

# National Testing: Lessons for America from Europe



*American advocates of a national test can find food for thought in the European experience.*

THOMAS KELLAGHAN AND GEORGE F. MADAUS

**D**uring the 1980s, fear that the United States was losing its competitive edge in world commerce led to a series of commissioned reports which blamed the schools for the country's economic problems and decreased competitiveness. The reports proposed a series of reforms, one of which has received increasing emphasis over the last two years: a national test or system of tests.<sup>1</sup> By the turn of the decade, the proposal had emerged as a major weapon in the armory of educational reformers.

Because many countries in Europe have a long tradition of non-school-based examining (the preferred word for testing in Europe), here we consider the recent proposals for national testing in light of the that experience (Madaus and Kellaghan 1991).

## Proposals for National Testing

The most recent proposal, from President Bush, is *America 2000: An Education Strategy* (1991). The President proposes a "voluntary" national testing system, as part of an overall strategy to move the nation toward achievement of the six major national goals outlined by himself and the state governors. Tied to "world-class standards," the American Achievement Tests—for grades 4, 8, and 12—would measure achievement in the 5 core subjects mentioned in goal 3: English, mathematics, science, history, and geography.

A perusal of the various proposals for a national test or system of tests reveals great variety in their details: the purpose of the testing, whether it will be voluntary, the age/grade at which

students will be tested, the subject areas to be tested, testing methods, who will be responsible for testing, and the question of financing. All the proposals, however, share one characteristic: vagueness.

While several proponents justify the need for national testing by arguing that the United States is alone among industrialized countries in not having external examinations, their proposals are unclear. Since examination systems vary greatly from one country to another, American proposals will have to identify the foreign system which they regard as worthy of emulation, or at least the characteristics of systems which seem appropriate for import. Not only is this not done, but proponents also fail to point out that some of their recommendations actually deviate from practice in other industrialized nations.

Further, proposals invariably fail to acknowledge the problems and undesirable effects that other countries have experienced with examination over a considerable period of time—the recognition of which has led to reforms of those systems.

Before Bill or Annie actually sits down to answer examiners' questions, then, a number of issues will have to be addressed. Let's look at them.

## What Will the Purpose Be?

Proposals, including those of *America 2000*, envisage a variety of possible purposes:

- to foster good teaching and learning,
- to identify students for remedial intervention and grade promotion,
- to motivate students,
- to certify student achievement,
- to admit students to college or employment,

- to recognize publicly the performance of students who do well (in the five core areas delineated in the President's and the governors' national goals),

- to provide information to parents on students' performance and to the public on the performance of schools, school districts, states, and the nation.

The proposals are most compatible

## In Europe, examination results are not used to provide the public with information on the performance of individual schools.

with European practice when they emphasize the use of tests for selection. Indeed, the origins and tradition of examinations in Europe exhibit a major concern with selecting students for third-level education and certain jobs (Madaus and Kellaghan 1991). While today the certification function is also important, examination results are still used for selection, both inside and outside the school system.

In Europe, examination results are not used to provide the public with information on the performance of individual schools. There are, however, proposals to do so in Britain, where government policy currently places great emphasis on accountability and school choice. But the most usual way to deal with school accountability, in Britain and elsewhere in Europe, is through the system of school inspection.

Examinations in Europe are also not used to obtain information on performance of states or regions within a country or of the nation as a whole. When required, such information is

obtained from national assessments (similar to the National Assessment of Educational Progress in the United States); from participation in international studies (such as the International Association of Educational Achievement and the International Assessment of Educational Progress); or from surveys carried out by inspectors.

For example, in England and Wales, inspectors recently visited 3,274 primary schools, 2,436 secondary schools, 692 further and higher education institutions, and 529 independent schools. One of the study's conclusions was that, at both primary and secondary levels, about a third of schools were "good" or "very good," a third were "satisfactory," and a third were "poor." At a more specific level, one in every three lessons observed in primary schools in such fundamental subjects as English and mathematics was found to be unsatisfactory (Great Britain, Department of Education and Science, 1991).

### Will the Tests Be Voluntary?

*America 2000* is quite specific that the national examination system will be voluntary. However, it does not say whether or not individual students or schools will be free to or free not to take an examination. The use of the word *voluntary* is an attempt to assuage fears. However, something that is initially voluntary may over time cease to be so. One need only consider what has happened to the political safeguards that accompanied the launch of NAEP in this country in the 1960s.

Further, what is voluntary in theory may not be voluntary in practice. In Europe, students and even schools are free not to take external tests in the sense that nobody is obliged to take one. However, if a student wants evidence of his or her performance at school to show to a prospective employer or to apply to a university, he or she, except in exceptional circumstances, is not free to skip the examination. For obvious reasons, a school which does not prepare and present students for exami-

nations will not attract many students.

An examination certificate is not required for many jobs in Europe, such as that of Member of Parliament (the people decide). However, this did not prevent the educational career of John Major becoming a matter of intense interest to journalists when he replaced Margaret Thatcher as Prime Minister. Major, it seems, had not availed himself of the full benefits of the British educational system in his youth, giving rise to questions about the number of General Certificate of Examination subjects he had passed at ordinary level at age 16. Examinations may be voluntary, but taking them can help avoid embarrassing situations, even for a future Prime Minister.

### At What Grade Levels?

*America 2000* recommends that the American Achievement Tests be administered at grade levels 4, 8, and 12. Other American proposals, with some exceptions, envisage testing either at age 16 or at some point during the senior high school years, including at graduation.

Where *America 2000* departs most from European practice is in recommending the testing of 4th grade pupils. Formal external testing of this type no longer exists in Europe below the age of 15 or 16 years. A possible exception is Britain, where all children are tested at 7 and will in the future also be tested at 11 and 14. The tests at 7 and 11, however, are new developments (not yet complete) and are not part of the traditional external examination system. The tests are administered and scored by the pupils' own teachers.

Of the 12 European Community countries, 5 (Belgium, Greece, Luxembourg, Portugal, Spain) no longer administer a system of external examining at the end of junior high school or its equivalent (which can vary from grade 8 to grade 10). At the end of high school, 4 countries (Belgium, Greece, Portugal, Spain) operate a system of

school-based certification, not dissimilar to that in operation in the United States. The other countries vary in the amount of external testing they have for students at this age. In no country do all 18-year-olds take the same exam, as is suggested in the *America 2000* proposals.

### What Subject Areas?

American proposals generally focus on core subjects. *America 2000* specifies English, mathematics, science, history, and geography, while another proposal adds computer studies, and one specifies reading and writing in place of English.

However, it is not entirely clear at what level of achievement the American Achievement Tests proposed in *America 2000* will be pitched. On the one hand, *America 2000* speaks of world-class standards, pointing out that the tests must not simply measure minimum competencies, but also higher levels of reading, writing, speaking, reasoning, and problem-solving skills. Reference to European examinations designed for college-bound students adds to the belief that the tests will not be concerned with basics. Recently, Cheney (1991) provided some examples from such examinations. If these examples are to be models for American national tests, then by the year 2000, we can expect American students at the end of high school to be writing four-hour essays on such topics as: *How might one characterize rigorous thought? What does one gain by losing one's illusions? Can one say: "To each his own truth"?*

Writing on these topics seems a long way from other considerations in *America 2000*: that all students achieve set standards, that tests in the 3Rs be a priority, and that the problem of functional illiteracy be tackled.

Obviously, a consideration of the situation in Europe will not resolve this apparent contradiction in the proposals for an American examination. However, we will briefly outline the situation there.

In some European countries, a compulsory core of subjects is specified for examination at the end of grade 8, 9, or 10. In Denmark, students take Danish and mathematics; in France, students take French, mathematics, and history/geography. Other countries allow students a wider choice, providing examinations in core subjects and a wide range of additional optional

## In no European country do all 18-year-olds take the same exam, as is suggested in the *America 2000* proposals.

subjects (Britain, Ireland). In these countries, curriculums and examinations may be offered at different levels (for example, a higher-level and a lower-level examination in mathematics). To complicate matters further, in some countries, students in different types of school take different examinations (Germany, Netherlands).

By the end of high school, the examination experience of students is more differentiated. Most countries which have external testing offer one system of examinations for students pursuing academic curriculums and another for those following vocational curriculums. By this stage, many students are outside the educational and examination nets. Since the age for compulsory school attendance is 15 or 16, many students leave without completing the final grades of high school. Only 1 country in the European Community has a higher school participation rate for 17-year-olds—Belgium, at 93 percent—than the United States, at 89 percent. Only 5 more countries have participation rates in excess of 75 percent:

Denmark, France, Germany (if part-time students are included), Luxembourg, and the Netherlands. In Britain, the participation rate of 17-year-olds is only 52 percent (OECD 1990, Table 4.2).

Students who remain in school take a variety of examinations depending on the courses they follow, the schools they attend, and, in some countries, even the part of the country in which they live. In Germany, for example, students who take academic courses which prepare them for university take an examination (*the Abitur*), which is set by different state Ministries and offers a wide assortment of subjects, with different weights given to the results depending on the options chosen. The equivalent examination in France (*the Baccalauréat*), set by regional *académies*, is geared to a number of differentiated curriculums, though all students take a core of general education subjects. Four options in 1950 had grown to 38 by 1988 (Noah and Eckstein 1990). The failure rate on the examination is as high as 30 percent.

Given this variety, the European experience can contribute little to the design or implementation of proposals for a single external examination for all American students at grade 12.

### What Testing Method?

Proposals about examination methods in the American context range from standardized to "state of the art" tests. The search for new, improved tests reflects a growing dissatisfaction with multiple-choice varieties, which, because of their emphasis on lower level skills and memorization, are seen as limiting learning.

According to *America 2000*, the proposed American Achievement Tests "will measure higher-order skills (i.e., they will not be strictly multiple-choice tests" (p. 48). This position seems to confuse method of measurement with level of knowledge and skills measured. Higher-order skills can be measured by multiple-choice tests (though most such

tests may not do so) while other kinds of tests (for example, essay-type tests) may measure lower-order skills.

In America, there is much talk today of "authentic" testing, which would include performance, portfolio, and project examinations. Cheney (1991) points out that most advocates of national examinations stress the need for "performance testing," arguing that "students should demonstrate whether they can organize their thoughts, make analyses, and mount arguments; students should be tested to see whether they can use the facts they have learned" (p. 4). However, it is of interest that the test items from Japan, America's biggest competitor, reproduced in the Cheney (1991) report, are all multiple-choice.

Systems of examining in Europe can tell us little about the value of these proposed "authentic" measures, since European examinations are based mostly on written essays or short-answer questions. In some systems, even the traditional oral component has been curtailed, mainly because of the cost. Apart from obvious areas such as music and woodwork, practical examinations are rare.

If European examinations are as satisfactory as Cheney (1991), for example, seems to believe, why have they been the subject of reform in many countries in recent decades? The most widespread reform, in fact, has involved a reduction in the external element and an increase in the school-based element. One reason has been the recognition that many aspects of student learning cannot be assessed adequately in a one-time terminal examination, whether it be written, practical, or oral.

One development of particular interest to American educators centers on Standard Attainment Tasks, which are being developed in Britain to assess student achievement at 7, 11, 14, and 16 (National Curriculum, Task Force on Assessment and Testing 1987). So far, the tests, which teachers administer in their own classes, have been used only

with students at age 7. This experience has given rise to many complaints. For example, the tests occupy too much class time (three to four weeks), demand a great deal of extra work of teachers, are difficult to administer, and, despite the time and effort, tell teachers nothing new about pupils.

As a result of such experience, similar tests at higher age levels may

## **The *America 2000* position is quite different from the situation in Europe, where all educational systems (except Britain until 1988) have national- or state-level curriculums.**

not be developed as originally planned. Indeed, the official view in Britain is that testing at ages 11 and 14 will become less complex, more external, and based more on traditional pencil-and-paper approaches.

### **Who Will Control Testing?**

Proposals for national testing in the United States have so far avoided the question of control. Who will build the tests? Who will be responsible for administration, scoring, and reporting of results? Who will deal with questions of comparability from area to area, if different tests are administered in different regions, between subjects, or from year to year? Who will be responsible

for test security? What provision will be made for student appeals if results are regarded as unacceptable, and who will be the final arbiter?

These are all issues that have arisen in Europe and have been dealt with in different ways in different countries. As a result, the patterns of authority vary not only from one country to the next but according to the type of examination and the grade level of candidates.

At the end of junior high school or its equivalent, testing is completely under the control of individual schools in 5 of the 12 European Community countries. In only 3 countries (Britain, France, Ireland) are examinations for students at these grades almost entirely controlled by an agency outside the school.

At the end of high school, the external element in examinations becomes more prominent, though even at this stage, 4 countries (Belgium, Greece, Portugal, Spain) operate a system of school-based certification, not dissimilar to that in operation in the United States. In 3 countries (Britain, France, Ireland), testing is almost entirely under the control of an external body; that is, the tests are both set and scored by an agency outside the school (a ministry of education or examinations authority). In other countries, tests are set by an external agency but are scored by students' own teachers (Germany), or by other teachers as well as by students' own teachers (Denmark, Italy, Luxembourg), or by some other combination of school-based and external assessment (Netherlands).

The role of central government in the European context is most direct and influential in those countries in which the ministry of education sets, administers, and is responsible for marking examinations (as in Ireland). It is less direct when a ministry sets examinations but is not responsible for scoring (as in Germany). Another type of less-direct influence occurs when the central ministry has a general supervisory role in the activities of independent examining bodies (as in Britain).

In the larger countries, more than one body has responsibility for examinations. In Germany, it is the ministry of education in each state. In France, 23 regional académies assume this responsibility. In Britain, 8 regional examining groups are responsible for examinations administered to students at age 16, and 5 university-related boards are responsible for examinations at the end of high school.

Devolution of control gives rise to problems when different but supposedly equivalent examinations are set by different authorities. The inadequacy of general guidelines for examination administration, agreed by state ministers of education in Germany, is recognized in the questions which are raised "about the extent to which grading standards are kept consistent even within a given Land (state)" (Noah and Eckstein 1990, p. 90). Problems regarding comparability from one agency to another and over time have also been raised in Britain and France.

### **Will There Be a National Curriculum?**

*America 2000* says that the President and the governors oppose a national curriculum or federalization of the education system. Tests will examine the results of achievement, saying nothing about how the results are produced. The point is an important one, even if somewhat ingenuous. It is important because politicians have to be conscious of the long tradition of local control in American education. It is ingenuous if it implies that important tests or examinations do not come to drive curriculums.

The *America 2000* position is quite different from the situation in Europe, where all educational systems (except Britain until 1988) have national- or state-level curriculums. In theory, examinations are based on these curriculums, though the tradition of examinations over time comes to define the particular aspects of the curriculum

and student outcomes that are attended to in practice. Because of this situation, the first thing teachers want to see when a new curriculum is introduced is a sample examination paper. It would be surprising if the proposed examination for the United States did not come to define actual school curriculums in the way that examinations do in Europe.

If *America 2000*'s proposal that examining be independent of curriculums does indeed become a reality (though it is hard to see how), it will run completely counter to a new British system in which targets of attainment for different levels throughout the system are being developed. These targets are the basis of the new national curriculum, and testing will be geared to them.

The future of the American curriculum may well depend on how the new testing system is funded. If it is funded by the federal government, then states and localities that use the tests may find themselves implementing not only a national curriculum, but a federal one.

### **Disadvantages of External Testing**

Over time, a number of disadvantages have been associated with external examinations throughout the world (Kellaghan and Greaney 1991, Madaus 1991, Madaus and Kellaghan, in press). Here are some of them:

- The examinations narrow the curriculum in the sense that approaches to learning are limited and subjects and aspects of subjects not covered in the examinations are excluded or neglected.
- The cost of examining oral and practical skills is quite high, forcing most systems to rely on written tests.
- It is impossible to assess in a terminal examination such factors as student planning, perseverance, and adaptability in the execution of a project.
- External examinations promote the tendency to emphasize lower-order skills in teaching, since these are more easily examined than higher-order ones.

• Examinations are unsuitable for assessing some students (particularly lower achieving ones).

• Examinations engender a diminished professional role for teachers, since important curriculum decisions are in effect decided by the examination.

One final and insidious effect examinations may have on teaching and

**If the American testing system is funded by the federal government, then states and localities may find themselves using not only a national curriculum, but a federal one.**

learning is that teachers sometimes devote great efforts to teaching students strategies to answer examination questions. This alliance between teachers and students to "beat the system" has a long history. Matthew Arnold (a school inspector), in describing a high-stakes testing program in Great Britain in the 19th century, observed that it was "a game of mechanical contrivance in which teachers will and must learn to beat us. It is ... possible by ingenious preparation to get children through the (exams) in reading, writing, and ciphering without their really knowing how to read, write, or cipher" (quoted in Sutherland 1971, p. 52).

Serious consideration must be given to these disadvantages and every effort made to create conditions to limit them. If an examination system is imple-

mented, it must be carefully monitored to identify problems as they occur.

### What Testing System to Emulate?

As we have noted, examination systems in Europe vary considerably. Further, because systems are under continual review, practice today may no longer be practice tomorrow. How, then, is one to determine which model is most appropriate for the United States?

If one accepts that the economic power of a country is the result of the effectiveness of its educational system in general and of its examination system in particular (unlikely though this proposition is), then in the European context, one should surely look to Germany. Germany, however, does not have an external testing system similar to that envisaged in the American Achievement Tests. Examinations are primarily school-based; while exams may be set by state-level ministries, they are scored by students' own teachers.

A consideration of American proposals suggests that British rather than German practice is of prime interest. This is strange for a number of reasons. First, the British economy is in a state of severe depression, with extremely high unemployment figures.

Second, the British educational system is a continuing object of criticism and concern. For example, Richard Hoggart (author of *The Uses of Literacy*) observed that in Britain only a minority of people "read at all, 'if to read' means more than occasionally skimming over two-syllabled words about trivial matters" (1991). According to Hoggart, most leave school critically, "culturally, and imaginatively sub-literate."

Third, the British examination system is going through major reforms. American interest seems to lie more in the proposed new Standard Attainment Tasks, a system that may be changed substantially in the direction of more paper-and-pencil tests than in traditional

public examinations. What confidence can one have either in a system being discarded or in one which has yet to be fully developed and tested?

### Thoughts to Ponder

In the context of American proposals for a national test, European experience with external high-stakes testing is

## Will American reformers be satisfied if educational quality is equated with whatever it is that tests measure?

worth considering for several reasons. First, some advocates of national testing seem enchanted by European practice, and, second, European experience goes back to the 18th century. What issues are most pressing as deliberations proceed?

Determining the purpose of testing is a primary task. Until purpose is decided, we cannot anticipate who or what will be affected or how, nor can the American Achievement Tests or any other proposed national test be designed or validated. Purpose will, among other things, affect the content that will be tested. For example, the content of a test designed to make decisions about admission to higher education would probably differ considerably from the content used in a test to select students for semi-skilled or unskilled work.

If the content of national tests in the United States is to reflect the content of European tests, it has to be borne in mind that tests given at the end of high school in Europe are not taken by all

students. Even external essay tests of differing levels of difficulty, which it was hoped the vast majority of students at 16+ would take, have been unsatisfactory for some students in Britain and Ireland. Lower achieving students were not motivated by the tests and either avoided them by dropping out of school or failed them.

While in America it is proposed that pupils be tested at grade 4, several European countries, which at one time tested pupils at the end of elementary school, have all abandoned the practice. Teacher-controlled testing, which will influence teaching and learning in an immediate, frequent, and formative way, is regarded as more important than summative testing during the elementary grades—and, indeed, in several countries, during secondary schooling also. It is also regarded as important, particularly during the elementary years, that the curricular experiences of pupils not be limited by external examinations. Indeed, one reason for abandoning examinations at the primary level was the recognition that curriculum areas, skills, and knowledge, even though specified in curriculum documents, were being ignored or given little attention in schools if they were not included in examinations.

American reformers, too, must be sensitive to the implications of the European experience for the goals of education. Will they be satisfied if educational quality is equated with whatever it is tests measure? And if they accept that the results of education are to be valued primarily in terms of their measurability, will this not inevitably lead to a devaluation and even ignoring of school effects, which, though not measurable, are regarded as intrinsically valuable?

Finally, we learn from Europe that whoever controls the external examination system will control what goes on in schools. To move from a tradition of local control to one of federal control—certainly a possibility in the light of current proposals—is a major step.

which many Americans will wish to ponder seriously. □

'For example, *A Nation at Risk* (National Commission on Excellence in Education 1983), High School: A Report on Secondary Education in America (Boyer 1983).

## References

- America 2000: An Education Strategy. Sourcebook* (1991). Washington D.C.: U.S. Department of Education.
- Boyer, E. L. (1983). *High School: A Report on Secondary Education in America*. New York: Harper and Row.
- Cheney, L. V. (1991). *National Tests: What Other Countries Expect Their Students to Know*. Washington, D.C.: National Endowment for the Humanities.
- Great Britain, Department of Education and Science (1991). *Standards in Education*. London: Author.
- Hoggart, R. (June 14, 1991). "The Pursuit of Quality." *Bookseller* 4460: 1700-1703.
- Kellaghan, T., and V. Greaney. (1991). *Using Public Examinations to Improve Education: A Study in Fourteen African Countries*. Washington D.C.: World Bank.
- Madaus, G. F. (1991). "The Effects of Important Tests on Students: Implications for a National Examination or System of Examinations." Paper presented at the American Educational Research Association Invitational Conference on Accountability as a State Reform Instrument: Impact on Teaching, Learning, Minority Issues and Incentives for Improvement, Washington D.C.
- Madaus, G. F., and T. Kellaghan. (1991). "Examination Systems in the European Community: Implications for a National Examination System in the United States." Paper prepared for the Science, Education, and Transportation Program, Office of Technology Assessment, U.S. Congress, Washington, D.C.
- Madaus, G. F., and T. Kellaghan. (In press). "Curriculum Evaluation and Assessment." *Handbook of Research on Curriculum*, edited by P. W. Jackson. New York: Macmillan.
- National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform*. Washington D.C.: U. S. Government Printing Office.
- National Curriculum, Task Group on Assessment and Testing. (1987). *A Report*. (P. J. Black, Chair). London: Department of Education and Science and the Welsh Office.
- Noah, H.J., and M. A. Eckstein. (1990). "Trade-Offs in Examination Policies: An International Comparative Perspective." In *Changing Educational Assessment: International Perspectives and Trends*, edited by P. Broadfoot, R. Murphy, and H. Torrance. London: Routledge.
- Organization for Economic Cooperation and Development. (1990). *Education in OECD Countries 1987-88*. Paris:
- OECD.
- Sutherland, G. (1971). *Elementary Education in the Nineteenth Century*. London: Historical Association.
- Authors' note:* We wish to acknowledge the financial support of the Ford Foundation for the preparation of this paper. The views expressed are ours.

Thomas Kellaghan is Director, Educational Research Centre, St. Patrick's College, Drumcondra Rd., Dublin 9, Ireland. George F. Madaus is Boisi Professor of Education and Public Policy, Boston College, The Center for the Study of Testing, Evaluating, and Educational Policy, Chestnut Hill, MA 02167-3807.

**Ph.D.  
or Ed.D.  
FOR THE ACTIVE PROFESSIONAL**

- Complete your doctorate in a self-paced, personalized, accredited program without interrupting current work patterns.
- Address significant issues in conjunction with your professional responsibilities.
- Work with nationally recognized faculty toward your doctorate.
- Prerequisite: Masters or equivalent.

Walden University  
• 801 Anchor Rd. Drive, Naples, FL 33940  
• 415 First Ave. N., Minneapolis, MN 55401

Yes, Please send a catalog to:

NAME \_\_\_\_\_  
ADDRESS \_\_\_\_\_  
CITY, STATE, ZIP \_\_\_\_\_  
PHONE ( ) \_\_\_\_\_  
Mail to: Walden University, Process Center  
801 Anchor Rd. Drive, Naples, FL 33940

For information concerning programs in your area call  
**1-800-444-6795**

Copyright © 1991 by the Association for Supervision and Curriculum Development. All rights reserved.