TELEMARKS FORSKING

NOTODDEN

Gard Brekke

Diagnostic assessment

Assessment tools developed on the basis of the KIM project

Rapport 19/00
Telemarksforsking-Notodden
Desember 2000

Telemarksforsking - Notodden

ISBN: 82-7463-063-5

Det må ikke kopieres fra denne rapporten i strid med åndsverkloven
eller i strid med avtaler om kopiering inngått med KOPINOR,
interesseorgan for rettighetshavere til åndsverk

Content

This report is based on a guest lecture that was held at University of Melbourne, November 1999.

Diagnostic assessment. Assessment tools developed on the basis of the KIM project

Our understanding of what constitutes mathematical competence is changing, resulting in related curricular developments. Generally though, assessment is not keeping pace with these changes, although it has been a growing emphasise on research and development of assessment in mathematics education during the 90's. The aim of this discussion will be to focus on formative and diagnostic assessment as one particular aspect of assessment in mathematics. These types of assessment have the objective to inform teachers and the learners about the individual learner's progress and about the support needed in their learning process.

This introduction will describe a project based on diagnostic assessment, KIM, "Kvalitet i matematikundervisningen" (Quality in Mathematics Teaching), where assessment is used as a basis to aid conceptual development.

Objectives and material

The project was initiated by the Norwegian Ministry of Education in 1993 and is funded by the Ministry, and is conducted by Telemark Research Institute, Notodden and the Department of Teacher Education and School Development at University of Oslo. It involves mathematics from grade 1- 10, and has this year also been extended to the upper secondary schools.

The objectives of KIM are to:

- develop integrated test- and in-service training packages that can be used by teachers as part of their assessment practice
- develop a collection of test-instruments of a diagnostic character, which can be used as a starting point for teaching practice within various parts of the subject matter
- survey attitudes and conceptions that pupils have towards mathematics and the teaching of mathematics
- report the whole spectrum of pupil performance within the various areas of school mathematics, and not only on minimal competence
- survey pupils performance in relation to a broad spectrum of objectives specified in the current curriculum.

This presentation is limited to first three points above.

KIM develops sets of diagnostic test items. The different sets are linked to a specific part of school mathematics, and thus intend to cover most of the concepts of school mathematics when the project is finished. Written materials are produced for each set of diagnostic test items. The main focus of these materials is on reporting and discussing the extent of misconceptions and conceptual obstacles identified by the diagnostic test items. The items are developed to enquire into different aspects of the particular concept in question.

The first package developed, Numbers and number operations, consists of two diagnostic tests (each in three versions for grades 4, 6 and 8 respectively) and two booklets for teachers. The main mathematical focus of this material is on the decimal number concept as well as the mathematical operations on decimal numbers. The first booklet, Introduction to diagnostic teaching in mathematics, presents a brief discussion of mathematical competence and ways of working in the classroom. It advocates an increased emphasise on conceptual development in school mathematics. The second booklet is a guideline for teachers to the mentioned mathematical content area. The main content of all the guidelines are:

- analyses of the national data obtained from the test items according to different groups of misconceptions and identifications of conceptual obstacles
- examples of teaching materials which are thought to be helpful in overcoming misconceptions and conceptual obstacles observed by the diagnostic tests

The teaching activities are often designed to create a cognitive conflict that should be resolved through discussions and reflections amongst the pupils. Teachers are encouraged to give children the opportunity to stop and reflect on their actions and experiences in their process of developing a certain concept. The aim is that the children should become aware of their own learning process.

In addition to the mentioned package about decimal numbers and numerical operations similar packages related to the *graphical aspect of functions*, *algebra* and *measurement and units*. National standardisation of data for *Geometry* has taken place. In addition data is collected for three units for upper secondary school, *Numbers and operations*, *Geometry* and *Measurement and units*. Data related to students' and teachers' attitudes and beliefs in relation to mathematics, and mathematics teaching and learning is collected. A booklet for teachers is being produced based on these data.

Test development

The reports and discussions in the material for teachers mentioned above are based upon data obtained from a "national standardisation" by KIM. The data is based on written responses to diagnostic items from approximately 2000 children from 100 schools for each subject area and each grade level. This is a survey among children at different class levels based on the diagnostic items. In this section I will focus on methods used to develop single items as well as cluster of items, which makes it possible, to investigate into children's ideas linked to the concept in question.

The test items with the purpose described above have to be clustered to give valuable information about children's understanding of specific concepts. Clustering means here that the same aspect of a concept, or a known misconception, is investigated by varying the context in which the concept is embedded as well as the numbers involved. In this way the items in our test for the national standardisation had to be different from items used in assessment, which may have other main purposes. Since the main objective is to use the diagnostic items as a basis for teaching activities we are more interested in pupils problems, or unfinished understanding of a concept, than in the percentage of children who manage to solve a given problem correctly. A diagnostic test will because of this objective usually contain relatively many items on which pupils give wrong answers. The national standardisation will therefore not provide a picture of an overall "national standard of mathematical competence", but rather a description of difficulties or a diagnosis of mathematics. The national standardisation gives insights into areas of understanding in which teaching has to be carefully planned.

The procedure of item development

The first decision is to choose which aspects of the concept in question that we want to focus upon. In KIM it has been of importance to select items that have proved to give good diagnostic information as well as to develop new items to broaden the information. Previous used items, often from studies carried out in other countries, give information that also adds to the validity of the study. A general search in research literature on conceptual understanding of specific mathematical concepts has made a basis for the analysis of different aspects involved in such concepts.

In general this is not an easy task in the test development. Even though there has been carried out research on children's conceptual development in relation to many different concepts in mathematics, it is amazing to discover how little we still know about general ideas connected to these concepts. The study of concept development requires researchers to view a concept in relation to three different variables:

- A set of situations that makes the concept meaningful (the reality).
- A set of invariants (objects, properties and relationships) that can be recognised and used by the children to analyse and master these situations (representations).
- A set of symbolic representations that can be used to point to and represent these invariants and therefore represent the situations and the procedures to deal with them (also representations).

One difficulty is that a single concept does not refer to only one type of situation, and a single situation cannot be analysed with only one concept. An important restriction to item development in the national standardisation carried out by KIM is that both instruction and answers are communicated via writing.

After this rather lengthy process of collected research ideas categories of items linked to these aspects are chosen both on the basis of the curriculum, the textbooks and the research literature. This selection process is made in cooperation with a group of teachers. The collaborating teachers carry out several rounds of pilot studies of single items or small sets of items. The teachers and researchers used these rounds of pilots to validate if the single items or cluster of items provided information about the chosen aspect. These rounds of pilots were used to construct a set that could be answered by children in 40 minutes, again tested out in schools by the same group of teachers in other classes in their schools. After this round a larger pilot study of the sets of test items are carried out in 10 schools from different regions in the country. The results from this study of around 150 children on each class level are used for the final revision of the items and the test as well as a basis for the development of a scheme to record categories of answers given by the children.

The teaching material

One of the most important findings of mathematics education research carried out in the last twenty years has been that all pupils constantly "invent" rules to explain the pattern they see around them. For example, it is well known that many pupils think that by multiplying two numbers the answer gets larger. These pupils "over-generalise" their experiences from working on integers to operating also on decimal numbers and fractions. A decimal number consists of a pair of two whole numbers is another “invention” of an aspect of a concept based on generalisation from restricted experiences from decimal numbers used in measuring for example lengths of different kinds. In such situations it is an integer numbers of for example meters to the left of the comma and an integer number of cm to the right. We name such “inventions” for misconceptions.

A misconception is an incomplete or unfinished ideas linked to a concept. It is important to understand the difference between the errors that pupils make and the misconceptions they have. Misconceptions are not accidental. Behind these there is a particular thinking - an idea - that is used consequently. (See Brekke 1995a). Some other common misconceptions are:

- The longest number has the greatest value
- It is impossible to divide a small number by a large number
- One can only divide by whole numbers
- A graph is a picture of the situation it represents

Overcoming misconceptions present the teachers with a dilemma. For example, when we are introducing multiplication with integers it does not make sense for the pupils to make the point that: “later you will experience that multiplication can make numbers smaller”. This would be too far away from the problems these children are dealing with at the moment. Thus it is probably impossible to prevent that misconceptions of this kind develop.

Diagnostic items are items that are designed to:

- discover which ideas an individual have in relation to different concepts
- make a survey of difficulties or conceptual obstacles related different concepts
- help the teacher to plan teaching

An item is a “good” diagnostic item if it reveals misconceptions if children have developed incorrect or over generalised ideas, which they use consequently. For example will the items $0.24 \div 2$ and $0.12 \div 2$ give different information to the teacher. It is possible to get a correct answer to the first even if one consider a decimal number to consist of a pair of whole numbers, while one would get the answer 0.6 with this idea in the second item.

It is not a simple case to develop good diagnostic items. First of all one needs to have an overview of the most common misconceptions linked to different concepts of mathematics and next to make a cluster of items, varying the contexts and numbers involved to bring out the ideas in question.

There have been several studies that have focused on teaching directed towards conceptual development. I would like to refer to Bell (1993a) and (1993b) for a theoretical overview and to experiments done by Swan (1983), Onslow (1986) Brekke (1987) and (1991), Birks (1987), and Bassford (1988). All these studies have applied a method named diagnostic teaching, which have proved to result in long-term learning. The crucial idea in this method is to focus on intensive discussions related to the children’s’ understanding of the concept. To work intensively on a few, well designed, activities, is more effective than to work through a great number of routine activities. This method has got a prominent place in the resource material linked to the diagnostic test of the KIM-project. In this discussion I will just refer to the main points of diagnostic teaching:

- Identify misconceptions
- Prepare the teaching in such a way that misconceptions are exposed, for example by cognitive conflicts
- Solve the conflict by discussions and reflections
- Immediate feedback of correctness
- Use the new (or extended) concept in new "situations"

Examples of findings

The collected data for the study reported here is based on written responses to diagnostic items from 1953 children of average age of 16.5 from 114 schools. Schools and classes were chosen from a stratified sample of all Norwegian upper secondary schools, covering all branches of the Norwegian schools**.**

Background

The following discussion of proportional reasoning is based on the framework developed by Vergnaud (1983, 1988 and 1997). Vergnaud (1983) uses the term conceptual field to encompass situations, concepts and processes that are requisite to handle situations. He claims:

> *It is difficult and sometimes absurd to study separately the acquisition of interconnected concepts. In the case of multiplicative structures, for example, it would be misleading to studies on multiplication, division, fraction, ratio, rational number, linear and n-linear function, dimensional analysis, and vector space; they are no mathematically independent on of another, and they are all present simultaneously in the very firs problems students meet.* Vergnaud 1983 (p. 127)

The objective of the framework is focus the researchers awareness on that:

- Mathematical concepts are rooted in situations and problems.
- Such situations, as well as the procedure students use to deal with these situations, have to be classified and analysed.
- Students' ideas and competencies develop over a long period of time.

Teaching students at a particular grade requires that one have a fair idea of the steps they may or may not have gone through and the next ultimate steps one would like them to reach.

Proportional reasoning is contained in the part of the multiplicative conceptual field, which is named *Isomorphism of measures* by Vergnaud. Schematically one can represent isomorphism of measures by figure 1.

Multiplication

M_1	M_2
1	b
c	x

Partition

M_1	M_2
1	x
c	d

Quotition

M_1	M_2
1	b
x	d

Four-number problems

M_1	M_2
a	b
c	x

Figure 1. *Schematic representation of isomorphism of measure.*

This implies that, schematically, multiplication and division are just special cases of proportional reasoning. The word ratio is used for a comparison between quantities of like nature (measures) and rate of unlike nature.

Strategies

For the analysis of strategies used by students solving four-number problem one usually distinguishes between strategies applied inside or in between the measures. Vergnaud uses the notations scalar factor and functional operator. Referring to item 2 below, 2.5 is a scalar factor and $\frac{2}{3}$ or $\frac{4}{6}$ is a functional operator. These strategies will be analysed further in relation to ways used to compute the answers to the given problems. Figure 2 is a schematic representation of this classification of operators.

$$\begin{tikzcd} \underline{M_1\text{(persons)}} & \underline{M_2\text{(dl)}} \\ 6 & 4 \arrow[d, "\text{x } 2.5"] \\ 15 \arrow[r, "\text{x} \frac{2}{3} \text{ (eller 0.667)}"'] & ? \end{tikzcd}$$

Figure 2. Schematic representation of scalar and functional operators.

Test items

The students had to explain how they calculated or reasoned when solving items one to four. The answer to item five requires a single number and item six is a multiple-choice item. The relationship between the numbers are kept the same for the first four items, but changed around in the model (Figure 2) to keep the complexity of the computation at the same level for all of these items. In this way we can analyse strategies according to context, scalar or functional operator, and to some extent continuous or discrete measures.

1. 15 litres of berry weigh 10 kg. What is the weight of 6 litres of the same berries?

2. Lars uses this recipe for making rice cream for 6 persons. How much rice does Lars need to make rice cream for 15 persons?

Rice cream for 6 persons:
4 dl rice
7 dl milk
0,5 dl cream
2 spoons of sugar

3. A picture is 4 cm high and 10 cm wide. Lise enlarge this picture for here essay. The picture should be 6cm high in here essay. How wide will the enlarged picture be?

4. When Mari makes lemonade she uses 4 spoons of sugar and 6 spoons of lemon juice. Tor makes lemonade using 10 spoons of sugar and 15 spoons of lemon juice. Which lemonade is sweetest? (Ring your choice)

A Mari's lemonade is sweetest
B Tor's lemonade is sweetest
C They have the same sweetness
D It is impossible to decide which one is the sweetest from the given information

5. Brass is a made from zinc and copper in the ratio 1 : 4
How many kg of zinc is there in 40 kg of brass?

6. Kåre makes grey paint by mixing 1 litre of black and 5 litre of white paint.
How much black paint will he have to buy to make 30 litre of this grey paint?
Ring the expression, which fits this problem

30 : 5   30 · 5   30 : 6   30 · 1   30 · 6

Item 4 differs from the structure illustrated in figure 2, as it requires a comparison of two ratios (number of spoons). The structure can be represented schematically by.

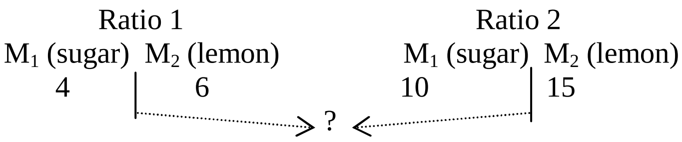


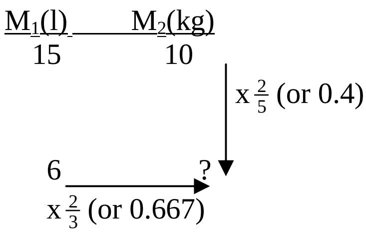
Figure 3: Schematic representation of a ratio compare problem.

Discussion

Correct answers and the most common incorrect answers to items 1a to 4a are given in tables 1 to 4. We notice a high occurrence of no answer to the open items and also a general low percentage of correct answers. The multiplicative structure is indicated for each item.

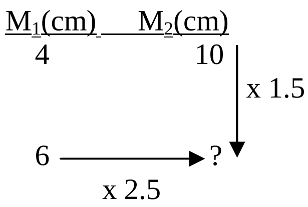
Item 1a	N	%
No answer	106	19.9
4 or 4 kg (correct)	203	38.2
11 or 1	10	1.9
3.6 or similar	30	5.6
9 or similar	128	24.1

Table 1. *Responses to item 1a*



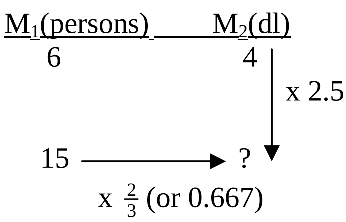
Item 3a	N	%
No answer	82	15.4
15 or 15 cm (correct)	236	44.4
12	148	27.8
13, 13.3 or 13.33	6	1.2
14	16	3.0
16	6	1.1

Table 3. *Responses to item 3a*



Item 2a	N	%
No answer	80	15.0
10 or 10 dl (correct)	302	56.8
9 or similar	52	9.8
22.5	22	4.1
11	11	2.1
60 or 6	10	1.9

Table 2. *Reponses to item 2a*



Item 4a	N	%
No answer	41	7.7
C (correct)	222	41.5
A	154	28.8
B	69	12.9
D	47	8.8

Table 4. *Responses to item 4a*

See figure 3 for the structure of this item.

We notice that both scalar factor and functional operator is less than 1 in item 1, which is usually considered to be more difficult than if they had been larger than 1. Similar only the functional operator factor is less than 1 in item 2, and none is less than 1 in item 3. This gives us the possibility to analyse how the preference of operator is effected by its numerical size. This study has too few items to analyse a possible hierarchy between this effect and the effect of other variables, such as familiarity of context, size and complexity of numbers etc. But the low percentage of correct answers to item 1 should indicate that the size of the scalar factor and/or operators effects the performance. The large proportion giving the answer 9 to item 1, which is probably due to applying the functional operator 1.5 to 6, also indicates this. We label this strategy "reversed rate" (See strategies applied below). If we only consider the size of the operators we might have believed that item 3 should have been easier than item 1 and 2, but in this item it is another relationship that plays an important role. See discussion of "wrong additive strategy" below.

To solve item 4 the ratios between sugar and lemon in each mixture have to be considered. Student's explanations showed that a variety of strategies, with cancelling the numbers in each ratio as much as possible as the most common strategy (12%). The most frequent wrong strategy was comparing the differences between the ratios in each mixture (2 and 5). This strategy was applied by 14%. See below for discussion of "wrong additive strategy".

The following categories are used to describe the strategies applied:

- *Ratio/rate equation*. An equation is formed by the given information, either by making a ratio between the numbers of the same measure (ratio) or between the measures (rate).
- *Unit factor (rate)*. It is well known from several studies, Hart (1981), Vergnaud (1983) and Kaput & West (1994) that people generally avoid multiplying with a fraction and even with a decimal number. One strategy is to search for a simpler multiplicative relationship, as for example in item 1, 1.5 litre per kg and the next step would than be to calculate (in different ways) how many kg relates to 6 litre.
- *Functional operator*. This is an operator between the measures, for example $\frac{10}{15}$, $\frac{2}{3}$, 0.667 or similar, is applied to the third number in the problem. This operator is a rate in items 1 and 2.
- *New unit*. This means that a new whole-number relation is established, for example the weight of 3 litre is 2 kg in item 1, and next a multiplication, repeated addition or a building up strategy is applied to this numerical relationship.
- *Quotition*. This means that a straightforward quotative division is applied.
- *Scalar*. This strategy is applied inside a measure, for example multiplying by $\frac{15}{6}$ or 2.5 in figure 2.
- *Build up/down strategy*. Often students prefer to avoid multiplication or division, if possible. One way of doing this is either to build up from the smallest to the largest number in one of the measure (or build down) and perform equivalent operations to the given number in the other measure. This may be an efficient strategy when the relations between the numbers are simple, and is also a possible way of discovering the multiplicative relationship.
- *Incorrect additive strategy*. Sometimes students consider differences between the numbers either inside or across the measures and apply this difference to the remaining number in the problem. This strategy also occurs when a building up strategy is used and the last "step" in this process is a fractional part of the building block.
- *Reversed rate*. This means that a rate is used, but in reversed order, which usually happens when the rate is less than 1.

Strategies	Item 1		Item 2		Item 3	
	Corr	Incorr	Corr	Incorr	Corr	Incorr
Ratio/rate equation	8.8	0.9	8.8		15.2	1.7
Unit factor (rate)	12.0	5.8	15.4	8.3		
Functional (rate)	2.6	0.2	5.6	0.6	2.1	0.2
New unit	3.6	0.4	1.7	0.2		
Quotition	4.5	0.2	0.8	0.8		1.9
Scalar	0.8		8.5		18.7	
Build up/down		1.1	10.2		0.6	
Incorrect strategies						
Additive strategy	1.9		5.2		21.7	
Reversed rate	22.9		5.1		0.6	
No answer	28.0		22.0		29.7	

Table 5. *Strategies applied, split by correct and incorrect answers for items 1b to 3b*

Table 5 shows that there is a great variation between strategies applied by students between these three items, which may be due to factors described above. It is interesting to notice the high percentage of the "building up" strategy, successfully applied in item 2. The context of this item is well known. In addition, one of the measures (people) consists of just integers, which makes it easy to apply this additive strategy. This example illustrates that the types of numbers involved in a problem also plays an important role in the choice of solution strategies.
It is also interesting to notice a more extensive use of a ratio/rate equation in item 3. Is this so because this strategy is introduced to the students in a context of geometrical enlargement or a ratio of some measurement done?

The problem of the wrong additive strategy is also illustrated by item 3. The numerical relationship in this item is relatively simple, and both operators are larger than 1, which should indicate a higher percentage of correct answers for the two items with which it is compared. The reason for the extensive use of this strategy in this case is probably due to lack of experience of what it means mathematically to enlarge to a similar shape. This illustrates the role of the context in choice of strategies. The wrong additive strategy applied in item 2 is mainly used for the last step of a building up strategy: 6 persons – 4dl, 12 persons - 8dl and 15persons – 11(8+3) dl.

Items 5 and 6 focus on the understanding of ratio. Item 5 does not requires an explanation, but this item still give important diagnostic information, especially when it is compared with the responses to item 6. That more than half of the students answers 10 or $\frac{40}{4}$ indicates that they consider 1:4 in this case to be the same as a fraction.

Item 5	N	%
No answer	128	23.9
8 or 8 kg (correct)	72	13.5
10 or $\frac{40}{4}$	274	51.2
160	22	4.1

Table 6. *Responses to item 5*

The ratio is given in another format in item 6. In addition both quantities are referred to the same unit (kg). These factors are probably both contributing to the higher percentage of correct answer to item 6. It is our belief that the students have problems interpreting the text where the information is given partly as a “part-part” ratio (1 litre black to 5 litre white) and partly as “part-whole” ratio.

Item 6	N	%
No answer	103	19.3
30 : 6 (correct)	169	31.6
30 : 5	108	20.2
30 · 6	34	6.4
30 · 5	72	13.5
30 · 1	38	7.1

Table 7. *Responses to item 6*

Conclusion

This study shows that one may, in an efficient way, analyse parts of student’s proportional reasoning by using relatively few well designed diagnostic items. It also illustrates that to do an analysis of what role each of the different factor plays in relation to choice of solution strategies to such problems would need a much larger test than would be possible in a in the setting of the KIM-project.

References

Bassford, D., (1988). *Fractions, A comparison of Teaching Methods.* Master of Education thesis. Shell Centre for Mathematical Education, University of Nottingham, England.

Bell, A., (1993a). Principles for the design of teaching. *Educational Studies in Mathematics* **24:** 5-34.

Bell, A, (1993b). Some experiments in diagnostic teaching. *Educational studies in Mathematics* **24:** 115-137.

Birks, D., (1987). Reflection: *A Diagnostic Teaching Experiment*. Master of Education thesis. Shell Centre for Mathematical Education, University of Nottingham, England.

Brekke, G., (1995a). *Introduksjon til diagnostisk undervisning i matematikk.* Nasjonalt læremiddelsenter, Oslo.

Brekke, G., (1995b). *Veiledning til Tall og Tallregning.* Nasjonalt læremiddelsenter. Oslo.

Foxman, D., Ruddock, G., Joffe, L., Mason, K., Mitchell, P. & Sexton, B. (1985). *Mathematical Development: A Review of Monitoring in Mathematics 1978 to 1982. Parts 1 and 2,* APU/Department of Education and Science, London.

Hart, K. 1981. *Childrens’s Understanding of Mathematics: 11-16.* John Murray. London.

Kaput, J. & West, M.M. 1994. Missing-Value Proportional Reasoning Problems: Factors affecting Informal Reasoning Patterns. In G. Harel & J. Confrey (Eds), *The Development of Multiplicative Reasoning in the Learning of Mathematics.* pp 235-287. State University of New York Press.

Onslow, B. (1986).*Overcoming Conceptual Obstacles concerning Rates: Design and implementation of a Diagnostic Teaching Unit.* Doctoral thesis. Shell Centre for Mathematical Education. University of Nottingham; England

Tinnes, M. Nedler 1998. *Proporsjonalitetstenking hos grunnkurselever*. Master Degree Thesis. Matematisk-naturvitenskaplige fakultet. Institutt for lærerutdanning og skoleutvikling. Universitet i Oslo.

Vergnaud, G. 1983. Multiplicative Structures. In R. Lesh & M. Landau (eds), *Acquisition of Mathematics Concepts and Processes.* pp 127-174, New York: Academic Press.

Vergnaud, G. 1988. Multiplicative Structures. In J. Hiebert, M. Behr (Eds)*, Number Concepts and Operations in the Middle Grades*. pp 141-161.Hillsdale, N.J., Erlbaum/Reston, VA: National Council of Teachers of Mathematics.

Vergnaud, G. 1997. Algebra, Additive and Multiplicative Structures. Is there any coherence at the early secondary level? In M. Hejný & J. Novotná. *Proceedings for The European Research Conference on Mathematical Education*. pp 33-45, Charles University, Czech Republic.