

# Trends in Large-Scale Testing Outside the United States

Richard P. Phelps

*Is standardized testing on the decline outside the United States? How much testing goes on internationally? What types of testing and for what purposes?*

or some time now, we have witnessed considerable interest in state-level or national systems of student examinations. Public opinion polls have revealed a large majority of adults and parents in favor of large-scale examination systems for decades (Phelps, 1998). Likewise, business associations have proffered proposals for such systems for over a decade (Phelps, 1998). First the Bush administration and then the Clinton administration made specific proposals for national examination systems.

## Background

As part of the discussion, both proponents and opponents of national testing in particular and standardized testing in general have drawn attention to testing regimes in other countries. Some proponents have argued that students in other countries learn more because they have the focus, motivation, and ordered curricula created by large-scale tests. Some opponents have argued that, while other countries may have employed large-scale testing more frequently than we have in the past, the trend is toward testing less.

It is that latter claim of a trend toward less large-scale testing in other countries that is examined here. The claim can be found in five publications: *Testing in American Schools: Asking the Right Questions* (U.S. Congress, 1992); a contractor report for the OTA (Madaus & Kellaghan, 1991); "Educational Testing Abroad and Lessons for the United States," a spin-off article in *Educational Measurement: Issues and*

*Practice (EM:IP)* by two of the OTA report authors, M. J. Feuer and K. Fulton (1994); and the chapter "National Curricula in European Countries" (Kellaghan & Madaus, 1995, pp. 91–92) and book "The Use of External Examinations to Improve Student Motivation" (Kellaghan, Madaus, & Raczek, 1996, p. 59) produced by the American Educational Research Association.

Though both the claims and the comparison group of countries varied a bit from one publication to another, essentially the major assertions about the trends in large-scale, external testing in other developed countries were:

- Other countries are dropping, not adding, large-scale, external testing because they no longer need it as a selection device. They have made (or are making) places in lower and upper secondary schools universally available and have widened admission to higher education. Thus, they no longer need tests to exclude people.
- External examinations are "no longer used to make decisions about students' educational paths during the period of compulsory education" (Kellaghan et al., 1996, p. 59). Adoption of large-scale, external examinations in the United States during the period of compulsory education would run counter to the opposite trend in other countries where "standardized national examinations before the age of 16 have all but disappeared" (U.S. Congress, Office of Technology Assessment, 1992, pp. 135, 144).

- The trend toward less external testing in other countries can be found at all levels of education "even at the postsecondary level" (U.S. Congress, Office of Technology Assessment, 1992, p. 144).
- The trend is universal across countries and unidirectional; large-scale, external tests are being "abolished" (U.S. Congress, Office of Technology Assessment, 1992, p. 143).

These assertions were supported by a few anecdotes here and there, and, without citation, a claim made that four European countries—Belgium, Greece, Portugal, and Spain—had eliminated national examinations.

## Purpose, Scope, and Method

In order to see more data on the topic of trends in large-scale, external testing across developed countries, a search was made for source documents that provided such information. The research universe includes developed countries—those belonging to the Organisation for Economic Co-operation and Development (OECD)—plus China and Russia. The OECD includes most of the countries of Europe (except those in the Balkans or those formerly belonging to the Soviet Union), Canada, the United States, Korea, Japan, Australia, and New Zealand. The universe of potential sources examined is listed in Table 1.

Table 1 includes the universe of potential sources, including many documents not used as sources. Indeed, only a small proportion of the several hundred documents examined actually contained mention of

---

*Richard P. Phelps is Senior Study Director at Westat, Rockville, MD. His specializations are economics of education, testing systems, and international comparisons.*

---

**Table 1****Sources Used in This Study, by Organization and Source Type**

---

- Organisation for Economic Co-operation and Development (OECD) publications
- All available country education studies ( $n \sim 12$  book-length reports)
  - All country education system profiles in the annual indicator publication, *Education at a Glance* ( $n > 25$  country profiles)
  - All OECD research publications related to curriculum, standards, and assessments across countries ( $n > 15$  reports)
  - All editions of the newsletter of Network A, the member country network concerned with testing in the Indicators of National Education Systems project (INES), which produces the annual OECD indicator publication, *Education at a Glance* ( $n = 8$  newsletters)
  - All country responses to an OECD survey of country testing practices, conducted in 1991. ( $n \sim 16$ )
- European Union education publications (EURYDICE/CEDEFOP)
- All country profiles in *Structures of the Education and Initial Training Systems in the European Union* ( $n > 20$ )
- United States Education Department publications
- All U.S. Education Department reports with international comparative information ( $n > 10$  reports)
- United States Congressional Research Agencies
- All documents related to international comparative education information from the General Accounting Office (GAO), Congressional Budget Office (CBO), and the Office of Technology Assessment (OTA) ( $n = 3$  reports)
- All reports from independent U.S.-based research organizations relevant to international education comparisons ( $n > 5$ )
- Organizations such as: The National Center for Education and the Economy; the Council for Basic Education; the Council of Chief State School Officers; Pelavin Research Institute; MPR Associates
- International Encyclopedia of National Systems of Education
- The relevant chapter for each country from the last three volumes (1996, 1992, 1988) of the Pergamon series ( $n > 75$  chapters)
- All book-length compilations of articles on comparative international testing ( $n = 6$  books)
- Such as those by Eckstein and Noah, Little and Wolf
- An ERIC search for articles and papers from all countries in the study universe using the parameters "testing" and "(country name)" ( $n > 100$  articles)
- All articles in *Educational Measurement: Issues and Practice* related to testing systems in other countries ( $n \sim 10$  articles)
- All articles in *Educational Assessment* related to testing systems in other countries ( $n = 1$  article)
- 

test adding or test dropping. Several thousand pages of text and tables were examined in a process that took, in absolute time, several weeks over the course of several years.

Sources were examined thoroughly and in detail. Even those providing mention of test adding or test dropping rarely provided it in readily accessible, easy-to-find

form. Usually, the mention was buried inconspicuously in long passages of text. The General Accounting Office study of trends in Canadian provincial testing represents a notable exception—a publication with a primary goal of accounting for test adding and test dropping over the course of the past few decades.

The two most typical sources of information were country education system studies, which might have a relevant mention of test adding or test dropping imbedded somewhere within them, and articles specifically devoted to standards, curriculum, or assessment systems in one or more countries that, likewise, might have a relevant mention of test adding or test dropping imbedded somewhere within them. The overwhelming majority of research articles on testing say nothing about the addition or deletion of large-scale, external testing or the timing of either. The most likely sources of information for this article, then, were not research articles in journals but documents from government and government-affiliated organizations, and organizations of governments, such as the OECD. Nonetheless, the search for information included research articles in journals.

The countries not included in the analysis are left out either because I could not confidently determine a trend; I could find no, or only incomplete, information; or there appeared to be no trend—the testing program remained the same as it was a quarter century ago (e.g., The French Community of Belgium, Italy, Switzerland).

Perhaps not all the information available on the subject has been found. Nonetheless, a great deal has been uncovered. All evidence that could be found on both test adding and test dropping is included here. Even if this study may not be absolutely thorough, it provides ample evidence to refute at least one of the hypotheses introduced at the outset of this article—namely, a unidirectional trend toward less large-scale, external testing.

The research method may be slightly biased toward the addition of tests, however. Certainly, all references found relating to the reverse trend—to countries or states dropping tests—have been included. However, it may be easier to find references for test adding than for test dropping, due to the nature of a minority of the source documents, which had the purpose of describing current testing programs and, thus, current tests. If a test had been

dropped in the past, it would not now be a current test. Conversely, if a test had been added in the past, it could now be a current test. Sometimes, the year of that addition is noted by way of explaining the background and intended purpose of current tests.

The period in question is the last quarter of this century, up to and including 1999. This period was chosen because the authors who claimed a decline in testing used a large time period from which to make trend judgments and because testing policies in other countries—perhaps unlike our own—are slow moving and changes in them sluggish. It takes some considerable perspective before one can call a change in testing policy a trend.

### Accounting for Test Adding and Test Dropping

Table 2 lists testing that has been added or dropped over the last quarter of this century across 31 countries and provinces. Only testing added, dropped, or changed is listed; testing that was present at the beginning of the time period and never dropped or changed is not listed because it represents no change. Testing is listed in Table 2 by country or province, along with its type and stakes, whether it represented an addition to a testing program, a deletion, or no net change. Finally, the table notes the net effect on the amount of testing in the country or province.

I have assembled a 20-page list of every testing program from Table 2,

with a citation to the source document(s) and the relevant text from the source document. This list is available on request. A briefer "List of Sources," organized by country, is included as an appendix in this article.

A brief examination of Table 2 reveals much more test adding than test dropping. If one counts countries and provinces by their net effects, 27 countries and provinces increased their amount of testing, while only 3 countries or provinces decreased it.

Figure 1 shows the cumulative net effect in the number of tests added or dropped across all the countries and provinces over the time period 1974–1999, utilizing information from Table 2 on the tim-

**Table 2**

*Addition, Deletion or Replacement of Large-Scale, External, Standardized Testing, by Type of Test, Stakes, and Country or Province, 1974–1999*

Country	Testing added or dropped	Type of test	Student stakes	Net effect
Australia	(–) Grade 10 exam in most states (+) Australian Vocational Certificate exam	Exit exam Exit exam	High High	Decrease
Belgium (Flemish)	(+) Constructing tests in math and Flemish for the final year of secondary education, 1996	Exit exam	?	Increase
Canada	(+) National assessments in School Achievement Indicators Program, 1993	Assessment	None	Increase
Canada (Alberta)	(+) Alberta Assessment Program, 1982 (+) Alberta Diploma Examination, 1984	Assessment Exit exam	None Medium	Increase
Canada (British Columbia)	(+) British Columbia Provincial Examination, 1983 (+) British Columbia Provincial Learning Assessment Program, 1976	Exit exam Assessment	Medium None	Increase
Canada (Manitoba)	(+) Manitoba Provincial Examination, 1991 (+) Manitoba Curriculum Assessment, 1979	Exit exam Assessment	Medium None	Increase
Canada (New Brunswick)	(o) New Brunswick Achievement Examinations, 1984 (+) New Brunswick 2nd Language Oral Proficiency Interview, 1978 (+) New Brunswick Provincial Examinations, 1994 (+) School-level Assessments, grades 3 and 6, 1995 (+) Student Assessments, grades 3 and 5, 1998 (+) Student Tests, grade 8 language arts and math, 1998	n/a Exit exam Exit exam Assessment Assessment Exit exam	n/a High High None Low High	Increase
Canada (Newfoundland)	(+) Newfoundland Public Examination, 1974 (+) Newfoundland Achievement Testing Program, 1974	Exit exam Achievement	Medium Low	Increase

(continued)

**Table 2 (Continued)**

Country	Testing added or dropped	Type of test	Student stakes	Net effect
Canada (Nova Scotia)	(+) Criterion-referenced assessment for grade 5 math, language arts, and writing, 1989	Assessment	Low	Increase
Canada (Ontario)	(+) Provincial Assessment, late 1980s	Assessment	None	Increase
Canada (Prince Edward Island)	(-) Norm-referenced testing program, 1992	Achievement	Low	Decrease
Canada (Québec)	(o) Variety of vocational tests consolidated from 104 to 21, 1990	n/a	n/a	Increase
	(+) Quebec Curriculum tests, 1992	End-of-course	Medium	
Canada (Saskatchewan)	(+) Curriculum Evaluation Program, 1991	Assessment	None	Increase
	(+) Provincial Learning Assessment, 1994	Assessment	None	
China	(-) Entrance exam to lower secondary school, late 1980s	Entrance exam	High	Increase
	(+) General high school completion test in some states, late 1980s	Exit exam	High	
Czech Republic	(+) Uniform university entrance exam, 1977	Entrance exam	High	
	(+) New, national secondary-school leaving examination, 1992	Exit exam	High	Increase
	(+) National assessments being developed, 1997	Assessment	None	
Denmark	(+) Higher Commercial Examination instituted nationwide, 1982	Entrance/exit	High	Increase
	(+) Higher Technical Examination established (1982), made nationwide, 1988	Entrance/exit	High	
	(+) Diagnostic tests in reading, spelling, and math introduced, 1980s	Diagnostic	None	
England & Wales	(+) Assessment of Performance Unit (APU), 1974	Assessment	None	Increase
	(+) National Assessment, 1991	Assessment	None	
	(o) General Certificate of Secondary Education replaces old Certificate of Secondary Education and O-levels exams, 1988	n/a	n/a	
	(o) General National Vocational Qualification replaces earlier vocational track exams, 1992	n/a	n/a	
	(+) Advanced Supplementary (AS) standardized end-of-course exams, 1987	End-of-course	Medium	
Finland	(+) National system of assessment in elementary, upper secondary, and adult education, 1998	Assessment	None	Increase
France	(+) National assessments, 1980	Assessment	None	Increase
	(+) <i>Brevet</i> , lower-secondary-school leaving exam reintroduced, 1990	Exit exam	High	
	(+) <i>Baccalauréat</i> for professional/commercial track added to already existing ones for the academic and technical tracks, 1995. There are now three general <i>bacs</i> , nine technical <i>bacs</i> , and a number of vocational <i>bacs</i>	Voc/tech exit	High	
	(+) Entrance examinations to the <i>grandes écoles</i> , 1990s	Entrance exam	High	
	(+) National assessment being developed for grade 1 (pilot test in 1998)	Assessment	None	
Germany	(+) Stiffened requirements for passing Abitur, 1989	Exit exam	High	Increase
	(+) With number of <i>abitur</i> recipients growing and exceeded the number of available places in universities, many universities introduce selective entrance examinations	Entrance exam	High	
Greece	(-) Eliminated entrance exams to all types of Lycae, 1980s	Entrance exam	High	Decrease

(continued)

**Table 2 (Continued)**

Country	Testing added or dropped	Type of test	Student stakes	Net effect
Hungary	(+) MONITOR program, administered every several years, 1986	Assessment	None	Increase
	(+) Basic Examination for end of grade 10 (end of lower secondary), 1994	Exit exam	High	
Ireland	(+) Diagnostic tests at ages 7 and 11, 1992	Diagnostic	None	Increase
Japan	(+) Joint First Stage Achievement Test (JFSAT) introduced, required for students applying to public universities, . . . in addition, each competitive university administered its own entrance examinations	Entrance exam	High	Increase
	(o) All students applying to public universities and some private universities must take the Test of the National Center for University Entrance Examination (TNCUEE), which replaces the JFSAT, 1990. . . . In addition to the TNCUEE, each competitive university sets and offers its own exam	Entrance exam	High	
	(+) National assessment in primary and secondary schools introduced, 1980s	Assessment	None	
Korea	(o) National system of upper secondary school entrance exams replaces local system, 1974	Entrance exam	High	No change
	(o) Preliminary Examination for College Entrance (PECE) replaced by Scholastic Achievement Test for College Entrance (SATCE), 1982	Entrance exam	High	
Netherlands	(+) One more subject added to secondary school exit exam, 1993	Exit exam	High	Increase
	(+) Two national assessments, covering all major subjects in primary school, 1980s	Assessment	None	
	(+) Comprehensive norm-referenced achievement battery for each primary grade introduced, 1980s	Achievement	High	
	(+) Subject-area tests for 7 subjects in both lower and upper secondary introduced, 1987–1994	End-of-course	Medium	
New Zealand	(+) New National Certificate program expands on old School Certificate program into more grades, wider scope, 1994	End-of-course	Medium	Increase
Portugal	(+) Final Written Examination (12th grade), 1995	Exit exam	Medium	Increase
	(+) National Comparative Assessment, grades 4, 6, and 9, 1995	Assessment	None	
Scotland	(+) Standard Course Examinations, end-of-course tests introduced, 1986	End-of-course	Medium	Increase
Spain	(+) National assessments at the primary and secondary levels and in the vocational track, 1996	Assessment	None	Increase
Sweden	(–) Final secondary school examination dropped, mid-1970s	Exit exam	High	Increase
	(+) Higher education Swedish Scholastic Aptitude Test, 1977	Entrance exam	High	
	(+) Expansion of coverage of Swedish SAT, 1991	Entrance exam	High	
	(+) The National Evaluation of Compulsory Education, 1989, a national sampling of classes in grades 2, 5, and 9 every three years.	Assessment	None	

*(continued)*

**Table 2 (Continued)**

Country	Testing added or dropped	Type of test	Student stakes	Net effect
	(+) Currently constructing national tests for five upper secondary school courses; added voluntary national tests in 3 subjects several years ago, late 1990s	End-of-course	Medium	
	(+) Criterion-referenced test at the end of compulsory (lower secondary) school, 1990s	Exit exam	High	
Total	(+) 27 countries and provinces (-) 3 countries and provinces (o) 1 country			Increase

No mention of test adding, dropping, or changing found in the time period for Austria, Belgium (French Community), Italy, Luxembourg, Mexico, Norway, Switzerland, or Turkey.

(+) New large-scale, external standardized testing introduced.

(-) Existing large-scale, external standardized testing dropped.

(o) No change in amount of testing, because test is replaced by another of same type, or change does not affect absolute amount of testing.

n/a not applicable.

ing of the test addition or deletion. Where the timing is not known precisely to the year but, say, to a several-year period, a median year is used to represent the timing in Figure 1.

Table 3 summarizes the Table 2 information by type of testing, with the number of countries or provinces adding or dropping each type of testing counted up in each row of the table. The number of types of testing added by all the countries and

provinces totaled 53. The number of types of testing dropped totaled only 5. The types of testing added totaled 53 among 31 countries and provinces because some countries and provinces added more than one type of testing.

One can observe in Table 3 that assessments were the most popular type of testing added across countries, with 18 countries adding and none dropping. (The term *assessment* is used here to represent examina-

tions used for system diagnosis and monitoring that do not have high student stakes.) Ranking next highest in frequency: 12 countries added upper secondary level exit exams (while 1 dropped); 5 countries added new university entrance exams; 6 countries added new subject-area end-of-course exams; 4 added lower secondary level exit exams (while 3 dropped); 3 added vocational/technical or commercial/professional track students to the upper secondary level exit exam system; 2 added primary or secondary level achievement testing programs (while 1 dropped such a program); and 2 added diagnostic testing.

The same data are summarized a bit differently in Table 4. Here, all added or dropped testing is counted up and summarized by type of testing. The total number of testing programs added (59) is larger than its counterpart total in Table 3 because some countries have added more than one test of a certain type.

Finally, Table 5 summarizes the same data by the stakes attached to each testing program. No-stakes tests bear no consequences to a student for poor performance on the test; most assessments have no student stakes. Low stakes generally identify tests that represent no unalterable, serious consequences—such as, retention in grade level or failure to graduate—just as

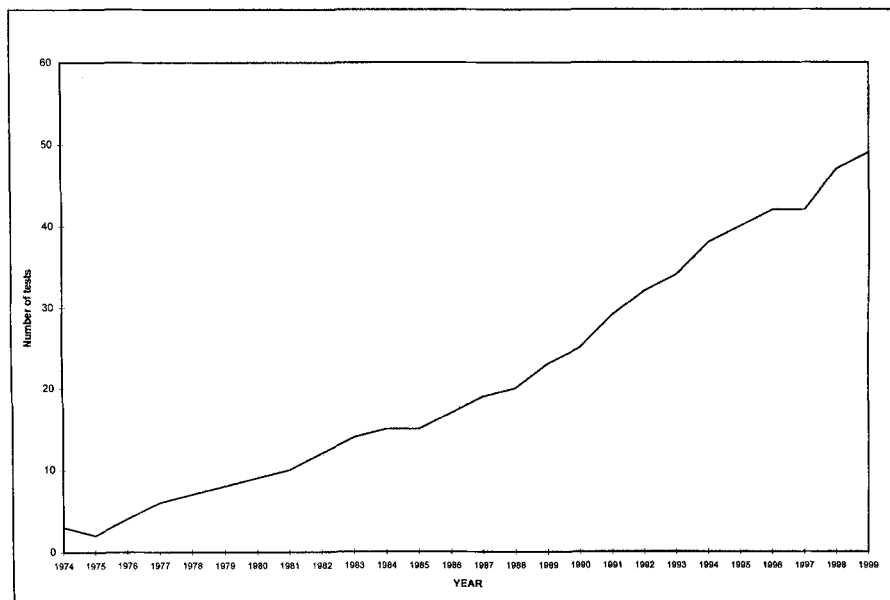


FIGURE 1. Cumulative net change in number of tests in 31 countries and provinces, 1974–1999

**Table 3**  
**Countries Adding or Dropping Large-Scale, External Testing, by Type of Testing: 1974–1999**

Type of testing	Number of countries or provinces	
	Adding testing	Dropping testing
Assessments	18	0
Upper secondary exit exams*	12	1
University entrance exams	5	0
Subject-area end-of-course exams	6	0
Lower secondary exit or entrance exams	4	3
Inclusion of vocational/technical or commercial/professional tracks in exit exam system	3	0
Primary/secondary-level achievement testing	2	1
Diagnostic testing	2	0
Total	53	5

\* In the Netherlands, one subject was added to a preexisting test.

signment to certain accelerated or slower tracks, perhaps. Medium and high stakes identify tests that do bear serious consequences—retention in grade level or failure to graduate. The term *medium* is commonly applied to systems where the stakes are high, but the test alone is only one of two or more factors considered in making the high-stakes decision; systems that mod-

erate or blend test scores with other measures have medium stakes by this definition, while those which do not moderate or blend have high stakes.

A glance at Table 5 will show that most (30) of the 54 added testing programs have medium or high stakes, while only 4 have low stakes, and 21 (mostly, the assessments) have no (student) stakes.

**Table 4**  
**Large-Scale, External Testing Added or Dropped Across 31 Countries and Provinces, by Type of Testing: 1974–1999**

Type of testing	Number of testing programs	
	Added	Dropped
Assessments	22	0
Upper secondary exit exams*	12	1
University entrance exams	7	0
Subject-area end-of-course exams	6	0
Lower secondary exit or entrance exams	4	3
Inclusion of vocational/technical or commercial/professional tracks in exit exam systems	4	0
Primary/secondary-level achievement testing	2	1
Diagnostic testing	2	0
Total	59	5

\* In the Netherlands, one subject was added to a preexisting test.

### Examining the Testing Trends Hypotheses Against the Evidence

In light of this evidence, several hypotheses made by testing opponents regarding trends in the amount of large-scale, external testing in countries outside the United States can be examined.

#### *Hypothesis 1: The Amount of Testing Is Declining*

The evidence clearly does not support this hypothesis. In 31 countries and provinces, 59 large-scale, external testing programs were added, and only 5 were dropped in the quarter century under consideration. It is certainly possible that *some* test adding or dropping in these countries for this period was missed in this search, but it is highly unlikely that more than 50 instances of test dropping and none of test adding were missed.

Those who believe the hypothesis to be true have also asserted that the trend toward less testing affects all levels of education—even postsecondary. The evidence suggests a decrease in testing at no levels of education, including postsecondary, where 7 new university entrance examinations systems have been introduced.

#### *Hypothesis 2: Decrease in Testing Due to Decreased Need for Selection*

It has been asserted that other countries used testing to select or exclude students at levels of education without enough places for all who wanted them. Since most developed countries now have enough places in both lower and upper secondary schools to accommodate all children in the relevant age groups, there is no longer any need to exclude anyone.

This hypothesis ignores two relevant issues: Selection can be for reasons other than exclusion from an entire level of education, and testing can be for reasons other than selection.

As one can see from the categories of test types in Tables 2–4, most testing has been added for purposes other than selection or exclusion from the next level of education: monitoring assessments (22 tests);

**Table 5**

*Large-Scale, External Testing Added or Dropped Across 31 Countries and Provinces, by Level of Student Stakes: 1974–1999*

Level of stakes	Number of testing programs added	Number of testing programs dropped
No stakes	21	0
Low stakes	4	1
Medium or high stakes	30	4
Total	54	5

subject-area end-of-course exams (6); primary/secondary level achievement tests (2); and diagnosis (2). In most cases, the upper secondary-level exit examinations (12 tests) do not serve to select and exclude students from places in the next level of education because most students do not continue on. For them, the exam provides a certification.

Moreover, selection may be for reasons other than gaining a place in the next level of education. In countries where students can choose their own school or program, there may be selection tests. In countries where students are assigned to different curricular tracks, there may be selection tests for that purpose. These purposes will remain, irrespective of the number of places available in any level of education overall.

Finally, while the number of places available in secondary education has grown to make places available to all who desire them, the same is not so at the higher education level, nor, probably, will it ever be. Taxpayers feel less obligated to support students past their post-compulsory years, and the students cost more to support. Five countries in the group examined here added some type of entrance testing for the higher education level.

#### *Hypothesis 3: Trends Are Universal Across Countries and Unidirectional*

Those who believe that the amount of testing is declining across countries have also asserted that the trend is universal—all countries share in the decline—and unidirectional—the amount of testing is

only declining and inevitably will decline more.

The information presented in Tables 2, 3, and 4 disputes this. The overall trend seems clearly toward more testing, not less, and most countries (27) have increased their amount of testing, while only 3 have decreased it.

Regarding the unidirectional hypothesis, some countries and provinces have dropped tests in the past (primarily in the 1960s and early 1970s) only to reintroduce them later (e.g., France and the *brevet*, Sweden with several tests, and Portugal, Alberta, British Columbia, and Manitoba with their secondary school exit exams). It would seem that other countries and provinces can go through cycles of favoring and not favoring tests, just as U. S. states can. All the aforementioned countries and provinces dropped high stakes examinations in the 1960s and early 1970s in the interest of making life easier for students and then reintroduced the tests later after deciding that students worked less and learned less with testing absent. Obviously, dropping a test does not necessarily mean it will remain dropped forever.

Sometimes, testing has been dropped at one level of government as part of an overall restructuring of governance regimes. Sweden, Spain, and Portugal, for example, have devolved authority over the past few decades from central to regional administrative authorities in education and other areas of governance. As the new governance regimes are settling in, new testing programs are being introduced that differ from those of the past. In Spain,

some of the Autonomous Communities are developing their own testing programs (OECD, 1996, Network A, Issue 3, p. 2).

#### *Hypothesis 4: No More Tests Before Age 16*

Those who have asserted the declining testing hypotheses have also declared that “standardized national examinations before the age of 16 have all but disappeared” (U.S. Congress, pp. 135, 144) and are “no longer used to make decisions about students’ educational paths during the period of compulsory education” (Kellaghan et al., 1996).

These statements simply are not true. There’s the French *brevet*. Italy, Iceland, Ireland, Denmark, New Zealand, Switzerland, Hungary, and the French Community of Belgium also have lower secondary school exit exams. Sweden is adding one. New Brunswick now has one in one subject and is adding more. Italy and the French Community of Belgium also have large-scale, external, exit exams from primary school. Quebec and Sweden have high stakes tests in several subjects in Grade 10. Japanese and Korean students must do well on exams at the lower secondary level in order to enter the upper secondary schools of their choice. Finally, there are the primary/secondary level achievement, or end-of-course, tests in several countries, including four included in this article, and the assessments, some of which are student-level assessments with student scores used in making recommendations for students’ educational paths.

Moreover, the assertion that tests at this level are not numerous ignores the fact that this level doesn’t really exist for some countries. There’s no lower secondary school exit exam in Germany, for example, because there’s no lower secondary school—students go straight from the end of primary to upper secondary graduation in one school, where they encounter a high-stakes exit examination.

#### *Reasons for Adding Testing*

Those who asserted the declining testing hypothesis suggested that there was only one purpose for tests: selection for an inadequate number



of places in the next level of education. They also suggested that the trend in testing was unidirectional and declining.

Both these beliefs appear to be incorrect. There are many purposes for tests, and the process of test adding and test dropping is dynamic. It can move in either direction depending on the place and time, and the prevailing, current international trend appears to be toward more testing.

There exist a variety of circumstances and factors—not just one—that determine whether a test is added or dropped. Among those mentioned in the source documents for test adding were:

- a concern about eroding standards;
- an increased incidence of geographical migration of students and their families and the consequent need to standardize curricula across regions;
- increased or decreased demand for places in selective secondary- or higher education-level programs or schools;
- a move to a higher level of government authority over education—such as, from cities to states—and a desire of those policymakers for more information;
- a devolution to a lower level of government authority over education, such as from country to states, and a desire of those policymakers to compare their system performance or the performance of their students to others; and
- a desire to use testing for more purposes than just end-of-level certification as in the past, to

take advantage of advances in testing theory and technique, and to imitate the successful use of testing for other purposes in other countries, such as the United States. Other purposes include: large-scale assessments, individual diagnosis, and end-of-course examination.

#### *International Assessments*

To be complete, it should be mentioned that many countries now also participate in international assessments. While there was one international mathematics and science assessment prior to the 1974–1999 period, it included only several countries and at only one grade level. International assessments have in more recent years become larger and more frequent. Mathematics and science assessments alone were administered in 1988, 1991, 1995, and 1999. The Third International Mathematics and Science Study (TIMSS) involved 45 countries at 3 grade levels in 1995 and is being repeated at one grade level in some of the countries in 1999. The Organisation for Economic Co-operation and Development (OECD) is sponsoring the Programme for International Student Assessment (PISA), an effort to assess students on a regular, every-few-years' basis. Though PISA will emphasize the major subject areas, it will focus on practical problem-solving skills. PISA is due to start within the next couple of years with dozens of countries participating.

Some countries rely on these international assessments for their internal information needs and not

just for international comparisons. Instead of developing the assessments themselves, they use the information from the international assessments to monitor and diagnose the performance of their school systems.

Were we to include countries' participation in international assessments among the tests added in Tables 2–5, the number of tests added in our quarter century time period would more than double. Of course, international assessments have, up to now, only been administered occasionally and without regular periodicity. All the testing counted in Tables 2–5 have regular, periodic (usually annual) administrations.

#### **Conclusion**

Comparing statistics on the addition or deletion of large-scale, external tests in 31 countries and provinces from 1974–1999 reveals a clear trend toward adding, not dropping, testing programs. Twenty-seven countries show a net increase in testing, while only three show a decrease. Fifty-nine testing programs have been added, while only four have been dropped.

Moreover, these statistics provide no support for any of the hypotheses posed at the outset of this article—namely, the amount of testing is declining; the declining trend is universal across countries and unidirectional; a decrease in testing is due to a decreased need for selection; and “standardized national examinations before the age of 16 have all but disappeared.” The evidence suggests the opposite in each case.

---

## **APPENDIX**

### *List of Sources*

#### **Australia**

National Center on Education and the Economy, “The International Experience with School Leaving Examinations,” April 1994.

#### **Belgium (Flemish)**

Organisation for Economic Co-operation and Development, Indi-

cators of National Education Systems project, Network A, “Review of Assessment Activities” Issue 3, Winter 1996.

#### **Canada**

Garet, M.S. and Moskowitz, J. “Multilateral Comparisons of National Assessments in Education,”

Pelavin Research Institute, March 1996.

#### **Canada—Alberta**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—British Columbia**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—Manitoba**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—New Brunswick**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April, 1993; and telephone conversation with Cary Grobe of the New Brunswick Education Ministry, April 21, 1998.

**Canada—Newfoundland**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—Nova Scotia**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—Ontario**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993; and telephone conversation with Ontario Education Ministry official, April 21, 1998.

**Canada—Prince Edward Island**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—Québec**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards,*

*Examinations, and Assessments*. PEMD-93-11, April 1993.

**Canada—Saskatchewan**

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations, and Assessments*. PEMD-93-11, April, 1993; and telephone conversation with Rick Johnson, Saskatchewan Education Ministry, April 21, 1998.

**China**

Eckstein, M.A. and Noah, H.J. *Secondary School Examinations: International Perspectives on Policies and Practice*, New Haven, Yale University Press, 1993.

Teng, T. "Peoples' Republic of China," in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 206-212.

Hong-wei, Meng. "Shifts towards monitoring in the world's largest education selection system: current Chinese experience." in Little, Angela and Alison Wolf, Eds. *Assessment in Transition*, pp. 154-171.

**Czech Republic**

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 8, Spring 1998

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 7, Autumn 1997

**Denmark**

Organisation for Economic Co-operation and Development, *Reviews of National Policies for Education: Denmark*. Paris: 1995.

Council of Chief State School Officers, *European Lessons from School and the Workplace*, 1991.

Jansen, N. and Kreimer, S. "Denmark," in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 266-274.

**England & Wales**

Gipps, C. and Goldstein, H., "Local and National Testing in the UK: The Last Ten Years," Paper presented at the Annual Meeting of the AERA, New Orleans, LA, 1984.

Gipps, C., "National Testing at 7: What Can It Tell Us?" Paper presented at the Annual Meeting of the AERA, San Francisco, CA, 1992.

Garet, M.S. and Moskowitz, J. "Multilateral Comparisons of National Assessments in Education," Pelavin Research Institute, March 1996.

National Center on Education and the Economy, "The International Experience with School Leaving Examinations," April 1994.

Stevenson, H.W. and Lee S., *International Comparisons of Entrance and Exit Examinations*, U.S. Department of Education, Office of Educational Research and Improvement, 1997.

**Finland**

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 8, Spring 1998.

**France**

Garet, M.S. and Moskowitz, J. "Multilateral Comparisons of National Assessments in Education," Pelavin Research Institute, March 1996.

National Center on Education and the Economy, "The International Experience with School Leaving Examinations," April 1994.

Stevenson, H.W. and Lee S., *International Comparisons of Entrance and Exit Examinations*, U.S. Department of Education, Office of Educational Research and Improvement, 1997.

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 5, Fall 1996.

## Germany

Eckstein, M.A. and Noah, H.J. *Secondary School Examinations: International Perspectives on Policies and Practice*, New Haven, Yale University Press, 1993.

Stevenson, H.W. and Lee S., *International Comparisons of Entrance and Exit Examinations*, U.S. Department of Education, Office of Educational Research and Improvement, 1997.

## Greece

SOURCE: Organisation for Economic Co-operation and Development, *Reviews of National Policies for Education: Greece*. Paris: 1997.

## Hungary

Dajka, B. Hungarian response to OECD "Survey on Education Standards," March 19, 1993, pp. 2-3.

Csapo, B. "Educational Testing in Hungary," *Educational Measurement: Issues and Practice*, v.11, n.2, Summer, 1992, pp. 5-8+.

## Ireland

Coolahan, J. "Ireland," in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 452-460.

## Japan

National Center on Education and the Economy, "The International Experience with School Leaving Examinations," April 1994.

U.S. Department of Education, Office of Policy and Planning, *Education Standards in the Asia-Pacific Region*, October 1992.

## Korea [unable to discern trend]

U.S. Department of Education, Office of Policy and Planning, *Education Standards in the Asia-Pacific Region*, October 1992.

Shin, Se-ho. "Korea" in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education* (1996), pp. 515-524.

## Netherlands

Thio, D.K. "The Educational Testing Movement in the Netherlands: Some Historical and Political Aspects," *Educational Measurement: Issues and Practice*, v.13, n.2, Summer 1994, pp. 21-23.

Keersemaeker, N. Dutch response to OECD "Survey on Education Standards," February 3, 1993.

Vuyk, E.J. "Netherlands," in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 699-708.

## New Zealand

Irving, J. New Zealand response to OECD "Survey on Education Standards," January 22, 1993. OECD, *Schools Under Scrutiny*, p. 34.

## Portugal

"Portugal," in *Structures of the Education and Initial Training Systems in the European Union*, EURYDICE/CEDEFOP, 1995.

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 5, Fall 1996.

## Russia [unable to discern trend]

Nikandrov, N.D. "Russia" in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 819-828.

## Scotland

Halls, W.D. "United Kingdom" in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 1025-1033.

## Spain

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 5, Fall 1996.

## Sweden

Marklund, S. "Sweden" in Postlethwaite, T.N., Ed. *International Encyclopedia of National Systems of Education*, pp. 940-947.

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 8, Spring 1998.

Wedman, I. "The Swedish Scholastic Aptitude Test," *Educational Measurement: Issues and Practice*, v.13, n.2, pp. 5-11.

Organisation for Economic Co-operation and Development, Indicators of National Education Systems project, Network A, "Review of Assessment Activities," Issue 1, Spring 1995.

## Notes

The author would like to thank Jay Moskowitz, Maria Stephens, Marilyn Binkley, Natalia Pane, John Bishop, and Laurent Allard for their help in providing source documents. The author retains all responsibility for errors.

## References

Feuer, M. J., & Fulton, K. (1994). Educational testing abroad and lessons for the United States. *Educational Measurement: Issues and Practice*, 13 (2), 31-39.

Kellaghan, T., & Madaus, G. F. (1995). National curricula in European countries. In E. W. Eisner (Ed.), *The hidden consequences of a national curriculum* (pp.79-118). Washington, DC: American Educational Research Association.

Kellaghan, T., Madaus, G. F., & Raczek, A. (1996). *The use of external examinations to improve student motivation*. Washington, DC: American Educational Research Association.

Madaus, G. F., & Kellaghan, T. (1991). *Student examination systems in the*

*European community: Lessons for the United States* (OTA Contractor PB92-127570). Washington, DC: U. S. Government Printing Office.

Phelps, R. P. (1998). The demand for standardized student testing. *Educational Measurement: Issues and Practice*, 17 (3), 5.

U.S. Congress, Office of Technology Assessment. (1992). *Testing in American schools: Asking the right questions* (OTA-SET-519). Washington, DC: U. S. Government Printing Office.