National Assessment of Basic Competencies in Hungary

Ildikó Balázsi

suliNova Kht., Center for Evaluation Studies

Abstract

In 2001 the Hungarian Ministry of Education initiated a large-scale study to assess school effectiveness in preparing students to the requirements of the labor market and lifelong learning. The National Assessment of Basic Competencies (National ABC) measures the abilities of students in reading and mathematical literacy in grades 6, 8 and 10 in every school. It also involves a short Student Background Questionnaire and a School Questionnaire to map the background characteristics of students and schools. The results are reported to schools and authorities on the basis of a sample of twenty students per school and these reports help schools to get an objective description of their achievements.

My paper and presentation will focus on the key properties of the study. To familiarize the reader with the National ABC I am going to shortly describe the framework of the cognitive domains and the defined proficiency scales, the methodological properties and the summarized basic results of the assessment.

Objectives of the National ABC

The National ABC study was launched in Hungary in 2001. In accordance with international assessment trends and the traditional national Monitor studies[1], the goal of the study is not to measure whether schools have conveyed the knowledge prescribed in the curricula, but whether students are able to use their skills and knowledge in everyday life and use it as a basis for lifelong learning. The achievement tests developed are therefore not to measure how well students do on curriculum related items, but how they can utilize what they have learned, how well they can perform on matters pertaining to real life situations. Until now the assessment has been implemented in the domain of reading and mathematics literacy, but our long-term goal is to include in the assessment science as well. Originally the assessment was implemented in grades 6 and 10, then in 2004 it was extended to grade 8 – currently we include three populations in the annual National ABC.

The conceptual work of the National ABC was carried out with regards to the following aims:

- Create and evolve the assessment and evaluation practices of the schools,

[1]National Monitor Studies based on a representative student sample in reading, mathematics and science. First study in 1995, carried out every two years by the Center for Evaluation Studies.

- Introduce the assessment methodology and the competency based achievement tests to the teachers and pedagogues,
- Provide schools with data and tools that enable them to objectively analyze their own performance,
- Provide school managements with data that are reliably comparable with the national performance data.

In order to ensure these goals, the Ministry of Education decided to implement the assessment in every school in Hungary, with all students in the populations assessed to be included. The assessment aims to provide schools with a data pool, and not to provide data about the individual students, the central analyses are carried out on a representative sample. Schools of course have the possibility to work with the student data not included in the sample. In order for schools to get an accurate picture of student performance, and rank themselves within the group of schools with similar attributes, the schools' principals and students were asked to fill in questionnaires as well. The questionnaires collect the necessary information on student family background and school attributes. Using this information, one can conduct more in-depth analyses of student performance, and can investigate how well each school has faired in educating their students using their resources, how much the educational added value is.

Achievement Tests

As the National ABC investigates whether students have the necessary skills to ensure their continuous development, the achievement tests do not measure how well students do on curricular knowledge, but how well they deal with real life problems and situations. Both domain tests focus on actual everyday situations and problems, and include common tasks known from ordinary events (e.g. filling out an order form, understanding an advertisement, shopping, travel or finance related tasks.).

The *reading tests* include narrative, expository and document type texts, students have to retrieve information, form a global understanding, and reflect on the content and style of the text.[2] The texts include various prose and document types, such as lists, forms, etc.

The *mathematics tests* assess those skills and competencies that show how well students can use their mathematical knowledge in everyday life. We seek to learn whether they can plan and carry out calculations based on a set of data provided for them, whether they can solve finance related problems, or make sense of information provided by the media (self-sufficient

[2] For more information see http://www.sulinova.hu/, National ABC Study.

understanding of tables, graphs, diagrams, the integrative comprehension of different representations of a certain phenomenon). We rarely deal with purely mathematical problems in everyday life, but use our mathematics skills and knowledge in practical situations. The National ABC includes the following mathematical topics in the assessment: quantities and operations, assignments and relationships, shapes in the plane and space, statistical characteristics and probability of events.

Items are arranged into two 45 minutes reading and two 45 minutes mathematics blocks, forming two three-hour long booklets, with different block orders. The booklets contain around 60 items for each domain. Since all students in the population participate in the study, their own teachers supervise the test administration. All teachers have to abide by the Survey Operations Manual provided them by the Center for Evaluation Studies (CES). After collecting the booklets the coding of the open-ended questions are carried on by trained coders in the CES.

Methodology

The National ABC assessment requires complex and diverse analyses. Every step in the construction, management and implementation of the study serves its primary goal, to be able to collect valid and reliable data to draw conclusions about the target populations' performance, to be able to provide policymakers and schools with useful information. This goal is supported by the modern methods used during data analyses that go beyond simple statistics, such as introducing the *performance model* and *benchmarking* student performance, using *scales* and calculating *empirical error***.**

Sampling

While all students in the assessed grades fill in the booklets and questionnaires, the national data analyses are based on the performance of 20 students from each school. We use stratified random sampling to decide which 20 students' assessment booklets and background questionnaires would be asked to be sent in to the CES.

After ranking students by their mathematics grades, we use simple random sampling procedure to select an initial value and based on the number of students in the school we calculate a step size. After this, the selection of students from the school is straightforward, we use those students who are assigned to the values on the ranked list we got in the previous step, in order to have better and weaker performing students alike in our sample. Using this

method instead of the simple random sampling, we get a better estimate of the schools' average performance.

We then analyze student performance using weights (proportionally to the number of the students at a given grade at the schools), therefore we get unbiased estimates for the student population as a whole.

Performance model

The abilities of students in reading and mathematics literacy, and the item difficulty and slope were calculated using a two-parameter model[3]. The two-parameter model used assigns a performance value (θ_i) to each student, and parallel to this, assigns two parameters to each one-point item: difficulty (b_j) and slope (a_j). Difficulty shows where each item is placed on the performance scale; the slope shows how fast the probability of a successful item solution is rising with the rise in student performance. With aforementioned parameters described, the probability of Student i's success in solving Item j is presented in the following formula:

$$P_{ij}(1) = \frac{1}{1+\exp(-1{,}7a_j(\theta_i - b_j))}$$

Apart from the difficulty and the slope, on all other than 0 scores on a multiple score value item a category parameter (c_{jv}) is assigned as well. Here the probability of reaching score k is described with the following formula:

$$P_{ij}(\mathrm{k}) = \frac{\exp\left[\sum_{v=0}^{k} 1{,}7a_j(\theta_i - b_j + c_{jv})\right]}{\sum_{c=0}^{m_j} \exp\left[\sum_{v=0}^{c} 1{,}7a_j(\theta_i - b_j + c_{jv})\right]},$$

where m_j is the maximum score, $c_{j0} \equiv 0$ and $\sum_{v=1}^{m_j} c_{jv} \equiv 0$.

We analyzed the student performance scores after standardization: the national average performance score and standard deviation were 500 and 100 standard points respectively, after carrying out the transformations.

[3] Parscale 4.2 software was used to calculate the parameters.

Calculating the errors

To calculate the errors of our estimates, out of all the distribution free sampling error calculating methods we use the Jackknife method, because the performance of the students attending the same school, having the same teachers, curriculum and shared learning experiences will be strongly connected, therefore the errors of the statistical software would not be the estimates of the actual errors.

The error of the school average was also incalculable using a traditional formula because of the small case number and the correlation between the students' performance attending the same class; hence the bootstrap[4] error calculation method was used.

Student Performance Levels

For analytical purposes, the performance levels defined by pedagogical and statistical criteria were of great importance. These levels are to group the students based on their performance so that we can state that students on a given level have *at least* a certain level of competencies, and these let us describe what areas of knowledge a certain level is lacking compared to those on a higher level. We defined four performance levels in both domains in the National ABC. The methodology of assigning students to levels was based on the methods used in the PISA 2000 Study[5].

Background variables influencing student performance

Analyzing the differences in student performance is aided by information on the students' family background as well as their schools' attributes. These altogether help us judge as to what degree home environment and school influence student performance. We collected data from students and from school principals using Student Background Questionnaires and School Questionnaires.

The family characteristics were combined in one common index, creating the students' home background index (HBI) based on the parents' educational attainments, material goods indicating financial status and those tools that aid student development. We applied multi-variable linear regression and factor analysis on the questions of the student background questionnaire to find the variables that has the highest influence on student performance. The HBI indicates the quality of the students' learning background, their possibilities, skills and learning motivation, all that they "take to school" with them. Based upon this we can use

[4] JUN SHAO – DONGSHENG TU: *The Jackknife and the Bootstrap*. Springer Verlag, 1995.
[5] RAY ADAMS, MARGARET WU (ed.): *PISA 2000 Technical Report*. OECD Publishing, 2002.

linear regression to determine how much schools apply these, how high the educational added value is.

Constructing trend data

Since all students in a population write the achievement tests, the tests become public after the testing sessions. Therefore we cannot carry out direct trend analyses due to the lack of common items. In order to have comparable data for each year, we have a representative sample (approximately 20 students from 150 schools) who writes the so-called Core test after the achievement tests, this is a confidential test that remains unchanged year after year. The items of the Core test create a link between assessments and enable us to project the data of the upcoming years to the scale of the previous ones allowing us to construct trend data.

Defining student performance levels and the distribution of the students on the various performance levels

The levels of student performance show us a sort of hierarchy of the competencies. *Those students who reach a certain level not only possess all skills and knowledge typical to that level, but also those on the levels below it.* Therefore a student whose performance ranks on level three will naturally have completed the requirements of the first and second levels. A student on a given level will probably be able to answer at least half of the items ordered to that given level.

The following two tables briefly summarize the expected processes and procedures on the respective levels. These definitions are valid for all the grades assessed, but naturally we took the age-specific knowledge and skills of the students on the given grade levels into consideration.

Performance level 1 Student reaching this level can retrieve one or more pieces of explicitly stated information based on given criteria. Student can recognize basic connections within the text; identify the main idea and the author's agenda based on specific criteria. Student can evaluate key points of the text based on own background knowledge.
Performance level 2 Student reaching this level can retrieve one or more pieces of explicitly stated information based on one or more given criteria, can connect obviously similar information. Student can recognize basic connections within the text; draw lower level conclusions based on a single or multiple sections of the text. Recognizes the main idea of the text, can interpret the text's key ideas, can identify the author's intent. Student can evaluate key points and elements of the text based on own background knowledge.
Performance level 3 Student reaching this level can retrieve information based on multiple criteria. Can assess the similarity of textual elements, is able to identify information relevant to his/her task.

Can draw identify connections, draw conclusions based on sections of the text or as its whole, can arrange the elements of the text. Using background knowledge student can explain and explore the less than ordinary meaning of textual elements. Student is able to evaluate and judge the content and the style of the text. Student can reflect on the text based on own background information and opinions.
Performance level 4 Student reaching this level can retrieve embedded information, which does not appear in the text word for word. Can retrieve information based on multiple criteria, can assess information relevant to his/her task, can deduce what information is essential to his/her task, can decide and pick the correct one amongst similar information. Can investigate complex connections, can recognize the connections between parts of the text and the text as a whole, can draw conclusions based on various textual elements and interpret them. Can interpret the entire text, a segment of the text from a given point of view, unexpected points of view, and those opposing his/her views. Using background knowledge student can critically reflect on the content and the style of the text, analyze style nuances, critically analyze the text as a whole, or in parts, can form hypothesis based on textural information. Student can utilize all this even when working with a text with totally alien form and content.

Table 1. Skills and knowledge of students reaching the various performance levels in reading

Performance level 1 Student reaching this level can solve problems that involve using the most common mathematics skills. Student must evoke well-known mathematics tasks and carry out simple calculations. Evaluation the context of the problems does not need deeper understanding. Student can generally solve the problem in one step.
Performance level 2 Student reaching this level has to carry out routinely practiced tasks. The tasks at hand require the student to use much practiced, well-known techniques, processes. Recognizing patterns and rules, using known and practiced calculus strategies and following and carrying out mathematics commands. Student works with data presented in various forms, such as tables, and graphs with generally known processes.
Performance level 3 Student reaching this level can self-sufficiently plan the solution of a problem and can connect various fields of mathematics while doing so. Can accurately understand, group and combine information gathered from a task. Student usually selects a known strategy when solving the problem, sometimes develops simple procedures. Student can transfer a not too complex problem into the world of mathematics. Level includes tasks from the previous two, while the context of a problem can be known, the problems at this level are least known.
Performance level 4 Student reaching this level can self-sufficiently plan and use various methods of solution, even if the context is unfamiliar. This level requires student to combine knowledge and skills from various fields of mathematics. Communication skills are in the foreground at this level, not only does the student have to analyze and solve a problem, but has to compose mathematical arguments, has to reason, and has to consider the validity of statements (from a mathematical point of view.) This level requires highly developed mathematical thinking, deductions and reasoning skills, generalizations as well. The problems are presented in such unusual contexts, that understanding them, transforming them into mathematics problems, and the understanding of the results requires above average mathematics skills.

Table 2. Skills and knowledge of students reaching the various performance levels in mathematics

The following table shows the distribution of students on the performance levels in 2004.

Reading			
		Points assigned to the level	Percentage of students (%)
Grade 6	Below level 1	Below 336 points	4%
	Level 1	336 to 426 points	17%
	Level 2	426 to 516 points	29%
	Level 3	516 to 606 points	32%
	Level 4	Above 606 points	18%
Grade 8	Below level 1	Below 374,5 points	11%
	Level 1	374,5 to 454 points	21%
	Level 2	454 to 533,5 points	31%
	Level 3	533,5 to 613 points	25%
	Level 4	Above 613 points	12%
Grade 10	Below level 1	Below 345 points	6%
	Level 1	345 to 445 points	24%
	Level 2	445 to 545 points	36%
	Level 3	545 to 645 points	26%
	Level 4	Above 645 points	8%

Table 3: Percentage of students on the various performance levels in reading.

Mathematics			
		Points assigned to the level	Percentage of students (%)
Grade 6	Below level 1	Below 397,5 points	14%
	Level 1	397,5 to 486,5 points	29%
	Level 2	486,5 to 575,5 points	32%
	Level 3	575,5 to 664,5 points	19%
	Level 4	Above 664,5 points	6%
Grade 8	Below level 1	Below 389 points	13%
	Level 1	389 to 471 points	26%
	Level 2	471 to 553 points	33%
	Level 3	553 to 635 points	19%
	Level 4	Above 635 points	9%
Grade 10	Below level 1	Below 357,5 points	8%
	Level 1	357,5 to 452,5 points	27%
	Level 2	452,5 to 547,5 points	35%
	Level 3	547,5 to 642,5 points	21%
	Level 4	Above 642,5 points	9%

Table 3: Percentage of students on the various performance levels in mathematics.

While defining the levels, we found that level 2 was the performance level that would be the minimal level of skills and knowledge necessary for the students to get by in everyday life and

continue their studies. Unfortunately, there was a rather high percentage of students in each of the three grades who had a level 1 or lower performance in both domains.

Between-school differences

Our previous assessments and the PISA study also show that the between-school differences at the secondary school level are especially high in Hungary, compared to the international data. The National ABC also shows that while in elementary schools (grades 6 and 8) about 20% of variance in student performance is due to between-school differences, this percentage is 40-50% in grade 10.

After finishing grade 8, students can decide to continue their education in one of three school types in the Hungarian educational system: grammar schools or vocational secondary schools ending with certificate of final examination, and vocational training schools ending with qualification in some profession. The results of the National ABC study support the fact that the differences between school types are alarmingly high. The difference between grammar school and vocational training school students' performance is more than one and a half times the standard deviation, 155 points in reading, and 153 points in mathematics. Although there is a notable difference between the achievements of grammar school and vocational secondary school students (68 points in reading and 62 points in mathematics), this difference is dissolved after controlling for students' home background, while the performance of the vocational training school students is still considerably lower. While half of the grammar school students, and a quarter of the vocational secondary school students perform on the upper two levels, less than three percent of the vocational training school students reach these levels, and over 70 percent perform below level 2.

The type of school chosen by the students is in high correlation with the HBI values. The majority of the students with the lowest HBIs attend vocational training school; those with the highest indices attend grammar schools. The 2003 National ABC study shows that while only 5.9 percent of the grammar school students indicated the final examination as their desired highest level of education, in contract to the 34.1 percent of the vocational training school students – the rest aim at completing tertiary education. As you can see, those with high HBI values, who wish to go to university chose grammar schools that provide them with better chances, the low index value students focus at learning a trade, while those with medium index values chose the secondary school's final examination and some sort of trade qualification, that leaves the door open for further studies in tertiary education for the future.

Future goals related to the development and use of the National ABC study

The National ABC study is currently the sole national student performance assessment project that plays a significant role in the Hungarian educational system. The development and perfecting of the assessment methods and framework is a continued effort.

One of our most important tasks is to educate schools on the assessment, gain their acceptance of it, have them integrate the results into their within school evaluation, because without these the assessment objectives would not be fully achieved. The School Reports, School Software along with the supplemental and support materials sent to schools all help the schools in getting a realistic evaluation of their own effectiveness. Another possible step in the development plans is to ensure that the results get enough publicity that could help inform those involved in the educational system; parents and students. In the future we would like schools to publish the results of the National ABC in their self-evaluation reports; this would lessen the dangers of a one-sided evaluation based on assessment data alone.

Conclusion

Studies assessing the effectiveness of the educational system, like the National ABC study are becoming more and more accepted by the general public and those working in the educational system in Hungary. These studies give accurate feedback to primary and secondary school teachers about how effective schools are in conveying the skills and knowledge necessary for integration in the society, and also help develop the evaluation practices of the teachers. The assessments are important for policy makers as well: data analyses provide them with a clear picture of the effectiveness of teaching and learning, the validity of the curriculum, so that they can make decisions on the changes deemed necessary.