

Educational Testing in Hungary

Benő Csapó

Attila József University

What has been the role of educational testing in Hungarian education? How are the political changes in the country expected to affect testing in the future?

Today Hungary faces one of the greatest challenges in its history: after 40 years of limited freedom the country has to find its own way of development. As one of the newly founded Central European democracies, Hungary is seeking its original traditions, trying to reconstruct its natural European relationships and to redefine its place in the modern world. This transitional process stimulates a great deal of discussion at every level of society, including general and specific questions of the educational system. While the contradiction between decentralization in the control of the educational system and the relatively high national standards cannot be handled without proper evaluation techniques, one of the crucial issues is what role evaluation and testing should play in the schools of the future.

This article reviews the development of educational testing in Hungary in the light of recent changes. The main evaluation tendencies will be presented within a historical framework. The state of testing in the last decade, present work, and predictable future developments will also be discussed. Although this article discusses the development of testing in a broad sense, including all kinds of systematic evaluation, the terms *test* and *testing* will be used in their narrow meaning. These terms refer to the use of formal measurement instruments to assess student academic achievement. Other forms of oral or written exams will be differentiated from them.

Historical Background

As a consequence of geographical location, Hungarian culture has traditionally been strongly influenced by the German-speaking countries. In certain historical periods and in some domains, French and, to a lesser extent, Anglo-Saxon influence was also detectable. While the origin of modern formal education goes back to the Austro-Hungarian empire, there are many German/Austrian traces in the present Hungarian educational system, and its development is very similar to the educational systems in other Central European countries.

The end of World War I became a turning point in the history of Central Europe. The Austro-Hungarian empire disintegrated and Hungary became an independent state. Between the two wars, schools in Hungary were dominated by religious, conservative, and national values. The idea of excellence and high standards have been guiding principles. Emphasis on the national culture and quality education were considered to be the means for maintaining or improving the position of the country within the region. The availability and quality of schooling was significantly improved in this period, but it was not provided at the same level for every young citizen. The distribution of schooling was highly uneven. A small portion of the population received a really outstanding education, while large masses were barely provided with elementary-level instruction. In order to preserve high quality schooling for a few students,

the educational system became highly selective. The selection took place by means of several examinations. Both oral and written exams played a significant role, but testing was almost unknown and never used in practice.

Origins of Systematic Evaluation

The most prestigious, regulated, and organized examination was (and still is) the *matura*, which was taken at the end of high school in order to apply for entry to a university. The *matura* was first introduced in Prussia in 1788. In Hungary, the first *matura* examinations were held in 1851. Their basic function was to limit the number of applicants to the universities. After World War I, the class-selection feature of the *matura* became dominant: it was a precondition for entering the higher social classes. Since its beginnings, the *matura* was always taken before a committee and it consisted of two parts: a written and an oral section. A regulation from 1884 defined the five subjects of the *matura* as language arts, Latin, history, mathematics and physics. This idea remained more or less unchanged until World War II.

The two main trends in thinking about education during this era might be characterized by the works of Ernő Fináczy, who represented a Herbartian, theoretical-philosophical approach, and by László Nagy, who emphasized psychological aspects of education and the empirical study of children. Fináczy was involved in devising and improving the examination systems. He worked on the more strict and detailed rules for the *matura*, while Nagy carried out a

Benő Csapó is a professor in the Departments of Education and Psychology at Attila József University, Petőfi sgt. 30-34, H-6722 Szeged, Hungary. He specializes in educational measurement.

large-scale survey using questionnaires on the effect of war on children's emotional development as early as 1916. One of L. Nagy's followers, Gábor Kemény (1934), wrote the first comprehensive book on educational evaluation in Hungary. In his work, based on previous empirical studies, he argued against the routine use of marks. He outlined the psychological principles of "proper evaluation," by which he meant a qualitative, verbal, motivating, and formative type of evaluation instead of the use of scales or numbers for grading.

The Post-War Period

From the point of view of the development of educational testing in Hungary, the post-war era can be divided into three major periods. Due to political overdetermination, the shift from one period to another roughly coincides with political events.

The short period of 1945–49 was the beginning of a pluralistic democracy. During the first years right after the war, representatives of reform pedagogy, liberal left-wing thinking, and humanistic approaches to schooling dominated the theoretical discussions. In practice, the founding of a new school system required great effort, and less attention was paid to such specific questions as evaluation or testing.

In 1949 the communist party took power, and this meant the beginning of the "dark fifties." Leading positions were occupied by party-backed apparatchiks and educational theorists, and researchers who did not accept the ideology of Marxism and Leninism were forced to leave the universities and research institutions or in some cases were even taken to prison. Soviet textbooks on education and psychology were translated and considered the only authentic sources of information. Selecting individuals by political or ideological principles became a general practice everywhere (including within school systems), and systematic evaluation procedures or tools without the "class concept" would have been disturbing. Testing, measurement, and experimentation in education, or even the empirical methods of research in social sciences, were declared to be tools of "bourgeois

objectivism" and were banned by every possible means. The Stalinist system was broken by the revolution of 1956. After the revolution had been suppressed and its participants punished, a slow opening to the West started at the end of the decade.

The beginning of the third period can be set in the early sixties. State control was gradually restricted to ideologically sensitive fields like theory and history of education, while the politically neutral problems of instruction, including testing and measurement, were left to the experts. Thus, after a while there emerged two more or less independent groups, which did not encroach on each other's territory. One worked under the prevailing ideological umbrella dealing with test theory, and the other moved toward western-oriented empirical research. From the mid-seventies, empirical work received more financial support, and probably this marked the renaissance of educational research in Hungary.

Research Projects in Educational Testing

In educational testing two major research centers have developed, and during the last decades most educational evaluation programs have been carried out or directed by these two centers. At the beginning, both groups were formed by individuals sharing similar research interests, and they created their institutional framework step by step. Today both groups work as formal institutes. Their development can be characterized by some of their programs and publications.

The first group was founded by Árpád Kiss in the National Institute for Education (NIE, formerly Scientific Institute of Pedagogy). In the post-war period Kiss first conceptualized the questions of measurement, and he organized the first large-scale testing of students' knowledge in Hungary (Kiss, 1961). In his papers he reported on a survey that was started in 1958 and involved the assessment of 4th, 6th, and 8th grade elementary school students (Kiss, 1960/61). Under his leadership, a group in the NIE joined the International Association for the Evaluation of Educational Achievement (IEA) in 1968. Since then,

Hungary has been participating in the major IEA studies (Science, Reading Comprehension, English, Second Mathematics Study, Written Composition, Second Science Study, Computers in Education, Pre-Primary Study, Reading-Literacy Study).

The IEA studies first put the achievements of the Hungarian educational system into an international context. This resulted in some interesting, sometimes surprising, findings and inspired further evaluation research and development projects in Hungary. For example, there was a large discrepancy between the mathematics and science scores on the one hand and the reading comprehension scores on the other. Hungarian students were usually ranked first in mathematics and science, but performed poorly on reading comprehension when compared on an international basis. Although the positive results received very little publicity for a long time, they helped educational experts to defend the real values of the educational system in several discussions with party bureaucrats. The high achievements have recently attracted the attention of researchers in highly industrialized countries, and cross-cultural research projects have been launched to find out how education can work so well in some fields while using such slender resources. On the other hand, the poor scores in reading highlighted the weaknesses in the teaching of reading. While the Hungarian language uses the phonetic way of writing and it is relatively easy to acquire basic reading skills, e.g., to read a word letter by letter, little emphasis was put on higher text-processing and comprehension skills. (Concerning the evaluation of reading performance, see Kádár-Fülöp, 1985.)

Later Zoltán Báthory (1972, 1973) became the leader of this group, which under different names (Department of Curriculum Theory, Center for Evaluation) formed a department of the NIE. A network of the County Institute for Education was formed in the 19 counties of the country, and with their assistance the Center for Evaluation of the NIE carried out several nationwide assessments (Báthory & Kádár-

Fülöp, 1985). A 1980 project, running parallel with the IEA data collection, aimed at providing proper feedback information on (a) the efficiency of the schools themselves, (b) specific problems of some content domains, and (c) relationships of social-economic status and school achievement (Báthory, 1983).

The so-called MONITOR program was launched in 1986, and it was the first step in a periodically repeated nationwide representative assessment project. It is planned that the program will be carried out with similarly composed samples of students in the same content domain to obtain data for a long-term analysis of educational progress (similar to the NAEP program in the U.S.). More than 12,000 students were involved in the 1986 survey in four age groups. Language arts and mathematics were the focal topics of the assessments. The results were published in professional journals and attracted little attention (Vári, 1989). The MONITOR surveys were repeated 5 years later in 1991. The first report of results indicated a decline in reading comprehension and a stagnation in mathematics. The release of information at a press conference in January, 1992, resulted in immediate and stormy discussions reported in the mass media.

The other research team was organized in the Department of Education at Attila József University (AJU) at Szeged by József Nagy in the early seventies. J. Nagy directed several projects that aimed at devising tests for the assessment of several components of knowledge from the most simple elementary skills to the more complex thinking operations. First, he carried out a representative survey on the development of elementary counting skills (Nagy, 1971) and then with his coworkers devised a series of test batteries for the measurement of other basic skills (Nagy, 1973a; Orosz, 1974). The next project was the development of standardized knowledge tests for five school subjects in the 5th to 8th grade of elementary school (Nagy, 1972).

Nagy developed a method of analyzing the thematic units or learning tasks of the textbooks. This analysis resulted in a complete list of basic

statements on the teaching material that could be directly transferred into test items. The main characteristics of this test-construction method were similar to those of early domain-referenced testing, but this test-construction method used a norm-referenced type of scoring system (Nagy, 1973b). This work resulted in a series of 17 volumes, which presented the tests, the coding instructions, the achievement level of a representative sample for

Around the beginning of the eighties, a new direction in testing was launched at AJU aimed at devising tools to assess developmental levels in operational thinking.

each item, and a standardized scoring system. The closing (18th) volume in this series summarized the theoretical output of the project based on the data of test analyses (Nagy, 1975). So far, this project has been the most comprehensive enterprise for developing tests in Hungary as the complete series covered the knowledge of main subjects in the 5th to 8th grade.

As a result of the research traditions at AJU, the Examination Center for Basic Knowledge was established in the Department of Education at AJU in 1991. In this center, during the first year of its operation, a nationwide network of experts was organized, test batteries were devised, and pilot studies on testing were conducted. These steps can lead to the introduction of examination systems at turning points in public education.

The first item-bank was also developed under Nagy's coauthorship for junior school word-problems (Csáki & Nagy, 1976). In the mid-seventies he developed a test battery for the assessment of school-entry compe-

tencies that was standardized on a large sample of over 10,000 pupils (Nagy, 1980, 1986). Around the beginning of the eighties, a new direction in testing was launched at AJU aimed at devising tools to assess developmental levels in operational thinking. Thinking operations were interpreted within the framework of Piagetian and Neo-Piagetian traditions and batteries of paper-and-pencil tests were constructed and used to outline the developmental trends of these operations within the age range of 6 to 17 years (Nagy, 1987).

The most recent work of Nagy's team at AJU is concerned with the devising of diagnostic tests for the practical needs of schools (Vidákovich, 1987, 1990). These tests have been developed by several groups of experts and teachers under the coordination and with the theoretical support of a university research team. Depending on the availability of the necessary facilities, regional or county evaluation centers receive camera-ready copies of these tests and print them for use in schools. After administering the tests, teachers use them to grade their students, and then data are entered into computers at schools or the tests are sent to local or regional centers. Where computers are available, results are analyzed and discussed by local experts and groups of teachers in order to improve their own work. The collected data are used for sophisticated test analysis and for the improvement of the tests. This work offers feedback at several levels in the educational system and at the same time aims at disseminating information on test construction and computerized test analysis.

The Present State in Testing

The state of the art in educational testing in Hungary can be best characterized through the kind and depth of knowledge about testing within the various groups of educationalists and the testing practice at the different levels of education.

Since testing (e.g., for selecting personnel for different positions, etc.) is not used as frequently in Hungary as in some highly industrialized societies, parents have little experience with it. Test construction or any information concerning test-

ing has not comprised part of teacher training; thus, practicing teachers have very limited theoretical knowledge about testing. They have used tests as they received them, and sometimes they tried to make new ones using other test as examples. There are very few manuals or technical guides (Báthory, 1973; Csapó, 1988; Joó, 1980; Nagy, 1972) available in Hungarian for those teachers who wish to construct their own measurement instrument, and this fact also limits the development of teacher-made tests.

There is a circle of educational researchers and professional test developers (at present about 20 persons within Hungary) who follow international literature and participate in cross-cultural investigations and who possess up-to-date knowledge that enables them to use the most recent methods and deal with the most sophisticated computer programs. There is a growing interest among practicing teachers about the problems of testing that will increase the communication between the test experts and practitioners. However, the number of papers discussing theoretical questions in Hungarian educational or psychological journals (e.g., Csapó, 1987; Horváth, 1985) is very low and the dissemination of information about testing is rather slow.

There are very few areas in the Hungarian educational system where usage of tests is a regular practice. Before entering school, there are school-readiness exams when psychological tests or batteries devised specifically for this purpose are used (Nagy, 1986). In the first year of elementary school, the form of evaluation is mostly qualitative and verbal. From the second year to the universities, a marking system is almost exclusively used that is based on a 5-grade scale: 1 means a fail, 2 is the lowest acceptable level, and 5 is the best achievement. In elementary and secondary schools, students receive about four to eight marks per semester based on their oral presentations, short quizzes, and written work, e.g., solving tasks in math and science or written compositions in the language arts. These marks determine their year's result. In written work some teachers may use tests where available. These tests

are often provided as a supplement to textbooks or as separate booklets. There are no definite national or local guidelines concerning the use of these tests; teachers are free to use them as they choose.

The only centrally organized testing process is part of the higher education entrance examination. Only one-third on average of the applicants can enter the universities or colleges, and these exams are very competitive. The tests are developed by small groups of experts and are kept secret until the day of testing. Thus, these tests are never tried out and are of poor quality from the point of view of test construction and reliability. In higher education, tests are relatively rarely used. They are used mostly in language teaching where foreign test banks are directly adaptable.

Directions for the Nineties

The most recent changes in the Hungarian educational system imply some new development in evaluation and testing. A new educational act is being developed and will probably be approved by the parliament in 1992. Its aim is to delineate what the legislative consequences of the change of the political system are for education and to give a legal framework for the evolution of education. Although several outlines of the planned law have already been discussed, both in professional and lay circles, the latest, fifth draft is probably not the last one ("Conceptions of the Law," 1992). Despite the active discussions, the main developmental tendencies have already emerged.

Instead of the totally regulated central curriculum, a more flexible national core curriculum is planned that leaves more autonomy for local and school-level decisions (Nagy & Szebenyi, 1990). The existing 8 + 4 (elementary + grammar school) and 8 + 3 (elementary + apprenticeship) structure of the school system might also be changed and replaced by other forms, e.g., 4 + 8, 10 + 2, or 6 + 6.

This greater variability in schools and teaching systems is expected to be controlled through an output oriented evaluation system. This means that local authorities could chose the structure of schools and

schools could chose the content and methods of teaching, but at certain ages, students would be expected to reach standards accepted nationwide. The achievement of students would be checked through centrally organized examinations, probably at the ages of 12, 14, 16, and 18. The exams at the ages of 12 and 14 play mostly a diagnostic role and might help the transition between school levels. The age of 16 marks the end of compulsory education, and the exams planned at this point are to close the compulsory schooling. Most of the actual research efforts in educational evaluation are concentrated to develop the proper test batteries for these exams. The traditional *matura* is also expected to be renewed. The written exams will mostly be replaced by standardized tests.

References

- Báthory, Z. (1972). Értékelés a pedagógiában [Evaluation in education]. *Pedagógiai Szemle*, 22, 212-222.
- Báthory, Z. (1973). *7 standardizált tantárgytest* [Seven standardized knowledge tests]. Budapest: Országos Pedagógiai Intézet.
- Báthory, Z. (1983). Az iskolai nevelés néhány összetevőjének vizsgálata egy felmérés tükrében [A study of certain factors in school education]. *Pedagógiai Szemle*, 33, 135-185.
- Báthory, Z., & Kádár-Fülöp, J. (Eds.). (1985). Educational evaluation studies in Hungary. *Evaluation in Education: An International Review Series*, Vol. 9.
- "Conception of the law of public education." (1992). Budapest: Ministry of Culture and Education (version of January 14, mimeographed).
- Csáki, I., & Nagy, J. (1976). *Alsó tagozatos szöveges feladatbank* [Item bank for junior school word-problems]. Szeged: Acta Paedagogica Series Specifica.
- Csapó, B. (1987). A kritériumorientált értékelés [Criterion-referenced evaluation]. *Magyar Pedagógia*, 87, 247-266.
- Csapó, B. (1988). *A tanulói teljesítmények értékelésének mérési módszerei* [Assessment methods for evaluating student achievement]. Budapest: MM. Vezetőképző és Továbbképző Intézet.
- Horváth, G. (1985). Tesztelmélet: problémák és perspektívák [Test theory: Problems and perspectives]. *Pszichológia*, 5, 53-78.

Continued on page 15

93% of administrators' surveys. These are the proportions of each group expressing an opinion on each individual item.

References

- Barrett, R., & Stevens, V. M. (1987, July 15). Districts in west valley analyze test scores. *Arizona Republic*, pp. 1E, 4E.
- Berk, R. A. (1988). Fifty reasons why student achievement gain does not mean teacher effectiveness. *Journal of Personnel Evaluation in Education*, 1(4), 345-364.
- Bishop, C. D. (1989, November 9). Presentation to the State Joint Legislative Committee on Goals for Educational Excellence, Phoenix, AZ.
- Brandt, R. (1989). On misuse of testing: A conversation with George Madaus. *Educational Leadership*, 46(7), 26-30.
- Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. *Educational Researcher*, 16(8), 16-20.
- Droege, R. C. (1966). Effects of practice on aptitude scores. *Journal of Applied Psychology*, 50, 306-310.
- Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, 39, 193-202.
- Freidly, R. (1989, October). Remarks presented at the symposium "How Arizona Schools Prepare and Use Standardized Tests and the Effects on Students," at the annual meeting of the Arizona Educational Research Organization, Mesa, AZ.
- Gallup, A. M. (1989). The 21st annual Gallup poll of public's attitude toward the public schools. *Phi Delta Kappan*, 71, 39-54.
- Haas, N. S., Haladyna, T. M., & Nolen, S. B. (1989). *Standardized testing in Arizona: Interviews and written comments from teachers and administrators* (Technical Report 89-3). Phoenix: Arizona State University West.
- Haertel, E. (1985). Construct validity and criterion-referenced testing. *Review of Educational Research*, 55(1), 23-46.
- Haertel, E. (1986). The valid use of student performance measures for teacher evaluation. *Educational Evaluation and Policy Analysis*, 8(1), 45-60.
- Haladyna, T. H., Haas, N. S., & Nolen, S. B. (1989). *Test score pollution* (Technical Report 89-1). Phoenix: Arizona State University West.
- Madaus, G. F. (1987). *Testing and the curriculum*. Chestnut Hill, MA: Boston College.
- Madaus, G. F. (1988). The influence of testing on curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum. Eighty-Seventh Yearbook of the National Society for the Study of Education* (pp. 83-121). Chicago: University of Chicago Press.
- Mehrens, W. A., & Kaminski, J. (1989). *Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? Educational Measurement: Issues and Practice*, 8(1), 14-22.
- Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 447-474). Washington, DC: American Council on Education and Macmillan.
- Richards, T. S. (1989). Testmania: The school under siege. *Learning*, 17(7), 64-66.
- Rothman, R. (1988, February 17). E. D. will prepare "consumer guide" on standardized tests. *Education Week*, pp. 16-17.
- Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership*, 46(7), 4-9.
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.
- State-Mandated Tests, 1986-7. (1988). *FairTest Examiner*, 2(3), 12.
- Taylor, C., & White, K. R. (1982). The effect of reinforcement and training on group standardized test behavior. *Journal of Educational Measurement*, 19, 199-210.
- Tuinman, J. J., Farr, R., & Blanton, B. E. (1972). Increases in test scores as a function of material rewards. *Journal of Educational Measurement*, 9, 215-223.

Testing in Hungary

Continued from page 8

- Joó, A. (1980). *A feladatkészítés kérdései* [Questions of item writing]. Budapest: Országos Oktatástechnikai Központ.
- Kádár-Fülöp, J. (1985). The CTD reading study. *Evaluation in Education: An International Review Series*, 9, 117-151.
- Kemény, G. (1934). *Iskolai értékelés és kiválasztás* [Evaluation and selection at school]. Budapest: Merkantil-nyomda.
- Kiss, Á. (1961). Docimológia, osztályozás, mérés [Docimology, classification, measurement]. In *Pszichológiai tanulmányok*, 3 (pp. 253-266). Budapest: Akadémiai Kiadó.
- Kiss, Á. (1960/61). Iskolai tanulók tudásszintjének vizsgálata, 1-4 rész [Examination of the knowledge of school children, parts 1 to 4]. *Pedagógiai Szemle*, 10, 194-206, 585-593, 775-784, and 11, 600-613.
- Nagy, J. (1971). *Az elemi számolási készségek* [Elementary counting skills]. Budapest: Tankönyvkiadó.
- Nagy, J. (1972). *A témazáró tudásszint-mérés gyakorlati kérdései* [Practical problems of thematic knowledge assessment]. Budapest: Tankönyvkiadó.
- Nagy, J. (1973a). *Alapműveleti számolási készségek* [Skills of basic counting operations]. Szeged: Acta Paedagogica Series Specifica.
- Nagy, J. (1973b). A standard osztályzat [Standard grades]. *Pedagógiai Szemle*, 23, 225-234.
- Nagy, J. (1975). *A témazáró tesztek reliabilitása és validitása* [Reliability and validity of knowledge tests]. Szeged: Acta Paedagogica Series Specifica.
- Nagy, J. (1980). *5-6 éves gyermekeink iskolakészültsége* [School readiness of 5-6-year-old Hungarian children]. Budapest: Akadémiai Kiadó.
- Nagy, J. (1986). *PREFER* [A test battery for assessment of entrants' entry competences]. Budapest: Akadémiai Kiadó.
- Nagy, J. (1987). *A rendszerezési képesség kialakulása: Gondolkodási műveletek* [The evolution of systematizing ability: Thought operation]. Budapest: Akadémiai Kiadó.
- Nagy, J., & Szebenyi, P. (1990). Hungarian reform: Towards a curriculum for 1990s. *Curriculum Journal*, 1, 247-254.
- Orosz, S. (1974). *A fogalmazástechnika mérésmethodikai problémái* [Problems of assessment written compositions]. Budapest: Tankönyvkiadó.
- Vári, P. (1989). A MONITOR '86 ismertetése [A review of MONITOR '86]. *Pedagógiai Szemle*, 39, 1123-1130.
- Vidákovich, T. (1987). *Innovatív célú diagnosztikus pedagógiai értékelés* [Innovative diagnostic evaluation in education]. Budapest: Közoktatási Kutató-sok Titkársága.
- Vidákovich, T. (1990). *Diagnosztikus pedagógiai értékelés* [Diagnostic educational evaluation]. Budapest: Akadémiai Kiadó.