to analogy.

"Ideational Confrontation" model (Champange, Gunstone and Klopfer 1985) pays attention to dialogue-based strategy. Model is specially designed to alter students' declarative knowledge. Each student develops an analysis that supports his predictions and presents it then to the class. Students attempt to convince each other of the validity of their ideas. Finally instructor demonstrates the physical situation and presents a theoretical explanation using science concepts, and further discussion allows students to compare their analyses with the scientific one.

Rowell and Dawson (1985) proposed a model in which resolution between student's prior ideas and new conceptions occurs after new conceptions have been introduced. The approach is based on the following premises. A prior theory is only replaced by a better theory and not discarded on the basis of contradictory evidence alone. The construction of a better theory need not involve an immediate confrontation with the knowledge that an individual spontaneously considers relevant. Therefore, the students' ideas are first established. The new theory is then linked to basic knowledge already available.

Tsai (2000); Tsai and Chang (2005) has proposed the use of ´conflict maps´ as a way of promoting both teaching and learning. The conflict map emphasizes the use of discrepant events to challenge students' alternative conceptions. According to Tsai students should resolve two conflicts during the process of conceptual change: one between the discrepant event and students' alternative conception, and the other between the alternative conception and the scientific explanation. Therefore, the conflict map of a phenomenon contains in addition to the student's alternative conception, the target scientific concept, a critical event or explanation and relevant scientific conceptions and other scientific supporting perceptions. The critical event tries to justify directly the conflict between students' alternative conception and the scientific conception whereas the discrepant event usually challenges students' alternative conception. Tsai (2000) stresses, that conceptual change can be achieved by using a suitable discrepant event and a critical event. To make the target conception intelligible also linguistic expressions, images, examples, or analogies should be used.

Cognitive conflict strategies do not always lead to conceptual change (e.g. Dreyfus, Jungwirth and Eliovitch 1990; Dekkers and Thijs 1998; Lee, Kwon, Par, Kim, Kwon and Park 2003). When students' ideas are confronted with contradictory information through instruction students may not at all recognize the conflict, or if a solution is proposed at a level which is beyond that of students it will remain meaningless to them and the effect of the conflict is lost, or that sometimes the contradictory information can even be threatening to students who do not have enough knowledge to solve the conflict. The cognitive conflict approach is not effective when students lack the foundation and tools to construct new, scientifically better ideas. The use of cognitive conflict as an instructional strategy fails also when the significance of the conflict is not apparent to the students (Vosniadou 1999; Limón 2001). However, if properly used the cognitive conflict approach creates stimulating and motivating learning events.

When teachers use anomalous data, they may fail as students who recognize the data have different ways to respond to it. Chinn and Brewer (1998) have listed eight possible ways how pupils deal with cognitive conflict. They suggest that science teachers should use the taxonomy to guide classroom discussion. Kang et al. (2004) found a significant correlation between the cognitive conflict induced by a discrepant event and the conceptual change. However, they warn that cognitive conflict is only one of the important factors to be considered in concept learning rather than a necessary prerequisite for it.

### 4.2.2 Analogies

The second group of teaching strategies builds on pupils' existing ideas. Analogies are particularly common in everyday speech and in science because of their power to compare one object or situation to another. Particularly, in many processes they transfer either details or relational information or both between each other (Gentner 1983; Curtis and Reigeluth 1984; Duit 1991; in finnish Saari 2000).

Glynn (1991) defined an analogy is as a process: it is a process of identifying similarities between different concepts. Then analogies are used often in scientific and technical contexts. Clement (1993) showed a basic model of analogical reasoning.

In the Teaching-with-Analogies model, an analogy is seen by identifying similarities between two concepts (Glynn et al. 1995). Ideas can be transferred from a familiar *analog* to an unfamiliar *target*. Both the analog concept and the target one have *features* or *attributes*. If the analog and the target have common or similar features, an analogy can be drawn between them. A systematic comparison between features is called a *mapping*.

In figure 4.2 appears that an abstract representation of an analogy consists of parts. The analog and target are examples of a *superordinate, higher order* or *cross-domain* concept.

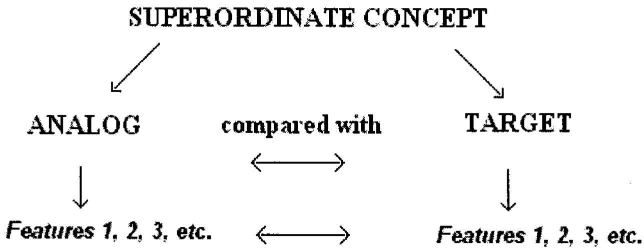


FIGURE 4.2. Abstract representation of an analogy, with its constituent parts. Analog features are mapped onto target features (Glynn et al. 1995).

No analog maps onto the target perfectly; if the target and analog were identical, they would be the same concept. Every analog breaks down at some point, and no two analogs are alike. Each analog has its corresponding and non-corresponding features; some analogs will be better for some purposes than others. For these reasons, teachers and authors should try to suggest several analogs to students.

Many researchers have studied analogies in different perspectives. According to Holyoak and Thagard (1995), scientific analogies have at least four distinguishable uses: discovery, development, evaluation and exposition. Discovery analogy contributes to the formation of a new hypothesis. Once a hypothesis has been formed, analogy may facilitate further theoretical or experimental development. Analogy can also serve to form arguments for or against a hypothesis' acceptance and then analogy can convey the new ideas to other people.

Brown (1993) and Harrison and Treagust (1993) considered the use of analogies beneficial for conceptual change in science learning. Duit (1991) also agrees that analogical reasoning can facilitate understanding and problem solving. Stavy and Tirosh (1993) suggest that salient external features of problems (similar structure, similar process, and figural similarity), as well as factors related to the solver (age and instruction) influence students' responses to the problems. The theoretical framework is not the most influential factor in determining students, responses to the problems.

Brown and Clement (1989) have developed an analogical teaching strategy which assumes that conceptual change can be encouraged by providing opportunities for students to build up qualitative-intuitive understandings of phenomena before mastering quantitative principles. Understandings are developed by forming analogy relations between a misunderstood '*target case*' and an '*anchoring example*', which wakes up intuitive knowledge held by the student. The use of '*bridging strategy*' has been proved to be successful in developing this relationship.

The bridging strategy (Brown and Clement 1989) consists of four steps:

- The students' misconceptions are made explicit by using a target question relating to the topic under consideration (e. g. forces acting on a book on a table. Majority of students consider the table as passive and believe that there is no upward force.)

- The instructor suggests then an analogous case, which is an anchoring example (e. g. a hand pushing down on a spring).

- The instructor asks students to make an explicit comparison between the anchor and target cases to establish the analogy relation.

- If the students do not accept the analogy, the instructor attempts to find a bridging analogy (or series of bridging analogies) between the target and the anchor (e. g. a book on the top of a spring, and so on...)

To overcome misconception about static force, frictional forces and Newton's third law, the use of bridging strategy has been seen as a successful teaching approach (Clement et al. 1989; Treagust et al., 1996; Savinainen et al., 2005).

Stavy (1991) reported a study which used students' intuitive perceptual knowledge. In this study (grades 5 and 6) the aim was to research how the matter is conserved in evaporation. She suggested that the use of an analogical relation between known and visible (gaseous iodine) and unknown (acetone invisible gas) can help students to adopt new, unknown information and discard or modify misconceptions. The intuitively understood, in perceptually supported tasks apparently served as an analogical example for students (see fig. 2.1).

The aim of Niedderer' model (1987) is not to replace students' theories by the scientific theory but allow students to arrive at conscious knowledge of both theory and simultaneously to learn scientific concepts and differences between everyday life thinking and scientific thinking. This strategy consists of six stages. The first *Preparation* stage is the preliminary teaching process. In this stage the pupils learn the tools which are needed for learning the main scientific concept. In the second *Initiation* stage the open-ended problem is posed. In the third *Performance* stage the students formulate questions or hypotheses; they plan and perform experiments, make observations, theoretical discussions and formulate findings. After that in *Discussion of findings* stage the class discusses the findings in small groups. In the fourth *Comparison with science* stage findings are compared with the same themes in historical or modern ideas (scientific knowledge). Differences are stated and possible reasons for those differences are discussed. At last in *Reflection* stage students are encouraged to look back on the process of *Performance* stage and to consider particular questions or difficulties, which have risen.

Wong (1993) considered that generative analogies are dynamic tools to facilitate understanding rather than representations of the correct and the static explanations or solution. He emphasizes the *generative* nature of analogies. Learning from analogies can be viewed as *generative* if analogical reasoning might occur in the absence of specific pre-existing schema. Prior knowledge is incomplete and poorly organized and conceptual growth emerges in these conditions from a continual refinement and synthesis of fragmented, incomplete knowledge. Typically in analogical reasoning a problem is defined by the researcher. When moving from teacher-directed to learner-directed analogical reasoning occurs, the learner assumes primary responsibility for creating analogies. In these situations, the learner must consider that all analogies are faulty in some respects and no analogy will completely and accurately represent the nature of the phenomenon. Learners must 1) evaluate their own analogies by identifying both similarities and differences between the two domains, and 2) construct a series of analogies in an effort to address the deficiencies of prior analogies.

Chiu and Lin 2005 designed their study to investigate how multiple analogies influence students learning of electronic circuits. The results demonstrated that using analogies not only promoted understanding of complex scientific concepts (electricity), but analogies also helped students overcome their misconceptions of these concepts.

However, there are some studies that conclude that findings in analogical reasoning are not especially promising because most students are unable to employ analogical reasoning to solve similar problems regarding different phenomena, and learners are not able to see the analogy (Gick and Holyoak 1983; Duit, Roth, Komorek and Wilbers 2001). Dagher (1994) reviewed several studies and comments on the role of analogies. She argues that although several studies claim conceptual change occurred, analogies act simply as references for initial explanations or conjectures rather than bringing fort a conceptual change. Chi (1992) argues that analogies are considered a way of assimilating new knowledge to an existing structure and thus that is not a conceptual change. A growing body of research shows that analogies may be powerful tools for guiding students from their pre-conceptual conceptions towards science concepts. But it has also become apparent that analogies may deeply mislead students' learning processes (Duit et al. 2001).

It is seen above that no ideal analogy exists, and each analogy has its limits. Therefore it is necessary to make use of diverse analogies for different purposes (Chiu and Lin 2005). Some represent the function of explanation, some point to new perspectives, and others allow for making more explicit targets. The more similar the representations are, the more easily the translations between representations occur. Spiro, Feltovich, Coulson and Anderson (1989) explained that additional analogies could promote conceptual understanding. They assigned following functions to multiple analogies: supplementation (with new analogy), correction (with new analogy), alteration (of an earlier analogy), enhancement (of an earlier analogy), magnification (or elaboration), perspective shift, competition, sequential collocation. They also believed that multiple analogies are all applied to the same stage of a process and all of the analogies in the set are partially correct. Dagher (1998) points out that what is the most important thing is the fact that the steps in selecting the set of multiple analogies are interlocking, because each new analogy is chosen to correct the negative aspects of the preceding analogies.

### 4.2.3 Variation theory

According to phenomenography a phenomenon can be experienced in a limited number of qualitatively different ways. The nature of phenomenography is basically descriptive and methodologically oriented. The theory of variation is an experiential theory that is derived from the characterization of learning in the educational context. Variation theory tries to answer the question: "How do different ways of experiencing something evolve?" Marton and Tsui (2004) consider variation corresponding to the critical aspects of the phenomenon i.e. the dimensions of variation within the framework of awareness.

*The structure of awareness*: Awareness is the totality of person's experiences of the world at each point in time (Marton and Tsui 2004). It changes dynamically all the time, and every situation is experienced against the background of previous experiences. A way of experiencing something is related to how a person's awareness is structured. It contains both a 'what' aspect, which corresponds to the object, and 'how' aspect, which is related to the act. Whatever an individual sees or experiences, he perceives a thematic whole, which is discernable from its surrounding context. All the time, certain aspects of the object come into focal awareness to constitute the theme or figure of particular understanding. In experiencing a phenomenon, certain aspects of the phenomenon are always brought up simultaneously in a person's focal awareness.

*Concepts of discernment*: An experienced variation is a necessary condition for discernment. Marton, and Booth (1997) pointed out that in order to bring a certain aspect of the phenomenon to a learner's focus so that it will be discerned the learner must experience variability. For according to Bowden and Marton (1998), discernment assumes experienced variation. This means that a learner is aware of the critical features so that s/he uses them either at different points in the time or s/he uses them simultaneously either at one specific time. Learning takes place through experiencing variation in the critical aspects of a phenomenon, which leads to a discernment of the critical aspects in the individual's focal awareness in a simultaneous manner. Marton and Tsui (2004) assert that we can experience simultaneously something that we can *discern*; we can only discern something that we experience to vary, and we can only experience *variation* if we have experienced different instances previously and are holding them in our awareness *simultaneously*. These three concepts of variation theory are closely related to each other. In classroom teaching it is very important to bring the critical features of the objects of learning into students' focal awareness, because arranging for student's learning implies arranging for developing students' ways of seeing or experiencing.

*Patterns of variations*: Starting point for the variation theory is that learning is based on the learner's dynamic structure of awareness, and it is related to *discernment, variation and simultaneity*. Marton, Runesson and Tsui (2004) noticed that relation to object of learning other necessary features of variation can be identified. First, in order to experience something pupils have to experience something else to compare it with (*contrast*). Secondly, pupils have to experience varying appearances in order to grasp the idea and separate it from the irrelevant features (*generalization*). Thirdly, in order to experience a certain aspect and in order to separate it from other aspects this aspect has to vary while other aspects remain invariant (*separation*). Fourthly, if there are several critical aspects that pupils have to take into consideration at the same time, they must all be experienced simultaneously (*fusion*).

*Wholeness, parts, and context*: The first thing pupils have to realize is that they have to pay attention to the behavior of the system as a whole, e.g. they have to notice that the whole in pulley system consists of bodies connected by an inelastic light string passing over a freely moving pulley instead of looking only at two separate bodies hanging on the string. Marton and Tsui (2004, p. 12) point out discerning the relation of parts within wholes and likewise discerning the whole from the context is an important aspect of discernment. It is equally important to discern the way the whole relates to the context. This is because the way the whole relates to the context shapes the discernment of the parts within the whole.

*Forming a conception and learning*: For students it is important what aspects of the situation they discern and focus on. Marton, Runesson and Tsui (2004, p.3-42) stress, that the teacher should define the objects of learning as clearly as possible, because learning is always the acquired knowledge from the objects. A particular way of experiencing a phenomenon can be understood in terms of either particular features of a person's awareness, or the dimensions of variation that are discerned and simultaneously focused on. The different ways in which different phenomena can be experienced by pupils reflect the differences in the structure and organization of awareness. The learner's dynamic structure of awareness is understood when a pupil changes the focal point of his attention. According to Pang (2003) when a pupil directs his focal attention to a certain aspect or aspects of a context or situation, a certain conception is formed. Learning is seen in phenomenography as a change in the learners' capability of experiencing a phenomenon in the world around them. It amounts to being able to discern certain aspects of the phenomenon, and to keeping them in focal awareness. According to Marton and Booth (1997) certain patterns of variation characterize certain ways of experiencing a phenomenon. To bring about a particular way of experiencing a particular phenomenon it is necessary to follow that very pattern of variation.

*Conceptual framework of Learning Study*: Here a *Learning study* refers to a *Lesson study* that is based on the Theory of variation. Theory of Variation postulates that learning is always directed at 'something', and it is related to how people make sense of it developing a certain skill, value or understanding in dealing with it. This is referred to as an 'object of learning' (Ling, Chik, and Pang 2006). According to Marton and Booth 1997, an object of learning could be thought of formed by two aspects: the content that is being learnt (e.g. Archimedes' principle), and the kind of understanding that the learner is able to develop as studying this specific content (e. g. being able to interpret and explain a concept with the use of an example, and being able to discern the object of learning in different situations or contexts (Marton and Tsui 2004, 4)).

A key feature of learning in the Theory of variation involves the idea that a person experiences a phenomenon in a new light; the phenomenon that previously was not on focus or which one took for granted is brought into one's focal awareness.

To experience a phenomenon in a particular way certain critical aspects must be discerned at the same time. For instance, to understand completely the behavior of Archimedes' principle one must be aware of the weight of a body immersed in water as compared to its weight when not immersed, and of the weight of the water displaced at same time in a simultaneous manner as a complete representation (Marton, and Booth 1997). Respectively, in order to understand the function of the pulley at balance the pupil has to realize that the both objects are influenced by equal forces as otherwise the bigger force would put the objects in accelerated motions. So a particular way of experiencing something is a set of related critical aspects focused at the same time.

There are studies concerning the framework of variation to analyze classroom teaching, and learning (e.g. Runesson 1999; Pang 2003; Ling, Chik, and Pang 2006). Runesson (1999) studied teachers' different ways of handling the same mathematical topic. The results of that study revealed that although teachers taught the same topic in a similar way, different teachers thematized certain aspects of the content and put other aspects into peripheral awareness. Implication of this study was that for learning to take place, critical aspects of the content and pupil's learning should be discerned simultaneously by teaching against a background of experienced variation of the aspects concerned.

Pang (2003) discovered that discerning any feature increased experiencing variation in a dimension, which corresponds to that feature. He was able to show that when simultaneous variation in critical aspects was approved, students reached a good understanding, twice as much as did students in other classes where simultaneous variations in the critical aspects of the object of learning were not given.

Ling, Chik, and Pang (2006) used the learning theory of variation advanced by Marton and Booth (1997) to improve teaching and learning. They also developed a Learning study, which is guided by a conceptual framework that is based on the Theory of variation. According to Ling et al. 2006 teachers have used in a Learning Study three types of variation:

1. Variation in students' ways of experiencing what is to be taught/learnt.
2. Variation in teachers' ways of dealing with the 'object of learning'.
3. The use of 'pattern of variation' as guiding principle of pedagogical design to enhance student's learning.

Ling, Chik, and Pang 2006 used analogies to guide pupils in Learning Study on the topic of 'the color of light'.

In this study I show first how phenomenography reveals the different ways how pupils experience the same phenomenon, and I give an example how the method is used. Then I describe the phenomenographic method to find pupils' conceptions in the pulley in balance, and thirdly I develop a teaching intervention using Variation theory together with cognitive conflict to help pupils to perceive weight as a force in the gravitation field. I am particularly interested in producing variations through demonstrations and with them to guide pupils' learning.

# EMPIRICAL PART

## 5 Research process

### 5.1 The research questions

There are many key ideas about the concept of force that students have to assimilate in order to understand the concept properly. These are, for example, force as the cause of acceleration, forces like gravity as an action-at-a-distance and contact force, force as an interaction between bodies, balanced and unbalanced forces and so on. Developing the concept of force takes time with many misconceptions (Galili and Bar 1992; Champagne, Klopfer and Anderson 1980; Minstrell 1982; Ioannides and Vosniadou 2002).

Many of the misconceptions common to students' in mechanics deal with weight and gravitation. In many cases these misconceptions are very difficult to separate from each other and from the scientific concept as they are so entangled and also, in many cases, so good for describing the world around us. There are few demonstrations which bring pupils misconceptions about weight or gravitation to light better those using the pulley. In pulley demonstrations pupils may argue that from the two objects the bigger, the larger, the lower, the higher, the upwards climbing or the downwards moving object is heavier than the other one. The point is that by using the pulley for demonstrations, one can recognize what kind of conceptions lie behind the arguments. There are some earlier studies in literature about using pulley in science education (Champagne et al. 1980; Gunstone and White 1981; Mohapatra and Bhattacharyya 1989). The novelty of the present study is its aim to find out by using the pulley what kind of mental models pupils have, how the models are changed, and how permanent are the changes obtained.

To promote discernment in their Learning Study, Ling, Chik, and Pang 2006 used analogies to guide pupils' thinking towards the scientific concept. In the present study I use successive conflicts to guide pupils' thinking from misconceptions to the scientific concept. The first research questions were:

1. *What are the pupils' mental models of weight when they see two bodies hanging in different positions in a pulley in balance?*

2. *How do the pupils change their mental models of weight from fifth grade to ninth grade?*

These questions were treated in the first study of this thesis, reported in "Pupils' mental models of a pulley in balance" (Hakkarainen, O. and Ahtee, M., 2005). Instead of using the term mental model in this connection I preferred the terms prediction category and explanation category (mental model). The answers to the question 1 are collected in table 7.3 in Chapter 7.1. The changes of pupils' mental models or explanation categories from the fifth grade to the ninth grade (question 2) are seen in the same table 7.3.

In the next study I started to investigate how these mental models, based on looking at the Figure in the questionnaire, differ from the mental models based on manual weight comparison:

3. *How do pupils estimate the weight of the standard mass and the equally heavy bag of steel wool when they weigh these objects in their hands?*

4. *What are pupils' reasons behind their conclusions and how do these change from the fifth grade to the ninth grade?*

I studied these in the second study "Mental models in manual weight comparisons between two objects of different size" (Hakkarainen, O., 2005). Answers to the questions 3 and 4 are presented in the table 7.4 in Chapter 7.1.

In the next study I tried to change these pupils' conceptions in pulley context. I planned a series of learning interventions based on conflicting pulley demonstrations. With the help of the demonstrations I tried to change pupils' opinions, and then to follow variations of the mental models:

5. *How do the pupils' statements about weight change when the order of the demonstrations is changed?*

6. *How do the pupils' argumentations change when the demonstration is changed?*

7. *What are the specific difficulties pupils encounter with a pulley in balance?*

8. *How does the order of pulley demonstrations effect the pupils' resulting concept development?*

Answers to these questions are presented in the article "Importance of the order of demonstrations in changing pupils' conceptions" (Ahtee, M. and Hakkarainen, O., 2005). The results for the questions 5, 6 and 8 are presented in Figures 7.5 and 7.6 in Chapter 7.2. Pupils' difficulties in interpreting demonstrations are treated in chapter 7.4.

Finally, the durability of mental models was studied:

9. *What are the fifth and/or ninth graders' conceptions after half a year's schooling* from *the pulley demonstration?*

    a. *How have the pupils' predictions changed?*
    b. *How have the pupils' argumentations changed?*

10. *Have the pupils made any desired progress or have they fallen back to their*

*preconceptions?*

Answers to these questions are analysed in the article "The durability of conceptual change in learning the concept of weight in the case of a pulley in balance" (Hakkarainen, O. and Ahtee, M., 2007). The results of questions are treated in Chapter 7.3 and they are shown in Figures 7.3 and 7.4.

In these studies, conceptual change is seen as a process that makes it possible for pupils to synthesize models in their minds (Vosniadou 1994). Conceptions are no more isolated because mental models help the pupils to explain groups of phenomena, visualized as similar ones (Greca and Moreira 2000). The pupils' perceptions are determined by analysing the answers they gave in the questionnaire form after seeing the objects or feeling the objects in their palms.

## 5.2 Research Design

The essential feature of experimental research is that the researcher controls and manipulates the conditions which determine the events in which he/she is interested. Cohen, Manion and Morrison 2000 present the following research designs for educational research: historical research, developmental research, surveys, case studies, ex post facto research, experiments, action research, accounts, role-playing, and interview. Most empirical studies in educational settings are quasi-experimental rather than experimental. The present study is seen as a quasi-experimental design. The most important difference between the quasi-experiment and the true experiment is that in the former case, the researcher undertakes his/her experiment with groups that have been constituted by means other than random selection.

| | | |
|---|---|---|
| **Experimental group** | X | $O_1$ |
| **Control group** | | $O_2$ |

FIGURE 5.1. The quasi-experimental design of this study (see Cohen, Manion and Morrison 2000, 214-215). The dashed line in the Figure indicates that the experimental and control group are not equated by random assignment.

In the Figure 5.1 **X** represents the exposure of a group to an experimental variable or event, the effects of which are to be measured and O refers to the process of observation or measurement. The dashed line separating the parallel rows in the diagram of the non-equivalent control group indicates that experimental and control groups have not been equated by randomisation. The equivalence of groups can be strengthened by matching, such key variables as socio-economic background of pupils, school size, organisation and teaching methods. Where matching is not possible, the researcher is advised to use samples from the same population or samples that are as alike as possible (Cohen, Manion and Morrison 2000).

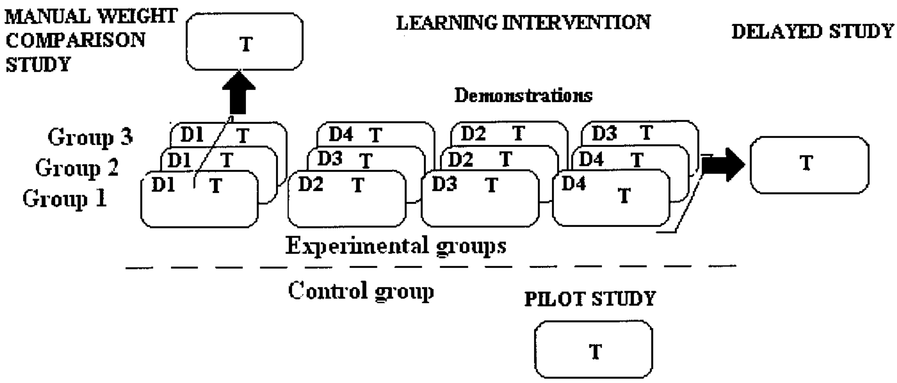


FIGURE 5.2. The analyzed groups in this research. The number of groups is equal in the 5th and the 9th grade. In the 7th grade only *Manual weight comparison study* and *Pilot study were analyzed.* One group contains about 100 pupils. The dashed line in the Figure indicates that the experimental and control group are not equated by random assignment. T refers to the fact that the group is tested.

The design of the research is presented in Figure 5.2. In this Figure there are three experimental groups and a control croup. The pupils in the experimental groups were the subjects of three studies: 1. *Manual weight comparison study,* 2. *Learning Intervention,* and 3. *Delayed study.* The first two studies were carried out in the autumn 1999 - 2001; the interviews took place in October 2000. *Delayed Study* was carried out using of some pupils from the *experimental groups* about six mouths after *Learning Intervention.* The pupils were mainly randomly selected for the *Delayed Study* but so as to ensure that there were representations in all three different experimental groups in the learning intervention and so that the total number of pupils in each grade and group was about one hundred. The research data for the fourth *Pilot Study* was collected in autumn 2000 at the same time as *Learning Intervention* in same schools but in different learning groups (classrooms). *Pilot Study* served as a control group for *Learning Intervention* and *Delayed Study.* It is important to note that every experimental and control group contained about 100 pupils and this research was not a follow-up study but the studies in fifth grade, seventh grade and ninth grade were carried out at the same time 1999-2001 with different pupils.

### 5.3 Design of the Learning study

Pang (2003) observed that, depending on whether certain aspects were varied or kept constant, the learning was either supported or hindered (see 4.2.3). Ling, Chik and Pang (2006) have used three types of variation in their Lesson Study, namely, variation in students' ways of experiencing what is to be taught/learnt; variation in teachers' ways of dealing with the "object of learning", and the use of "pattern of variation" as a guiding principle of pedagogical design to enhance students' learning. In this study I paid attention to the cognitive conflicts to guide pupils' thinking. I tried to carry out demonstrations as identically as possible in all groups. In order to find out the different ways in which pupils estimate the weights using the pulley I arranged a Pilot test (paper-and-pen test) with a group of pupils who did not take part in the Learning Intervention. In the pulley context pupils did not automatically discern the mobility of the pulley as a critical aspect, because in the pilot study it was seen that most pupils predict that the lower object is the heavier of the two objects of equal weight hanging in the pulley. The teacher has to pay attention to the ways in which the contents are dealt with and carried out in the classroom, and therefore the Learning Study in the pulley context was divided into two components: pre-activities, and the Learning Intervention.

*Space of Learning:* When a person tries to learn there are certain necessary conditions for succeeding. Of course, these conditions are not always sufficient, and it would not be certain that learning takes place even if they were fulfilled. 'Space of Learning' defines the limits of minimal circumstances in which learning is possible.

*Pre-activities*: I arranged pre-activities for those pupils who took part in the Learning Intervention. First I walked around in the classroom, and every pupil estimated with his/her hands which one of the objects was heavier and recorded this judgment and the reason for their opinion on the paper. I wanted to stimulate interest in the pupils to predict the weights of the objects for the Learning Intervention. Then I taught about the working of the pulley, so that they would hold its working in their minds simultaneously with the following pulley demonstrations. I showed pupils that the objects in the pulley were connected with an elastic string passing over a freely moving pulley. The role of the pulley was being a prosthetic devise for a resource for thinking (see Chapter 3.1).

*The Learning Intervention*: In the Learning Intervention an effort was made to create relevant patterns of variation. The mobility of the pulley was the critical aspect by varying the positions of the bag and the standard mass in the pulley, three successive variations was obtained. The scientific explanation of how the pulley operates (motionless in the case of the objects of the same weight) works in a wider area than those of students' distinct predictions of weight after the position of objects. I arranged cognitive conflicts to challenge pupils' thinking and I hoped that the pupils noticed and realized that they have to pay attention to the behavior of the pulley system as a whole instead of looking only at a separate body.

## 5.4 Methodology

### 5.4.1 Subjects

Basic education in Finland takes place in comprehensive school, which consists of primary school (lasting six years for children aged between 7 and 13) and lower secondary school (lasting three years for pupils aged between 13 and 16). Teaching groups are formed according to yearly classes. During the first six years instruction is given by a class teacher, who teaches most subjects, especially science (three or four hours per week, 25 % of which is chemistry and physics), mother tongue and mathematics. Teachers of Grades 7-9 are normally subject teachers and physics or chemistry are taught two hours per week. The National Board of Education creates and decides the core curriculum, which is the basis for local curricula (OECD-PISA 2000).

TABLE 5.1. Subjects of the research.

| *Pilot study* | *Girls* | *Boys* | *Total* |
|---|---|---|---|
| 5th grade | 46 | 51 | 97 |
| 7th grade | 49 | 49 | 98 |
| 9th grade | 58 | 51 | 109 |
| *Total Pilot study* | 153 | 151 | 304 |
| *Manual comparison and Interventions* | | | |
| *Group 1* | Girls | Boys | Total |
| 5th grade | 46 | 51 | 97 |
| 7th grade | 45 | 41 | 86 |
| 9th grade | 50 | 50 | 100 |
| *Total Group 1* | 141 | 142 | 283 |
| *Group 2* | | | |
| 5th grade | 46 | 45 | 91 |
| 7th grade | 71 | 35 | 106 |
| 9th grade | 46 | 42 | 88 |
| *Total Group 2* | 163 | 122 | 285 |
| *Group 3* | | | |
| 5th grade | 42 | 46 | 88 |
| 7th grade | 43 | 53 | 96 |
| 9th grade | 40 | 53 | 93 |
| *Total Group 3* | 125 | 152 | 277 |
| *Total Interventions* | 429 | 416 | 845 |
| *Delayed study* | *Girls* | *Boys* | *Total* |
| 5th grade | 38 | 59 | 97 |
| 9th grade | 54 | 57 | 111 |
| *Total Delayed study* | 92 | 116 | 208 |

It is well known that almost all Finnish children complete comprehensive school and in every part in Finland education in comprehensive schools is very similar (OECD-PISA 2000). Therefore, the classes which took part in this study were normal comprehensive school classes. The classes used were from the

Greater Helsinki City Area (6 classes) and from Central Finland (44 classes), ensuring a variety of backgrounds (urban 24 classes and rural 20 classes) among the pupils. The total number of participants in interventions was 845 pupils with 429 girls, and 416 boys. In fifth grade the number of subjects was 276 pupils, of which 134 were girls and 142 were boys, in the seventh grade 288, with 159 girls and 129 boys, and in the ninth grade 281, with 136 girls and 145 boys. Some of these pupils took part in the Delayed Test. In the fifth grade, total number was 97 pupils (38 girls and 59 boys), in ninth grade 111 pupils (54 girls and 57 boys). Altogether the participants in these four studies are shown in the table.

In the study *Mental models in manual weight comparisons between two objects of different size* the grouping differed from the one above. The three groups in each grade build up into one big group. The total number of participating pupils in this study from the fifth grade was 272, of which 130 were girls and 142 boys, from the seventh grade 288, with 161 girls and 127 boys, and from the ninth grade 276, with 135 girls and 141 boys.

### 5.4.2 Collecting data

The subjects in this study were comprehensive school pupils (age eleven to fifteen). At this age the pupils have already mastered reading and writing skills. Therefore, the data gathering method of the Pilot Study was a paper-and-pencil test as presented in the study of Mohapatra and Bhattacharyya (1989).

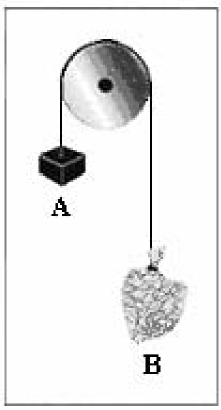


FIGURE 5.3. A standard mass and a bag are hanging in a pulley. The pulley can move freely and the string is very light. What can you say about the weight (mass) of the standard mass A and the bag B compared to each other? Give the reason why you think so?

The paper-and-pencil test was done in autumn after the summer holiday (about 2.5 months) during science lessons at about the same time as the instructions were done with the other pupils (1999 - 2001). The pupils have had no instructions for the subjects in question.

The pupils at the pilot study were asked to compare the weight (mass) of a small standard mass and a clearly bigger bag hanging on a pulley (Figure 5.3) and judge which one of the objects is heavier. Then they wrote the reason for their answers on the same paper. The latter part of their answer was open-ended. To ensure that the pupils understand what the pulley is, there was a demonstration the functioning of a real flywheel fixed on its axis and moving around in both directions so that all pupils could see how the pulley was working and that it was moving freely. This was done so that all pupils would understand what a pulley is. During the test the pupils were not allowed to talk to each other and it was stressed that they should give their own reasons for their answers. I carried out the test myself in the primary school classes (fifth grade). In the lower secondary schools (seventh and ninth grade) the local physics teacher carried out the test during a lesson with following written instructions Figure 5.3.

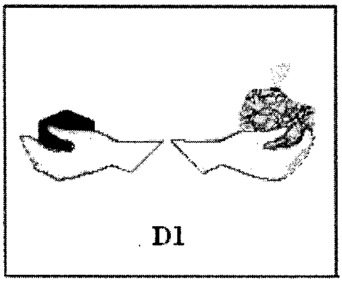


FIGURE 5.4. In the first part D1 of Learning intervention, the pupils judged in their hands which one of the objects was heavier, the little standard mass or large steel wool bag.

Before the Learning Intervention with the pulley demonstrations I introduced the preconception of heaviness of objects to the pupils. I walked around in the class, with the standard mass (weight 20 g and volume about 2 $cm^3$) and steel wood bag (weight 20 g and volume 200 $cm^3$), stopped in front of each pupil, and then I gave the objects to the pupil and pupil judged in his palm which one of the objects was heavier or if they were perhaps equally heavy (Figure 5.4). Then the pupil recorded her/his choice in multiple choice part of test-paper. After that the open-ended question was, "What are the reasons for your choice?" After that pupil had written her/his answer, I collected the paper and I went on to the next pupil repeating the performance. It was emphasized that the pupils should give their own opinions, not being influenced by those of others, and that cooperation and discussion was forbidden. I made these tests after the summer holiday with help of the local science teacher at about same time as the pre-test but with other pupils in different groups and, partly, in different schools. The total number of participating pupils from the fifth grade was 272, of which 130 were girls and 142 boys, from the seventh grade 288, with 161 girls and 127 boys and from the ninth grade 276, with 135 girls and 141 boys.

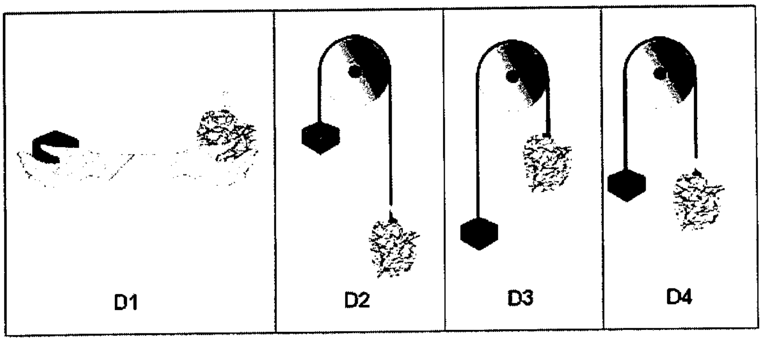


FIGURE 5.5. Demonstrations used in studies. D1. Compare with your hands the weights of a standard mass and a bag. D2 - D4: What can you say about the weight of the standard mass and the bag compared to each other? Why do you think this is so?

The latter part of the second study consisted of the demonstrations like Gunstone and White (1981). The aims of demonstrations were to move the pupils' conceptions of weight towards scientific concept. Therefore, the Learning Intervention continued with three demonstrations (D2-D4). Pupils had to compare the weight of the same objects suspended in a pulley in three different positions (see Figure 5.5).

D2 The bag hangs lower than the standard mass.
D3 The standard mass hangs lower than the bag.
D4 The bag and the standard mass hang on the same level.

In each case pupils were asked to write down what they thought the weight of the standard mass and the bag would be compared to each other, and the reasons why they thought so. Then the pupils recorded their choice in the multiple choice part of test-paper. After that the open-ended question was repeated,"What are the reasons for your choice?" After the pupils had given the answer, I collected the papers, distributed the next test-paper and showed the next demonstration. I repeated this twice.

I supposed also that the order of demonstrations may affect the comprehension. Therefore the order of the demonstrations D2 to D4 changed in different teaching units. In group I the order was D2, D3, and D4, in group II D3, D2, and D4, and in group III D4, D2, and D3. The number of pupils in the different groups was: fifth grade: group I 97, group II 91, group III 88, seventh grade: group I 86, group II 106, group III 96, and ninth grade: group I 100, group II 88, group III 93.

The Delayed Study was carried out about six months later. Some of these fifth and ninth grade pupils, who were taught with the Learning Intervention, were given the same paper-and-pencil test used in the Pilot Study. The aim of that test was to examine how durable the possible new concepts are. The participants in that study were in fifth grade, 97 pupils 38 girls and 59 boys, in ninth grade 111 pupils, 54 girls and 57 boys.

Besides the paper-and-pencil Pilot Study twelve seventh graders were studied by means of short oral interviews to increase the reliability of the results. The interviews carried out a few days after the Pilot Study when the pupils remembered their own responses.

### 5.4.3 Data analysis

In phenomenography the dominant method to obtain data has been to interview people about a phenomenon, to let them tell us as freely as possible about their conception. These interviews are then read through many times so that the researcher starts to see the different aspects and interpretations that people have given to the phenomenon or the special concept. Also group interviews, observations, drawings, written responses and historical documents have been used. It is also possible to study artifacts - historically or comparatively, for example - from the point of view of the different ways of understanding the world around us that are embedded in those artifacts (Marton 1994). Particularly in experimental phenomenography it is possible to use experimental/control design and to utilize quantitative research methods. It reflects a controlled approach. A typical approach involves participants being given a text to read. The analysis determines how the text is understood or what the participants have learned (Hasselgren and Beach 1997). I used mainly written responses under the pulley (an artefact) context but I also interviewed some of the pupils in the Pilot Study to determine what the participants understood about the pulley context and what they have learned from the demonstrations. In order to be successful it is important that the researcher understands the world of the people in the study group so that his interpretation is valid. The researcher is especially interested in the so-called second order perspective i.e. in the relationship between how persons see, explain the phenomenon and how they have experienced the phenomenon. This referential perspective refers to the general combination of features discerned and focused upon by the subjects (the first order perspective). The basic assumption is that different people give different value to the phenomenon according to their experiences and awareness (Marton and Booth 1997). The main results from the research are the qualitatively differing categories. These categories of descriptions will contain information how different people are thinking.

The following data categories were deduced from the phenomenographic analysis by two researchers and then used in both the Pilot Study and the Delayed Study. The same classification was also carried out in the Learning Intervention part of the study "Importance Of The Order Of Demonstrations In Changing Pupils' Conceptions". The pupils' answers were first classified into four main categories:

I Standard mass is heavier,
II Bag is heavier,
III Bag and standard mass weigh the same.
IV Rejected

If the answer contained no reasoning it was rejected and thus not taken into account in the categories I to III. These categories correspond to the referential aspect of the concept of Weight.

Secondly, the structural aspect deals with how the referential aspects are understood. The interrelated and hierarchical nature of meanings can be determined by analysing the complexity of the structural aspects across experiences. The answers in the three first referential categories were further divided into the structural aspects of the conception i.e. the explicit variations of how the pupils reasoned their choice. Categorisation of the reasons was checked out with another experienced researcher. After negotiation, as a compromise, the following four structural categories were then accepted and they are the same in all three studies (Pilot Study, Learning Intervention and Delayed Study). Then the structural categories are placed in hierarchical order according to the abstraction level. If a pupil gave more than one reason the answer was placed according to the reason that gives the higher category in the categorisation. The pupils' reasoning in a certain referential category differs slightly from each other. Examples of the pupils' answers are given in each case.

1. *Motion*. The scientifically correct argument is based on the movement of the bag, the standard mass or the flywheel. This idea includes some notion of the idea of effects of gravity i.e. of the force concept.

   *The bag goes upwards and the standard mass downwards* (I). *The bag has moved downwards* (II). *The bag and the standard mass stay at their positions* (III). *The masses do not move (III).The standard mass does not go upwards and the steel wood does not go upwards (III). The standard mass or the bag goes not downwards in the pulley (III). Because it goes not down(III). Because they do not move (III).*

2. *Position*. The argument is based on the positions of the bag and the standard mass. Some pupils gave also a quantitative estimate of the weights.

   *The standard mass hangs higher (I). It is lower than the upper edge of the bag (I). Because the standard mass is lower, it is therefore heavier (I). The* *standard mass because, it is lower from the pulley (I). The bag hangs lower (II). The steel wood is lower than the standard mass in the balance (II). The bag B is heavier than the standard mass A because the bag B is nearer the floor (II).*

3. *Appearance.* The pupils pay attention to the concrete appearance of the bag and the standard mass.

   *The standard mass looks heavy (I).A, because it looks heavier (I). The bag is larger (II). The standard mass A is lighter than the bag B because bag B is considerably larger (II). The bag is heavier because it is larger (II).*

4. *Material.* The pupils give concrete properties to the bag and the standard mass.

   *The standard mass is of metal and the bag is of plastic (I). The bag contains something (II). Because it is heavier and metallic (I).The standard mass is solid and metallic (I).The bag B because there is something inside (II). Because in the bag B there are something cables inside (II).*

5. *No argument or confusing idea.* In most cases the pupils only stated their thoughts about the weights of the bag and of the standard mass compared to each other.

   *The bag is lighter than the standard mass (I).The bag is heavier than the standard mass (II). Because the bag is larger than the standard mass gravitation affects it more and therefore it is lower than the standard mass even though it weighs the same (III).The bag B (II). The bag B, I cant' give any reason (II). The bag B, I guessed (II). I do not know anything about things like that, thanks.*

The same categories as found in the Pilot Study, are used in the Learning Intervention and also in Delayed Study; namely 1. Motion; 2. Position; 3. Appearance; 4. Material; 5. No argument or confusing idea. However, there were fairly few responses in the categories Appearance and Material so that I decided to join these categories to a new category 3. Concrete Properties. This was done to the results in the studies "Importance of the order of demonstrations when changing pupils' conceptions" and "The durability of conceptual change in learning the concept of weight in the case of a pulley in balance". Thus, I was left with three categories only. In the categories of Concrete Properties, Position and Motion pupils pay attention correspondingly to the size or material, position, and movement of the bodies.

There was a different classification in study of "Mental models in manual weight comparisons between two objects of different size". The answers were first classified according to the choices, the referential categories:

A. the bag is heavier,
B. the standard mass is heavier,
C. the bag and the standard mass are equally heavy.

Then the pupils' arguments for the choice were classified, the structural categories. Three structural categories were found. The placement of the arguments into structural categories was performed several times with another researcher checking a sample of answers until a satisfying grouping was achieved. If a pupil gave more than one explanation, the answer was categorized according to the explanation which was higher in the hierarchy given below. In the following definitions for class categories are given together with some examples of answers.

1. *Pressure.* Pupils determine the weight of the objects by their size, area or pressure.

   *Because there is more of steel wool and it has larger area than the standard mass (A). The standard mass has a smaller area (B). Since the weight of the standard mass is distributed over a smaller area, the standard mass feels heavier although it probably isn't heavier (C). Steel wood is on a wide area, but the standard mass in a small volume (C).*

2. *Material.* Pupils determine the weight of the objects based on their material, density or content.

   *The standard mass is heavier, since metal is denser in the standard mass than in steel wool (B). Because in the standard mass material is denser, than in the steel wool bag (B). It is of denser steel (B). If steel wool were melted, the same weight would be obtained because there is much more of it (C). The standard mass is solid. While the steel wood contains of many little wires (B). Steel wood is not solid, therefore it is lighter (B). There is air the bag (B).*

3. *Feeling.* Pupils indicate that the bag or the mass feels heavier. In the choice standard mass is heavier, feeling means a sensory perception. In the other two choices, Bag is heavier and The bag and the mass are equally heavy the decision is made against the sensory perception and feeling means the result of a contemplation process.

   *It feels and looks heavier than the standard mass (A). It feels heavier (B). The standard mass felt heavier (B). I felt it (B). I see it clearly (B). I don't know, I just felt like it (C). It felt heavier (B). They feel equally heavy (C). The standard mass felt in the hand heavier (B).*

4. No explanation or a nonsensical explanation.

*If steel wool is compressed, it may weigh more (A). I don't know... (B). Guess (C). Well, it only is (B). It is only heavier (B).*

In the preceding paragraphs I have described in detail the subjects, collecting the data, and forming categories of how pupils experience the concept of weight in the pulley context. I have collected, in Appendix 1, a table that shows the placements of the pupils in the different categories in the different phases of the successive demonstrations.

*Interviews:* Twelve seventh graders were studied by means of short oral interviews a few days after they had done the Pilot Study in autumn 2000. The aim of this was to increase the reliability of the research. The questionnaire was the same as in the Pilot Study. I will now present some of these interviews. First I showed the questionnaire and the pupil expressed her/his opinion as to which one of the objects, the standard mass or the bag was heavier. Then I asked for the pupil's explanation for that argument. Finally, I asked if s/he knows of any apparatus which operates like a pulley.

Pupil 1
Pupil: *The bag B is heavier than the standard mass A.*
Teacher: *How do you know it is heavier, do you think about an apparatus?*
Pupil: *The bag is larger. I don't think about any apparatus.*
Teacher: *Do you know any apparatus which works in this way?*
Pupil: *I don't know.*

Pupil 2
Pupil: *The bag is heavier than the standard mass, because it is lower.*
Teacher: *What is the reason you know it is heavier?*
Pupil: *I don't know after all.*
Teacher: *Have you meet an apparatus which resembles that?*
Pupil: *No at all.*

I think that these pupils had no model for their reasoning. The reasoning happened spontaneously, on the spot.

Pupil 3
Pupil: *A is smaller, B is larger.*
Teacher: *Why you answered in this way?*
Pupil: *Because B is low.*
Teacher: *Had you an apparatus as a model?*
Pupil: *No I had not.*
Teacher: *state an apparatus which works like pulley.*
Pupil: *Steelyard.*
Teacher: *Have you used a steelyard as model?*
Pupil: *No.*

This pupil used two explanations for estimating the weight of objects (appearance and position). S/he also thinks that a pulley works like a steelyard. S/he doesn't use however any conscious apparatus to estimate weights of objects.

Pupil 4

| | |
|---|---|
| Pupil: | *One is solid and the other is in separate pieces.* |
| Teacher: | *Why it is heavier?* |
| Pupil: | *It is lower, so it must be.* |
| Teacher: | *What was the model?* |
| Pupil: | *There was no model; it reminds me of a steelyard.* |
| Teacher: | *Did you stop to think how it works?* |
| Pupil: | *I stopped to think. I used as a model a steelyard.* |

This pupil used also two explanations for estimating the weights of objects (material and position explanations). At first s/he denies the help of model but later confesses the importance of model thinking.

Pupil 5

| | |
|---|---|
| Pupil: | *They are equally heavy. Otherwise the pulley cannot move freely.* |
| Teacher: | *Why do you think so?* |
| Pupil: | *I don't know.* |
| Teacher: | *Which apparatus works in this way?* |
| Pupil: | *A clock.* |
| Teacher: | *What kind of a clock?* |
| Pupil: | *A cuckoo clock.* |

This pupil predicts that the objects are equally heavy, but with wrong argument. S/he also thinks that a cuckoo clock works like pulley.

*Summary*: The interviews give very similar results as the paper-and-pencil tests. Only one pupil changed his prediction *the bag is heavier* to the opinion *the objects are equally heavy*. He, however, could not give reasons for his argument. The most popular model for explanations was Position Model. Pupils used also the Material and Appearance Model in their arguments. The explanations with Motion Model were lacking. It was remarkable that in paper-and-pencil tests there was no spontaneous mention about an apparatus which works like a pulley. In the interviews in which I asked to give a description such one, pupils showed an apparatus nearer a balance than a pulley.

## 6 Overview of the articles

**I Hakkarainen, Olavi and Ahtee, Maija (2005). Pupils' Mental Models Of A Pulley In Balance. Journal of Baltic Science Education, 2005, 2(8), (pp. 26-34).**

In the literature of science education there is no exhaustive research on how pupils judge the weights of two objects hanging from the pulley. Pupils from $5^{th}$, $7^{th}$ and $9^{th}$ grades compared the weight of a small standard mass and a big bag hanging in a pulley in different positions in equilibrium. In all three age groups the majority (about 70%) of the pupils stated that the lower hanging bag is heavier. Less than 15% of the pupils in the $7^{th}$ and $9^{th}$ grades understood that the bodies must have the same weight. Pupil$^{s}$' reasoning was classified into four categories that were interpreted according to Vosniadou's framework theory of mental models. Only about 5% of the seventh graders and 10% of the ninth graders seem to grasp the scientific Motion model that leads to the concept of gravity. The notable result was that there are some different mental models which pupils can use to estimate weights and the scientific model was not usual in any grade. Therefore, it is possible to use these models in the pulley demonstrations in the Learning Intervention changing pupils' conceptions of weight from the fifth to the ninth grade. These three studies were Pilot Studies of the Learning Intervention.

**II Hakkarainen, Olavi. Mental Models In Manual Weight Comparisons Between Two Objects Of Different Size. Themes in Education,Vol. 6, No 2, (pp. 151-167).**

This was the first research on the Learning /intervention. The aim of manual weight comparisons was that pupils grow awareness because of the Learning /Intervention. With the help of these studies it was possible to compare statistical compatibility of three different demonstration groups. Pupils in the fifth, seventh and ninth grades (836 pupils in total) were asked to compare manually the weight of two objects of different size but of the same weight and to give the reason for their estimation in writing. The majority of pupils believed the standard weight to be heavier than the larger bag of wool. The number of correct conclusions tripled from grade 5 to grade 9. Pupils' explanations could be divided into three categories: Pressure (weight concluded through size, area or pressure of the objects in the hand); Material (weight concluded through the content or density of the objects) and Feeling. These categories correspond to the mental models of Vosniadou's theory (1994) and they are in accordance with the development of the force concept. It is important that teachers are aware of the mental models of their pupils and how they evolve to be able to support pupils' development. Based on these findings, implications are drawn regarding more productive means of assessing students' understanding in the classroom. The main result of this research is that the pupils do not succeed in learning the conception of weight and pressure without the help of teacher.

**III Ahtee, Maija and Hakkarainen, Olavi (2005). Importance Of The Order Of Demonstrations In Changing Pupils' Conceptions. NorDiNa Nordic Studies in Science Education, 1, (pp. 31- 42).**

This research includes the results of the Learning /interventions. Conceptual change is linked with the question of how a teacher can connect pupils' prior knowledge with the new content to be learned. Fifth and ninth graders were asked to compare the weights of two objects suspended in a pulley in balance in three different positions. Pupils' conclusions and arguments were classified after each demonstration. The proportion of the pupils who concluded after the last demonstration that both objects weighed the same depended on the order of the demonstrations and varied from 60% to 90%. Teachers may consider the demonstrations to be alike but pupils experience them as different. Pupils tend to rely on the concrete and familiar models they have used earlier. The result was that one demonstration is not enough for successful learning. It is also possible that demonstration only creates a new misconception to replace the old ones. In other conditions it is possible that the same demonstrations may promote learning and in another order prevent it. Therefore, it is important that teachers help pupils to look for more general explanations behind the phenomenon, and that the pupils are aware of their preconceptions.

**IV Hakkarainen, Olavi and Ahtee, Maija. The Durability Of Conceptual Change In Learning The Concept Of Weight In The Case Of A Pulley In Balance. International Journal of Science and Mathematics Education, 5, 3, (461-483).**

This study deals with the question of how teachers can help pupils to perceive the concept of weight (gravitation). Fifth and ninth graders were asked in a paper-and-pencil test to compare the weight of two objects suspended in a pulley in balance half a year after the Learning /intervention consisting of three successive pulley-in-balance demonstrations. The understanding that some of the fifth graders seemed to have formed from the weight concept after the teacher-independent learning intervention, had almost totally vanished in the Delayed Study. In the demonstrations the pulley served as an efficient prosthetic device for thinking helping pupils to pay attention to the behaviour of the whole system instead of looking only at the separate objects. In the ninth grade about 45% of the pupils achieved general understanding about the pulley in balance by transferring the scientific explanation from the pulley demonstrations also to the paper-and-pen test. It is proposed a learning-with-conflict model that is based on pupil's alternative explanations about a discrepant event. These explanations are then challenged with a conflicting event. When pupils see and understand how the concept works in different contexts it is possible for the pupils to reach a conceptual change independent of context.

# 7 Results and discussion

In this chapter I have collected the answers to the research questions so that in section 7.1 I give results in the form of categories about pupils' predictions, and explanations regarding the predictions publications I and II). In the section 7.2 I consider how pupils' prediction categories and explanations are changed by the Learning Intervention (publication III). Then I study the development and the durability of Explanation categories which are now studied with the theory of Vosniadou's Mental Models (Publication I, Publication III and Publication IV). Finally, I interpret pupils' conceptions in the pulley demonstrations using the variation theory.

## 7. 1 Prediction and Explanation Categories

The main results of a phenomenographic study are the categories of description. In this work, there are two kinds of categories. First, the prediction categories that tell what kind of predictions pupils give the demonstration of the pulley in balance, or how pupils estimate the weight of two objects in their hands (referential categories). The second kind of categories are based on the arguments how the pupils justify their predictions, explanation categories as structural categories.

The Prediction Categories or referential categories are the same in all the papers: I The bag is heavier, II The standard mass is heavier, and III The bag and standard mass are equally heavy. The name of the Prediction Category varies slightly with the papers. In the publication I it is called Main Category, and in the Delayed Study in the publication IV the name is the Prediction Category. In the table 7.1 are compared the 5th, and 9th graders' initial predictions obtained from the paper-and-pencil test (publication I), and the predictions that the pupils who took part in the Learning intervention with successive pulley demonstrations gave six months later (Delayed Study, publication IV).

TABLE 7.1. Prediction categories in the Pilot study (publication I) and in the Delayed study (publication IV).

| Category | 5th grade | | 9th grade | |
|---|---|---|---|---|
| | Pilot study n = 90 % | Delayed study n = 94 % | Pilot study n = 90 % | Delayed study n = 111 % |
| Mass heavier | 21 | 12 | 11 | 5 |
| Bag heavier | 68 | 39 | 67 | 36 |
| Equal heavy | 4 | 45 | 14 | 59 |
| Rejected | 7 | 4 | 8 | 0 |
| Total | 100 | 100 | 100 | 100 |

The distributions on the Delayed Study are significantly different from those obtained after the Pilot Study (5th grade: $\chi^2(3) = 44.5$, $p<0.001$, 9th grade $\chi^2(3) = 41.6$, $p<0.001$). In the Pilot Study the most popular category in both grades are "The bag is heavier", but after the Delayed Study it was "Equal weights". The proportion of this category was in the 9th grade 59% and in the 5th grade a little less 45%. It is noticeable that total distributions on the Delayed Study in 5th and 9th grade do not differ ($\chi^2(2) = 4.64$, $p < 0.1$) whereas on the Pilot Study the total distribution of the fifth graders differed significantly from that of the ninth graders ($\chi^2(3) = 8.26$, $p<0.05$).

In the table 7.2, there are the pupils' predictions of manual weight comparison in publication II. The distributions in the fifth grade and in the ninth grade differ from each other ($\chi^2(2) = 47.8$, $p<0.001$). The most differences are found in the category A (Standard mass is heavier), and in the category C (The bag and standard mass are equally heavy). Ninth graders estimated equal weights more often than fifth graders ($\chi^2(1) = 45.8$, $p<0.001$), and the fifth graders estimate that the standard mass is heavier more often than the ninth graders ($\chi^2(1) = 43.7$, $p<0.001$).

TABLE 7.2. Distribution Of Pupils' Prediction Categories in Manual Weight Comparison (Publication II).

| Category | 5th grade<br>n = 272<br>% | 9th grade<br>n = 276<br>% |
|---|---|---|
| A. Mass heavier | 91 | 68 |
| B. Bag heavier | 3 | 5 |
| C. Equally heavy | 6 | 27 |
| Total | 100 | 100 |

In the publication I, I named the Explanation Category as Subcategory. In the publication II, and in the publication III it was called an Argument Category, and finally in the publication IV it was called an Explanation Category. In publications I, II and IV there are three explanation categories: 1. *Motion*, 2. *Position*, and 3. *Concrete* category. In the publication II Explanation Categories are: 1. *Pressure*, 2. *Material*, and 3. *Feeling*, because pupils estimated the weights with their hands, whereas in the former cases the pulley context was used.

In the table 7.3 is the distribution of Explanation Category in Pilot Studies and Delayed Studies in the fifth and the night grade. The distributions in the Pilot Studies between 5th grade and 9th grade differ considerably from each other ($\chi^2(3) =20.9$, $p<0.001$), but in both grades the position category is the most popular. In ninth grade the second popular category is the motion category, but in fifth grade it is concrete category. In Delayed Studies the situation has changed. The distributions between categories differ further from each other ($\chi^2(3) = 37.7$, $p<0.001$), but now the order of popularity with categories has changed. In 5th grade "No argument and confusing idea" category is the most popular, and the concrete category is the next one, but in 9th grade the motion category was the most popular, and the motion category the next one.

TABLE 7.3. Explanation Categories in Pilot Study And in Delayed Study.

| Category | 5th grade | | 9th grade | |
|---|---|---|---|---|
| | Pilot study n = 90 % | Delayed study n = 94 % | Pilot study n = 95 % | Delayed study n = 111 % |
| Motion | 7 | 11 | 27 | 45 |
| Position | 40 | 26 | 43 | 31 |
| Concrete | 34 | 28 | 13 | 8 |
| No argument | 11 | 32 | 8 | 16 |
| Total | 92 | 97 | 91 | 100 |

The Explanation Categories in the Manual Weight Comparison (publication II) in different grades differed also significantly from each other ($\chi^2(3) = 22.4$, $p<0.001$). The order of the categories changes also in different grade. In the 5th grade the most popular category was Material, but in 9th grade it was Feeling Category. The greatest difference between the 5th grade, and the 9th grade was in Pressure Category 10% units ($\chi^2(1) = 15.6$, $p<0.001$).

TABLE 7.4. Distribution of the pupils' answers in the different Explanation Categories in the Manual Weight Comparison.

| Category | 5th grade n = 272 % | 9th grade n = 276 % |
|---|---|---|
| Pressure | 5 | 15 |
| Material | 43 | 30 |
| Feeling | 41 | 47 |
| No explanation | 11 | 8 |
| Total | 100 | 100 |

*Summary*: As shown above pupils have several misconceptions in pulley context.

1. The denser the material the heavier the object (*Material category*).
2. The bigger the object the heavier it is (*Appearance category*).
3. The lower the object hangs the heavier it is (*Position category*).

With the aid of the pulley the teacher can demonstrate that these conceptions are not valid, and then pay attention to motion of the flywheel and objects if the objects are of different weights. Only a few of the pupils understood that the bag and the standard mass are of equal weight if they hang in different heights on pulley. For this they had first of all to realize that in the case of the pulley in balance neither the size nor the position of the hanging bodies had any effect on the weight. The weight of the object is the gravitational force exerted on the object by the earth.

In summary, there is observed a little development towards Scientific Category during the comprehensive school time. Therefore, demonstrations in comprehensive school with the pulley are very well suited to study pupils' changes of understanding of concept of gravitational force or weight.

### 7. 2 Development of explanation categories in the Learning intervention

Learning Intervention took place in nine groups (each about 100 pupils) in the 5th, 7th and 9th grade. I had three different series of pulley demonstrations, which I performed one in a different group and different grade (see Chapter 5.2). This summary contains only the demonstrations in series D2 D3 D4 (D2. The bag hangs lower, D3. The standard mass hangs lower, and D4. The bag and the standard mass hang on the equal level), and in ninth grade the series D4 D2 D3 (D4. The bag and the standard mass hang on the equal level, and D2. The bag hangs lower, D3. The standard mass hangs lower). The results of all other demonstrations resemble these examples.

The Figures 7.5 and 7.6 represent the improvement of Prediction and Explanation Categories in Learning Intervention. The Prediction Categories contain the explanations which are based on predictions (motion, position and concrete). Note that predictions that consist less than 5% of the answers are left out to make the figure clearer. At the bottom is given the total number of answers in percentage grouped with the three demonstrations. Before every demonstration series the pupils compared the weights of the bag and the standard mass in their hands (D1).

*Categories in demonstration series D 2D3 D4 in fifth grade*: After D1 I presented the first demonstration D2, *the bag hangs lower* (Chapter 5.2). The results in 5th grade are in Figure 7.5. After the first demonstration D2 the predictions split up in all three *Prediction Categories*. It was remarkable that all *Prediction Categories* won popularity (44%, 33%, and 24%) but the majority of pupils predicted that the bag is heavier 44% (*Motion* 8%, *Position* 14%, and *Concrete* 12%). It was totally different from the prediction compared with a short time earlier in manual weigh comparison.

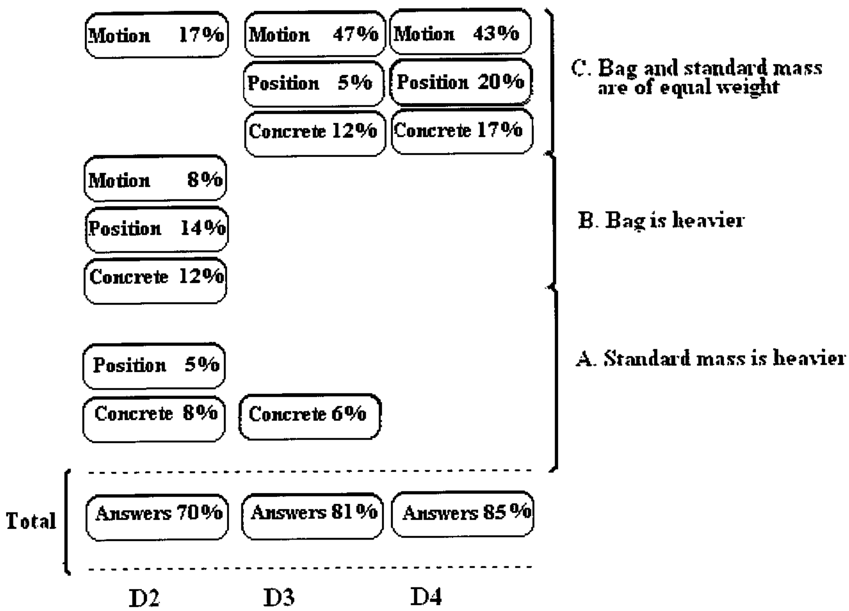


FIGURE 7.5. Improvement of Prediction and Explanation Categories in Learning Intervention during the pulley demonstration D2 D3 D4 in the 5th grade: D2. The bag hangs lower, D3. The standard mass hangs lower, and D4. The bag and the standard mass hang on the same level. The Explanation categories that consist of less than 5% of the answers are left out to make the figure clearer.

*In Manual Weight Comparison* (Publication 2); about 90% of the 5th graders considered that the standard mass was heavier. Even so the most popular explanation after the first demonstration D2 was *Motion Category* (17%) in the category of *Bag and standard mass are of equal weight* (33%), which was the scientific explanation.

The next demonstration D3 (The standard mass is lower) produced the biggest changes. The Prediction Category *Bag and standard mass are equal weight* became the most popular (82%), and *Motion* (47%) attracted the largest Explanation Category. In the same Prediction Category *Position* category increased to 5%, and *Concrete* category to12%.

After the last demonstration D4 (*The bag and the standard mass are in equal level*) there were Explanation Categories only in the scientifically accurate Prediction Category (94%) *Bag and standard mass are of equal weight*. The right alternative *Motion* remained to 43%, but *Position* category increased to 20%, and *Concrete* category also to 17%. It is remarkable that the number of total answers increased during the demonstrations from 70% to 85%, but the number of Explanation Categories in different Prediction Categories decreased from six to three.

*Categories in demonstration series D4 D2 D3 in ninth grade*: The pulley demonstrations began with the demonstration D4 (*The bag and the standard mass hang at the same level*). The Prediction Category *Bag and standard mass are of equal weight* was now the most popular category 77%. In this Prediction Category all three Explanation Categories contained the answers: *Motion* 24%, *Position* 23%, and *Concrete* 16%. In other Prediction Categories only in *Standard mass is heavier* category (20%) contained *Concrete* category (8%) which exceeded the limit of 5%.

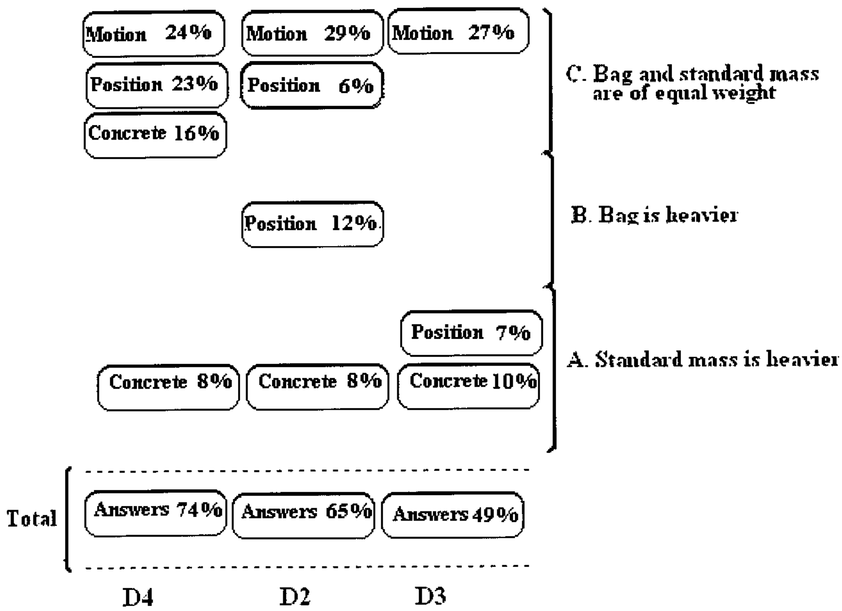


FIGURE 7.6. Improvement of categories in the Learning intervention during the pulley demonstration D4 D2 D3 in ninth grade: D4. The bag and the standard mass hang on the same level. D2. The bag hangs lower, and D3. The standard mass hangs lower. The Explanation Categories that consist of less than 5% of the answers are left out to make the figure clearer.

After the demonstration D2 (*The bag is lower*) the number of answers diminished from 74% to 65%. The most popular Prediction Category was *Bag and standard mass are of equal weight* (59%) with Explanation Category of *Motion* 29%, and *Position* 6%. In other Prediction Categories *Bag is heavier* (17%) contained only *Position* (12%) Category, and in *Standard mass is heavier* (24%) contained only *Concrete* category (8%).

After the last demonstration D3 (The standard mass hangs lower) the most popular Prediction Category was *Bag and standard mass are of equal weight* (62%). The only Explanation Category in there was *Motion* 27%. Over the limit of 5% researched *Position* (7%) and *Concrete* (10%) both in *Standard mass is heavier* Prediction Category. The point here was that the number of total answers to Explanation Categories (Motion, Position, and Concrete) decreased to 49% and the answers outside these to *No argument or confusing idea* category increased to 51%. This tendency was seen all the time in demonstrations in series D4 D2 D3.

### 7.3 Explanations as mental models

The role of mental model is to account for the individuals' reasoning both when they try to understand discourse and when they try to explain and to predict the physical world behaviour (Sasse 1997). In this study the situation is just like that. Pupils try to understand how a pulley works, and then they write their predictions and arguments about it on the paper. For example Vosniadou 1994; Ioannides and Vosniadou 2002; Coll and Treagust 2003; Schoultz, Säljö and Wyndhamn 2001 have mapped pupils' mental models. According to Vosniadou 1994 mental models are unstable, but it is possible that some mental models, which have proven useful in the past, are stored, and then retrieved from long-term memory when needed. The properties of mental models are treated in Chapter 3.4.

In the preceding paragraph I treated Prediction, and Explanation Categories which were constructed by pupils during a very short time (less than 20 minutes). During so short a time it is hardly possible that the new mental models are developed. In this chapter the developing time of the explanation categories of pulley working, in pupils' minds, in the same grade was about six months. Therefore, I speak now of mental models. I shall show the mental models in Pilot Study (Publication I), after the Learning Intervention (Publication III, and in the Delayed Study (Publication IV), and show how permanent the models are.

In the Pilot Study (Publication I), and the Delayed Study (Publication IV) there was the same paper-and-pencil test seen the first time by the pupils. In the Learning Intervention the summary contains mental models of the 5$^{th}$ and 9$^{th}$ graders. In both grades there is one common summary of results of the last demonstrations of all three groups in Figures 7.7 and 7.8. In these pictures the Explanation Models are also compared with Vosniadou's mental models.

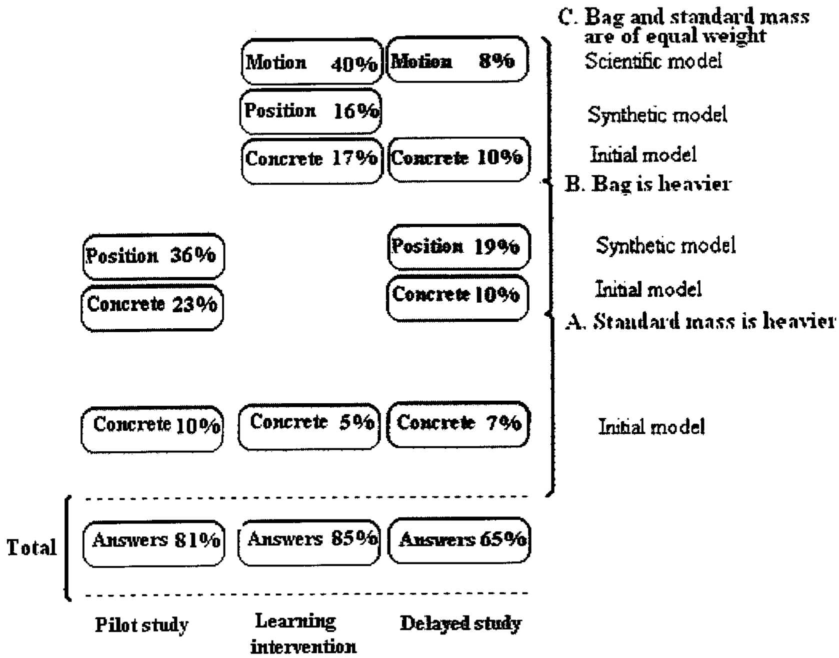


FIGURE 7.7. The development of mental models of 5th graders from Pilot study to Delayed study. The models that consist of less than 5% of the answers are left out to make the figure clearer.

*Development of mental models in 5th grade*: In Pilot Study the most popular Prediction Category in 5th grade was *Bag is heavier* (68%). The explanations were concentrated on *Position* model 36% and *Concrete* model 23%. In the Prediction Category *Standard mass is heavier* was the next popular 21%, and only one Explanation Model is seen *Concrete* 10%. After Learning Intervention there was majority of Prediction Category *Bag and standard mass are of equal weight* (83%). The Explanation Categories in this area were *Motion* 40%, *Position* 16%, and *Concrete* 17%. In Delayed Study all Prediction Categories included answers with mental models. The most popular was still C. *Bag and standard mass are of equal weight* with the proportion of 45% with *Motion* model 8%, and *Concrete* model 10%. The second popular Prediction Category was *Bag is heavier* 39%. There were two groups of mental models: *Position* 19%, and *Concrete* 10%. The proportion of *Standard mass is heavier* was 12%, with one group of model *Concrete* 7%. It was remarkable that 32% chose the alternative *No argument or confusing idea.*

*Development of mental models in the 9th grade*: In the 9th grade in Pilot study the most popular Prediction Category was *Bag is heavier* (67%) just like in the fifth grade. The category was split up into three groups of mental models: *Motion* 16%, *Position* 37%, and *Concrete* 11%. The proportions *Bag and standard* *mass are of equal weight,* and *Standard mass is heavier* were almost of equal size 14% and 11%. The former included the model *Motion* 9% and the latter *Position* 6%.

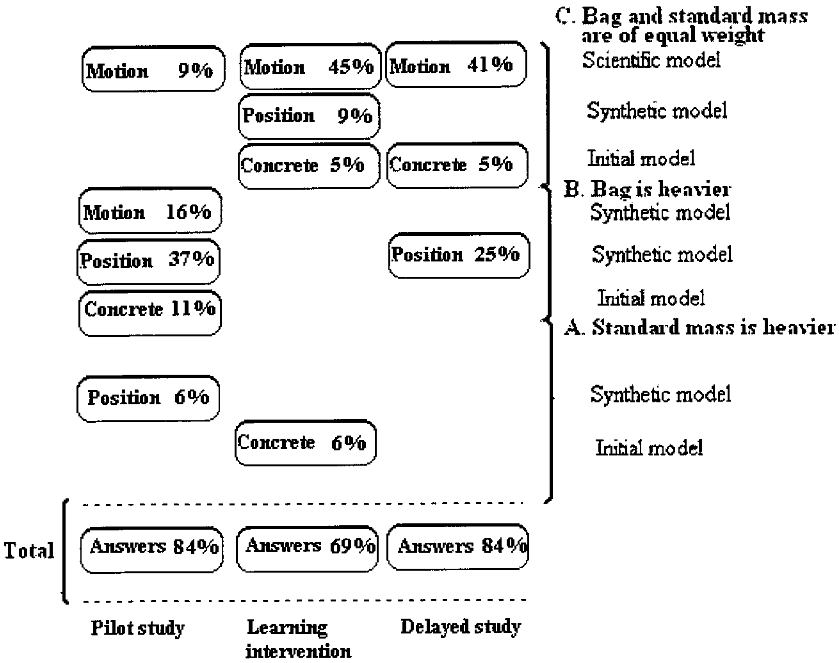


FIGURE 7.8. The development of mental models of the 9th grade from Pilot study to Delayed study. The models that consist of less than 5% of the answers are left out to make the Figure clearer.

Learning Intervention was concentrated in Prediction Category *Bag and standard mass are of equal weight* (78%), with three groups of mental models: *Motion* 45%, and *Position* 9%, and *Concrete* 5%. There is a small proportion of *Concrete* model (6%) in category of *Standard mass is heavier.* In Delayed Study the answers were divided into two *Prediction* categories *Bag and standard mass are of equal weight* 59%, with the groups of mental models *Motion* 41% and *Concrete* 5%, and *Bag is heavier* 36%, with *Position* model 25%.

It is seen that the mental models of the 5th grade and the 9th grade differed in some respects from each other. In Pilot Study in the 9th grade mental models are split up into 5 categories, two more than in the 5th grade. After Learning Intervention in the 9th grade the Prediction category *Bag and standard mass are of equal weight* was more coherent than that of the 5th graders. Besides in the 9th grade proportion of *Motion* model was 54%, in the 5th grade 43%. The biggest differences were in Delayed Study. It was seen that the 9th graders retained their development, but the fifth graders were on the decline from the Learning

Intervention. *Motion* model decreased from 43% to 8%, and the number of Explanation categories increased remarkably. If we take off the answers in *Bag and standard mass are of equal weight* from the data we see that the distribution is the same as in Pilot study ($\chi^2(2) = 0.519$, $p>0.1$). Similar regression is seen after Delayed Study in the 9th grade in Prediction Category of *Bag is heavier*. The proportion of *Position* model increased to 25%.

## 7.4 The discernment of variation in the successive demonstrations

In this study the positions of a bag and an equally heavy standard mass were varied in relation to each other in a pulley. When I changed the positions of the hanging objects, most pupils faced conceptual conflicts, except the *Concrete* properties (appearance and material) did not vary, so pupils did not perceive any systematic conflict either. The results in Pilot Study (publication I) and demonstration series D2 D3 D4 are shown in table 7. 5 (Pilot Study, D2 the bag hangs lower, D3 the standard mass hangs lower, and D4 the bag and the standard mass hang on the same level).

TABLE 7. 5. The results of the pilot study, the demonstrations D2 D3 D4 and Delayed Study.

| | Model | Pilot study % | D2 % | D3 % | D4 % | Delayed study % |
|---|---|---|---|---|---|---|
| 9th grade n= 100 | Motion | 9 | 24 | 48 | 54 | 41 |
| | Position | 43 | 32 | 16 | 16 | 31 |
| | Concrete | 14 | 14 | 9 | 9 | 9 |
| 5th grade n=98 | Motion | 2 | 17 | 47 | 43 | 8 |
| | Position | 40 | 20 | 10 | 20 | 26 |
| | Concrete | 34 | 23 | 21 | 21 | 28 |

Note: Motion Model contains pupils' responses only in the prediction category of the bag and the standard mass are of equal weights. Position and Concrete Models contain answers in all three categories.

*Motion* explanation model increased in the 9th grade after the first demonstration (D2 the bag hangs lower) from 9% to 24% and after the second demonstration from 24% to 48%. Correspondingly results in the 5th grade are from 2% to 17% and 17% to 47%. The results were very much alike. It is remarkable that after the demonstration D4 (the bag and the standard mass hang on the same level) no bigger increase or decrease was noticed. The results are therefore statistically consistent ($\chi^2(3) = 3.98$, $p> 0.1$).

*Position* explanation model decreased at the same time; first in the ninth grade in Pilot study the proportion of *Position* model was 43% and then after D2 32%, after D3 16% and after D4 16%. In the fifth grade correspondingly 40%, 20%, 10% and 20%. These four results are statistically inconsistent ($\chi^2(3) = 11.63$, $p<0.01$), but the three first results are consistent ($\chi^2(2) = 1.03$, $p> 0.1$).

In those two categories (Motion and Position) were seen changes caused by succeeding demonstrations. The situation is different in Concrete Categories. In ninth grade in Pilot study 14% pupils used Concrete explanation model, after the demonstrations D2 14%, D3 9%, D4 9% and after Delayed Study 9%. These all were all consistent with each others ($\chi^2(4) = 3.39$, $p > 0.1$). In fifth grade the proportions were correspondingly: 34%, 23%, 21%, 21%, and 27%. These also were consistent with each others ($\chi^2(4) = 5.95$, $p > 0.1$).

As stated above Motion explanation models in all studies and Position explanation models in three first studies in the fifth grade and in the ninth grade are very much alike. Therefore, later in this section only pupils in the ninth grade are included.

*The role of cognitive conflicts*: In order to make the pupils pay attention to the concept of weight they were first asked to compare with their hands the weights of the small standard mass and the bigger bag. The results of this work are reported in the publication II. This pre-activity had an effect on the results.

After the first demonstration D2 with the pulley 24% of the ninth graders had reached the scientific conception (table 7.5) whereas only 9% could give the corresponding argument when not shown the manual comparison of the objects (Publication I). So I interpret it that these pupils when recognizing the cognitive conflict produced by the disturbing conclusion from the manual comparison (the standard mass is heavier) based on the material of the standard mass or the sensory perception feeling, and the new conclusion from the first pulley demonstration (the lower hanging big bag is heavier) were able to pay attention to the working of the pulley.

Atkinson and Peijnenburg (2004) describe the hallmark of a remarkable thought experiment as "it is simultaneously destructive and constructive; it destroys an old theory and at the same time establishes a new one". Respectively here, when the pupils had these ideas simultaneously in their minds, they had first to get rid of the old notion that the hanging objects had different weights. After that they could solve the cognitive conflict by noticing that because there is no motion in the pulley system the hanging objects have to be equally heavy. According to Schoultz, Säljö and Wyndhamn (2001) the pulley system acts here also as a prosthetic device for thinking. After finding the satisfactory explanation these pupils applied this conception to the next demonstrations (Figure 7.9).

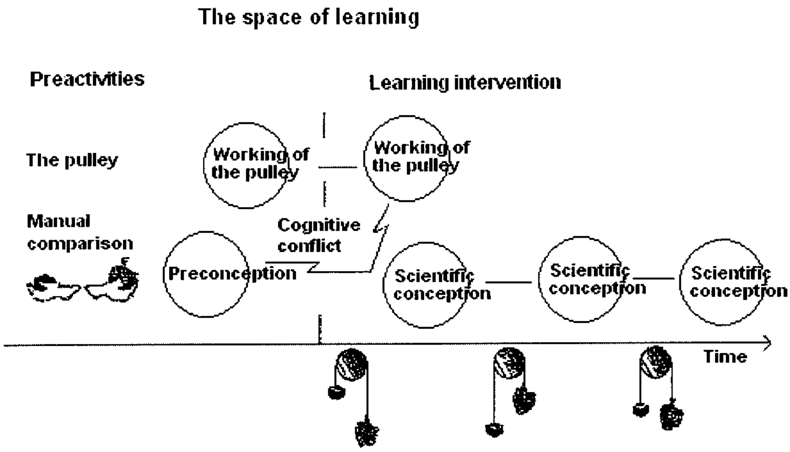


FIGURE 7.9. The space of learning showing how some of the ninth graders end up with the scientific conception after the first pulley demonstration.

The majority (60%) of the pupils state, after the first pulley demonstration, that the objects have different weights. About 32% of all pupils solve the cognitive conflict between the manual comparison and the first pulley demonstration using the position model (table 7.5). When, in the next demonstration, the bag is moved up and the standard mass down, most of these pupils recognize their erroneous thinking and they abandon the position model. The decrease in the use of the position model among the pupils stating that the objects have different weights is near the percent unit as the increase in the use of the motion model, 16% and 24% (table 7.5 and Appendix I).

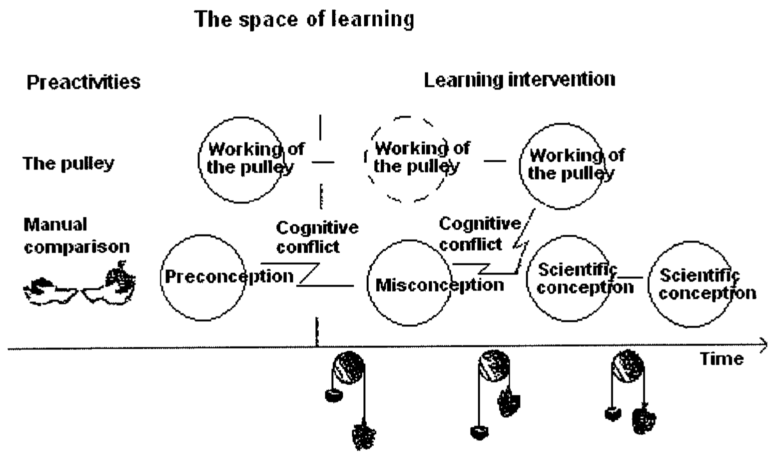


FIGURE 7.10. The space of learning showing how some of the ninth graders end up with the scientific conception after the second pulley demonstration.

As they give up their earlier wrong conclusion about the different weights they are open to look for new explanations, and they take up the notion about the immobility of the system (Figure 7.10). It is important that the pupils have the possibility to check their new prediction with the working of the pulley in the next demonstration. In Delayed Study (Publication IV) about half a year after the learning intervention study the proportion of the ninth graders using the motion model stayed on the same level. These pupils had changed their thinking.

As seen from Figure 7.5, still 15% of the pupils base their prediction after the last demonstration D4 on the fact that the objects have the same weight in the *Position* model (Appendix I). Obviously these pupils did not recognize during the demonstrations the critical detail that the system is not moving. Now the third demonstration in which the objects were on the same level strengthened their notion. In the earlier demonstrations it was easier to observe the immobility as most of the pupils assumed the objects to have different weights and then it is natural to expect that they start to move. Those pupils who based their arguments on the appearance or the material of the objects, i.e. on the concrete models could not notice any variation in these demonstrations. Therefore, it is understandable that their proportion stayed more or less the same throughout the demonstrations (table 7.5).

## 8 Credibility of the study

In the following section I evaluate the reliability and validity of the results of this study. First I introduce a statistical method, the Chi-square test, which I have used to estimate the reliability the study. Then experimental validity is evaluated in the role of internal validity and external validity. Finally, I answer the questions of Ahonen (1996), which are the criteria of validity in a phenomenographic study.

### 8.1 Reliability of the study

*Statistical method*: The only inferential statistical test done in this study was the Chi -square test. This test can be used to test independence between two variables arranged in a contingency table. The use of test regards the following limitations (Key 1997 in Savinainen 2004):

1. Data must be in frequency form (nominal data).
2. The sample must be representative (random).
3. The individual observations must be independent of each other.
4. The sample size must be adequate:
   a) In a 2 x 2 table, the chi-square test should be used if n is less than 20.
   b) In a larger table, no expected value should be less than 1, and not more than 20% of the variables can have expected values of less than 5.
5. The distribution basis must be decided on before the data are collected.
6. The sum of observed frequencies must equal the sum of that of the expected frequencies.

In this study the data met the other requirements except randomness of the samples. Therefore the samples were matched so that the subjects of these studies were chosen from normal comprehensive school classes in the Greater Helsinki City Area and from central Finland, ensuring the variety of urban and rural backgrounds among the pupils. It is known that in Finland the education in comprehensive schools does not differ much (OECD-Pisa, 2000).

*Experimental reliability*: In the experimental approach, reliability is the extent to which the measurements of a test remain consistent over repeated tests of the same subject under identical conditions. An experiment is reliable if it yields consistent results of the same measure. It should be noticed that only these Chi-square tests were carried out where the data met the above-mentioned requirements. In this study I try to find differences in pupils' thinking caused by the Learning Intervention. Therefore, I have written down only those results in which differences of the two data are of statistical significance.

I have used Chi-square test (the goodness-of-fit test, and test of independence) to test groups between categories in the same grade, in different grades, and different studies. In the publication I *Pupils' Mental Models Of A Pulley In Balance* (Hakkarainen and Ahtee 2005) all three main categories (prediction categories) and also pupils' arguments (explanation categories or mental models) were compared. In publication II *Mental Models In Manual Weight Comparisons Between Two Objects Of Different Size* (Hakkarainen 2005) statistical comparisons were made between fifth, seventh, ninth graders but also between boys and girls in the ninth grade.

The three groups before learning intervention in publication III *Importance Of The Order Of Demonstrations In Changing Pupils' Conceptions* (Ahtee and Hakkarainen 2005) were tested with the Chi-square test, and no statistical difference was found. This means that all three groups in each grade were very much alike. During Learning Intervention pupils were tested after every demonstration to find statistical differences between pupils before and after every demonstration. After the Learning Intervention Chi-square test is used to check differences between categories in different learning groups, and also between the categories of the first test and the last test in the same group. This all was done both in the *prediction categories* and *explanation categories*.

Chi-square test was used also between the different studies. The study *The Durability Of Conceptual Change In Learning The Concept Of Weight In The Case Of A Pulley In Balance* (Hakkarainen and Ahtee 2006) includes also Delayed Study. Comparisons are then made between Delayed Study and the data of the last demonstration of the Learning Intervention both in prediction categories and explanation models. The same comparisons are made also between Delayed Study, and Pilot Study. Finally in Delayed Study the fifth graders' and ninth graders' explanation models and categories are compared with each other.

In the cases above statistical reliability is the consistency of a set of measurements. Experiments are reliable if they yield consistent results of the same measures. In this study Chi-square test was used to check data, and only the results which have statistical significance are documented.

### 8.2 Validity of the study

Validity refers to getting results that accurately reflect the concept being measured, in other words if it appears to measure what it is supposed to measure. Experimental validity is divided to several categories, of which internal validity and external validity play a significant role in proving validity of a research. A research possesses internal validity if it properly addresses a causal relation between two variables. External validity is a form of experimental validity. A study possesses external validity if the results of experiments hold across different experimental settings, procedures and participants. If a study possesses external validity, its results can be generalized to the larger population.

This research yields internal validity. In this study the teacher first manipulated the pulley demonstration (the independent variable, the "cause") and then the effect of demonstration was seen as a change in pupils' thinking (the dependent variable, the "effect"). This is seen in the data of study "Importance Of The Order Of Demonstrations In Changing Pupils' Conceptions" (Ahtee and Hakkarainen 2005).

The most common loss of external validity comes because experiments using people often employ small samples from a single location. In my study, the classes which took part were normal, comprehensive school classes in the Greater Helsinki City Area and from central Finland ensuring the variety of backgrounds among the pupils.

Ahonen 1996 produces criteria for validity in phenomenographic study. The main aspect is that both data and the categories must correspond to the ideas of the subjects (authenticity), and at the same time conclusions from the data and the categories must be connected to the theoretical framework of the research basis (relevance).

TABLE 8.1. The criteria of validity in phenomenographic study (Ahonen 1996).

| | Raw data | Categories |
|---|---|---|
| Authenticity | Are the raw data the same to the researcher and to the subjects? | Are the meanings of categories the same as those of the subjects? |
| Relevance | What is the relevance of the data from the viewpoint of the theory of the research? | What is the relevance of categories from the viewpoint of the theory of the research? |

In the following I give answers to questions in the table above, 8.1.

*Research design:* The design of this research is seen in the Section 5.2. The design in this study does not follow the commonly used quasi-experimental design because in this research there were no pre-tests in the control group and the experimental groups. The reason is that the students were very interested in tasks of the testing and I thought that testing was a real risk to the objectivity of the research. If the test takes place before the intervention with the same subjects as in *Learning Intervention*, it is possible that pupils would deal with the tasks during the time before *Delayed Study* and that would naturally affect the results of *Delayed Study*. Answers to two short equally similar questions are very easy to clear up and then recall in *Delayed Study*. That was the reason why I tested only the pilot test and the delayed test group. I wanted the subjects in *Delayed Study* and in *Pilot Study* to see the questions in control test for the first time.

*Maturation:* Between any two observations subjects change in a variety of ways. Such changes may produce differences that are independent of the experimental treatments. Maturation is more acute in educational studies of longer duration than in brief laboratory experiments. Maturation has an effect on my work in Pilot Study from fifth grade to ninth grade during four years, but hardly at all in studies in *Learning Intervention* and *Pilot Study* taking place about same time, so different are distributions in Figures 7.5 and 7.6 from Figures 7.7 and 7.8.

*Explanation of data gathering process:* The data gathering process is described in Chapter 5.4.2. The pupils' backgrounds are given and also the guiding factors how the pupils were chosen. No pupils with special backgrounds, (like groups chosen with science or mathematics tests, or unisex groups), were taken into the study. There are also seen further observations such as how choosing the time, the place, and the duration of the *Learning Intervention*.

*Evaluation of the validity of the data:* The authenticity of data requires that the researcher and subjects see that the data in question come from the same phenomenon. First ensuring the validity of the data, the apparatus always was shown to all of the pupils to ensure that they would understand what a pulley is. To build up a confident atmosphere the data gathering situation was kept similar to a normal lesson so that the familiar teacher was present with the researcher. The authenticity in the pupils' arguments is seen in the pupils' answers. The answers they gave are expressed in connection to categories (see chapter 5.4.3).

*Evaluation of the validity of the conclusions:* Conclusions are expressed in the categories in which the data are classified. The different categories include ontologically different kinds of answers (see Chapter 5.4.3).The lowest category, Concrete, is a combination of the Material and Appearance category, and the highest is the scientifically accepted Motion category.

*Validity of collecting data*: It is possible after having studied the description of the outcome, another researcher should be able to judge what categories of description apply to each individual case in the material in which the categories were found (Marton, 1994). The distribution of pupils' written arguments was read and discussed with another researcher an expert in this research area. Also the distribution of pupil's arguments into the categories was carried out several times until satisfactory grouping was achieved. I have described the categories based on pupils' reasoning so that another researcher could be able to recognize instances of different ways of experiencing the phenomenon of question, and after having studied the description of the outcome space the researcher would be able to judge what categories of description apply in each individual case to the material in which the categories are found. I have done these in the Publication I (Hakkarainen and Ahtee 2005) and in Chapter 5.4.3. There are also examples of pupils' answers given for each case.

### 8.3 Relevance of the models

In order to show the relevance of the mental models that I found in this work, I have studied the corresponding models of cognition in Conceptual Change Theories to compare them with models in this work. I have found some theories of conceptual change, similar to my categories or models. In Table 8.2 I compare these (weight or gravitation) in pulley context.

Chi et al. (1994) assume that an entity of the world belongs to different ontological categories such as Matter (like objects) and Processes (like force). The ontological status of pupils' initial conceptions and the scientific conception determine the difficulty of learning. In my study the Concrete Properties Model corresponds to the category of Matter and category Motion corresponds to the category of Processes. The fifth graders look at the concrete properties of the objects so that they connect weight to the category of Matter. However, in the case of the pulley in balance pupils should consider weight as an interaction between the object and the Earth i.e. belonging to the category of Processes. According to Chi et al. (1994) the stability of the 5th graders' Material model is thus due to their concentration on the properties of the objects without taking into account the functioning of the pulley (see Chapter 3.2)

TABLE 8.2. Categories of the Pulley model with Models from Conceptual Change Theories.

| Models in this study | Bao and Redish's model (2003) chapter 3.1 | Vosniadou's model (1994) chapter 3.2 | Chi's categories (1994) chapter 3.2 |
|---|---|---|---|
| Incoherent | Pre-naïve | | |
| Concrete | Naïve | Initial | Matter |
| Position | Mixed | Synthetic | |
| Motion | Expert | Scientific | Process |

Based on empirical observations of students' conceptual development process Bao and Redish (2003) describe college students' learning of a particular concept topic with four stages. The first pre-naïve stage represents the situation where students by themselves are not able to develop a logically consistent model of the concept topic. Many students can't provide any kind of reasoning. It corresponds to the Incoherent Stage in my study in which pupils have a confusing idea or no argument at all about the weights of the pulley. Vosniadou (1994) argues that children do not separate weight from force, and weight or force is an inner property of big and large objects. Thus the initial model of Vosniadou's framework theory is Concrete Model of this study. So in my study, the Concrete Properties Model would correspond also to Chi's Matter category, Bao and Redish's naïve stage and Vosniadou's initial model (chapter 3.2).

The third one in Bao's and Redish's conceptual development process is the mixed model stage in which students consistently hold on to one or several models of the concept topic. These models are often developed by students based on their real-life experience and previous school learning and they can produce correct results in a limited set of cases. This stage like my Position Model corresponds to a synthetic model in Vosniadou's framework theory (1994) because it represents pupils' attempts to reconcile the information they have received from their environment and certain basic presuppositions and beliefs (the lower one is heavier) with the successive demonstrations. These two models correspond to Position Model of this research. These pupils try to explain the functioning of the pulley in balance with the aid of the two-armed scale. Here Bao and Redish point out that it is better if the teacher first helps students learn some new ideas in reasoning. In our case this would mean that the teacher suggests to pupils to pay attention to the behaviour of the whole system instead of looking only at the separate objects (see 3.1 and 3.2).

My Motion Model is near Bao's, and Redish's Expert Model Stage in which students start to understand coherent use of the scientific model. Motion Model is also near Vosniadou's scientific model because pupils' arguments are based on the movement of objects (see 3.1 and 3.2).

The same tendency is seen in the history of science. In the Ancient and Middle Ages people generally used concrete properties of an object (volume, appearance, and material) to estimate the weight or mass of an object (Concrete). Aristotle claimed that in the case of free fall the weight of an object increases as the object gets closer to its natural place (Position), and Galileos' idea was that weight was a characteristic of a single object, a one-way downward pull which is the cause of motion (Motion), see 2.1.1 and 2.1.2.

## 9. Conclusions

*Quantitative comparisons of studies*: In science literature there are only a few studies that address the pulley in science education (e.g. Champagne, Klopfer and Anderson, 1980; Gunstone and White 1981; Mohapatra and Bhattachaayya, 1989; McDermott, Scaffer and Somers, 1994. These studies are mainly quantitative studies. When the results of these studies are compared with the quantitative parts of recent research, the results are very similar. Champagne et al. (1980) noticed that about 80% of physics students in college believed that being "lower (close to the earth) implies heavier" for objects suspended on the pulley. In Mohapatra's and Bhattacharyaa's (1989) paper-and-pencil study about 60% of the 9$^{th}$ graders stated that the downward force on the lower hanging body was bigger than that on the higher hanging body even though it was mentioned in the question that the two bodies were of equal mass. In the Pilot Study (Publication I) of this research the corresponding value in the fifth grade was 68%, and in ninth grade 67%, which are between the values of those two earlier studies.

Gunstone and White 1981 carried out research on the same topic with a bicycle wheel mounted as a pulley. They found that 27% of first-year university students reasoned the lower hanging body to be heavier than the higher hanging body (of same mass). In table are the results of Gunstone's and White's research. In the same table there is the 9$^{th}$ graders' distribution in the publication IV of this research. According to Chi-square test that distribution is similar, $\chi^2(2) = 0.519$, $p < 0.1$.

TABLE 10.1. Students' conceptions about weight of body in pulley in balance.

| Study | | Weights equal | Lower heavier | Higher heavier |
|---|---|---|---|---|
| Gunstone&White (1981) | n | 320 | 122 | 17 |
| | % | 70 | 26 | 4 |
| Publication IV | n | 65 | 40 | 6 |
| | % | 59 | 36 | 5 |

*Pupils' Conceptions Of Weight In The Pulley Context*: In the Publication I (the Pilot Study) the aim was to find out how the pupils give reasons for their arguments of weights of two bodies hanging in pulley in balance. The data of this part of research were collected and categorized with the methods of *Experimental Phenomenography*. The results then were analyzed with the theory of Vosniadou 1994, and Vosniadou and Ioannides 1998.

There seems to be a spontaneous development in pupils' mental models about the behaviour of a pulley in balance starting from the concrete model Material up to the scientific model Motion. In the initial mental model, Material is formed early and pupils have applied it successfully to many situations. It contains the belief that metallic objects are heavy and those of soft material are light and thus it leads to an intuitive conclusion that the standard mass is heavier. About 8% of the 5th graders still think like this (Publication I in Fig. 4).

Everyday common experience that the weight of an object depends on its size (volume) joined with the intuitive rule the bigger-the-heavier (Stavy, and Tirosh, 1996) led some pupils to conclude that the big bag is heavier and thus apply the spontaneous model Appearance (Publication I in fig. 4). However, only 9% of the ninth graders use it anymore. Later I used the term Concrete model for both Appearance and Material Model.

Some of the pupils perceive to connect the weight of an object to its position. This leads to the dominant Position Model that is applied when understanding of the behaviour of a beam balance. It reinforces the wrong conclusion- the big bag is heavier- made spontaneously on he basis of the Appearance Model coming first into mind. Position Model has also been supported by physics teaching in the lower secondary school especially in the connection of measurements with beam balances and investigation of the laws of moments.

There are a few pupils who connect the weight and the functioning of the pulley, which the teacher showed at the beginning of the test by rotating the easily moving flywheel in the Motion Model. According to Vosniadou and Ioannides (1998) it is a synthetic model because it is a product of physics teaching, but it is also near the scientific model because it includes the idea that weight is a force and it causes (accelerated) motion.

No pupil used gravity or gravitational force in their answers. Although gravitation is shortly mentioned in the physics textbooks, it is not in everyday use.

*Resolution of cognitive conflict*: A common instructional strategy to foster conceptual change is to confront pupils with discrepant events that contradict their existing conceptions (see e.g. Stavy and Berkovitz 1980; Dekkers and Thijs 1998; Kang, Scarmann and Noh 2004). This is intended to invoke a cognitive conflict that induces students to reflect on their conceptions as they try to resolve the conflict.

Teaching strategies that are based on cognitive conflict and its resolution (see 4.2.1) can be divided in two groups: *discrepant events* and *conflict between ideas*. Strategies using *discrepant events* involve situations where the student's existing ideas are made explicit to the student through an exposing discrepant event and then challenged in order to create a state of cognitive conflict through *pupil's explanation* and *science explanation* about that event. Finally teacher encourages and guides the pupil to accommodate and to develop a new conceptual model consistent with the scientific view (Nussbaum and Novick 1982). According to Tsai 2000, students should resolve two cognitive conflicts during the process of conceptual change: one between the discrepant event and students' alternative conception, and the other between the alternative conception and the scientific explanation. In both cases there are critical events between conceptual frameworks of pupil alternative conceptions familiar to pupil and a scientific target concept unknown to pupil. Tsai and Chang (2005) claim, those discrepant events which the pupils often confronted are often unfamiliar, new findings to pupils, whereas critical events are already well-known to the pupils.

I have used in my study the discrepant event in another meaning than the researchers above. In my study that event reveals to the learner and to the teacher how the learner thinks about that event. So I renamed it the Revealing Event. In the Publication IV the Revealing Event (manual weights comparison) and Critical Event are familiar to pupils because I showed before the Learning Intervention how the pulley works. From the Publication I, I gained the knowledge of how the pupils meditated in pulley context. From that knowledge I provided the circumstances where pupils predict weights of two objects hanging in the pulley and then gave reasons for their interpretations. The hanging objects are then put in different heights as in the earlier episode. The old interpretations were so challenged by new different positions of the same objects. It is very important to pupils that both events (a revealing event and a critical event) are familiar to the pupils, so pupils see that their interpretations are not valid in all successive events. If they do not see that, they do not experience the cognitive conflict either.

In this study in Chapter 7.4 in Figures 7.9 and 7.10 there are two models of how the pupils meet a cognitive conflict. In my study all pupils made the same observation about the weights of the objects. This was done with the manual comparison (see Publication II). Pupils are familiar with this Revealing Event, because it is an initial way how a child gets to know weight as pressuring force (see chapter 2.2). Most pupils felt the standard mass to be heavier than the bag. Before the Learning Intervention I showed pupils also the working of the pulley, but I did not use it as a scale. From the Publication I, I got the knowledge that most pupils think that in a pulley the object which hangs lower is heavier (position model), so I decided that in the Demonstration II the bag has to be set so that it hangs lower. Although these two events (manual comparison and the first pulley demonstration) were in different contexts, after the first pulley demonstration 15% more of the ninth graders than in the Publication I answered that objects had the same weight. They utilized now the working of pulley, and transferred working of the pulley to the learning episode (see Figure 7.10) and explained that the objects have the same weight because they are not moving.

The 'revealing event' to the demonstration D3 (the bag is higher) is the preliminary demonstration D2 (the bag is lower) with the argument 'lower is heavier'. The demonstration D3 is then a 'conflict event' to a larger part of pupils (24%), who noticed that their old prediction 'lower is heavier' is not valid after the demonstration D3. They solved the cognitive conflict so that they spontaneously paid attention to the immobility of objects and transferred to use the Motion Model (Figure 7.10). In these two cases the pulley acts as a prosthetic device for thinking (see Scholtz, Säljö, and Wyndhamn 2001).

# 10 Implications for teaching

The starting point of this research was to study and to improve teaching and learning in classroom context. My opinion is that the focus of learning and teaching lays on how teachers make a specific content of learning available to students. I carried out the research in classroom context as the design of the Learning Study (Ling, Chik, and Pang, 2006) based on the learning theory of the Variation (Marton and Booth, 1997). Therefore I think at least parts of the results of my research could be useful for teaching and learning in a classroom.

The starting point to design the space of learning (see Chapter 5.3) is the conditions of learning according to the Marton's variation theory (see Figure10.1). I paid attention to the use of cognitive conflicts to guide pupils' thinking to the more scientific direction. Also the cognitive conflict approach stresses that the learner has to experience discernment, variation, and simultaneity about the objects of learning. In addition to discerning features and values of features one has to be able to discern parts within wholes (e.g. universal explanation that is valid in the pulley context) as well as wholes from

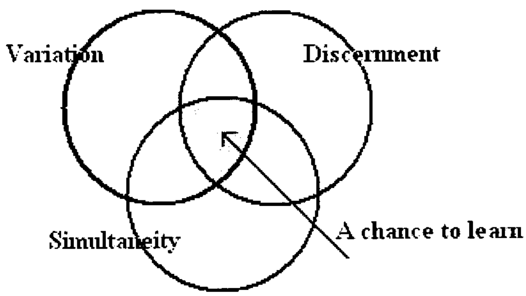


FIGURE 10.1. The conditions of learning according to the variation theory (Hakkarainen and Ahtee 2007).

their context (manual comparison, the pulley positions, and the beam balance). I emphasize especially that the teacher has to know what the pupils think about the learning subject, which are the pupils' misconceptions, and a possibility first to create to pupils a uniform preconception, which then is easier to change with the help of Marton's variation theory as a group with many different misconceptions. Pang (2003) observed that according to whether certain aspects were varied or kept constant, the learning was either supported or hindered. In the present study the results are parallel. In the same lesson if pupils on the Concrete Models did not notice any variation in the demonstrations, there was no learning during the learning episode. Those pupils who noticed variation on the Position Model changed their conceptions to the Motion Model.

I suggest that the Space of learning should be divided into two parts: Pre-activities and the Learning Intervention. In the part of the Pre-activities attention should be paid to pupils' prior knowledge. It is important that the teacher becomes aware of pupils' conceptions. Phenomenography reveals the different ways how pupils experience the same phenomena. In many cases there are such studies in literature, but if not it is possible to carry out one. The main result from the Phenomenography study is the qualitatively differing categories but it is also useful to take with the frequencies of the categories because if you try to change pupils' arguments towards the scientific trend it is sensible to start from the misconception with a great number of pupils. It is also important that pupils taking part in the Learning Intervention become aware of their own conceptions. Therefore, a revealing event that is familiar to the pupils should be used. This revealing event can act later to create the first cognitive conflict. The revealing event provokes also interest for pupils in the learning object. The teacher also has to show the working of a prosthetic device to help pupils to test their thinking in a cognitive conflict.

In the part of the Learning Intervention attention should be paid to pupils' teaching and learning. According to the Variation Theory, to discern a feature the pupil has to experience variation in that feature. In the Learning Intervention I used two types of variation: movement of the pulley (freely moving compared to not moving at all), and the changes in the position of the hanging objects. A teacher has to choose the variation so that the pupils who would use scientific model (Motion Model) in their predictions would not meet a cognitive conflict, but those pupils with a specific misconception (Position Model) would meet it. So the teacher would guide pupils with the help of successive cognitive conflicts to discern the critical aspect, and then solve the cognitive conflicts. Also a teacher at school can use this method in his ordinary work. Two events are not enough, instead according to this study there must be a series of familiar events or demonstrations for pupils. An obvious addition to my study is to change the sizes and materials of the objects. This is easier to do with an ordinary scale than in the pulley context. Through creating cognitive conflicts the teacher can help pupils to discern variation in the critical aspect.

Sometimes pupils do not recognize the conflict or a solution is proposed at a level which is beyond their understanding so that it will remain meaningless to them, and/or the effect of the conflict is lost, or the contradictory information can even be threatening to pupils, who do not have enough knowledge to solve the conflict. The use of cognitive conflict as an instructional strategy fails when the significance of the conflict event is not apparent to the pupils (Vosniadou 1990, Limon 2001). A firm misconception may hide the right information.

Let us consider the situation where a teacher uses a prosthetic device in his teaching and chooses an event that a pupil interprets according to his old and familiar idea. When the pupil's idea seems to give a satisfactory explanation he does not pay any more attention to other possibilities. It is, however, important that the pupil should test his thinking with other kind of situations like taking into account the working of the device. The teacher must create events so that testing is possible. If not the learning fails like with the 9th graders in demonstration series D1 D4 D2 D3 (see Publication III).

## REFERENCES

Ahonen, S. 1996. Fenomenografinen tutkimus. In L. Syrjälä, S. Ahonen, E. Syrjäläinen ja S. Saari. Laadullisen tutkimisen työtapoja, 113–160. Kirjapaino Oy West Point: Rauma. In Finnish.

Ahtee, M. 1992. Pupils' conceptions on optical phenomena and their utilization in teaching. Research reports no. 102. Helsinki, Finland: University of Helsinki, Department of Teacher Education. In Finnish.

Ahtee, M. 1994a. Pupil's conceptions on heat and temperature. Research reports no. 126. Helsinki, Finland: University of Helsinki, Department of Teacher Education. In Finnish.

Ahtee, M. 1994b. Secondary pupils' conceptions of dissolving and burning. Research reports no. 136. Helsinki, Finland: University of Helsinki, Department of Teacher Education. In Finnish.

Ahtee, M. and Hakkarainen O. 2005. Importance of the order of demonstrations in changing pupils' conceptions. NorDiNa, Nordic Studies in Science Education, 1, 31- 42.

Arons, A. 1997. Teaching Introductory Physics. New York: John Wiley and Sons.

Asunta, T. and Hakkarainen, O. 1996. Kokeellista fysiikkaa ja kemiaa peruskoulun yläasteelle. Journal of Teacher Researcher, 96, 1. In finnish.

Atkinson, D. and Peijnenburg, J. 2004. Galileo and prior philosophy. Studies in History and Philosophy of Science, 35, 115-135.

Bao, L. and Redish, E. F. 2003. Educational assessment and underlying models of cognition. In W. E. Becker and M. L. Andrews (Eds.). The scholarship of teaching and learning in post-secondary education: The contributions of research universities. Bloomington, IN: Indiana University Press.

Bar, V., Zinn, B., Goldmuntz, R. and Sneider, C. 1994. Cildren's concepts about weight and free fall. Science Education, 78, 2, 149-169.

Berg, T. and Brouwer, W. 1991. Teacher awareness of student alternative conceptions about rotational motion and gravity. Journal of Research in Science Teaching, 28, 3-18.

Bliss, J. and Ogborn, J. 1994. Force and motion from the beginning. Learning and Instruction, 4, 7-25.

Bowden, J. and Marton, F. 1998. The University of learning. Beyond quality and competence. London: Kogan Page.

Brown, D.E. 1993. Refocusing core intuitions: A concretizing role for analogy in conceptual change. Journal of Research in Science Teaching, 30, 1273-1290.

Brown, D.E. and Clement, J. 1989. Overcoming misconceptions by analogical reasoning: abstract transfer versus explanatory model construction. Instructional Science, 18, 237-261.

Carey, S. 1985. Conceptual change in childhood. Cambridge, MA: MIT Press.

Champagne, A., Klopfer, L. and Anderson, J. 1980. Factors influencing the learning of classical mechanics. American Journal of Physics, 48, 1074-1079.

Champagne, A.B., Gunstone, R.F. and Klopfer, L.E. 1985. Effecting changes in cognitive structures among physics students. In L. West and A. Pines (Eds.). Cognitive Structure and Conceptual Change. Orlando: Academic Press.

Chi, M.T.H. 1992. Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.). Cognitive models of science, 129-186. Minneapolis, MN: Minnesota Press.

Chi, M.T.H. 2005. Commonsense conceptions of emergent processes: Why some misconceptions are Robust. The Journal of the Learning Sciences, 14, 4, 161-199.

Chi, M.T.H. and Slotta, J. D. 1993. The Ontological Coherence of Intuitive Physics. Cognition and Instruction, 10, 2-3, 249-260.

Chi, M., Slotta, J. and de Leeuw, N. 1994. From things to processes: A theory of conceptual change for learning science concepts. Learning and Instruction, 4, 27-43.

Chinn, C.A. and Brewer, W.F. 1998. An empirical test of a taxonomy of responses to anomalous data in science. Journal of Research in Science Teaching, 35, 623-654.

Chiu, M-H. and Lin, J-W. 2005. Promoting Fourth Graders' Conceptual Change of Their Understanding of Electric Current via Multiple Analogies. Journal of Research in Science Teaching, 42, 4, 429-464.

Clement, J. 1993. Using Bridging Analogies and Anchoring Intuitions to Deal with Students' Preconceptions in Physics. Journal of Research in Science Teaching, 30, 10, 1241-1257.

Clement, J., Brown, D.E. and Zietsman, A. 1989. Not all preconceptions are misconceptions: finding 'anchoring conceptions' for grounding instruction on students' intuitions. International Journal of Science Education, 11, 554-565.

Cohen, L., Manion, L. and Morrison, K. 2000. Reseach Methods in Education. Great Britain: Routledge Falmer.

Coll, R.K. and Treagust, D.F. 2003. Learners' mental models of metallic bonding. Science Education, 87, 5, 685-707.

Cosgrove, M. and Osborne, R. 1985. Lesson Frameworks for Changing Children's Ideas. In R. Osborne and O. Freyberg (Eds.). Learning Science: The implications of children's science. Auckland: Heinemann.

Curtis, R.V. and Reigeluth, C.M. 1985. The use of analogies in written text. Instructional Science, 13, 2, 99-117.

Dagher, Z.R. 1994. Does the use of analogies contribute to conceptual change? Science Education, 78, 601-614.

Dagher, Z. R. 1998. The case for analogies in teaching science for understanding. In J.J. Mintzes, J.H. Wandersee, and J.D. Novak (Eds.). Teaching science for understanding: A human constructivist view, 195-211, Boston: Academic Press.

Dekkers, P.J.J. and Thijs, G.D. 1998. Making productive use of students' initial conceptions in developing the concept of force. Science Education, 82, 31-51.

Demastes, S., Good, R. and Peebles, P. 1996. Patterns of conceptual change in evolution. Journal of Research in Science Teaching, 33, 407-431.

diSessa, A. 1993. Toward an epistemology of physics. Cognition and Instruction, 10, 2-3, 105-225.

Dreyfus, A., Jungwirth, E. and Eliovitch, R. 1990. Applying the "cognitive conflict" strategy for conceptual change: Some implications, difficulties, and problems. Science Education, 74, 555-569.

Driver, R. 1983. The pupils as scientist? Milton Keynes, UK: Open University Press.

Driver, R., Guesne, E. and Tiberghien, A. 1985. Children's Ideas in science. Milton Keynes, UK: Open University Press.

Driver, R., Squires,A., Rushworth, R. and Wood-Robinson, V. 1994. Making Sense of Secondary Science: Research into Children's Ideas. London, UK: Routledge.

Driver, R., Leach, J., Millar, R., and Scott, P. 1996. Young People's Images of Science. Buckingham, UK: Open University Press.

Duit, R. 1991. On the role of analogies and metaphors in learning science. Science Education, 75, 649-672.

Duit, R. 2006. Bibliography - STCSE. Students' and Teachers' Conceptions and Science Education. On line at http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html. Accessed 29.1.2007.

Duit, R., Roth, W-M., Komorek, M. and Wilbers, J. 2001. Fostering conceptual change by analogies - between Scylla and Chafybdis. Learning and Instruction 11, 283-303.

Duit, R. and Treagust, D. 2003. Conceptual change. A powerful framework for improwing science teaching and learning. International Journal of Science Education, 25, 671-688.

Dykstra, D. A. 1992. Studying conceptual change: Constructing new understandings. In R. Duit, F. Goldberg and H. Niedderer (Eds.). Research in physics learning: Theoretical issues and empirical studies. Proceedings of an international workshop, 40-58, Kiel, Germany: IPN - Leibniz.

Dykstra, D.I., Boyle, C.F. and Monarch, I.A. 1992. Studying conceptual change in learning physics. Science Education, 76, 615-652.

Enderstein, L. and Spango, P. 1996. Beliefs regarding force and motion: alogitual and cross-cultural study of South African school pupils. International Journal of Science Education, 18, 4, 479-494.

Fensham, P.J., Gunstone, R.F. and White, R.T. 1994. The content of science: A constructivist approach to its teaching and learning. London: Falmer Press.

Fraser, B.J. and Tobin, K. G. 1998. International Handbook of Science Education. London: Kluwer Academic Publishers.

Galili, I. 2001. Weight versus gravitational force: historical and educational perspectives. International Journal of Science Education, 23, 10, 1073-1093.

Galili, I. and Bar, V. 1992. Motion implies force: where to expect vestiges of the misconceptions? International Journal of Science Education, 14, 1, 63-81.

Galili, I. and Bar, V. 1997. Children's operational knowledge about weight. International Journal of Science Education, 19, 3, 317-340.

Gentner, D. 1983. Structure mapping: theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gentner, S. and Stevens A. L. 1983. Mental models. Hillsdale, NJ: Erlbaum.

Gick, M.L. and Holyoak, K.J. 1983. Scema induction and analogical transfer. Gognitive Psychology, 15, 1-38.

Glynn, S.M. 1991. Explaining science concepts: a teaching-with-analogies model. In S. Glynn, R.Yeany and B. Britton (Eds.). The Psychology of Learning Science, 219-240. New Jersey: Erlbaum.

Glynn, S.M., Duit, R. and Thiele, R.B. 1995. Teaching Science w ith Analogies: A Strategy for Constructing Knowledge. In S.M. Glynn and R. Duit (Eds.). Learning Science in the Schools, Research Reforming Practice, 247-273. New Jersey: Erlbaum.

Greca, I. M. and Moreira, M. A. 1997. The kinds of mental representations-models, propositions and images-used by college physics students regarding the concept of field. International Journal of Science Education, 19, 6, 711- 724.

Greca, I.M. and Moreira, M.A. 2000. Mental models, conceptual models and modelling. International Journal of Science Education, 22, 1, 1-11.

Greca, I.M. and Moreira, M.A. 2002. Mental, Physical, and Mathematical Models in the Teaching and Learning of Physics. Science Education, 86, 1, 106-121.

Greenslade Jr, T.B. 1985. Atwood's machine. The Physics Teacher, 23, 1, 24-18.

Gunstone, R. and White, R. 1981. Understanding of gravity. Science Education, 65, 3, 291-299.

Hakkarainen, O. 1995. Spektrin valokuvaaminen. Dimensio, 59, 6, 38. In Finnish.

Hakkarainen, O. 2000. Demonstraatioiden järjestyksen vaikutus oppimiseen. Dimensio, 64, 4, 34-37. In Finnish.

Hakkarainen, O. 2005. Mental models in manual weight comparison between the objects of different size. Themes in Education, 6, 2, 151-167.

Hakkarainen, O. and Ahtee, M. 2005. Pupils' mental models of a pulley in balance. Journal of Baltic Science Education, 2005, 2, 8, 26-34.

Hakkarainen, O. and Ahtee, M. 2007. The durability of conceptual change in learning the concept of weight in the case of a pulley in balance. International Journal of Science and Mathematics Education, 5, 3, 461-483.

Halloun, I. and Hestenes, D. 1985. Common sense concepts about motion. American Journal of Physics, 53, 11, 1056-1065.

Harrison, A.G. and Treagust, D.F. 1993. Teaching with analogy: A casa study in grade-10 optics. Journal of Research in Science Teaching, 30, 1291-1307.

Hasselgren, B. and Beach, D. 1997. Phenomenography a "good-for-nothing brother" of phenomenology? Higher Education Research &Development, 16, 2, 191-202.

Havu, S. 2000. Changes in Children's conceptions through Social Interaction in Pre-School Science Education. University of Joensuu. Joensuu.

Hecht, E. 2006. There is No Really Good Definition of Mass. The Physics Teacher, 44, 40-135.

Hewson, M.G. and Hewson, P.H. 2003. Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. Journal of Research in Science Teaching. 40, supplement, s86-s98.

Hewson, P.W. and Hewson, M.G. 1992. The status of sudents' conceptions. In R. Duit, F. Goldberg and H. Niedderer (Eds.). Research in physics learning: Theoretical issues and empirical studies, 59-73. Kiel: IPN.

Hewson, P.W. and Thorley, N.R. 1989. The conditions of conceptual change in the classroom. Internal Journal of Science Education, 79, 147-166.

Holyoak, K.J. and Thagard, P. 1995. Mental leaps: Analogy in creative thought. Cambridge, MA: The MIT Press.

Ioannides, C. and Vosniadou, S. 2002. The changing Meanings of Force. Cognitive Science Quarterly, 2, 1, 5-62.

Ivarsson, J., Schultz, J. and Säljö, R. 2002. Map reading versus mind reading: Revisiting children's understanding of the shape of the earth. In M. Limon and L. Mason (eds.). Reconsidering conceptual change: Issues in theory and practice. Kluwer Academic Publishers.

Jackson, P.W. 1966. The way Teaching is. Washington: National Education Association.

Jammer, M. 1957. Concepts of force. Cambridge, Massachusetts: Harvard University Press.

Jammer, M. 1961. Concepts of mass in classical and modern physics. Cambridge, Massachusetts: Harvard University Press.

Jammer, M. 1997. Concepts of mass in classical and modern physics. Mineola, New York: Dower Publications.

Jammer, M. 1999. Concepts of force. Mineola, New York: Dover Publications.

Johnson-Laird, P.N. 1983. Mental models. Cambridge, Great Britain: Cambridge University Press.

Kang, S., Scharmann, L.C. and Noh, T. 2004. Reexamining the role of cognitive conflict in science concept learning. Research in Science Education, 34, 71-96.

Kansanen, P. 1999. Teaching-Studying-Learning process. Scandinavian Journal of Educational Research, 43, 1, 81-89.

Kuhn, T.S. 1970. The Structure of Scientific Revolutions. Chicago: The University of Chicago Press.

Kurki-Suonio, K. 2005. Massa opetuksen näkökulmasta. Arkhimedes, 4, 12-15. In Finnish.

Kurki-Suonio, K and Kurki-Suonio, R. 1995. Vuorovaikuttavat kappaleet, 143-144, and 323-326. Helsinki: Limes ry. In Finnish.

Key, J.P. 1997. Research Design in Occupational Education: Module S7-Chi Square. Oklahoma, USA: Oklahoma State University.

Lee, G., Kwon, J., Park, S.S., Kim, J.W., Kwon, H.G. and Park, H.C. 2003. Development of an instrument for measuring cognitive conflict in secondary-level science classes. Journal of Research in Science Teaching, 40, 585-603.

Limón, M. 2001. On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. Learning and Instruction, 11, 357-380.

Ling, L.M., Chik, P. and Pang, M.F. 2006.Patterns of variation in teaching the colour of light to Primary 3 students. Instructional Science, 34, 1-19.

Linder, C.J. 1993. A challenge to conceptual change. Science Education, 77, 293-300.

Magnani, L and Nersessian, N. 2002. Mode-Based Reasoning, Scientific Discovery, Technological Innovation, Values. New York: Kluwer Academic/Plenum Publishers.

Malinen, P. 1997. Tutkitaanko koulussa oppimista vai opiskelua? Kasvatus, 28, 2, 191-196. In Finnish.

Maloney, D.P. and Siegler, R.S. 1993. Conceptual competition in physics learning. International Journal of Science Education, 15, 283-295.

Marton, F. 1994. Phenomenography. The International Encyclopedia of Education. Second edition, Vol. 8. (Eds.). Torsten Husén and T. Neville Postlethwaite. Pergamon 1994, pp. 4424–4429. On line at: http://www.ped.gu.se/biorn/phgraph/civil/main/2res.appr.html

Marton, F. and Booth, S. 1997. Learning and awareness. New Jersey: Lawrence Erlbaum.

Marton, F. and Tsui, A.B.M. 2004. Classroom discourse and the space of learning. New Jersey: Lawrence Erlbaum.

Marton, F., Runesson, U., and Tsui, M. 2004. The space of learning. In F. Marton, and B.M. Tsui (Eds.). Classroom discourse and the space of learning. Mahwah: Lawrence Erlbaum Associate.

Mayer, R.E. 2002. Understanding conceptual change: A commentary. In M. Limón and L. Mason (Eds.). Reconsidering Conceptual Change Issues in Theory and Practice, 101-111. Printed in the Netherlands: Kluwer Academic Publishers.

McDermott, L.C. and Redish, E.F. 1999. Resource Letter: Per-1: Physics Education Research. American Journal of Physics, 79, 9, 755-767.

Mildenhall, P.D. and Williams, J. S. 2001. Instability in students' use of intuitive and Newtonian models to predict motion: the critical effect of the parameters involved. International Journal of Science Education, 23, 6, 643-660.

Minstrel, J. 1982. Explaining the 'at rest' condition of an object. The Physics Teacher, 20, 10-14.

Mohapatra, J. V. and Bhattachaayya, S. 1989. Pupils' and teachers' induced incorrect generalization and the concept of force. International Journal of Science Education, 11, 4, 429-426.

Niaz, M. 1995. Relationship between students performance on conceptual and computational problems of chemicals equilibrium. International Journal of Science Education, 17, 3, 343-355.

Niedderer, H. 1987. A teaching strategy based on students' alternative frameworks -theoretical conceptions and examples. In J.D. Novak (Ed.). Proceedings of Second International Seminar Misconceptions and educational strategies in science and mathematics. Vol. II. Cornell University, Ithaca, NY, USA, 360-367.

Niedderer, H. and Goldberg, F. 1994. An individual student's learning process in electric circuits. On line at:

http://www.ind.unibremen.de/bubs/Niedderer/1994-NARSLS.pdf
Accessed 21.2.2007

Niedderer, H and Schecker, H. 1992. Towards an explicit description of cognitive systems for research in physical learning. Research in Physics Learning -Theoretical Issues and Empirical Studies, 74-98. Proceedings of an Internal Workshop in Bremen, Kiel: IPN. On line at:

http://www.idn.uni-bremen.de/pubs/Niedderer/1992-HB-HN-HS.pdf
Accessed 24.2.2007.

Noce, G., Torosantucci, G. and Vicentini, M. 1988. The floating of objects on the moon: prediction from a theory or experimental facts. International Journal of Science Education, 10, 1, 61-70.

Novak, J.D. 1987. Misconceptions and educational strategies in science and mathematics. Ithaca, NY, USA: Cornell University.

Novak, J.D. Eds. 1993. Proceedings of Third International Seminar Misconceptions and Educational Strategies in Science and Mathematics. Ithaca, NY, USA: Cornell University.

Nussbaum, J. and Novick, S. 1982. Alternative frameworks, conceptual conflict and accommodation: toward a principled teaching strategy. Instructional Science, 11,183-200.

OECD-PISA 2000. 2002. The Finnish success in Pisa and some reasons behind it. Jyväskylän yliopisto: Kirjapaino Oma Oy.

Osborne, H. and Freyberg, P. 1985. Learning in Science: The implications of children's science. Aucland: Heinemann.

Palmer, D. 2001. Students' alternative conceptions and scientifically acceptably conceptions about gravity. International Journal of Science Education, 23, 7, 691-706.

Pang, M. F. 2003. Two Faces of Variation: continuity in the phenomenographic movement [1]. Scandinavian Journal of Educational Research, 47, 2, 2003.

Pfundt, H. and Duit, R. 2001. Bibliography: Students' alternative frameworks and science education. Kiel, Germany: IPN at the University of Kiel.

Piaget, J. 1972. The Child's Conception of Physical Causality. Totowa, NJ: Littlefield, Adams.

Piaget, J. and Inhelder, B. 1974. The Child's Construction of Quantity. London: Routledge and Kegan Paul.

Pintrich, P.R. 1999. Motivational Beliefs's as Resources for and Constrains on Conceptual Change. In W. Snottz, S. Vosniadou and M. Carretero (Eds.). New Perspectives on Conceptual Change, 33-50. Oxford: Pergamon and Earli.

Pintrich, P.R., Marx, R.W. and Boyle, R.A. 1993. Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. Review of Educational Research, 63, 167-199.

Posner, G., Strike, K., Hewson, P. and Gretzog, W. 1982. Accommodation of a scientific conception: Toward theory of conceptual change. Science Education, 66, 211-227.

Rowell, J.A. and Dawson, C.J. 1985. Equilibration, conflict and instruction: A new class-oriented perspective. European Journal of Science Education 4, 4, 331-344.

Runesson, U. 1999. Teaching as constituting a space of variation. Paper presented at the 8th EARLY- conference, Göteborg, Sweden. On line at: http://www.ped.gu.se/biorn/phgraph/civil/graphica/ur.pdf

Saari, H. 2000. Oppilaiden käsitykset malleista ja mallintaminen fysiikan peruskouluopetuksessa. University of Joensuu, Joensuu. In Finnih.

Saari, H and Viiri, J. 2003. A research-based teaching sequence for teaching the concept of modelling to seventh-grade students. Internal journal of science education, 25, 11, 1333-1352.

Sasse, A. 1997. Eliciting and Describing Users' Models of Computer Systems. Dissertation, Birmingham, England: University of Birmingham. On line at: http://www.cs.ucl.ac.uk/staff/a.sasse/Frontpage

Savinainen, A. 2004. High School Students' Conceptual Coherence of Qualitative Knowledge in the Case of the Force Concept. University of Joensuu.

Savinainen, A., Scott, P. and Viiri, J. 2005. Using a bridging representation and social interactions to foster conceptual change: Designing and evaluating an instructional sequence for Newton's third law. Science Education, 89, 2, 179-195.

Scott, P., Asoko, H.M. and Driver, R.H. 1991. Teaching for conceptual change: A review of strategies. In R. Duit, R., H. Golgberg, and H. Niedderer (Eds.). Research in Physics Learning. Theoretical Issues and Empirical Studies. Kiel, Germany: IPN at the University of Kiel.

Schnotz, W., Vosniadou, S. and Carretoro, M. 1999. New Perspectives on Conceptual change. Advances in learning and instruction series. Oxford: Pergamon, Earli.

Schoultz, J., Säljö, R. and Wyndhamn, J. 2001. Heavenly talk: Discourse, artifacts, and children's understanding of elementary astronomy. Human Development, 44, 103-118.

Sequira, M and Leite L. 1991. Alternative Conceptions and History of Science in Physics Teacher Education. Science Education, 75, 1, 45-56.

Shulman, L.S. 1986. Paradigms and research Programs in the Study of Teaching: A Contemporary Perspective. In M.C. Wittrock, (Ed.). Handbook of Research on Teaching. A project of the American Educational research association, 3-36. Canada: Macmillan Publishing Company.

Shulman, L.S. 1999. Knowledge and Teaching: Foundation of the New Reform. In J. Leach and B. Moon (Eds.). Learners and Pedagogy, 61-77. London: Open University.

Sinatra, G.M. and Pintrich, P. R. 2003. The role of intentions in conceptual change learning. In G.M. and P.R. Pintrich (Eds.). Intentional Conceptual Change. Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, R.G. and Peacock, G. 1992. Tackling contradictions in teachers' understanding of gravity and air resistance. Evaluation and Research in Education, 6, 113-127.

Smith, S., Carey, S. and Wiser, N. 1985. On differentiation a case study of development of the concept of size, weight and density. Cognition, 21, 177-277.

Spelke, S.E. 1991. Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey and R. Gelman (Eds.). The epigenesis of mind: Essays on biology and cognition, 33-170. Hillsdale, NJ: Lawrence Erlbaum Associates.

Spiro, R.J., Feltovich, P.J., Coulson, R.L. and Anderson, D.K. 1989. Multiple analogies for complex concepts: Antidotes for analogy-induced misconceptions in advanced knowledge acquisition. In S, Vosniadou and A. Ortony (Eds.), Similarity and Analogical Reasoning, 498-531, New York: Cambridge University Press.

Stavy, R. 1990. Pupils' problems in understanding conservation of matter. International Journal of Science Education, 12, 5, 501-512.

Stavy, R. 1991. Using analogy to overcome misconptions about conservations of matter. Journal of Research in Science Teaching, 28, 503-313.

Stavy, R. and Berkovitz, B. 1980. Cognitive conflict as abasis for teaching quantitative aspects of the concept of temperature. Science Education, 64, 679-692.

Stavy, R. and Tirosh, D. 1993. When analogy is Perceived as Such. Journal of Research in Science Teaching, 30, 10, 1229-1239.

Steinberg, M. S., Brown, D. E. and Clement, J. 1990. Genius is not immune to persistent misconceptions: conceptual difficulties impeding Isaac Newton and contemporary physics students. International Journal of Science Education, 12, 3, 265-273.

Strike, K. A. and Posner, G. J. 1982. Conceptual change and science teaching. European Journal of Science Education, 4, 231-240.

Tao, P-K and Gunstone, R.F. 1999. The process of conceptual change in Force and Motion during computer supported physics instruction. Journal of research in Science Teaching, 36, 859-882.

Treagust, D.F., Harrison, A.G. and Venville, G.J. 1996. Using analogical teaching approach to engender conceptual change. International Journal of Science Education, 18, 213-229.

Treagust, D. Duit, R. and Fraser, B. 1996. Improving Teaching and Learning in Science and Mathematics. New York: Teacher College Press.

Trumper, R. 1997. Applying conceptual conflict strategies in the learning of the energy concept. Research in Science Technology Education, 15, 5-18.

Tsai, C.C. 2000. Enhancing science instruction: the use of 'conflict maps'. International Journal of Science Education, 22, 285-302.

Tsai, C.C. and Chang, C.Y. 2005. Last effects of instruction guided by the conflict map: Experimental study of learning about the causes of the seasons. Journal of Research in Science Teaching, 42, 1089-1111.

Tyson, L.M., Venville, G.J., Harrison, A.G. and Treagust, D.F. 1997. A multidimensional framework for interpreting conceptual change events in the classroom. Science Education, 81, 387-404.

Viennot, L. 1979. Spontaneous reasoning in elementary dynamics. European Journal of Science Education, 1, 205-221.

Viennot, L. 2001. Physics Education Research: Inseparable Contents and Methods- The Part Played by Critical Details. In M. Ahtee, O. Björkqvist, E. Pehkonen and V. Vatanen (Eds.). Research on Mathematics and Science Education, 89-100. University of Jyväskylä, Jyväskylä.

Viiri, J. 1995. Teaching the Force Concept: A Constructivist teaching Experiment in Engineering Education. University of Joensuu, Joensuu. In Finnish.

Vosniadou, S. 1994. Capturing and modelling the process of conceptual change. Learning and Instruction, 4, 45-69.

Vosniadou, S. 1999. Conceptual change research: State of the art and future directions. In W. Schnotz, S. Vosniadou and M. Carretero (Eds.). New perspectives on conceptual change, pp. 3-13. Amesterdam: Pergamon.

Vosniadou, S. 2001. Mental models in conceptual development. In L. Magnani and N.J. Nersessian (Eds.). Model-based reasoning: Science, technology, values. Dordrecht: Kluwer Academic Publishers.

Vosniadou, S. 2003. Exploring the relationship between conceptual change and intentional learning. In G. M. Sinatra and P. R. Pintrich (Eds.). Intentional Conceptual Change, 377-406. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Vosniadou, S. and Ioannides, C. 1998. From conceptual development to science education: A psychological point of view. International Journal of Science Education, 20, 1213-1230.

Vosniadou, S. Ionnides, C., Dimitrakopoulou, A. and Papademetriou, E. 2001. Designing learning environments to promote conceptual change in science. Learning and Instruction, 11, 381-419.

Watts, D.M. 1982. Gravity: don't take it for granted. Physics Education, 17, 119-21.

Watts, D.M. and Zylbersztajn, A. 1981. A survey of some children's ideas about force. Physics education, 16, 360-365.

Wong, E.D. 1993. Understanding the Generative Capacity of Analogies as a Tool for Explanation. Journal of Research in Science Teaching, 30, 19, 1259-1272.