# QUANTITATIVE PEDAGOGICAL ESTIMATION

József Nagy

Because of various reasons, experimental pedagogy in Hungary regressed after the First World War and was not able to come to the fore. Social conditions developed only in the middle of the 1960's to grant civil rights to assessment, and the methods and means of quantitative pedagogical estimation. This situation demanded the catching-up on a large backlog of material, because we had to study the results of roughly half a century of research work accumulated in the international special literature and, at the same time, adapt it to our socialist conditions.

It was to our advantage that the beginning of the renewal of quantitative pedagogical estimation, that is, the elaboration of notions such as oriented criterium formative estimation, total aquisition, the development of many other new trends, and the emergence of dissatisfaction with traditional standardized tests came about in this period. Thus it was possible to avoid the early inflexibilities and exaggerations of quantitative pedagogical estimation. At the same time, the succesful effort of László Kalmár to build up the Cybernetic Laboratory /computer centre/ in Szeged also occurred at this time. It gave us the opportunity of working with large samples and of applying such methods of estimation which had been available for a long time but with large samples and manual methods had been practically unusable.

We can count as favourable the fact that there was a great interest on the part of the pedagogical public; hundreds and thousands of teachers took part in the work without any compensation. Simultaneously, a period of quick learning started in other pedagogical institutes, too.

It is characteristic of the situation that the mathematical statistical handbook written for psychologist appeared only in 1968.

The first handbook written in Hungarian for teachers, entitled "Assessment Methods in Pedagogy" /Ágoston-Nagy-Oross

1971./, was written by us on the basis of our early experiences.

On the strength of our initial results we met with good-will and gained support from those responsible in the Ministry for the development of the research work.

Our research can be broken up into two periods: the work carried out from the middle of the 60's to the middle of the 70's and the new 10-year period begun in the middle of the 70's. For the sake of simplicity we present our efforts and results in accordance with these two periods.

I.

Taking international experiences indicating problems into consideration, we started from the principle that a quantitative pedagogical estimation is needed such as can help both the practice and theory of pedagogy and the development of the system of public education. Therefore, we intended to develop tests and to carry out representative investigations which, at the same time, would serve several /all the basic/ functions. In our day, these investigations are called multifuncitonal estimations /and the means are called multi-functional tests/.

Since we have little experience, first of all we worked in the field of knowledge and skills. Within this, we chose two fields: /concerning certain skills/ the mapping of the development of skills and the elaboration of tests which are used at the end of a thematic unit to estimate the efficiency of the treatment of the theme.

In accerdance with the facts mentioned above we wished to achieve the following aims in these two fields:

- to bring about types of tests such as are suitable for these two kinds of estimation, to elaborate methods for constructing tests;
- to give the teachers such tests on the contents in question which are usable in pedagogical practice;
- to reveal reliably and in great detail the level of attainment in Hungary on the basis of a representative investigation made by tests on a

country-wide scale;
- with the help of tests measuring skills, to learn the developmental process taking place in the population /practically it means, that from the first year of teaching to the end of the twelfth school-year, with independent samples and with the same series of tests in every grade, we sized up the level of development of a given skill/.

As a starting hypothesis in constructing tests we introduced the principle of structural totality which we developed step by step on the basis of our experiences.

First of all, this principle directs that the matter to be assessed must be submitted to a careful structural analysis to obtain a system of theme, knowledge, notion and skill which is complete, logically closed and which is described in detail.

For all the "logical peaks, points" of the structure revealed in this way, we perform suitable work.

Thus, for an average theme /about 10-12 lessons/ 100-200 /200-400 items/, for an averagely complicated skill some dozens, and for more complicated ones some hundred items of work requiring much time, are made.

If we want to learn the level of attainment in a given theme in details, that is, in such a way, that we may be able to establish a diagnosis which will help in the development of curricular design and the directing of the pedagogical process, we cannot be satisfied with the starting theses of the traditional theory of tests. According to the latter, we retain samples from the possible matter, as population of the test /if it is possible, with the help of the method of accidental choice/ and we suppose that these elements represent the whole matter to be assessed. This method may be suitable for making distinctions among pupils and for qualifying the pupils' achievement. But we know nothing about the level of attainment in elements not occuring in the test,

therefore these tests are less suitable for diagnosis, for helping the directing of the study process, developing curricular design and directing and developing the educational system.

Thus, the principle of structural totality goes beyond the traditional theory of tests and makes the tests suitable for directing estimation.

Naturally, it is impossible for one pupil to solve the exercises in all the elements to be assessed. However, it is not necessary. We can separate the exercises into parallel-tests. If these test-variants are measured on independent samples, we get the necessary values for the level of attainment of every exercise /every logical meeting-point of the structure in question/.

If we express the values obtained in every parallel-test on a standard scale, the results of all the test-variants can be compared.

The concept of a problem-bank /item bank/ was still unknown at the time of the introduction of the principle of structural totality. It came about later, that this principle led us to form the special notion of a problem-bank.

Then we carried this recognition consciously so as to reveal all the possible types of elementary exercises with text /384 types of exercise/. From these we made up 48 parallel tests, and, with all of them, on an independent sample /that is, every pupil got different tests which began with fixed serial numbers/ we carried out a representative investigation in the country from the fourth to the twelfth grade in every class. As a result, we could see the developmental process of every type of exercise. We think, there is no need to account for the diagnostical power and ruling role of these tests and investigations. /Nagy-Csáki: Junior School Bank of Exercises with Text, 1976./

Otherwise, the teacher can choose as he likes from the problem-bank and can establish means of estimation because, with a simple calculation, he can work out the national level for comparison.

Beyond this theme mentioned as an example, we explored the developmental process of skills in the following fields.

As antecedents of solving verbal problems, we /Nagy, 1971/ investigated the elementary skills in calculation and the skills in the four kinds of calculation, including the skills of writing numbers and changing measurements, too. /Nagy, 1973./

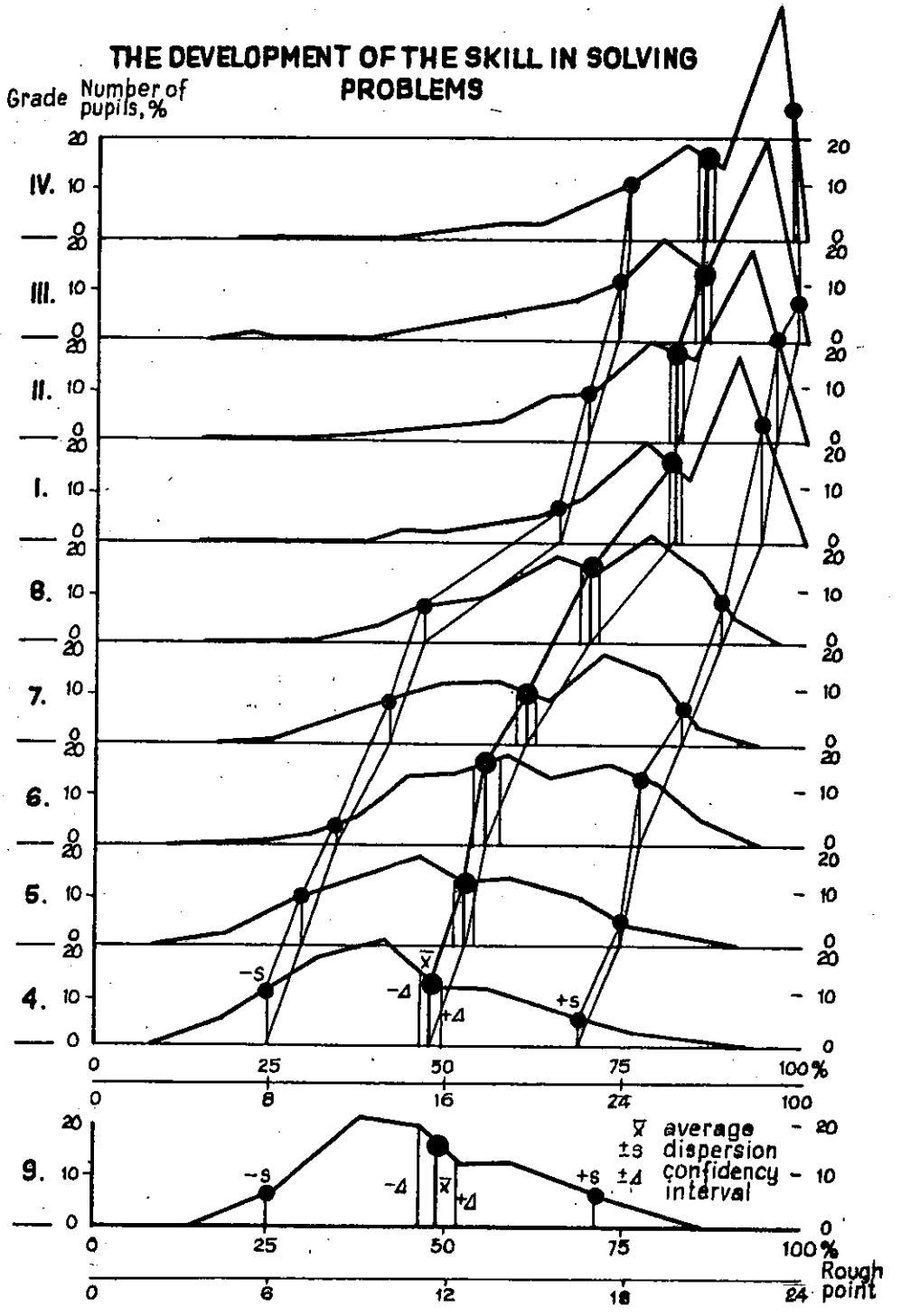
In the field of the mother tongue the subjects of our research were skills in composition-technique /Orosz, 1972/ and spelling /Orosz, 1974./. /The mapping of the development of skill in writing was also in our plans but the colleauge who worked on this theme, because of other committments could not finish the research work./

The special characteristic features of tests measuring skills, the methods of their construction, the special sampling and treating procedures resulting from the structural totality were formed in these fields.

The tests /that is, the test-variants brought together by means of suitable methods from the problem-bank/ show the degree of total attainment. On the one hand, this degree depends on the matter determined by the structural totality, on the other, from the 100 % achievement obtainalde in the test we deduct points for mistakes which can be made. /We do it in the light of a reliability-index and on the basis of an analysis of distributions./ For example, the total attainment in the change of measuring units is shown by, a minimum 90 % achievement and that in division an 85 % achievement.

The developmental process taking place among the pupils over the course of years is described by the series of distributions. /A typical illustration of this can be seen in Fig. 1./

The tests used at the close of a thematic unit were made in five subjects /Hungarian grammar, mathematics, physics, chemistry, natural history/ for all themes in all




Note: The tests in the 3rd grade are different, therefore they cannot be compared to those of other grades.
Source: Nagy - Csáki: Junior School Bank of Exercises, 1976

grades in the senior section of the eight-year general school. We elaborated, at least, four, in rare cases five or six, parallel tests for each theme.

The questions of the construction and usage of tests were described in the book "The Practical Questions of Assessing Attainment When Completing a Thematic Unit" /Nagy, 1972./. We summarized the theoretical experiments of the work of many years in "The Reliability and Validity of Tests Used at the End of a Thematic Unit" /Nagy, 1975./.

The tests in the five subjects and the results of the national survey were published in 17 volumes written by the leaders of the subject-researching teams: four volumes by Sándor Orosz on Hungarian grammar /5-8th classes/, books by István Gazsó on mathematics /5-8th classes; because of the death of the author, the last volume was published under the editorship of Mrs. Kunstár/; four books by Géza Dobó on natural history, three volumes by János Veidner on physics /6-8th classes/ and two books by Elemér Kunsági and Mrs. Vida on chemistry /7-8th classes/.

To illustrate our work we show variant D, and the survey results belonging to it, from the six parallel tests on a theme in chemistry /Kunsági - Mrs.Vida, 1973./.

It is revealed in the test, that it is not the technique of choosing answers we used. It has reasons of validity which we cannot touch upon here. This problem is solved by the principle of adequacy, the functional validity. "Functional validity is trivial only then when the problem-solving activity is adequate to the activity of realizing the working of the psychical structure to be assessed." /Nagy, 1975. p. 32., see further information here/

The other speciality of the test is that we divide every problem into, so-called, alternative elements about which only two definite qualifications can be given when it comes to estimation: 1. good, perfect, 2. imperfect, bad, insufficient. This is what, with the help of a key, makes possible an objective estimation which, otherwise, must be checked using suitable methods in the course of making up the test.

Work sheet measuring achievement at the end of the theme — Variant D

Eight-year general school — Name:

Chemistry, form 7 — Class:

## THE MOST IMPORTANT CARBON COMPOUNDS

1. We heat wood /or coal/ in an air-proof space. What is this process called?

   a/

   Name the products too!

   The name of the product: state:

   b/ c/

   d/ e/

   f/ g/

| a | b | c | d | e | f | g | |
|---|---|---|---|---|---|---|---|
| 2 | 2 | 2 | 1 | 1 | 1 | 1 | |

2. Complete the following:

   The state of carbon dioxide: a/ a/

   Its specific weight compared to air: b/

   c/

   Its influence on the organism: d/

| a | b | c | d | |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | |

3. 


   Which important characteristic feature of carbon monoxide gas does this figure express?

   a/

   Describe this process! Let the oxygen be in its atomic state:

   ..... + ..... = ..... + heat
   b,c d,e f,g

   Let the oxygen be a two-atom molecule! Describe the process as an equation!

   ..... + ..... = ..... + heat
   h,i j,k l,m

| a | b | c | d | e | f | g | h | i | j | k | l | m | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 4 | 3 | 2 | 3 | 2 | |

4. Which gas comes into being when destroyed organisms, having got into the depths of the earth, decay?
   a/
   Which group of compound matters does this gas belong to?
   b/
   The name of its main component: c/
   Its formula: d/

5. What is petrol used for?
   in traffic: a/
   in therapy: b/
   and in the household: c/

6. The following table summarizes the characteristics of lipoids and oils. Complete it:

| | LIPOIDS | OILS |
|---|---|---|
| Building atoms: | a/ | f/ |
| state on room-temperature: | b/ | g/ |
| specific weight compared to water: | c/ | h/ |
| its dissolvent: | d/ | i/ |
| usage: | e/ | j/ |

7. What kind of chemical process is: 
   the carbonization of sugars: a/
   its product: b/
   and what kind of chemical process is the burning of sugars: c/
   its product: ..... + ..... = .....
   d/ e/ f/

8. What kind of organism only can produce protein?

a/

Complete the following:

b/ ..... protein  c/ ..... the protein characteristic of the human organism

| a | b | c | |
|---|---|---|---|
| 2 | 2 | 2 | |

9. Name three industrial raw materials containing protein and name the industry which uses each.

| The name of the raw material containing protein | Which industry uses it ? |
|---|---|
| a/ | b/ |
| c/ | d/ |
| e/ | f/ |

| a | b | c | d | e | f | |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | |

10. Name the oxides of carbon!

a/ .....  b/

Write down the formulae of the oxides

c/ .....  d/

In what kind of chemical process can they come into being?

e/ .....  f/

| a | b | c | d | e | f | |
|---|---|---|---|---|---|---|
| 3 | 2 | 4 | 4 | 5 | 6 | |

Achievement: %score

VOLUNTARY TASK

11. We enumerate materials and characteristics relating to materials. Write the number of the proper characteristic beside the name of the material!

petroleum: a/
protein: b/
carbohydrate: c/
lipoids, oils: d/

1. Nutriments
2. Combination of hydrocarbons
3. Medium of life
4. Its becoming rancid: decay

| a | b | c | d | |
|---|---|---|---|---|
| 2 | 2 | 2 | 2 | |

The value of the voluntary task:

.....%score

MARK:

THE MOST IMPORTANT CARBON COMPOUNDS — Variant D

1. a/ dry distillation
   b/ wood gas
   c/ aeriform
   d/ coke /wood-charcoal/
   e/ solid
   f/ tar
   g/ liquid

2. a/ aeriform
   b/ heavier
   c/ does not feed
   d/ poisonous

3. a/ combustible gas
   b/ 1
   c/ CO
   d/ 1
   e/ O
   f/ 1
   g/ $CO_2$
   h/ 2
   i/ CO
   j/ 1
   k/ $O_2$
   l/ 2
   m/ $CO_2$

4. a/ natural gas
   b/ combination
   c/ methane
   d/ $CH_4$

5. a/ fuel
   b/ cleaning of wounds
   c/ detergent

6. a/ C, H, O
   b/ solid
   c/ lighter
   d/ petrol, alcohol
   e/ food
   f/ C, H, O
   g/ fluid
   h/ lighter
   i/ petrol
   j/ soap making, paint making

7. a/ dry distillation /decay/
   b/ sugar coal
   /synthetic coal/
   c/ oxidation
   d/ $CO_2$
   e/ $H_2O$
   f/ heat

8. a/ vegetable
   b/ vegetable
   c/ animal

9. a, b, c, d, e, f
   counting from the beginning, we leave uncancelled as many letters as correspond with the right answers given by the pupil

10. a, b, c, d, e, f
    counting from the beginning, we leave uncancelled as many letters as correspond with the right answers

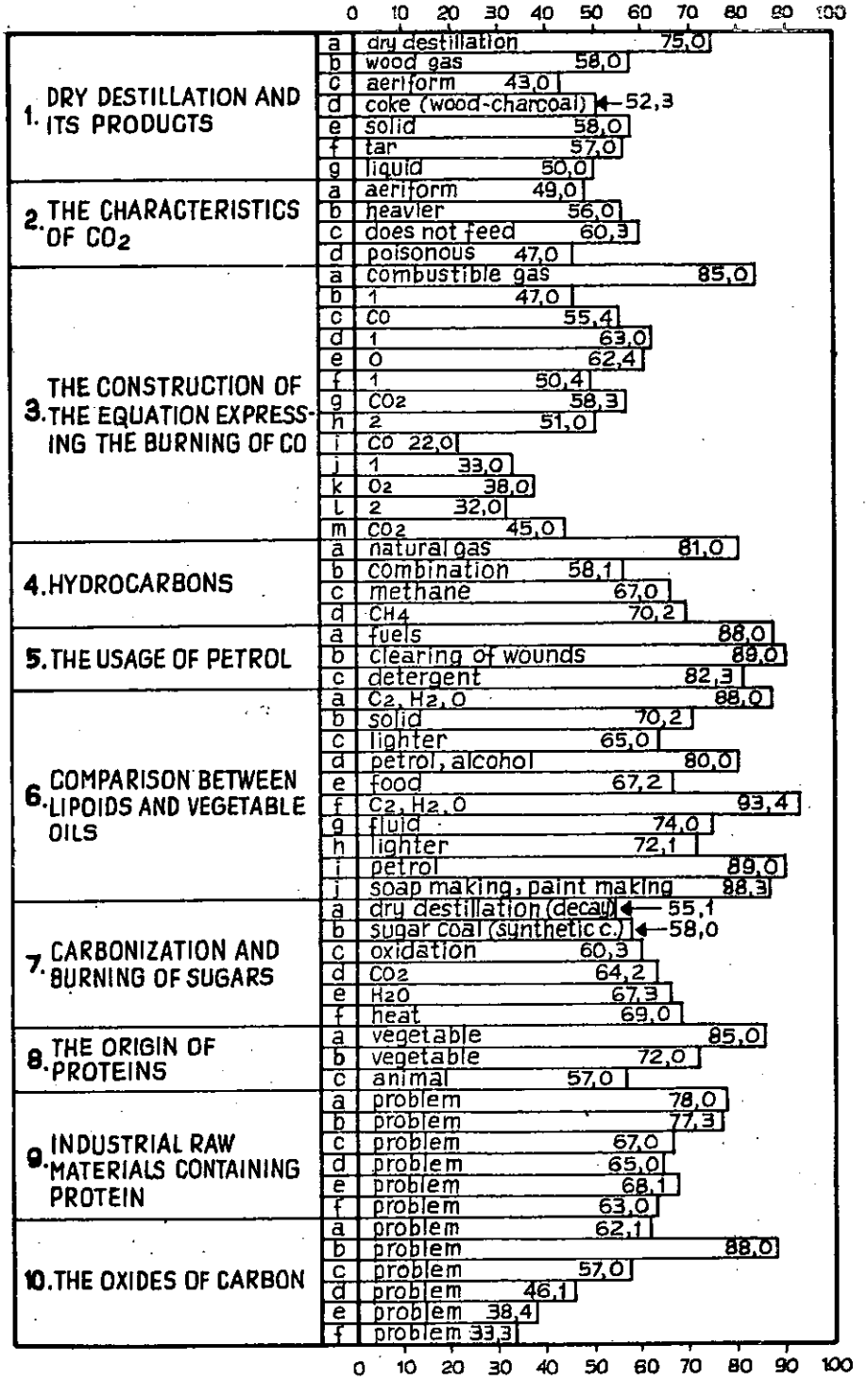
VOLUNTARY TASK

11. a/ 2
    b/ 3
    c/ 1
    d/ 4

VARIANTION IN MARKS

excellent from 84,6 to 100.0
good
satisfactory
pass mark
fail

# THE RESULTS OF VARIANT III/D



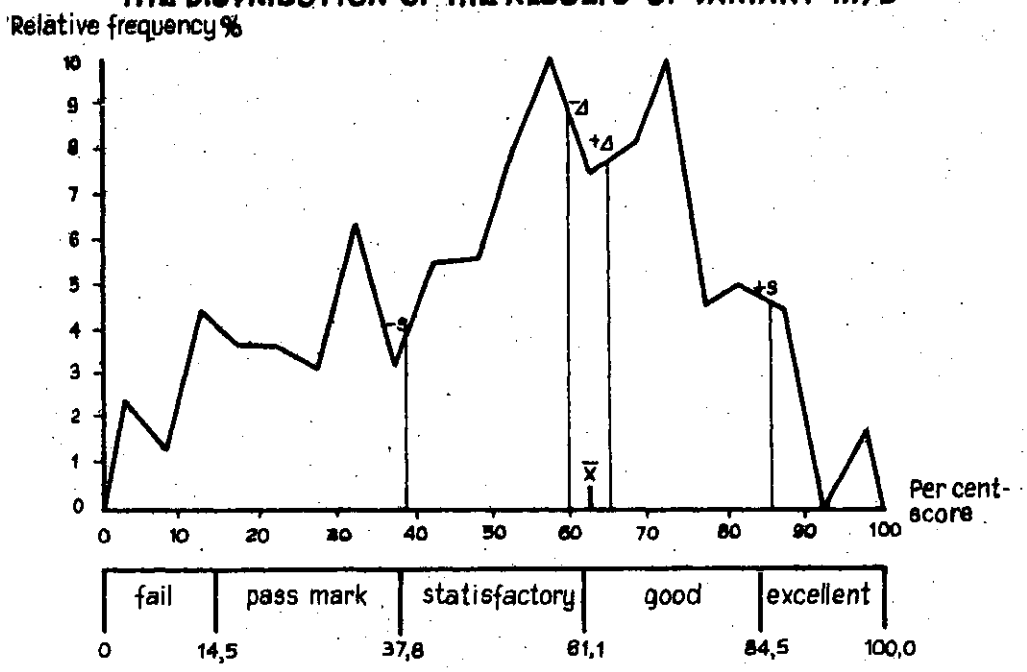
Summary data of variant III/D

| | | |
|---|---|---|
| Number of pupils | | 158 |
| Average | $\bar{x}$ | 62,9 |
| Confidency interval | $\pm\Delta$ | $\pm$3,6 |
| Requirement of accuracy | | $\pm$ 6,9 |
| Dispersion | $\pm s$ | $\pm$ 23,3 |
| Relative dispersion | | 44,4 % |

Distribution

| % score | Pupil /%/ |
|---|---|
| 0,1 - 5,0 | 2,5 |
| 5,1 - 10,0 | 1,2 |
| 10,1 - 15,0 | 4,4 |
| 15,1 - 20,0 | 3,7 |
| 20,1 - 25,0 | 3,7 |
| 25,1 - 30,0 | 3,1 |
| 30,1 - 35,0 | 6,3 |
| 35,1 - 40,0 | 3,1 |
| 40,1 - 45,0 | 5,6 |
| 45,1 - 50,0 | 5,6 |
| 50,1 - 55,0 | 8,2 |
| 55,1 - 60,0 | 10,1 |
| 60,1 - 65,0 | 7,5 |
| 65,1 - 70,0 | 8,2 |
| 70,1 - 75,0 | 10,1 |
| 75,1 - 80,0 | 4,4 |
| 80,1 - 85,0 | 5,0 |
| 85,1 - 90,0 | 4,4 |
| 90,1 - 95,0 | 0,0 |
| 95,1 -100,0 | 1,8 |

THE DISTRIBUTION OF THE RESULTS OF VARIANT III/D



In the first half of the 70's, in possession of our accumulating experiences, we ventured to undertake the solution of a complex task. To our research work serving the modernization of the schooling system we added a special system to investigate the children's maturity before school /PREFER, Nagy, 197 / with the help of which we carried out a country-wide scale representative survey among ten thousand, 5-6-year-old children. This system of examination differs from the usual tests in the fact that it does not set out to assess abilities, intelligence and behavioural difficulties but tries to be a pedagogical examination. That is, it estimates the most varied states of readiness concerning the knowledge and behaviour necessary for starting school; it estimates only such characteristics which can be developed and taught by pedagogical activity.

The national survey also had the aim of getting to know, in consideration of various factors, the level and process of development of the aptitudes, skills, knowledge and behaviour of the two age-groups awaiting school. /The publication of the results is set out in the book: "Our 5 and 6 Year-old Children's Readiness for School/

II.

The fundamental characteristic feature of the new period of research beginning in the middle of the 70's is that we intend to help the guiding of the learning process still more directly. However, we continued with our estimations which help the teaching of thematic units and help to direct learning process which lasts many years.

Concerning the teaching of themes, the aim, using the notion of mastery learning, is to elaborate such means of estimation which are necessary for the realization of the so-called theme-compensation education, but it cannot be education itself. See the study on the plan and the results of the preparative experiments in "The Directing of

Learning With the Help of Problem-banks", 1977.

The essence of theme-compensation education lies in the fact that, before teaching the theme, we carry out a so-called pre-compensation to bring to the same level those conditions of knowledge which are necessary for the teaching of the theme. After developing the theme we set up post-compensation during which those who have reached the required levels get complementary, deepening tasks, while in the others the teacher makes up the missing knowledge.

To make this system work, pre- and post-tests are needed. However, to direct learning these traditional means are less suitable.

If we expect teachers to make them up, the tests can be more adaptable to the concrete conditions in a suitable manner as opposed to standardized tests prepared in advance/ but such an expectation is Utopian, and what is more, the level of the means coming into being in this way is, often, questionable; there is no basis for comparison.

If the work sheets for the pre- and post-survey are made centrally, the elements to be achieved must be selected from a certain point of view, which may cause a dangerous one-sidedness, while the restriction /standardized tests prepared in advance cannot be changed/ may involve education in formalism.

The solution of the problem may lie in the concept of the problem-bank which we elaborated in the previous phase of our research work and which we have already characterized briefly. For the purpose of estimation, both at the beginning and at the end of the theme, it becomes possible to construct tasks involving all the preconditional elements of knowledge and for all the elements of the theme to be achieved; it also makes possible for us to carry out the necessary experiments and surveys on them and, in the form of a problem-bank, to make the tested tasks showing national levels available to the teachers /concerning the form, it seems expedient that a problem-collection should be made which bears serial numbers and from which the

exercises marked by the teacher are done by the pupils in their exercise-books/.

The research will be finished in 1981. The probable final result will be a symposium which describes the methods of the preparing and usage of these means and shows the working of theme-compensation education. Besides that, we shall publish samples from some subjects with the results of the experiment.

In possession of the experiences of estimation concerning skills and proficiences in the learning processes lasting many years, we can undertake bolder plans. We endeavour to elaborate the system of the, so-called, operational skills and to explore their development during school-age /that is, in national representative samples/. We do this work in the hope that we can contribute to the conscious development of these essential skills /until we do not know the skills and their development, their conscious progress cannot be thought of/.

On the basis of the preparatoty works, we now think that three or four dozen operational skills are at work in the thinking activity of an intellectually cultured man. It has also became obvious that these skills cannot emerge without a proper level of development of either linguistic structure, or the so-called fundamental notions. Therefore, operational skills may and must be studied together with the development of certain general linguistic structures and fundamental notions.

This research is planned over several periods. The first period will come to an end in the middle of the 80's with the elaboration of the system of operational skills, with the experimental working out of the tests, problem-banks being necessary for measuring their development, and with the exploration of the national level of the state of development in some age-sections /probably, in the ages of 8-10 and 14-16 years/. The further research will lead to a more detailed revelation of the developmental process and a more direct help for conscious progress.