4

Diagnostic Assessment Frameworks for Mathematics: Theoretical Background and Practical Issues

Csaba Csíkos
Institute of Education, University of Szeged

Benő Csapó
Institute of Education, University of Szeged

Introduction

The main aim of this chapter is to link the previous three theoretical chapters and the detailed descriptions of content to be presented in the next part. We also intend to deal with the particular features of the frameworks and the reasons why we have chosen and applied certain solutions.

In the first two chapters, the outcomes related to the development of mathematical thinking and generally to the role of mathematics in developing thinking are outlined based on international research and primarily applying the results of developmental psychology. In Chapter 2, the external aims of mathematics education were presented making use of international research results on the application of knowledge. In Chapter 3, the traditions of Hungarian mathematics education and its curricular characteristics were shown as well as the practice the diagnostic system is to be adapted to.

The diagnostic assessment system is elaborated in three main domains parallel with each other according to the same principles[1]. Putting reading, mathematics and science in the same frameworks is supported by several

1 This chapter also contains sections which can be found in the chapter of the same function in the three volumes.

psychological principles as well as practical educational needs. A proper level of reading comprehension is required to have access to mathematics and science and also in turn learning mathematics and science improve the skills of reading comprehension of texts different from literary readings. The logic of mathematics and language can mutually reinforce each other. Science seems to be the most suitable field to apply the relationships acquired in mathematics in practice. In the beginners' stage of education it is particularly important to rely on and to make use of the various systems of relationships, when students' cognitive development is particularly fast and highly sensitive to stimulating effects.

The parallel treatment of the three domains can be beneficial to each other and the ideas and solutions can be very well used in the other field as well. Writing items and developing assessment scales, data analysis and systems of feedback make it necessary to treat the three fields parallel according to certain shared principles. However compromises are to be accepted, as the same principles can be applied in the same way only to a certain degree in the three fields. In order to be consistent, the three-dimensional approach is preserved and applied. Still the peculiarities of the fields will be taken into consideration in the description of various dimensions.

Another benefit of the work carried out parallel is the complementary effect. The theoretical backgrounds to the three domains are presented in nine chapters. In outlining the structure of the various chapters, we were not keen on the strict parallels thus it was possible to set forth one issue in one particular domain or to highlight another issue in another field. By way of illustration in the volume on reading in Chapter 1, the aspects of developmental psychology and cognitive neuroscience are much more emphasized which are also relevant for mathematics and science. Some of the cognitive abilities are described in detail in Chapter 1 of the volume on science but these abilities are also to be developed in mathematics. The second chapters of the volumes are devoted to the issues of how knowledge is applied and the general conclusions of any chapter can be used in the other two assessment domains. Chapter 3 also deals with practical, curricular issues in every field related to historical traditions of the Hungarian education system and the current practice. At the same time, in the selection and the arrangement of the content of education there seems to be a demand for following advanced international trends and making use of results achieved in other countries.

In line with these principles, the nine theoretical chapters altogether are considered to be the theoretical foundation of the diagnostic assessment system. In all three domains, we can draw on the background knowledge provided in the theoretical chapters without going into details of the identical issues in the parallel chapters.

The main aspects of the development of the frameworks are outlined in the first part of this chapter. First, the means used to describe the objectives of education and the content of assessments is described then our solution to the detailed description of the content of diagnostic assessments is presented. Further on it is shown in what ways these principles are applied in the elaboration of the frameworks of mathematics.

Taxonomies, Standards and Frameworks

We have made use of various resources during the creation of the frameworks for diagnostic assessments. We have attempted to give a clear-cut definition of the educational objectives and the contents of the assessments. First an overview of the systems describing contents is given then the methodology we have applied is shown related to these systems.

Taxonomies

The attempt to define educational objectives dates back to the 1950s. It was this time that the taxonomies by Bloom et al. were published, which later on have predominantly exerted an influence on pedagogy. What actually triggered the interest in taxonomies was the widespread discontent with the lack of clarity and precision in the description of curricular objectives on the one hand and the cybernetic approach making its way into education on the other. The demand for regulation has emerged, which required feedback and the prerequisite of feedback is to measure the distance between the aims that were set and the results achieved. By comparing the aims and the actual situation the deficiencies can be revealed and intervention can be planned accordingly. During the same period of time pedagogical evaluation became more pronounced and the widespread use of tests also required a clear definition of the object to be measured.

Taxonomy as a matter of fact is a structural framework, which serves as a kind of guide-line to arrange, to systematize and to classify things, in this particular case knowledge to be attained. It is like a chest of drawers with labels indicating what to put into them or like a table in which the headers are filled in showing what in the various columns and lines can be found. In comparison with the earlier general descriptions, making plans based on standardized systems represented a great progress and those involved with outlining concrete objectives to be attained in various subjects were in a way forced to consider their expectations resulting from instruction.

What exerted the greatest impact was the taxonomic system of the cognitive system published first by Bloom et al., 1956, which opened up new perspectives for designing curriculum and evaluation systems. The behavior patterns expected from students is described in terms of concrete and observable categories by the taxonomic system. What emerged as the greatest novelty was the six levels of frameworks built on each other, which could be uniformly used in every area of knowledge. Besides going into details, getting down to facts and accuracy represented enormous steps forward compared to earlier methods. Making use of the same detailed description both for planning learning processes and devising assessment devices implied a further advantage. This way it is taxonomy of objectives and evaluation.

It was in the US that the Bloom taxonomies exerted the first direct influence and later on they served as a basis of the first international IEA surveys. The hierarchy of knowledge assumed in the taxonomy was not confirmed by the empirical studies in every detail. Moreover, the underlying theory of the Bloom's taxonomy, the behaviorist approach was pushed into the background in the psychological foundation of education giving way to other paradigms, primarily cognitive psychology. Thus the initial cognitive taxonomies were rarely applied. The similar taxonomies in affective and psycho-motor areas were elaborated later on and even if they had been applied in several areas, their impact was not that predominant as that of the cognitive ones.

Taxonomies as principles of systematization are "empty systems" which are not concerned with concrete content. In taxonomy handbooks content is presented only by way of illustration. Thus, for example, when the six levels in *Bloom* taxonomy, *knowledge, comprehension, application, analysis, synthesis and evaluation* are used to describe objectives to be attained in a particular field of geometry, then it has to be clearly given what level of knowledge, comprehension and application is expected in geometry, etc.

Following the original taxonomies or revising them, currently more up-to-date handbooks have been produced in order to assist the description of latest systems of objectives (Anderson & Krathwohl, 2001; Marzano & Kendall, 2007). These works carry on the tradition established by Bloom, the operationalization of the objectives, and breaking down of knowledge into concrete measurable units. Methodologies established during the elaboration of taxonomies can serve as useful methodological resources for setting up standards.

Standards in Education

An impetus was given to setting up standards in the 1990s, especially in English-speaking countries where the normative documents regulating the content of education did not actually exist. In some countries, to say the least, in every school it was taught what was decided locally. Under these circumstances the options of educational policy narrowed down and the chances for improving the output of the school systems seemed rather poor. This is why the centralizing efforts were gaining ground and the aims and objectives of education were centrally set up either at local or at national level.

Educational standards are detailed descriptions of what students are expected to master, which in contrast to taxonomies as systems are concerned with concrete content. The standards of different domains are generally prepared by different professional teams and therefore according to the particular branch of knowledge, diverse solutions of forms can be applied.

Standards are normally devised by teams of specialists, who draw on the most up-to-date theories and scientific achievements. In the US for instance the professional association of mathematics teachers, the National Council of Teachers of Mathematics, (2000) created the standards for the 12 grades in public education. These standards generally describe what level of knowledge is to be achieved by students in various subjects after completing a grade.

The elaboration of standards went hand in hand with their application, similarly to the taxonomies both in assessment and the educational process. Several handbooks have been published which provide detailed methodologies of setting up and applying standards. However other aspects are emphasized than those in the taxonomies. Standards exert an influence directly in education (see Ainsworth, 2003; Marzano & Haystead, 2008) and evaluation seems to be only secondary to them (for example, O'Neill & Stansbury, 2000; Ainsworth & Viegut, 2006). What standard-based education actually means is that there are detailed, standardized requirements whose attainment can be expected from students of certain age groups.

Standards and standard-based education are not a complete novelty to Hungarian and other experts in education who have gained ample experience in centralized education systems. In Hungary, prior to the 1990s the contents of education were determined by a central curriculum and all the text-books were based on it. Every student of the primary schools studied the same teaching material and in theory every student had to meet the same requirements. In some fields, such as mathematics and science the standardized curricula had been the outcome of experiences of several decades, whereas other fields were exposed to political and ideological pressure. The trends of the 1990s were highly influenced by the earlier Anglo-Saxon model and the pendulum swung back, so the National Core Curriculum contained only minimal central requirements. This process was actually in contrast with what happened in other countries in the same period of time. By way of comparison it is worth mentioning that the printed volume presenting American mathematics standards (National Council of Teachers of Mathematics, 2000) is about the same in size than the first version of the Hungarian National Core Curriculum defining all domains of education published in 1995. In the meantime, the Hungarian national Core Curriculum has become even shorter.

Standards and standard-based education imply not merely standardization or centralization, but the professional arrangement of the contents of education based on research results. In this respect it is different from the earlier Hungarian central regulation to which this description applies only in some respects. The new kind of standards has become accepted even in countries where there were central curricula even earlier. For example, in Germany where the educational contents were determined at the level of lands earlier research into the development unified standards has begun (Klieme et al., 2003). The solid theoretical background is considered to be the essential feature of standards. Thus the elaboration of standards and standard-based education has triggered world-wide research and development.

During the preparation of the frameworks of diagnostic assessments we have relied on the theoretical implications of standard-based education on the one hand and contents and forms available in particular standards. In line with the traditions of setting up standards the characteristics of various content and assessment fields have been taken into consideration and we did not attempt to come up with totally identical solutions in forms in the description of the contents of reading, mathematics and science.

Our frameworks are different from standards because they do not set requirements and expectations. What the frameworks and standards have in common are the detailed and concrete description and the demand for solid theoretical background.

Frameworks

In line with the English usage the term framework is used for the detailed descriptions outlined in the present project. The frameworks of the assessments are similar to standards as they also provide detailed and systematic descriptions of knowledge. In contrast with traditional curricula it is not determined in the frameworks what and at what level should be taught and learned. Even the requirements to be achieved are not defined, although the content descriptions implicitly imply what can or should be achieved at the given domain.

The best-known frameworks have been prepared for international surveys. In case of assessments carried out in several countries laying down the requirements was of course out of the question. In this case what is presented by frameworks is what can be assessed and what is actually worth assessing. In outlining the content various aspects can be taken into consideration. In early IEA assessments the curricula of the participant countries were the starting point, namely what actually was taught in the given field.

In the main domains of the PISA assessment frameworks, it is presented what applicable knowledge is needed by the fifteen year-old young people of modern societies. In this respect the application of knowledge, and the needs of modern societies as well as the typical contexts of application play an essential role in the elaboration of frameworks related to the application of knowledge in various disciplines and school subjects.

Thee third approach mainly relies on research results related to learning and knowledge based on developmental and cognitive psychology. This aspect has been prevalent in the cross-curricular areas which were concerned with not one (or some) school subject. This kind of assessment was for instance in the fourth area of PISA 2000 where learning strategies and self-regulating learning were placed in the foreground, whose frameworks were essentially based on psychological research results related to learning (Artelt, Baumert, Julies-McElvany & Peschar, 2003). Students' attitudes can also be described according to psychological principles. The study of attitudes was carried out in most of the international assessments, and it was highly important in the PISA 2006 science survey (OECD, 2006). Similarly, psychological research has revealed the structure of problem-solving, which was the innovative assessment area of the PISA 2003 (OECD, 2004). The latest cognitive research results are to be used in the PISA 2012 dynamic problem-solving assessment.

The frameworks outlined for diagnostic assessments (see Chapter 5) have drawn on the frameworks of international assessments. They resemble PISA frameworks (for example, OECD, 2006, 2009) in a way that by focusing on three main assessment areas they lay down the foundations of the assessments of reading, mathematics and science. Whereas PISA frameworks have focused on one age-group, namely the fifteen-year-old young people, our frameworks focus on six grades, younger students, and developmental aspects are much more emphasized.

PISA frameworks have been prepared for a given assessment cycle, and although the frameworks are updated in every period the consecutive assessment cycles overlap. The PISA frameworks comprise the overall assessment process, from defining the domain, and organizing the domain to reporting scales showing the results. Relying on the above assessment process our frameworks imply the definition of the domain, the organization of the domain and the detailed description of the contents. The main dimensions of the assessments and the assessment scales are shown but in the current phase we are not concerned with the levels to be achieved on the scales and the quantitative issues of scales. With regard to developmental aspects drawing up scales needs further preliminary studies and empirical data.

Multidimensional Organization of the Assessment Contents

Over the past decade, innovations in education have been mainly integrative. Competencies themselves can be seen as the complex units of various elements of knowledge supplemented with further affective elements according to some interpretations (see e.g. Hartig, Klieme, & Rauch, 2008).

Competency-based education, project method, problem based learning, inquiry based learning, content-based development of skills, content-based language teaching and many other learning and teaching methods fulfill several objectives at the same time. Knowledge gained in this integrative way can be assumed to be more easily transferable and applied in a wider range. Summative achievement tests as well as PISA tests and competency assessments in Hungary are based on similar principles.

Other assessments are needed when learning problems are to be prevented and deficiencies hindering students in future achievements should be identified. When assessment results are used to identify the necessary intervention, it is not sufficient to prepare tests providing general indicators of students' knowledge. It is also not sufficient to find out whether the learners are able to do a complex task. The reasons for eventual failure should also be revealed as to learners lack in basic knowledge or the operations of thinking are not completely available, which are required for turning pieces of knowledge into logical chain of conclusions.

Diagnostic assessments need a detailed description of students' knowledge, which is why an analytic approach was used in contrast with the integrative approach in instruction. At the same time, the assessments assisting learning should also be adapted to the concrete educational processes. In line with these requirements the techniques of diagnostic and formative assessments are being worked out, which also draw on the results of summative assessments based on large samples. Moreover several fresh elements are introduced into the assessment techniques (Black, Harrison, Lee, Marshal, & William, 2003, 2005; Leighton & Gierl, 2007).

Several lessons can be drawn from the earlier achievements in similar fields for outlining the frameworks of diagnostic assessments, especially assessments conducted in early childhood (Snow & Van Hemel, 2008) as well as formative techniques for the initial stages of education (Clarke, 2001, 2005).

What we have considered most relevant is the multi-faceted, analytic approach emphasizing the psychological and developmental principles. The earlier paper-based formative and diagnostic tests have had their limitations. We use on-line tests which make it possible to conduct more detailed assessments more frequently, and so the frameworks should also be adjusted to these possibilities.

The Aspects of Organization of the Content to be Assessed

The content of the assessments can be organized according to three aspects. This way the content of the assessment makes up a three-dimensional cube, which is shown in Figure 4.1. The detailed description of the content however needs the linear arrangement of the blocks of this three-dimensional cube. The elements of the block can be listed in various ways depending on the way it is cut, according to which dimensions are cross-section produced first, the second and the third time. Here the aspect is shown, which seems to be more suitable for the requirements of diagnostic assessment.

The aspect highlighted first is itself a multi-dimensional system, which represents the three main dimensions of our analysis, namely the psychological (thinking), social (application) and disciplinary (subjects) dimensions. The developmental scales have been set up for these dimensions in the three main areas of assessment, such as reading, mathematics and science. (A more detailed description can be found in the next section.)

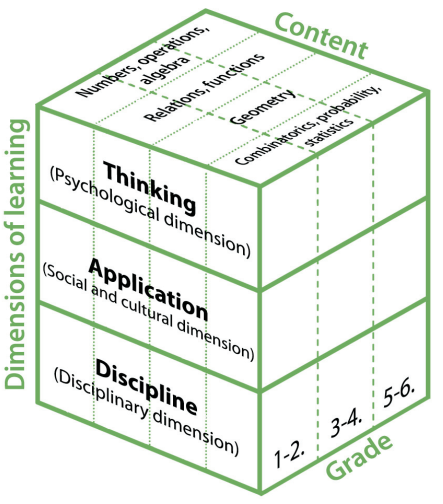

Figure 4.1 The multidimensional organization of the content of assessments

The second aspect is development. Accordingly the six grades were divided into three two-year stages such as grade 1-2, 3-4 and 5-6. Since the six grades are seen as a uniform developmental process, this division serves only as a practical solution to organize the content. Without empirical evidence, attachment to age group (grades) can only be hypothetical and approximate.

Finally a third aspect is the range of contents available in the given assessment area. The blocks of contents broken down in this way provide the units of detailed frameworks. Due to combining various aspects the increase in the values of certain aspects may result in a combinatorial explosion. Thus we have to be careful with the number of concrete assessment contents. Differentiation according to three dimensions of learning, three age-group blocks and four main contents in case of mathematics gives 36 blocks. This number would rapidly increase by differentiating further areas.

Scales of Diagnostic Assessments, Psychological, Application and Disciplinary Dimensions

Based on empirical results of earlier studies, a model was created, whose three dimensions are in accordance with the three main objectives of education. These objectives can be traced back to the history of schooling and are actually in line with the main trends of assessment of school achievements (Csapó, 2004, 2006, 2010).

Cultivating the mind and the development of thinking are kind of objectives which imply not external contents but rather internal characteristics. To put in modern terms, this is the psychological dimension. As it has been mentioned above, in the PISA surveys this dimension is also included. In several assessment areas the content of assessment was interpreted in psychological terms. In this respect in mathematics it is examined whether mathematics is developing thinking, general cognitive abilities to a degree it can be expected.

Another age-old objective is that usable knowledge should be provided by education which can be applied not only in schools. This aspect is the social dimension, and it covers the utilization and application of knowledge. The term has a lot in common with the transfer of knowledge which implies that knowledge gained in one context can also be applied in another one. The degrees of transfer can be characterized by the transfer distance. Any solution of task is the application of mathematical knowledge when what has been learned in one field of mathematics is applied in another field (near transfer) or beyond mathematics in other subjects or in practical context (distant transfer).

The third frequently mentioned objective is that students should obtain the essential components of the knowledge accumulated in science and art. This objective is being attained when the learning process is determined by the aspects and values of a particular branch of science, and this is the dimension of school subjects and disciplines. Over the past few years there have been attempts in education to counterbalance the earlier one-sided approach based on the branches of science. The competency-based education and large scale assessments focusing on application have pushed the considerations of the branches of science in the background. However in order that the teaching material should be a coherent, standardized and in this way available system in accordance with branch of science, the components of knowledge are also to be acquired, which contribute to application and the development of thinking indirectly. The establishment of the scientific validity of concepts and exact definitions are parts of this kind of knowledge.

The three-dimensional model implies that the same content, perhaps with a shift of focus, can be used for writing tasks in each dimension. By way of illustration we refer to the combinatorial reasoning, whose basic level is being formed in childhood through the interaction with the environment. This kind of thinking is improved by school activities and thus the levels of the ability of combination, a part of cognitive development can be measured. Or tasks can also be prepared in which combinative thinking and knowledge gained in combinatorics are to be applied in everyday situations. Finally it can be checked whether students are aware of variation and combination and they also know how to obtain formulas required for their calculations. This latter is a kind of knowledge that cannot be attained by activities encouraging cognitive development but only through proper knowledge of mathematics.

Among the three assessment areas mathematics is rather special with respect to the fact that the evolvement of mathematic thinking, especially in the early stages of education is closely related to the general cognitive development. In every field of mathematics, carrying out operations and thinking have an essential part, which is why the three dimensions are not separated from each other so much as it happens in the other areas of assessments. Thus it frequently occurs that the aspects of various dimensions emerge even in the same task.

Abilities in Mathematical Reasoning

In Chapter 1 of the present volume the components of mathematical reasoning are outlined on the theoretical base provided by Piaget and Vygotsky. Moreover a system of skills is proposed, which is comprehensive and at the same time specifically mathematical. Being comprehensive means that the mathematic thinking processes frequently designated in various ways are described according to four fundamental abilities in mathematical thinking. The system of abilities outlined in the chapter is still specifically related to mathematics. Independently of the structure of the science of mathematics and also social expectations, integers, rational numbers, the complete system of additive and multiplicative thinking patterns provide the description of mathematic thinking. In the following a further link will be set up between the theoretical chapters and the detailed frameworks through two sets of concepts, which have had a key role in international research over the past two decades.

Problem-Solving

In a major part of research on mathematical reasoning, it is considered to be a field of general problem-solving. Normally a task is called a problem if no algorithm is available to follow in the solution process, but the task requires conscious planning, monitoring and checking. When answering questions such as "how can you measure the length of the school-yard by using your steps?" or "how many liters of water goes in your bathtub?" students break down the problem into parts and analyze it and sum up the steps of problem-solving and solve the problem. The awareness of mathematical concepts and symbols and the development of mathematical skills and abilities are the precondition and assistance for the development of problem-solving in this sense.

Apart from mathematical abilities described in Chapter 1, other reasoning processes are also involved in mathematical problem-solving. This is what makes it possible to apply mathematical operations in new, unfamiliar situations in solving problems different from the routine tasks which need a lot of practice.

In the study of mathematical problem-solving, mainly word problems are used and the process of their solution is analyzed (Csíkos, Kelemen, &

Verschaffel, 2011). Word problems as the natural means of the development of problem-solving are actually used as early as in grades 1-2. However the precondition of solving word problems is proper comprehension, thus the length and the linguistic complexity of the text should be adjusted to developmental level of the learners. Initially word problems are described by two or three simple sentences. Another stage is when task is done not after listening but reading. As it can be seen in grades 3-4 word problems, originally meant to assist problem-solving by the various means of mathematical modeling of reality, more and more have become the means of improving counting ability. This is why understanding the text of the task and exposing the problems in an accessible way are essential.

An important method for improving problem-solving is putting the thinking processes into words, and asking "why" questions (Pólya, 1945). Discussing internal, mental images and drawings learners produced about the task (with great variety, according to their mental models) may contribute to the development of strategic and meta-cognitive components of problem-solving.

Mathematical Skills and Abilities

Over the past two decades, major efforts have been made to define skills and abilities related to mathematics either in terms of development or assessment. One of the trends in the research into intelligence has attempted to reveal the structure of intelligence based on the differences in the measurable psychological characteristics. It was Carroll (1993) who summarized the findings of this line of research and then he made an attempt to describe mathematical skills in the system of abilities of intelligence. According to Carroll (1998), several components of fluid intelligence can be directly detected in mathematical achievements, for instance general sequential reasoning, quantitative reasoning and the so-called Piagetian reasoning. Carroll also has highlighted several components of crystallized intelligence as the key elements of mathematical abilities.

Linguistic development is also highly relevant as people tend to count in a particular language and the designation of numbers in various languages influence counting skills as well. Oral and written comprehension skills are obviously needed to understand mathematical word problems. Some of the components of counting ability can be found in the factors of intelligence, such as the general processing speed and number facility, which the measurable components of thinking are as a matter of fact.

In Hungary, there are research programs, which aimed at studying the establishment and early development of counting skills: PREFER and DIFER (Nagy, 1980; Nagy, Józsa, Vidákovich & Fazekasné, 2004).

Reasoning abilities are highly relevant in mathematical thinking. Hungarian diagnostic assessment programs have been involved with the following five thinking abilities: inductive (Csapó, 2002), deductive (Vidákovich, 2002), systematizing (Nagy, 1990), combinative (Csapó, 1998), and correlative (Bán, 2002). These abilities, which are also important in mathematics, can be assessed and developed in varied mathematical contents as well.

The Domains of Applying Mathematical Knowledge

During the development of mathematical concepts there is a continual interaction between the observed phenomena and the emerging mathematical concepts. As Rényi (2005, p. 39) puts it:

"When children are taught to count first they count pebbles or sticks. Only if children are able to count pebbles or sticks will they be able to realize that not only two pebbles and three pebbles are five, but two something and three something are always five something, i.e. two and three are five."

Understanding and using mathematical concepts can happen at several levels. In cognitive approach, understanding is quite often seen as eliminating cognitive dissonance (Dobi, 2002). The level of understanding, when mathematical concepts themselves are considered (as Rényi did), represents a higher level of understanding. This higher level of understanding is labeled 'reflective thinking' by Skemp (1975).

When the application of mathematical knowledge is assessed, children solve word problems in which concepts and mathematical phenomena known from everyday experience also appear. Since ancient times at least three functions of the word problems have been intertwined.

(1) Putting mathematical knowledge into words, namely improving and practicing mathematical skills and abilities through word problems. In this case the wording of the tasks is familiar and student-friendly but not necessarily is concerned with practical problems. These kinds

of tasks are called "educational examples" by Julianna Szendrei, 2005). Routine examples or as we may call them practicing word problems primarily aim at developing and practicing mathematical skills and abilities and the concrete content of the wording can be almost freely replaced.

(2) Mathematical word problems can serve as a means of mathematical modeling of the world. The clerks in ancient Egypt and the merchants in Venice during the Renaissance were trained through mathematical tasks containing everyday situations and problems they had to cope with in real life.

(3) Recreational and riddle-like mathematical word problems also date back to many thousand years. The archetype of this kind of word problems is the nonsense sort of riddle asking "how old is the captain" and also the riddles in which the person trying to find the solution is expected to figure out what actually was meant by the task. The obviously incomplete tasks and also what is called insight problems in the psychology of creativity also belong to this field (Kontra, 1999).

The above-mentioned functions of word problems are quite often interrelated. It can happen that for some students in a given age-group a task can be a word problem to be routinely solved, while for others it can be a means of mathematical modeling of the world. Moreover depending on the way the task is set, students can consider the same word problem a practicing word problem or a problem whose solution mobilizes their resources and makes it possible to choose various mathematical models. In the chapter dealing with the application of mathematical knowledge several cases are shown in which learners who failed to realize that a word problem is actually a routine example got into disadvantageous situation. Thus it can be said that information on the process of the solution of word problems is part of mathematical knowledge. The role of teachers is emphasized by the fact that students in various grades acquire different socio-mathematical norms as to the nature, process and ritual of solving tasks.

Realistic Word Problems

As it has been described in the theoretical chapter on the application of mathematical knowledge, there is a group of word problems whose primary function is not the wording of some mathematical operation or knowledge components but rather assisting the mathematical modeling of familiar components of knowledge which are accessible in their imagination and experience. Where can the the border between routine examples and realistic word problems lie?

Actually no tasks can be called realistic or not realistic. To make a distinction between them several factors need to be taken into account. It takes at least three factors to consider a word problem realistic.

(1) The role of things and properties in the word problems: if the characters, phenomena and properties are essential part of the word problems inasmuch as their changes lead to essential changes in the process of solution, then probably it can be considered a realistic word problem.

(2) The relationship of the things in the word problem and students' knowledge. The adjective 'realistic' in the original sense refers to the fact that the things in the word problem are imaginable. It is not a requirement that only everyday objects should be found in realistic word problems. Even a combinatorial word problem involving the seven-headed dragon can be realistic.

(3) In a classroom setting it is the socio-mathematical standards that determine to what degree the ritual and regulation is binding in the steps of the solution of word problems. From this point of view what may indicate that a word problem is realistic is that the 'familiar' algorithm fails to work in the solution.

It frequently happens in case of realistic word problems that even gathering data and selecting the operations to be carried out requires finding a mathematical model implying planning, monitoring and control processes.

Authentic Word Problems

Within the set of realistic word problems a specific group called authentic word problems can be found. In the theoretical chapter the features of authentic word problems were defined and in the detailed assessment frameworks they are described and illustrated by examples.

Authentic word problems, which are often intransparent problems, rely on l students' experience and activity. It is a specific feature of authentic word problems that by emphasizing students' activities they are much more motivated and involved. One superficial characteristic is that the texts in authentic word problems are longer, in which a real problem situation is described from mathematical point of view by means of either redundant or missing data. It may also be a characteristic of authentic word problems that students are asked to set a task related to the problem described. In the process of problem solving what is essential is the fact that in authentic word problems there is no direct, obvious algorithm to the solution, thus real mathematical modeling is needed, during which different activities take place. Activities which can also be observed are gathering data from external resources or by means of discussion, measurements, debates and conversations drawing on students' knowledge gained previously.

It can happen several times, as it happens in authentic every problems as well, that there is no single, well-defined solution to the problem, however from pedagogical point of view the process of dealing with mathematics, including planning, monitoring and control can be quite often considered as the solution of authentic problems. The solution of authentic word problems sometimes takes place in noisy team work, which can be very different from the traditional mathematical classroom setting approved by both lay people and teachers.

It was George Pólya (1945, 1962) who came up with one of the first models of mathematical problem-solving. The steps of successful mathematical problem-solving described by him can mainly be found in the solution of realistic and also authentic problems. The questions raised by Pólya, which these days might as well be called meta-cognitive questions, apart from the mathematical characteristics of the problems are concerned with the relationship of the person solving the problem and the mathematical problem. "Could you describe the problem in your own words?" "Can you come up with a figure or diagram that could be conducive to the solution of the problem?"

Domains according to the Science of Mathematics

In the diagnostic assessment of mathematical knowledge tasks are normally related to various domains of mathematics. As it was presented in chapter 3 the domains of mathematics in education are in accordance with the current structure of mathematical science. Different domains are put in the foreground in different grades and the historical development of some of the domains is diverse in Hungarian public education.

Numbers, Operations and Algebra

Numbers, operations and algebra are considered the basic foundations of teaching mathematics. In grades 1 and 2 the most time and energy is devoted to the development of counting skills. This domain includes the development of the concept of numbers, the extension of numbers, the acquisition of the four basic mathematical operations, and also the preparation for algebraic thinking by using signs instead of numbers. Moreover the requirements to apply mathematical knowledge also include modeling multitude observable in reality and everyday phenomena described in terms of the basic mathematical operations.

It is also essential to take into consideration the pedagogical implications of Dehaene's triple-code theory (1994). The names of numbers (primarily natural numbers), the written form of Arabic numbers and the interrelationships of mental quantity representation attached to a given number make it possible for students to obtain an established number concept. Even before they go kindergarten children know the name of some numbers and in case of smaller quantities they use them in a meaningful way, for instance two ears, three pencils. The written form of numbers is attached to numerals mostly at school.

Research results concerning the development of quantity representations attached to numbers show that in grade 2 the mental number line in case of natural number less than 100 is rather exact and linear (Opfer & Siegler, 2003). This makes it possible that by the end of grade 2 dealing with the set of numbers less than hundred that the written form of numbers, the oral naming of numbers and some sort of quantity representation are related to each other.

Basic mathematical operations are described in terms of the general principles of developing and improving skills. The stages of development are well-known, the familiar breaking-down points which hamper students from getting to 7 after 6 or , to 17 after 16 (Nagy, 1980). It is also shown by data that sometimes the basic counting skill can become too automatic in the lower primary grades. This fact can be traced back in the quantitative comparison of word problems and reality, or in the lack of it. When algebraic signs are introduced to this age group simple geometric shapes, such as squares, circle, semi-circle are used to designate unknown quantities.

Relations and Functions

Recognizing rules and patterns in the surrounding world is considered as one of the characteristics of reasoning. In the field of mathematical thinking the recognition and description of relationships can belong to various areas depending on the data and phenomena and whether the relationship is seen as deterministic or of probabilistic nature.

In the mathematical definitions of relations and functions, sets and mappings can be found. Both sets and mappings are basic mathematical concepts, that is why it is highly important to attach them to students' everyday experience, ideas and basic concepts. In dealing with this topic special difficulties may arise as to what degree the abstract mathematical concepts of relations and functions can be associated with visual images such as the tables of "machine-games" or the curves in the Cartesian two-dimensional coordinate system.

In the National Core Curriculum most of the requirements related to functions have not been linked to particular age groups. In terms of our assessment framework it implies that the development students' thinking is to be assessed through a well-defined system of tasks. For instance the requirement in National Core Curriculum namely "recording the pairs of data, triplets of data of quantities changing simultaneously: creating and interpreting functions and sequences based on experience" applies to every age-group of public education. Regarding assessment frameworks however it should be decided how to operationalize the components of knowledge based on each other and into which age-group sections they should be put. With respect to this requirement the following questions seem to be relevant. What kind of quantities simultaneously changing should be included into the tasks? In which grades the pairs of data and triplets of data should be presented? In what ways students are expected to provide the relationship between the data? What kind of vocabulary is expected in various grades to describe character and closeness of the relationships between the variables? Besides having raised these specific questions we still consider the topic of Relations and functions highly important in the development of proportional reasoning and more generally of multiplicative mathematical reasoning.

Geometry

Geometry, like the topic of *Numbers, operations and algebra,* has been traditionally embedded in the curriculum. As it is shown by the IEA international comparative curriculum analysis, in Hungary a large section is devoted to geometry in the mathematics curriculum (Robitaille, & Garden, 1989).

In the National Core Curriculum, it is for example, orientation which is emphasized among the objectives, values and competencies, and it can be defined as a sub-section of geometry. Geometry and the topic of measurements are suitable for attaining the objectives to orientate in space and in quantitative relationships of the world.

Every aspect of cognition is touched upon when geometry is covered, and the various approaches of creativity and activities that are autonomous, based on students' own plans, in line with given conditions, especially in the initial stage of geometry teaching are to be highlighted. Furthermore, creative activities contribute to the development of cooperation and communication.

Space and plane geometric perception is being formed by children's concrete activities and also through materials gained by various techniques representing the world, as well as models, for example, objects, mosaics, photos, books, videos, computers.

In the NCTM standards mentioned above in all grades the field of measurement is separated from geometry. In our view the principles and requirements related to measurement should be included into geometry. The two lists below, in which the NCTM major requirements are pointed out, make it clear that in the Hungarian mathematics education the two areas are getting along well with each other under the umbrella of 'geometry'.

The aims and objectives in geometry outlined by NCTM standards for this age-group:

(1) The characteristics and the recognition of two-and three-dimensional geometric shapes, their designation, building, drawing, description, discussion skill in geometry.

(2) Orientation in space and time, description, designation and interpretation of relative positions in space, application of knowledge.

(3) Recognition and application of transformations such as shifting, revolving, reflection as well as recognition and creation of symmetric shapes.

(4) By means of spatial memory and visual memory producing mental images of geometric shapes, recognition and interpretation of shapes represented in various perspectives, and making use of geometric models in problem-solving.

In the topic of measurement the objectives and requirements of NCTM Standard are rather similar to that of the framework curricula.

(1) Understanding the measurable properties of objects, units, systems and processes, such as length, weight, mass, volume, area and time, comparison and organization of objects according to these properties, measurement by means of accidental and standard units, selection of the tool and unit suitable for the properties.
(2) Using techniques, tools and rules suitable for the definition of sizes (comparison of measurement, selection of units, and the use of measurement tools).

Combinatorics, Probability and Statistics

In the first six grades of public education, the objective of teaching combinatorics, probability and statistics is mainly to gain experience. Accordingly the curriculum requirements focus on the development of basic skills rather than expanding knowledge. The concrete, future-built knowledge in this domain needs the foundation of combinative and probabilistic reasoning based on experience.

In the initial stage combinative thinking is primarily established by acquiring the importance of systematization. At first what children are interested in is not the number of options, but rather seeking and coming up with options. We attempt to be demanding in two ways. Focusing on particular conditions throughout the task on the one hand and checking what they have done, whether they have produced anything like this, whether the new thing is different from the rest. In secondary school curricula and GCSE requirements probability is much more in the foreground than before. This is why the topic needs more thorough preparation in lower primary grades. However there is a great difference between *development of probabilistic approach and the calculation of probability.* Theoretical calculations are highly different from experience gained from experiments. The growing awareness of putting down data is indispensable for the deeper understanding of the topic of statistics. At the beginning what happens more frequently is considered to be more probable and only at a later stage is it modified and seen that what can happen in several ways is more probable even if it is not supported by experimental data. Accordingly the curriculum development and assessment are also based on gaining experience. The concepts of "certain", "not certain", "probable" and "possible" can be best established through games, activities and gathering everyday examples.

Summary and Future Objectives

The detailed frameworks of mathematics are as a matter of fact the starting point for the elaboration of diagnostic assessment system. This is the initial stage of a long developmental process, in which the conception of assessment has been outlined, the research results have been summarized and the contents needed for the tools of assessment have been described in detail.

The theoretical background and the detailed frameworks can eventually be upgraded relying on various resources. At this point, due to time limitations it was not possible to organize discussions with out-of-school agents (decision makers, local stakeholders). These Hungarian and English language volumes will be available for the wide scientific and professional public. In the first stage of upgrading the feedback from professional circles will be processed and used.

The second, so to speak continuous stage is relying on the latest scientific results. In some fields rapid progress is taking place, for instance the research into early childhood learning and development. Understanding and operationalization of knowledge, skills and competencies can be found in several research projects. Similarly intensive work has been going on in the field of formative and diagnostic assessment. These research results can be used for reshaping the theoretical background and improving the detailed descriptions.

The development of the frameworks can actually make use of their application. Data, produced continually by the diagnostic system can be used for the study of theoretical framework as well. The system outlined at this point is based on our current knowledge and the organization of the content and linking age-groups roughly to it can be said to be a hypothesis. It will be shown by assessment data what students know at what age and further experiments are needed to find out what level students can achieve by means of efficient organization of learning.

Analyzing the connections between various types of tasks can reveal the connections between different measuring scales used for monitoring students' development. In the short run, the tasks can be analyzed which of them belong to specific scales and which are the ones that can belong to several dimensions. Diagnostic assessment data are also important in connecting the results longitudinally. In the long run the diagnostic value of the particular problems can also be investigated which knowledge domains can determine students' achievement in the future.

References

Ainsworth, L. (2003). *Power standards. Identifying the standards that matter the most.* Englewood, CA: Advanced Learning Press.

Ainsworth, L., & Viegut, D. (2006). *Common formative assessments. How to connect standards-based instruction and assessment.* Thousand Oaks, CA: Corwin Press.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing.* New York: Longman.

Artelt, C., Baumert, J., Julius-McElvany, N., & Peschar, J. (2003). *Learners for life. Student approaches to learning*. Paris: OECD.

Bán, S. (2002). Gondolkodás a bizonytalanról: valószínűségi és korrelatív gondolkodás. [Thinking about the uncertain: probabilistic and correlational thinking] In B. Csapó (Ed.), *Az iskolai tudás*. [School knowledge] 2nd edition. (pp. 231–260). Budapest: Osiris Kiadó.

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning. Putting it into practice.* Berkshire: Open University Press.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: the classification of educational goals. Handbook I: Cognitive Domain*. New York: Longmans.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge: Cambridge University Press.

Carroll, J. B. (1998). Matematikai képességek: A faktoranalitikus módszer néhány eredménye [Mathematical abilities: Some results from factor analysis]. In R. J. Sternberg, & T. Ben-Zeev (Eds.), *A matematikai gondolkodás természete* [The nature of mathematical thinking]. (pp. 15–37). Budapest: Vince Kiadó.

Clarke, S. (2001). *Unlocking formative assessment. Practical strategies for enhancing pupils learning in primary classroom.* London: Hodder Arnold.

Clarke, S. (2005). *Formative assessment in action. Weaving the elements together.* London: Hodder Murray.

Csapó, B. (1998). *A kombinatív képesség struktúrája és fejlődése* [The structure and development of combinatorial skills]. Budapest: Akadémiai Kiadó.

Csapó, B. (2002). Az új tudás képződésének eszközei: az induktív gondolkodás [Tools of generating new knowledge: inductive reasoning]. In B. Csapó (Ed.), *Az iskolai tudás*. [School knowledge]. 2nd edition. (pp. 261–290). Budapest: Osiris Kiadó.

Csapó, B. (2004). Knowledge and competencies. In J. Letschert (Ed.), *The integrated person. How curriculum development relates to new competencies.* (pp. 35–49). Enschede: CIDREE.

Csapó, B. (2008). A tanulás dimenziói és a tudás szerveződése. [Dimensions of learning the organisation of knowledge] *Educatio*, *2*, 207–217.

Csapó, B. (2010). Goals of learning and the organization of knowledge. In E. Klieme, D. Leutner, & M. Kenk (Eds.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. 56. Beiheft der Zeitschrift für Pädagogik*. (pp. 12–27). Weinheim: Beltz.

Csíkos, Cs., Kelemen, R., & Verschaffel, L. (2011). Fifth-grade students' approaches to and beliefs of mathematics word problem solving: a large sample Hungarian study. *ZDM – The International Journal on Mathematics Education*, *43*, 561–571.

Dehaene, S. (1994). *Number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dobi, J. (2002). Megtanult és megértett matematikatudás [Mathematics knowledge learned and understood]. In B. Csapó (Ed.). *Az iskolai tudás*. (pp. 177–199). Budapest: Osiris Kiadó.

Hartig, J., Klieme, E., & Rauch, D. (Eds.) (2008). *Assessment of competencies in educational context*. Göttingen: Hogrefe.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H. E., & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards*. Bundesministerium für Bildung und Forschung, Berlin.

Kontra, J. (1999). A gondolkodás flexibilitása és a matematikai teljesítmény. *Magyar Pedagógia*, *99*(2). 141–155.

Leighton, J. P., & Gierl, M. J. (Eds.) (2007). *Cognitive diagnostic assessment for education. Theory and applications*. Cambridge: Cambridge University Press.

Marzano, R. J., & Haystead, M. W. (2008). *Making standards useful in the classroom.* Association for Supervision and Curriculum Development, Alexandria.

Marzano R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives.* (Second ed.) Thousand Oaks, CA: Corwin Press.

Nagy, J. (1980). *5–6 éves gyermekeink iskolakészültsége* [School readiness of 5–6-year old children]. Budapest: Akadémiai Kiadó.

Nagy, J. (1990). *A rendszerezési képesség kialakulása* [The emergence of organization skills]. Budapest: Akadémiai Kiadó.

Nagy, J., Józsa, K., Vidákovich, T., & F. Fenyvesi, M. (2004). *DIFER Programcsomag: Diagnosztikus fejlődésvizsgáló és kritériumorientált fejlesztő rendszer 4–8 évesek számára* [DIFER Program Package: Diagnostic assessment and criterion referenced developing system]. Szeged: Mozaik Kiadó.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

O'Neill, K., & Stansbury, K. (2000). *Developing a standards-based assessment system.* San Francisco: WestEd.
OECD (2004). *Problem solving for tomorrow's world. First measures of cross-curricular competencies from PISA 2003*. Paris: OECD.
OECD (2006). *Assessing scientific, mathematical and reading literacy. A framework for PISA 2009 Assessment Framework. Key competencies in reading, mathematics and science.* Paris: OECD.
OECD (2009). *PISA 2009 Assessment Framework. Key competencies in reading, mathematics and science*. Paris: OECD.
Opfer, J. E., & Siegler, R. S. (2007): Representational change and children's numerical estimation. *Cognitive Psychology*, *55*. 169–195.
Pólya, G. (1945). How to solve it, A new aspect of mathematical method, Princeton: Princeton University Press.
Pólya, G. (1962). Mathematical discovery. On understanding, learning, and teaching problem solving. New York: John Wiley and Sons.
Rényi, A. (2005). *Ars Mathematica*. Budapest: Typotex Kiadó.
Robitaille, D. F., & Garden, R. A. (1989). *The IEA Study of Mathemtics II: Contexts and outcomes of school mathematics*. Oxford: Pergamon Press.
Skemp, R. R. (1971). The Psychology of learning mathematics. Middlesex, UK: Penguin.
Snow, C. E., & Van Hemel, S. B. (Eds.) (2008). *Early childhood assessment.* Washington DC: The National Academies Press.
Szendrei, J. (2005). *Gondolod, hogy egyre megy? Dialógusok a matematikatanításról tanároknak, szülőknek és érdeklődőknek*. Budapest: Typotex Kiadó.
Vidákovich, T. (2002). Tudományos és hétköznapi logika: a tanulók deduktív gondolkodása. [Scientific and everyday logic: students' deductive reasoning] In B. Csapó (Ed.), *Az iskolai tudás.* [School knowledge] 2nd edition. (pp. 201–230). Budapest: Osiris Kiadó.