Csapó, B. (1997): The validity of school knowledge. In: Lehmann, R. H., Venter, Gy., van Buer, J., Seeber, S. és Peck, R. (szerk.): Erweiterte Autonomie für Schule. Bildungscontrolling and Evaluation. Studien zur Wirtschafts- und Erwachsenenpädagogik aus der Humboldt-Universität zu Berlin. Band 13,2. 129-140.

Benö Csapó

The Validity of School Knowledge

1 Introduction

Cultivating the mind or, in other words, developing higher order cognitive skills and transmitting applicable knowledge have always been declared to be primary goals of school education. However, one of the most frequent criticisms of school instruction is that these goals have seldom been attained. Both the difficulties of operationalizing the goals and the insufficiency of the teaching methods have been broadly discussed in the literature. The efforts of those reformers who have been trying to focus school instruction on improving thinking abilities were often oppressed by the overwhelming facts and figures that were poured into the subject matters. In this century, especially in the sciences and mathematics, the amount of knowledge transmitted in school increased significantly and students demonstrate impressive achievement when their knowledge is measured in the school context. On the other hand, a large portion of what students have mastered in school is „inert" knowledge that can rarely be utilised in contexts different from the one in which it was learned.

Several cross-national studies revealed that the achievements of Hungarian students were amongst the highest when their science knowledge was compared. The best results were measured in the mid 1980s, when science and mathematics achievements of 14-year-olds were especially high (see KEEVES 1992). However, several other experiences indicated that students' knowledge was not outstanding at all when the application of knowledge and their reasoning skills were considered.

Recent cognitive research has redefined the problem of inert knowledge by pointing out that most cognitive skills are context-bound and that the transfer of skills from one content area to another is limited. Applying these new approaches in instructional psychology by the early 1980s, a new conception of knowledge emerged that emphasises the relationship between propositional and procedural knowledge and interprets thinking as the use of an existing knowledge base. In Glaser's words „a major component of thinking is seen to be the possession of accessible and usable knowledge" (GLASER 1984). In the past decade, a large number of studies appeared that supported this conclusion and proposed integrating the teaching of thinking skills into subject matter instruction (RESNICK 1987; PERKINS & SALOMON 1989; RESNICK & KLOPFER 1989; NISBET 1993).

We have been studying the development of cognition in this way (CSAPÓ 1997a) and we have carried out several projects to improve thinking by using the contents of teaching materials (CSAPÓ 1990, 1992, 1995 & 1997b). These projects produced further questions that require careful theoretical analysis as well as empirical research.

First, we have to reformulate our goals concerning improving cognitive functioning. If efficient thinking can be defined in terms of accessibility and usability of knowledge, then the accessibility and usability, or, simply the organisation of knowledge is what we have to modify. Therefore, a better understanding of the way students' knowledge is organised is required. Second, new methods of evaluating students' thinking must be devised. If good thinking means intelligent use of the existing knowledge base then new ways of measuring this attribute of thinking must be discovered. Third, if students' thinking in a school context is to be improved, a better understanding of how schools work and what mechanisms of existing practice result in the present knowledge structure will be required. If the elements of instructional practice can be linked to the attributes of knowledge the ones we have to modify may be identified.

Following the logic of this research strategy, in the present project we examined what school does and does not do in the terms of developing thinking and providing students with an applicable knowledge base, or, in general, how valid is the knowledge that is the outcome of the present schooling. Thus, we translated these questions into the categories of the empirical investigations and we administered a survey of the structure of knowledge that is the product of the present school. To maintain the link with the previous research, we used some existing measurement tools, and devised some new ones to reflect the new orientation. In this paper I present preliminary results of this survey. Some results have already been communicated at conferences (CSAPÓ 1996; CSAPÓ & KOROM 1996; CSAPÓ & NÉMETH 1996; VIDÁKOVICH 1996; KOROM 1997) and have been published in Hungarian (KOROM & CSAPÓ 1997). A more comprehensive publication of the results of the entire study is in preparation (CSAPÓ, in press).

2 Methods

2.1 Subjects

The samples for the study were drawn from Szeged and its metropolitan area. Szeged is one of the main cities in the South-East part of Hungary with a population of 180,000. National surveys show that student achievements in the schools of the city are near the national average, sometimes surpassing it. Whole school classes were the sampling units that were proportionally selected from the schools of the area for a proper representation of the different school types.

To study the age differences, two samples were selected. The seventh grade sample was drawn from the (eight year) elementary schools and it can be considered as representative for the whole population. The eleventh grader sample was proportionally selected from the two types of four year secondary schools (gymnasiums and technical schools) that prepare students to enter higher education. Since vocational school students were not involved in the survey, the eleventh grade sample represents only students attending four-year-long high school. The number of students in each sample exceeded 500.

2.2 Instruments

The system of instruments used for data collection was not based on a firm taxonomy. However, an attempt was made to represent the types of propositional and procedural knowledge that comport with the aims of the study. We devised a model in which four levels of achievement indicators are utilised. As for the concrete data collection tools, we tried to find a balance between the previously used instruments and newly developed ones. The system of instruments used in the survey is depicted in figure 1.

At the first level of the knowledge model of this study we simple collected students' grades assigned by teachers for the major school subjects. These grades are indicators of students' school achievements and largely determine their school carrier and their possibilities for further education. However, not only their subject matter knowledge determines the grades, but also the effects of social and verbal skills as well as the teachers' subjective judgements and biases moderates the grades. Thus, the grades can only be regarded as the representations of the „local value“ of student knowledge.

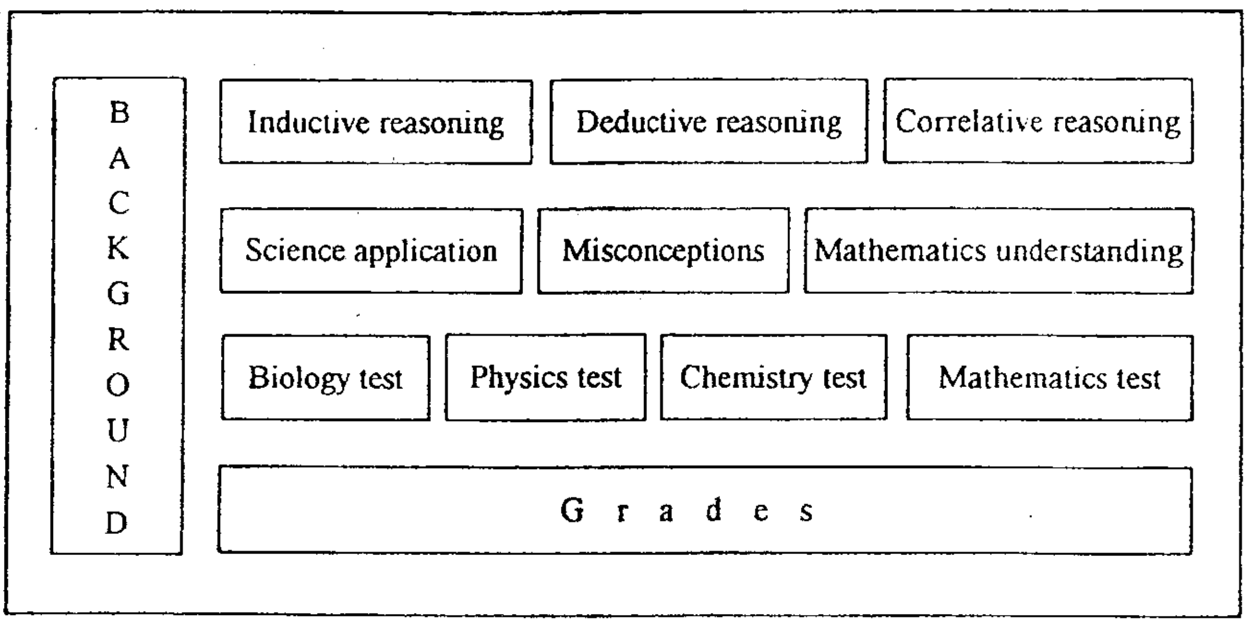
At the second level, standardised tests were administered to measure student knowledge of the three major science subjects and mathematics. These tests covered the knowledge taught during an entire school year. Since the same tests were administered to all students in the given age group (and separate ones to the two age groups), these tests measured subject matter knowledge on the same scale and so achievements of the students attending different schools are comparable.

At the third level, the quality of knowledge was assessed: how well was it understood and how could it be applied. At this level we administered a science application test that was already used in our previous studies (see CSAPÓ 1997a) and was further developed for this survey. It consists of 35 items that can be solved by applying knowledge learned in science subjects to everyday problems (for a description of the test items, see NÈMETH and CSAPÓ 1996). The mathematics comprehension test contains selected problems that can be solved only if students have a deeper understanding of the corresponding areas of mathematics. A task battery was used to study science misconceptions. Among the seven tasks, there are two known from the literature and five devised for this study (see KOROM & CSAPÓ 1996; KOROM 1997). Because the qualitatively different tasks do not cover a consistent body of knowledge these tasks do not form one single test. However, relationships between the achievements in these tasks were close enough to represent these task with one single variable for further computations.

At the fourth level, reasoning skills were assessed with tasks different from the ones that are usually used in the context of school instruction. The content of the tasks is not directly related to the content of teaching. Three tests (number analogies, verbal analogies and number series) developed earlier (see CSAPÓ 1997a) were used to evaluate inductive reasoning. The deductive reasoning test was reconstructed on the basis of previous experiences with the earlier versions of these tests (see VIDÁKOVICH 1996). The deductive tasks were designed to study the structural characteristics of deductive reasoning, but they can also be utilised to produce a quantitative measure for deductive reasoning. In the present analysis, only those tasks that were administered to both age groups will be used.

Beyond these cognitive variables, the study involved a number of background variables. Among others, we assessed student attitudes toward school subjects, their satisfaction with their achievements and their academic self-concepts. Data were also collected about the students' intention of further learning as well as about the level of education of their parents.

Figure 1: Levels and components of knowledge. A system of the measurement instruments used in the study.



2.3 Data collection

The data were collected in the end of the school year within a period of four weeks. All tests were administered to the entire classes, one test at a time. The four knowledge tests were placed into the regular evaluation processes of schools, so they were also graded and they counted in the students' final results. In this way, it was ensured that students were motivated to achieve well in these tests, as well as in any other evaluation situation at the schools.

3 Results and discussion

The data collection instruments were designed for precise analyses of the knowledge of the given domains. These detailed results will be published elsewhere (CSAPÓ, in press). In this paper I will present a general overview of the results and illustrate the main findings with some examples.

3.1 Grades as indicators of achievements

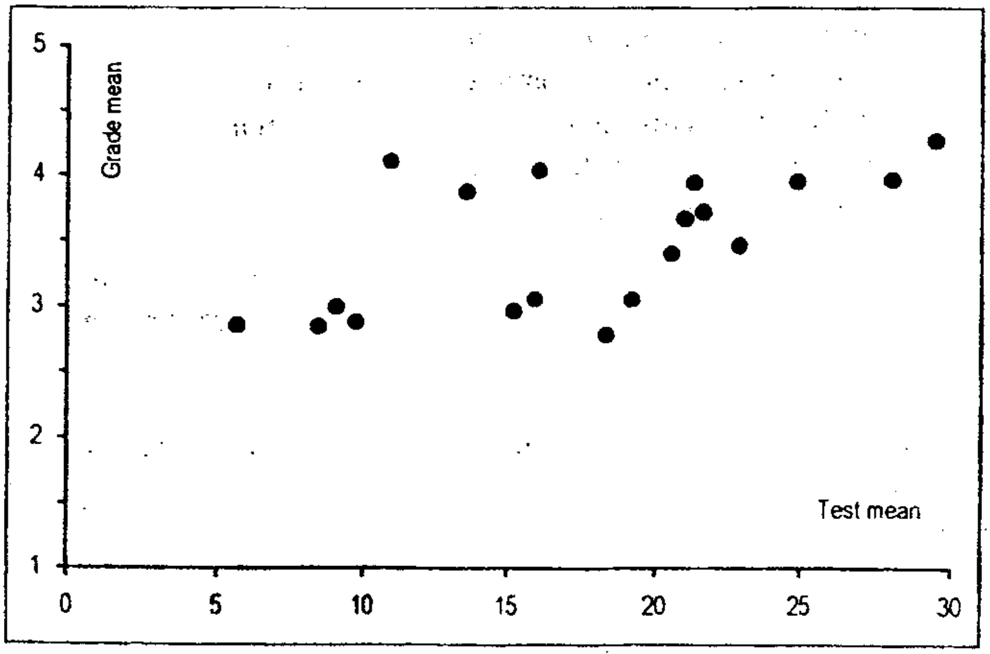
In this study - similarly to the results of several other surveys - we have found, that the grades given by the teachers' grading practice will be examined here:

(1) how well different teachers' in different schools and classes use the same standard when evaluating their students' knowledge,

(2) how consistent is the teachers grading practice when they grade the students in the same class, and

(3) how well the grades represent the developmental level of students' higher order thinking skill. Results of simple correlational analyses will be represented here. More sophisticated techniques of data analysis resulted in similar conclusions (See CSAPÓ 1977; CSAPÓ, in press).

Figure 2: The relationship between the class mean of the test scores and class mean of grades in seventh grade physics.



If we compute the means of test scores for the participating classes, we can see the differences between the average achievements of the students at the different schools. If teachers used the same standard for grading their students, these differences should appear in the differences of the class means of grades. The reality sometimes is far from this ideal situation. The problem is illustrated in figure 2, where class means of test scores and class means of grades are plotted for the seventh grade physics. Among the examined school subjects, this discipline represents the worst situation: teachers use very different standards in grading. On the figure, each point represents one school class. As it can be seen on the figure, in some classes scores below 15 are enough for a mean grade of 4, while in other classes with about the same mean grade students achieved scores over 25.

This observation can be quantified if the correlation coefficients between class means grades and test scores are computed. Table 1 presents these correlation coefficients for most of the subjects separately for the two age groups. (Since biology and chemistry are not taught in every 11th grade class, these are missing from the table.)

Table 1: Correlations between the class means of grades and class means of test scores

Subject	**r**
Biology 7th grade	0.67
Physics 7th grade	0.59
Chemistry 7th grade	0.63
Mathematics 7th grade	0.75
Physics 11th grade	0.61
Mathematics 11th grade	0.92

As the data show, the correlations are low for physics in both ages, and high for mathematics, for both ages. The correlations is especially high for the eleventh grade mathematics and this indicates that it is possible to share the same standards in grading. However, this case is rather exceptional and teachers of other disciplines grade their students according to very different standards.

The teacher' grading practice is not much better if the within-class consistency is considered. Characteristics of within-class correlations are presented in table 2, separately for the two ages.

Table 2: ,The characteristics of the within-class score-grade correlations

	7th grade			**11th grade**		
	Biology	**Physics**	**Chemistry**	**Mathematics**	**Physics**	**Chemistry**
Lowest value	0.21	0.07	0.27	0.02	0.02	0.21
Highest value	0.90	0.84	0.87	0.82	0.70	0.71
Mean value	0.49	0.53	0.60	0.55	0.37	0.49

The correlation coefficients vary in a broad range. In some classes the correlation is very high (above 0.80), so some teachers grades his or her students very consequently, assigning higher grades for higher achievement. The grading practice of many other teachers lacks such consistency.

Finally, we can examine, how well the grades represent the developmental level of students' higher order thinking skills. If the grades do not correlate very highly with students' disciplinary knowledge, we may presume that teachers evaluate something else when they assign grades, for example, more or less consciously, they asses their students' application skills or thinking abilities. Table 3 shows the correlations between grades and the application and reasoning test scores.

Table 3: The correlations between grades in the four subjects and the application and reasoning skill tests

	7th grade				11th grade			
	Math.	**Physics**	**Chemist.**	**Biology**	**Math.**	**Physics**	**Chemist.**	**Biology**
Misconceptions	0.12	0.10	0.14	0.07	0.31	0.30	0.10	0.29
Science applic.	0.38	0.35	0.37	0.29	0.34	0.35	0.23	0.40
Math. underst.	0.50	0.46	0.52	0.44	0.45	0.39	0.27	0.35
Inductive	0.59	0.52	0.60	0.48	0.25	0.27	0.30	0.37
Deductive	0.19	0.13	0.18	0.09	0.18	0.13	0.17	0.20
Correlative	0.23	0.23	0.21	0.16	0.16	0.17	0.08	0.21

The correlations for the two age groups were separately computed here, too. The coefficients are usually low, only inductive reasoning and mathematics understanding show closer relationships with the grades, especially in the seventh grade. Thus, the quality of students' knowledge does not play and important role in determining grades.

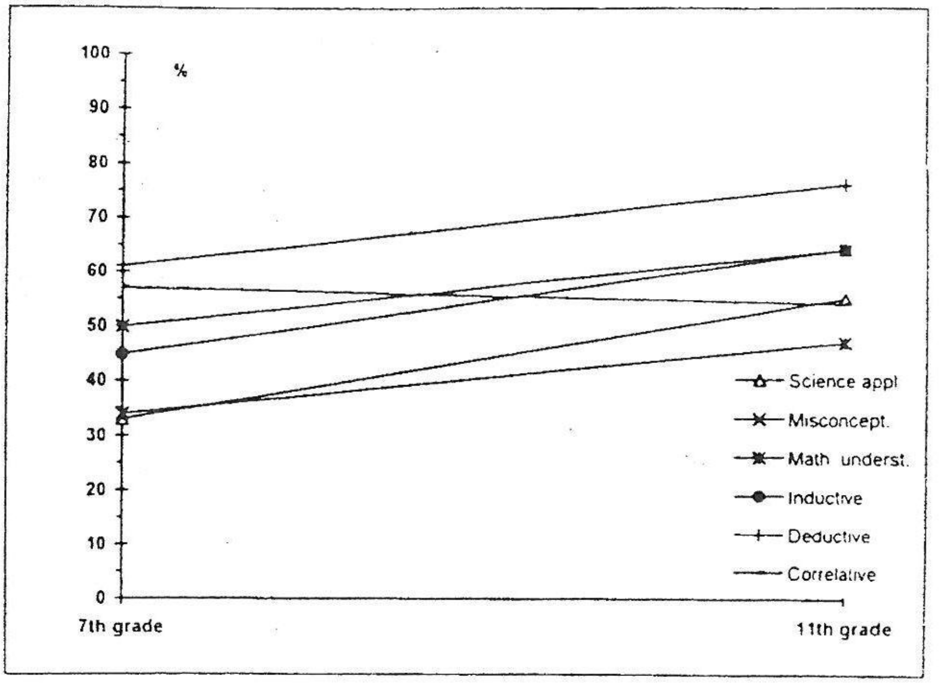
The feedback that the teachers give to students by evaluating them, especially by grading their achievement guides students as to what and how they are expected to learn. The analyses presented here indicate that the feedback given by the teachers' grading does not orient students toward meaningful learning. In serveral subjects (e.g. in seventh grade physics) we have not found strong correlation between the grades and other cognitive variables, so, on the basis of the present data we cannot explain, what teachers really grade. We may hypothesise that other variables that are beyond the cognitive domain have a major impact on the students' school grades, presumably their attitudes, their social skills and their behavior in general.

3.2 Changes in students' knowledge

At the first (grades in the school subjects) and second level (standardised texts) of knowledge indicators, the two age groups cannot be directly compared. The grades are distributed almost the same way in the two age groups. Since different knowledge tests were administered to the two age groups, the differences in the knowledge learned in school cannot be directly measured. However, comparison of the curricula on which the tests are based reveals a large increase in student knowledge. At the third level, the same three tests were administered in the same form to both age groups. At the fourth, the same version of inductive reasoning test was taken by both groups. Although the deductive and correlative reasoning tests used different designs to determine the structure of that reasoning forms, they also have some common core tasks that made comparison possible. Thus, we have six tests that make it possible to analyse what is the impact of four years of school learning on the students` applicable knowledge and reasoning skills

The results of the survey demonstrated (see Figure 3) that the differences between the achievements of the two age groups are not as large (only ca. one half standard deviation) as we may expect. In other words, students' applicable knowledge, the quality of their reasoning does not change very much during that period of schooling which, for most of them, is the time of the most intensive learning. The school obviously does not do as good a job in educating the mind in a broader sense as it does in transmitting curricular knowledge.

Figure 3: The achievements on the tests in the two age groups



In the case of correlational reasoning, students do not have a firm concept of correlational relationship. In a former study we found that older students were less likely to accept a correlational relationship as a relationship than the younger ones. Inspired by these results, in this study a more elaborated task battery was used that disclosed that students had a very slight idea of correlational relationship. In this test older students have lower scores, so this survey confirmed again that schooling has a negative effect on the students` correlational reasoning.

3.3 Structure of knowledge

Several ways exist to present the relationship (and lack of them) that were found among the large number of variables. In this paper I illustrate only the most general findings.

Reasoning skills are both 'instruments' of learning and the products of learning. For example, when students master factual knowledge in a meaningful way they use the procedures of inductive and deductive reasoning. This practice may contribute to the development of reasoning skills as well. Accordingly, when regression analysis models are constructed, the use of the examined components of cognition, sometimes as dependent variables, sometimes as independent, is plausible. In this way, by viewing the same variable from different angles, we learn a great deal about the role of reasoning skills in the cognitive processes. Therefore, one way of indicating how well a variable can be described with the other variables in the study is the square of the multiple regression coefficient (R Square). The R Square for the six examined variables (when every available and relevant variable is entered into the regression model) are summarised in table 4, separately computed for the two age groups.

Table 4: Squares of multiple regression coefficients when the test score is used as dependent variable

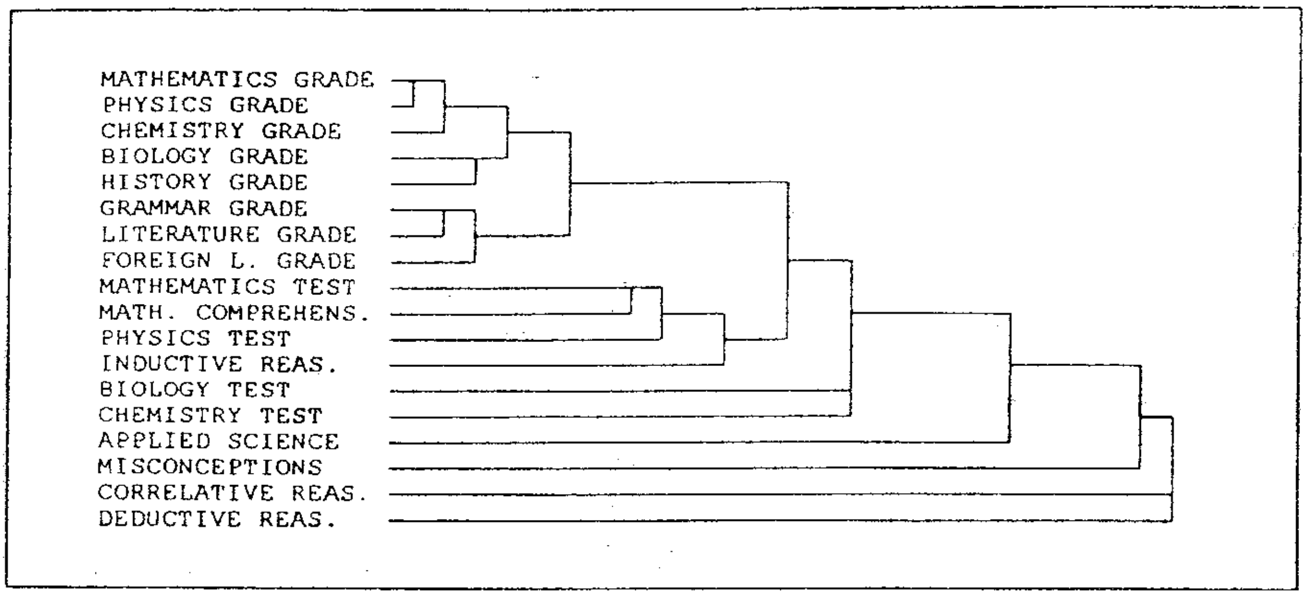
Test	**7th Grade**	**11th Grade**
Inductive reasoning	0.579	0.612
Deductive reasoning	0.217	0.404
Correlative reasoning	0.172	0.256 (n.s.)
Science misconceptions	0.250	0.805
Science application	0.469	0.708
Mathematics understanding	0.513	0.788

The data in the table clearly demonstrate that students' knowledge is organised in different ways in the two ages. In the seventh grade, inductive reasoning has 'the greatest relationship' to the other variables, e.g. its variance can best be explained by the other variables. By the eleventh grade, the relationship between the examined cognitive functions are increasing in general, especially those representing the quality (accessibility, applicability) of knowledge. Their R square are almost exactly the same, 0.80. In summary, these relationship indicate that students' knowledge becomes more consistent by grade eleven than in grade seven. On the other hand, we again found that correlational reasoning is uncertain.

These observations can be supported by other analyses as well. A cluster analysis (based on the correlation coefficients) that involves variables from each of the four levels of achievement indicators also demonstrate that knowledge is becoming more organised as the students age. Furthermore, this analysis highlights the problems of school grading practices.

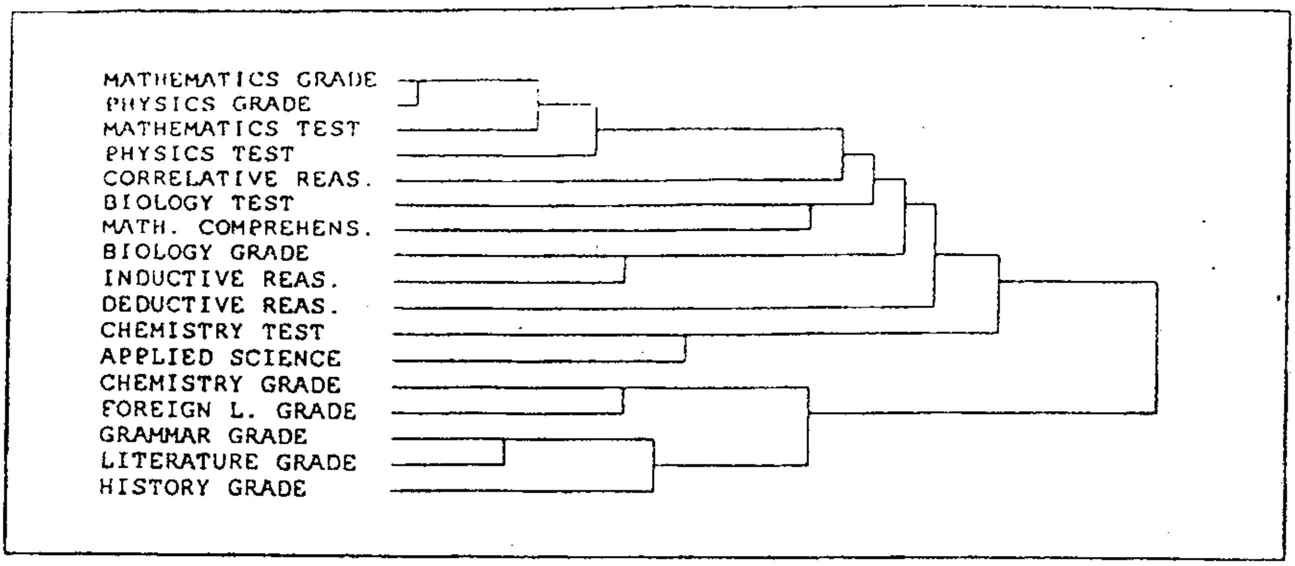
In the Figure 4, the dendrogram of school grades, knowledge tests and the six other measurement instruments is depicted. As can be seen, most of the variables fall into easily identifiable groups. The school grades from the most compact group and this group of variables is clearly separated from the others. Some knowledge tests (mathematics, physics) mix with the mathematics comprehension and inductive reasoning tests. The science application, misconceptions, correlational reasoning and deductive reasoning tests join with the weakest relationships to the other variables. The separa tion of the knowledge tests from the grades indicate that besides student knowledge, several other factors many influence what grade the teacher assigns to a certain student.

Figure 4: The dendrogram of the cluster analysis of the variables for the seventh grade



The dendrogram of the same variables for the eleventh grade is presented in Figure 5. Here the knowledge tests are closer to the corresponding grades. For example, the closest relationship can be found between mathematics and physics grades, but immediately after this the tests of mathematics and physics join the grades. In this older age, only the grades of the humanities (history, grammar, literature, foreign language) from a more or less independent group, and interestingly chemistry belongs to this group. The reasoning tests are somewhere at the medium range of 'relationship richness'.

Figure 5: The dendogram of the cluster analysis of the variables for the eleventh grade



4 General conclusions

The findings of our study are consistent with the main statements of educational and cognitive psychology. Our results confirmed that the application of knowledge to new situations does not come automatically with mastering the knowledge base; that students can achieve well at school just by performing routine studies without a deeper comprehension and that misconceptions are very persistent structures.

Although the changes in student reasoning (measured with our instruments) are relatively small during the four years of schooling, the fact that students become more efficient thinkers is unquestionable. The analysis of the relationships between the variables in the study reveals that parallel with the accumulation of propositional knowledge and the development of cognitive skills, a third, qualitative change takes place: students' knowledge becomes more consistent and more rich in relationships. Thus, when searching for new ways for evaluating the effectiveness of teaching thinking, we should find methods for assessing the consistency of knowledge at the individual level.

This study again confirmed the centrality of inductive reasoning. It usually produced stronger relationships with the other variables than deductive reasoning. Although we cannot prove with such correlational studies that improving inductive reasoning would improve the application skills or the quality of knowledge in general, these results are at least encouraging and support the need for continued experimentation in this field.

Literature

NÉMETH, M. & CSAPÓ, B.: Learning for the schools or learning for the life? Poster presented at 'The Growing Mind'. International Conference in Honor of the Centennial of Jean Piaget's Birth. Geneva 1996.

CSAPÒ, B.: Integrating the development of the operational abilities of thinking and the transmission of knowledge. In: MANDL, H., DE CORTE, E., BENNETT, N. et al. (eds.): Learning and instruction. European research in an international context. Volume 2.2. Analysis of complex skills and complex knowledge domains. Oxford 1990. 85-94.

CSAPÒ, B.: Improving Operational Abilities in Children. In: DEMETRIOU, A., SHYER ÉS, M, EFKLIDES, A. (eds.): Neo-Piagetian Theories of Cognitive Development. Implications and applications for education. Routledge, London 1992. 144-159.

CSAPÒ, B.: Improving inductive reasoning through the content of teaching materials. Paper presented in the symposium 'Teaching Intelligence'. Sixth European Conference for Research on Learning and Instruction. University of Nijmegen. 1995.

CSAPÒ, B.: Educating the mind: What school does and does not do. Paper presented in the symposium „Shaping the mind through education". The Growing Mind'. International Conference in Honor of the Centennial of Jean Peaget's Birth. Geneva. 1996.

CSAPÒ, B.: Development of inductive reasoning: Cross-sectional measurements in an educational context. International Journal of Behavioural Development. 20(4). 1997a. 609-626.

CSOPÒ, B.: Operational Enrichment: Improving operational reasoning through the content of teaching. In: HAMERS, J. & OVERTOOM, M. (eds.): Teaching thinking in Europe. Inventory of European Programmes. Utrecht, Sardes. 1997b. 235-239.

CSAPÒ, B.: Az iskolai tudás [The school knowledge]. Osiris Kiadó, Budapest. In press.

GLASER, R.: Education and thinking. The role of knowledge. American Psychologist. 39(2). 1984. 93-104.

KEEVES, J.: The IEA study of science III. Changes in science education and achievement. 1970 to 1984. Oxford 1992.

KOROM, E.: The effects of science education on children's naive Theories. Paper presented at the 7th European of Conference for research on Learning and Instruction. Athens 1997.

KOROM, E. & CSAPÓ, B.: Development of children's misconceptions of the world. Poster presented at ' The Growing Mind'. International Conference in Honor of the Centennial of Jean Piaget's Birth. Geneva 1996.

NISBET, J.: The thinking curriculum. Educational psychology. 13(3-4). 1993. 281-290.

PERKINS, D. & SALOMON, G.: Are cognitive skills context bound? Educational Researcher. 18(1). 1989. 16-25.

RESNIK, L.: Education and learning to think. Washington, D.C. 1987.

RESNIK, L. & KLOPFER, L. (eds.): Toward the Thinking Curriculum: current cognitive research. Association for Supervision and Curriculum Development. Alexandria 1989.

VIDÁKOVICH, T.: Development of deductive reasoning. Poster presented at 'The Growing Mind'. International Conference in Honor of the Centennial of Jean Piaget's Birth. Geneva 1996.