

Maija Ahtee is Ph.D. professor (emerita) in mathematics and science education in the Department of Teacher Education, University of Jyväskylä. She has taught mathematics and physics in senior secondary and secondary schools before becoming a lecturer in physics, in the department of physics and then a lecturer in physics and chemistry education, department of Teacher Education in the University of Helsinki. She was nominated as professor in mathematics and science education, Department of Teacher Education, University of Jyväskylä and retired from that position.

Olavi Hakkarainen is M.Sc. student in the Finnish Graduate School of Mathematics, Physics, and Chemistry Education, University of Jyväskylä. He works as the lecturer in mathematics, physics, and chemistry in the secondary school in Tikkakoski.

MAIJA AHTEE

Department of Applied Sciences of Education, University of Helsinki
maija.ahtee@helsinki.fi

OLAVI HAKKARAINEN

Department of Physics, University of Jyväskylän yliopisto
Olavi.Hakkarainen@jklmlk.fi

Importance of the order of demonstrations in changing pupils' conceptions

Abstract

Conceptual change is concerned 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 for 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. 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.

INTRODUCTION

George Atwood (1746-1807) developed an apparatus, the so-called Atwood's machine, to demonstrate Newton's second law of motion to students. Atwood's machine is in principle a pulley that shows the effect of gravity on the accelerating motion. In this study we have used a pulley to help pupils to become aware of their misconceptions related to the effects of gravity and thus to promote their conceptual change towards scientific thinking.

There are many studies that show the various conceptions pupils bring with them to the classroom when they start learning science (Pfundt & Duit 2001). Teaching methods based on e.g. conceptual conflict and analogies have been developed to change these pupils' ideas (Driver, Asoko, Leach, Mortimer & Scott 1994). Changing pupils' preconceptions is not easy and it takes a long time. It has also been pointed out that cognitive conflict strategies do not always lead to conceptual change (Dekkers & Thijs 1998). Familiar and often used explanations come first to mind when pupils are making predictions of new events or in new situations.

REVIEW OF LITERATURE

According to the constructivist view of learning the pupil is seen as an active builder of her knowledge. In order for learning to take place it is necessary to connect new knowledge with the student's existing knowledge structure. Scientific explanations of physical phenomena often differ from and are in conflict with intuitive ideas based on everyday experiences. Conceptual change is related to how a teacher can help students to connect their prior knowledge with the new content to be learned. Theoretical models to explain conceptual change have been developed (Posner, Strike, Hewson & Gertzog 1982; diSessa 1993; Chi, Slotta & de Leeuw 1994; Vosniadou 1994; Vosniadou & Ioannides 1998). According to these models learning requires the significant reorganization of existing knowledge structures. Cognitive conflict, i.e. the realisation of the need to change existing ideas, is the first step toward conceptual change. In addition, students need to be convinced that the scientific conceptions are more intelligible, plausible, and fruitful than their own conceptions (Posner *et al.* 1982). Tao and Gunstone (1999) found, when studying the process of conceptual change in force and motion using computer simulation programs, that conceptual change for many students was context dependent and unstable. In addition, it has been stressed that also motivational and affective factors as well as recognition that science is socially constructed has to be taken into account in models of conceptual change (Pintrich, Marx & Boyle 1993; Tyson, Venville, Harrison & Treagust 1997; Lee, Kwon, Park, Kim, Kwon & Park 2003).

Vosniadou (1994, 2001) defines mental models as a form of mental representation that preserves the structure of that which it represents. These dynamic and evolving models can be manipulated and changed by the individual. Pupils use them particularly in situations where implicit physical knowledge needs to be exploited for the purpose of answering questions, solving problems or in order to understand incoming information. Vosniadou assumes that most mental models are created on the spot to deal with the demands of specific problem-solving situations. However, some mental models that have proven useful in the past are stored as separate structures and retrieved from long-term memory when needed. Chinn and Brewer (1998) have pointed out that anomalous data are a key component of many instructional methods of a constructivist nature in fostering knowledge change in science. They have also stated that anomalous data often fail to promote conceptual change as students may discredit or discount them.

Understanding physics demands concentration and hard work from a pupil because of the reorganization of existing conceptual structures and the creation of new representations. Bao and Redish (2003) point out that different learning contexts may lead to different mental processes, either to activation or to creation. If the setting of a context is familiar to pupils from their experiences, this context can activate existing knowledge that is then applied to form a response. If, however, there are no good matches with the experiences, it is then more likely that the pupils will create a new type of explanation on the spot. The dominant process depends on the familiarity of the content knowledge.

Earlier fixed conceptions hinder a new way of thinking. When people meet a new phenomenon they generally try to apply old and familiar ideas to explain it. When an idea seems to be able to give a satisfactory explanation of at least some features attention is not paid to other possibilities any more. It is important to test explanations in many different situations and change the variables affecting the phenomenon. In this study we wanted to look at how existing and familiar explanations are used when pupils try to find a more proper explanation for a phenomenon.

Pupils' conceptions of mass, weight and force

Children acquire a simple conception of weight at a very early stage of their development by picking up, holding and dropping different objects. They do not distinguish mass from weight until they have been taught that the weight of an object means the gravitational force exerted by the earth on the object. On the other hand, common sense conceptions of weight for example as a pressing force on one's palm are learned through personal experiences without any deliberate teaching.

Gradually, with maturation, the conception of weight as quantity of matter develops (Bar, Zinn, Goldmuntz & Sneider 1994). For many pupils in comprehensive school the more undifferentiated concept of heaviness is closer to their understanding of weight (Hewson & Hewson 2003). Some objects like brick have weight while others like pin or hair do not. Children have also other misconceptions like if the shape of an object changes then also its weight changes, or the amount of matter depends on the form of the container inside which the matter is (Piaget & Inhelder 1974).

According to Galili and Bar (1997) younger children emphasize the *active* nature of gravity, which manifests itself as a sensation of weight (a pressing force) and as the ability of an object to perform certain actions such as breaking other objects when placed or dropped upon them. 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. Accordingly pupils equate the weight of an object with a sensory experience of weight (such as a pressing force or the feeling of increased weight caused by acceleration). Pupils may also equate the sensation of weight with the result of weighing an object on scales. In this case some pupils consider the weight, which is sensed, to be equal to the experimentally measured weight. Galili (2001) argues that normal instruction can not resolve the fundamental confusion between weight, gravitational force and weighing results.

The development of pupils' force concept has been studied by Ionniades and Vosniadou (2002). They grouped the observed meanings into two categories: those that appear to be based on everyday experience (initial theories) and show no influences from Newtonian theory, and those that have been influenced by instruction or by scientific theory (synthetic theories, Vosniadou 1994). The younger children thought that force is an internal property of big and heavy objects related to their weight (internal force meaning), while the 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. The *hybrid* meanings of force are created spontaneously not influenced by instruction. About at the age of fifteen as a result of instruction, various synthetic meanings of force were created through assimilation of the notions of gravitational force and the push/pull force.

A common misconception concerns the relationship between motion and force. It has been shown in many studies that even university students follow the Aristotelian idea that a continuous action of force is needed to keep an object in motion (McCloskey 1983; Clement 1982; Jimoyiannis & Komis 2003).

Young children when comparing two objects of equal weight but different volume will claim that the larger (less dense) object is lighter. This everyday conception was challenged with the aid of a pulley in balance. The subsequent demonstrations, however, contain other common preconceptions for example that the lower position indicates the heavier of the two bodies suspended in the pulley. Gunstone and White (1981) found with a bicycle wheel fastened as a pulley that 27% of first university physics students reasoned a lower hanging block of wood to be heavier than a higher hanging bucket of sand (of the same mass). Some of these students even draw inappropriate analogies to seesaws or beam balances. In Mohapatra's and Bhattacharyaa's (1989) pencil-and-paper test about 60% of the ninth graders stated that the downward force on the lower hanging body was more than that on the higher hanging body even though 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. Hakkarainen and Ahtee (2005) have explored how pupils in 5th, 7th and 9th grade understand forces in equilibrium using the pulley in balance demonstration. They found that less than 10% of the pupils understood that the hanging bodies had the same weight.

These pulley studies give information about pupils' misconceptions but they do not suggest how to change these conceptions. In this study we tried to find out how pupils' thinking works when they try to figure out an explanation and also to find factors that promote or hinder learning. We wanted also to find out if it is possible to promote more versatile pupils' thinking with a series of demonstrations.

The research questions were:

1. How does the order of the pulley demonstrations affect the pupils' resulting conception?
2. How do the pupils' statements about weight change when the order of the demonstration is changed?
3. How do the pupils' argumentations change when the demonstration is changed?
4. What are the specific difficulties the pupils encounter with a pulley in balance?

METHOD

Altogether 558 pupils from fifth and ninth grades (average age 11 and 15 years respectively) took part in this study. During a science lesson four demonstrations were introduced. In the first one (D1) pupils were asked to compare with their hands the weight of a small standard mass (20 g, about 2 cm³) and a bigger bag (20 g, about 200 cm³). In the next three demonstrations (D2 – D4) pupils had to compare the weight of these two objects suspended in a pulley in three different positions as seen in figure 1.

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 at 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 also the reasons why they thought so. The order of the demonstrations D2 to D4 was, however, 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 87,

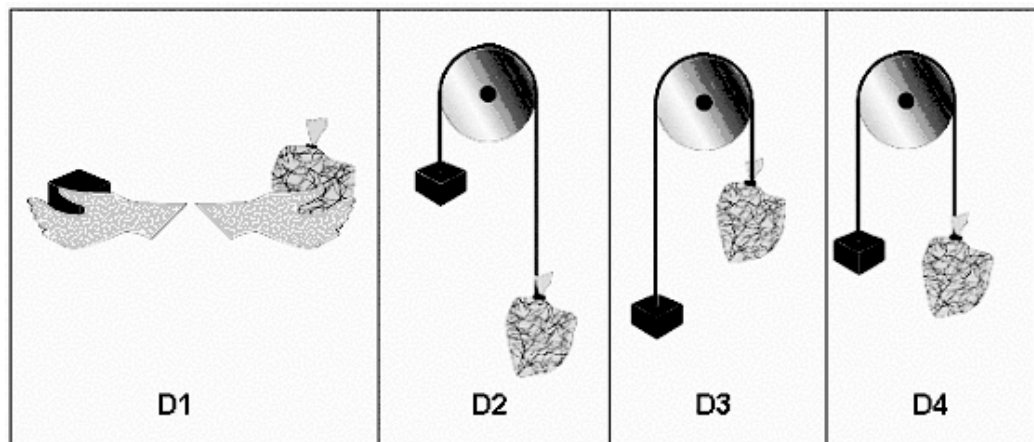


Figure 1. Demonstrations used in this study. D1: Compare with your hands the weight of a standard mass and a bag. D2 – D4: What can you say about the weight of the standard mass and the bag compared to each other? Why do you think so?

and ninth grade: group I 100, group II 88, group III 89. The purpose of the manual comparison in demonstration D1 was firstly to get the pupils to start thinking about the properties of the concepts *weight* and *volume* and to be aware of their own thinking and secondly to find whether there are any differences between the groupings in the fifth and ninth grade. A *chi-squared test* was used and it showed that all three groups in each grade were very alike (in the fifth grade: $\chi^2(4) = 4.61$; $p > 0.1$; in the ninth grade: $\chi^2(4) = 5.36$; $p > 0.1$).

In the different groups the pupils' responses were first classified after each demonstration to the following categories: A. Bag weighs more; B. Standard mass weighs more; C. Standard mass and bag weigh the same. The arguments were further divided into subcategories: 1. Motion; 2. Position; 3. Appearance; 4. Material; 5. No argument or confusing idea. Hakkarainen and Ahtee (2005) found these categories when they analysed the 5th, 7th and 9th graders' arguments about two balanced bodies hanging in a pulley. The category Material containing concrete properties of the hanging bodies was found only among the 5th graders. In the categories of Appearance, Position and Motion the pupils paid attention correspondingly to the size, position and movement of the bodies. In the following, short descriptions of the categories are given together with some examples of pupils' answers. Pupils' reasoning in a certain subcategory differs slightly according to the main category.

1. **Motion.** The scientifically nearly 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 (A). The bag has moved downwards (B). The bag and the standard mass stay at their positions (C).
2. **Position.** The argument is based on the positions of the bag and the standard mass.
The standard mass hangs higher (A). The bag hangs lower (B). The standard mass pulls the bag to the same level (C).
3. **Appearance.** Attention is paid to the concrete appearance of the bag and the standard mass.
The standard mass looks heavy (A). The bag is large (B).
4. **Material.** Concrete properties are given to the bag and the standard mass.
The standard mass is of metal and the bag is of plastic (A). The bag contains something (B).
5. **No argument or confusing idea.** In most cases pupils only stated their thought about the weights of the bag and of the standard mass compared to each other.
The bag is lighter than the standard mass (A). The bag is heavier than the standard mass (B). Because the bag is larger than the standard mass gravitation affects more on it and therefore it is lower than the standard mass even though it weighs the same (C).

In this study the changes in the pupils' answers were followed through the four successive demonstrations.

According to the earlier work by Hakkarainen (2005) the arguments based on the manual comparison were classified into four categories: 1. Pressure, 2. Material, 3. Feeling, 4. No argument or confusing idea as shown in table 1. The answer was classified under the category Pressure when the pupil seemed to notice the idea that the size or the area of the object against the palm has an effect to their estimation: *Since the weight of the standard mass is distributed over a smaller area its mass feels heavier although it probably is not heavier* (9th grader). When pupils based their estimation on the material or the content of the bag the answer was classified to the category Material. The category Feeling contains two ideas, the sensation of weight: It feels heavier (standard mass), or a result of a contemplation process: *It looks heavier* (the bag).

Table 1. The distribution of the 5th and 9th graders' arguments for their estimation after the manual comparison, in percentages. The number of pupils in the 5th grade was 272 and in the 9th grade 276.

Argument categories	5 th graders, %	9 th graders, %
Pressure	6	15
Material	43	30
Feeling	41	47
No argument	11	8

RESULTS

Pupil responses

The changes in the fifth graders' choices as to the weights of the hanging bodies after the demonstrations are shown in figure 2.

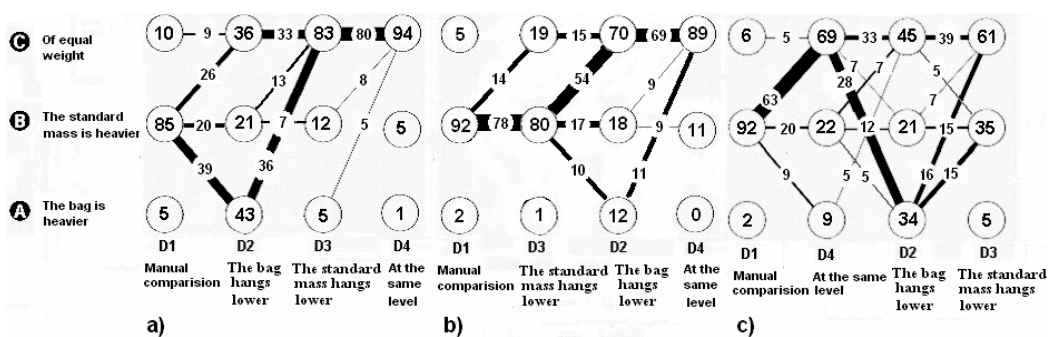


Figure 2. The distribution of the fifth graders' answers I) in group I (97 pupils) with the order of the demonstrations D1D2D3D4, II) in group II (91 pupils) with the order of the demonstrations D1D3D2D4, and III) in group III (87 pupils) with the order of the demonstrations D1D4D2D3. Numbers represent per cents of all the pupils. Only the choices exceeding 5% are shown.

After the manual comparison altogether 85% of the fifth graders in group I qualitatively estimated the standard mass to be heavier than the bag (figure 2). This is very natural as the pressure of the smaller object on the palm is higher than that of the larger object. When this demonstration series was planned it was assumed that after showing demonstration D2 many pupils may change their mind as they may think that the bigger the object the heavier it is. And nearly 50% of the pupils who had estimated the standard mass heavier changed their mind so that 43% of all the pupils in group I now stated the bag to weigh more. (The distribution of their arguments is shown in figure 4.) However, also the amount of the pupils who now concluded that the bodies must be of equal weight grew clearly. The fifth graders' choices after demonstration D2 changed significantly ($\chi^2(2) = 80.7$; $p < 0.001$) from the original estimations after manual comparison. After the last demonstration D4 nearly all the pupils (94%) seemed to be convinced that the standard mass and the bag weighed the same.

In group II the order of the demonstrations D2 and D3 was changed (figure 2). After demonstration D3 only some pupils realized that the bag and the standard mass were of equal weight whereas 80% of all the pupils in this group did not find any conflict between their original choice and the demonstration in which they saw the standard mass to hang lower than the bag. The main change

in the pupils' choices happened after demonstration D2. 70% of the pupils in group II concluded that the bag and the standard mass weighed the same. The pupils' choices after demonstration D2 differed significantly from their earlier D3 choices ($\chi^2(2) = 72.00$; $p < 0.001$). After the last demonstration D4 altogether 89% of the pupils made the correct choice. There was no difference between the choices of the pupils in groups I and II after the last demonstration ($\chi^2(1) = 1.401$; $p > 0.1$).

When demonstration D4 was shown as the first one (figure 2) almost 70% of the pupils were convinced that the bag and the standard mass weighed the same. However, almost 40% of them changed their mind after demonstration D2. It is interesting to notice that after demonstration series D4D2D3 less than two-thirds of the pupils reached the correct answer while one-third still believed that the standard mass is heavier than the bag. When the pupils' choices are compared after the last demonstration D4 in the series D2D3D4 and D3D2D4 with those after the last demonstration D3 in the series D4D2D3 the difference is significant ($\chi^2(1) = 37.46$; $p < 0.001$).

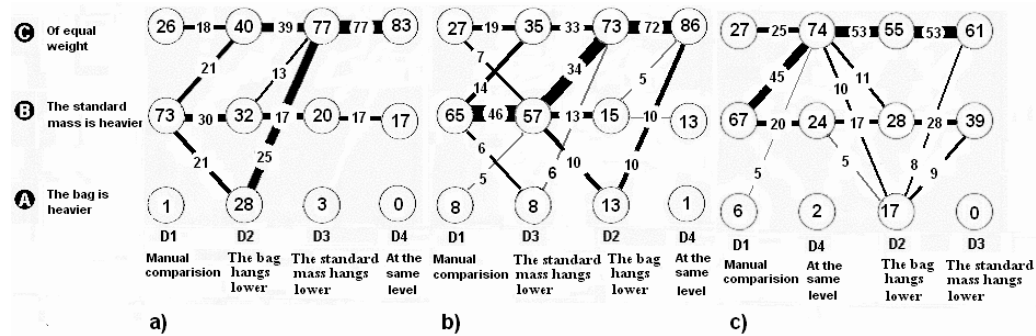


Figure 3. The distribution of the ninth graders' answers I) in group I (100 pupils) with the order of the demonstrations D1D2D3D4, II) in group II (88 pupils) with the order of the demonstrations D1D3D2D4, III) in group III (89 pupils) with the order of the demonstrations D1D4D2D3. Numbers represent per cents of all the pupils. Only the choices exceeding 5% are shown.

As seen in figure 3 the changes and the final choices in grade nine were much the same as in grade five in spite of the fact that about a quarter of the ninth graders estimated in the manual comparison D1 the standard mass and the bag to have the same weight. The difference between the fifth and ninth graders' choices after the manual comparison was significant ($\chi^2(2) = 34.11$; $p < 0.001$). This is understandable, because the concepts of weight and volume are more familiar to the older ninth graders. However, after going through all the demonstrations with the pulley the ninth and fifth graders ended up with more or less similar conclusions. Their choices after the last D4 demonstration in the series D2D3D4 and D3D2D4 are very alike ($\chi^2(1) = 0.404$; $p > 0.1$). The ninth graders' choices after demonstration D4 in these two series and after the last demonstration D3 in the series D4D2D3 differed significantly ($\chi^2(1) = 91.56$; $p < 0.001$) like the fifth graders' choices.

Pupil arguments

In the following the models Appearance and Material are joined together as Concrete model. Next the changes in the pupils' arguments were followed through the demonstration series. The results obtained about the changes in the fifth and ninth graders' arguments are collected in table 2 in which it is also shown the amount of the percentage of the correct choices the pupils made after each demonstration in the different demonstration series. The same thing is also shown in figure 4 in which it is easier to follow how the fifth and ninth graders' arguments, respectively, change

Table 2. Comparison of the number of correct choices after each demonstration and the distribution of the explanations at which the fifth and ninth graders arrived after the last demonstration in groups I, II and III.

	Amount of the correct choices, %								
	Group I			Group II			Group III		
	D2	D3	D4	D3	D2	D4	D4	D2	D3
5 th grade	36	83	94	19	70	89	69	45	61
9 th grade	40	77	85	35	73	86	74	56	61

	Distribution of explanations after the last demonstration, %					
	D2D3D4		D3D2D4		D4D2D3	
	group I		group II		group III	
Models	5 th grade N = 97	9 th grade N = 100	5 th grade N = 91	9 th grade N = 88	5 th grade N = 87	9 th grade N = 89
Motion	43	54	36	55	41	27
Position	20	15	22	11	6	1
Concrete	17	4	24	7	6	4
None	13	10	10	17	8	29

when the order of the demonstrations is varied in groups I, II and III in the case of the conclusion that the bag and the standard mass have the same weight. In groups I and II the biggest change in the arguments happens after showing the demonstration D3 in group I and D2 in group II. Then the pupils start to realize that they have to pay attention to the fact that the objects are not moving (Motion argument, 1) and not only to their position. However, the number of no arguments increases both among the fifth and ninth graders indicating that the pupils are not too sure about their thinking. In the fifth grade there is also an increase in the number of Appearance and Material arguments when the pupils tend to go back to the earlier concrete explanations. It is also interesting that after seeing the last demonstration D4 in these series the proportion of the pupils using the Position argument (2) increases both in the fifth and ninth grade whereas the proportion of Motion argument (1) increases in the ninth grade but decreases in the fifth grade.

After the demonstration D4 in group III almost a third of the fifth graders tend to prefer either the Appearance or Material argument whereas the ninth graders' explanations split evenly between the three arguments (figure 4). However, almost 15% of the ninth graders do not give any argument to support their choice, and after the next demonstration D2 (the bag hanging lower than the standard mass) the proportion of the amount of no argument increases even though about 20% of the pupils had changed their choice. After the last demonstration D3 (the bag hanging higher than the standard mass) in this series two thirds of the fifth graders who had ended up with the correct conclusion also ended up with the Motion arguments compared to less than half of the ninth graders. More than half of the ninth graders did not give any argument.

About 40% of the fifth graders perceive the idea of no movement, the Motion model independent of the order of the demonstrations (see table 2). A third of the fifth graders who had made the correct choice after the series D4D2D3 had not, however, understood the phenomenon properly. About half of the fifth graders who had been shown the demonstration series D2D3D4 or D3D2D4 gave improper explanations. Nearly a fifth explained their choices using the concrete models, Material and Appearance, and another fifth used the Position model. These fifth graders thus paid attention to the distinct and familiar features in the demonstration.

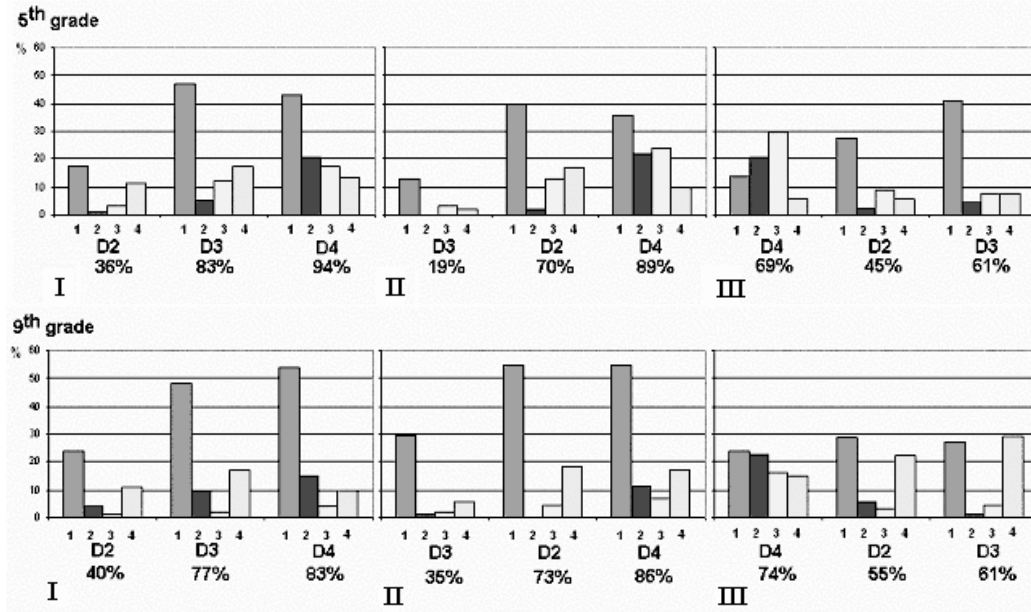


Figure 4. The changes in percentages in the fifth and ninth graders' arguments for the conclusion: The standard mass and the bag have the same weight in group I after the demonstrations D2, D3, D4, in group II after the demonstrations D3, D2, D4, and in group III after the demonstrations D4, D2, D3. The percentage underneath indicates the amount of the correct choice after each demonstration.

The arguments: 1. Motion, 2. Position, 3. Material + Appearance, 4. None.

D2: The bag hangs lower than the standard mass, D3: The bag hangs higher than the standard mass, D4: The bag and the standard mass hang at the same level.

After the demonstration series D2D3D4 and D3D2D4 half of the ninth graders perceived the Motion model. This is significantly more than the third of the ninth graders after the demonstration series D2D3D4 in group III ($\chi^2(1) = 10.46$; $p < 0.01$). However, also the Position model was more popular among the ninth graders in groups I and II than in group III. Slightly fewer ninth graders made the correct choice after the demonstration series D2D3D4 than the fifth graders (table 2; $\chi^2(1) = 5.73$; $p < 0.05$). Their explanations did not differ significantly from those of the fifth graders otherwise that the ninth graders used the concrete models Material and Appearance less than the fifth graders (4% and 8% compared to 17% and 22%, table 2; $\chi^2(1) = 4.96$; $p < 0.05$). After the demonstration series D4D2D3 less of the ninth graders compared to the number of the fifth graders used the Motion model to support their correct choice (28% compared to 41%, table 2; $\chi^2(1) = 4.86$; $p < 0.05$). However, almost 30% of these ninth graders did not give any reasoning for their choice compared to less than 10% of the fifth graders ($\chi^2(1) = 4.01$; $p < 0.05$). Altogether the difference between in the explanation given by the fifth and ninth graders in groups III after the last demonstration D3 was significant ($\chi^2(2) = 17.68$; $p < 0.001$).

The most striking difference in the ninth and fifth graders' results is in the groups III. In the fifth grade the use of the Motion argument increases steadily from 14% to 41% whereas it increases only 3% in the ninth grade from 24% to 27% as seen in figure 4. However, the amount of the ninth graders who do not give any argument is fairly high and increases from 15% to 29%. We interpret these results as some ninth graders in group III realising after demonstration D2 that the Position argument cannot be used and when they have no sensible explanation to offer they do not give

any answer. The force concept is not so well formed in their minds that they could rely on it. On the other hand, nearly 40% of the ninth graders in this group reason using the Position argument that the standard mass weighs more than the bag. Altogether the Position argument seems to hinder the ninth graders to start looking for an explanation to their observations.

Summing up the main results, it can be seen from table 2 that the order of the demonstrations has an effect on the choices pupils make. When the first demonstration (D2 or D3) coincides with some familiar way to think and is then shown wrong with the second demonstration nearly 90% of the pupils both the fifth and ninth graders reach the final correct conclusion. Whereas when the pupils are first shown an obvious case (D4) and after that the seemingly contradicting cases only 60% of the pupils draw the correct conclusion. However, among the fifth graders (90%) who drew the correct conclusion after the series D2D3D4 and D3D2D4 are respectively more pupils (30%) who used scientifically non-acceptable explanations (Position, Concrete, None) than among those fifth graders (60%) who followed the demonstration series D4D2D3. The ninth graders seem to respond to some extent negatively to the demonstration series D4D2D3 as half of those who draw the correct conclusion do not give any explanation to support their choice.

DISCUSSION AND CONCLUSIONS

The results show that one successful demonstration that points out the conflict in pupils' ideas is not enough to cause the conceptual change. In many cases it can happen that such a demonstration only activates another misconception. In both 5th and 9th grades the great majority of the pupils clearly ended up with the correct conclusion when the demonstration D4 was shown last. However, there are many factors which pupils have to take into account when they try to make sense from simply looking at everyday phenomena. It is well known that pupils have a variety of beliefs which they try to apply. Teachers easily underestimate the complexity of the process to reach the scientific explanation. For example, we thought that showing first demonstration D4 (both objects at the same level) would help pupils to see that both objects weigh the same. However, the demonstrations with the pulley in balance brought up other kind of ideas that confused pupils' thinking like that the weight of an object depends on its position, size and material. After manual comparison the majority of both the fifth and ninth graders estimated that the small standard mass of metal weighed more than the large bag of steel wool. However, the estimations based on the manual comparison did not seem to disturb the pupils' thinking. About 70% of the fifth graders and 50% of the ninth graders were ready to change their conclusion and argument after they had seen the first demonstration using the Motion model as explanation. This is in accordance with the suggestion that pupils' conceptions have different status according to how plausible and fruitful the conception was to the pupils when they formed it (Hewson & Thornley 1989).

When pupils think the situation self evident they tend to use simple models more easily. This can be seen in figure 4 for example in the case of the demonstration series D4D2D3 when both the fifth and ninth graders offer after the first demonstration concrete Appearance and Material models, and Position model. There is even a fair amount of pupils who seem to be fixed with the concrete models even though the ninth graders seem to be readier to drop these models when they have found a more plausible model.

The fifth graders are more flexible than the ninth graders in changing their thinking and looking for other explanations as can be seen from figure 4. On the other hand, the ninth graders use less arguments based on concrete and position model than the fifth graders. Very few of the ninth graders argued according to the concrete Material and Appearance models. They were inclined to use the Position model after the demonstration D4 but when they realized that the Position model was no good they seemed to become frustrated and did not give any reasoning. The different behaviour of the fifth and ninth graders is especially obvious in group III following the demonstration series

D4D3D2. While the number of the fifth graders using the Motion arguments increases the number of ninth graders stays more or less the same while the number of those ninth graders who refuse to give any argument increases. This leads to the conclusion that it is harder for the ninth graders to find a new explanation. Also 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. Vosniadou, Ioannides, Dimitrirakopoulou and Papademetriou (2001, p. 384) stress scientific explanations of physics phenomena often violate fundamental principles of intuitive physics and therefore learning physics requires then reorganization of existing conceptual structures and not just the creation of new, qualitatively different representations. This is hard work and it may lead to that pupils start to see physics difficult and incomprehensible which is seen in the change in pupils' attitudes towards physics (Häussler, Hoffman, Langeheine, Rost & Sievers 1998).

When pupils try to explain a phenomenon there are many factors to be taken into account at the same time and for a novice it is very difficult to see what factors are more important as they do not have the knowledge of the experts who look at the situation from the theory's point of view (Chi, Glaser & Farr 1988). One has to learn to be sensitive to details and differences in various phenomena to be able to find the general explanation behind what can be seen. In planning a demonstration the teacher has to be aware of the various alternative conceptions on which pupils are relying in their reasoning. Demonstrations should be such that they reveal to the pupils their preconceptions for example by causing a conceptual conflict. It is important to show a surprising demonstration first so that pupils really concentrate and start to form explanations. Also it is very important to start in lower grades. In the ninth grade pupils have more fixed ideas and when these do not work pupils seem not to be ready any more to look for new possibilities. When their explanation does not function they start to think that physics is difficult and impossible and turn off from it. Of course, good teachers are needed to guide pupils through the different possibilities with discussions and other demonstrations.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the reviewers for their valuable comments and careful checking of the manuscript.

REFERENCES

- Bao, L., & Redish, E. F. (2003). Educational assessment and underlying models of cognition. In W. E. Becker, & 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., & Sneider, C. (1994). Childrens' conceptions about weight and free fall. *Science Education*, 78, 149-171.
- Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chi, M.T.H., Slotta, J. D., & de Leeuw, N. (1994). From things to processes: a theory of conceptual change for learning science concepts. *Learning and Instruction*, 4, 27-44.
- Chinn, C. A., & 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.
- Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics*, 50, 66-71.
- Dekkers, P. M., & Thijs, G. D. (1998). Making productive use of students' initial concept of force. *Science Education*, 82, 31-51.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10, 105-225.

- Dreyfus, A., Jungwirth, E., & Eliovitch, R. (1990). Applying the "cognitive conflict" strategy for conceptual change: Some implications, difficulties, and problems. *Science Education*, 74, 555-569.
- Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23, 5-12.
- Galili, I. (2001). Weight versus gravitational force: historical and educational perspectives. *International Journal of Science Education*, 23, 1073-1093.
- Galili, I., & Bar, V. (1997). Childrens' operational knowledge about weight. *International Journal of Science Education*, 19, 317-340.
- Gunstone, R. F., & White, R. T. (1981). Understanding gravity. *Science Education*, 65, 291-299.
- Hakkarainen, O. (2005). Mental models in manual weight comparison between the objects of different size. *Themes in Education*, in press.
- Hakkarainen, O. & Ahtee, M. (2005). Pupils' mental models of a pulley in balance. Unpublished manuscript.
- Häussler, P., Hoffman, L., Langeheine, R., Rost, J., & Sievers, K. (1998). A typology of students' interest in physics and the distribution of gender and age within each type. *International Journal of Science Education*, 20, 223-238.
- Hewson, M. G., & Hewson, P. W. (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., & Thornley, R. (1989). The conditions of conceptual change in the classroom. *International Journal of Science Education*, 11, 541-553.
- Ioannides, C., & Vosniadou, S. (2002). The changing meaning of force. *Cognitive Science Quarterly*, 2, 5-62.
- Jimoyiannis, A., & Komis, V. (2003). Investigating Greek students' ideas about forces and motion. *Research in Science Education*, 33, 375-392.
- Lee, G., Kwon, J., Park, S. Kim, J.-W., Kwon, H.-G., & Park, H. K. (2003). Development of an instrument for measuring cognitive conflict in secondary-level science classes. *Journal of Research in Science Teaching*, 40, 585-603.
- McCloskey, M. (1983). Naïve theories of motion. In D. Gentner, & A. Stevens (Eds.), *Mental Models*. Hillsdale, NJ: Erlbaum.
- Mohapatra, J. V., & Bhattachaayya, S. (1989). Pupils' and teachers' induced incorrect generalization and the concept of force. *International Journal of Science Education*, 11, 429-436.
- Pfundt, H., & Duit, R. (2001). *Bibliography: Students' alternative frameworks and science education*. Kiel, Germany: IPN, University of Kiel.
- Piaget, J., & Inhelder, B. (1971). *The child's construction of quantities: Conservation and atomism*. London: Routledge & Kegan Paul.
- Pintrich, P. R., Marx, R. W., & 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., & Gertzog, W. (1982). Accomodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211-227.
- Tao, P., & 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.
- Tyson, L. M., Venville, G. J., Harrison, A. G., & Treagust, D. F. (1997). A multidimensional framework for interpreting conceptual change events in the classroom. *Science Education*, 81, 387-404.
- Vosniadou, S. (1994). Capturing and modelling the process of conceptual change. *Learning and Instruction*, 4, 45-69.
- Vosniadou, S. (2001). Mental models in conceptual development. L. Magnani, & N. J. Nersessian (Eds.), *Model-Based Reasoning: Science, Technology, Values*. Dordrecht: Kluwer Academic.
- Vosniadou, S., & Ioannides, C. (1998). From conceptual development to science education: A psychological point of view. *International Journal of Science Education*, 20, 1213-1230.
- Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction*, 11, 381-419.