University of Pennsylvania
ScholarlyCommons

CPRE Research Reports

Consortium for Policy Research in Education (CPRE)

4-1998

Do Curriculum-Based External Exit Exams Enhance Student Achievement?

John Bishop

Follow this and additional works at: http://repository.upenn.edu/cpre_researchreports

 Part of the Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, and the Education Policy Commons

Recommended Citation

Bishop, John. (1998). Do Curriculum-Based External Exit Exams Enhance Student Achievement?. *CPRE Research Reports.* Retrieved from http://repository.upenn.edu/cpre_researchreports/91

View on the CPRE website.

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cpre_researchreports/91
For more information, please contact libraryrepository@pobox.upenn.edu.

Do Curriculum-Based External Exit Exams Enhance Student Achievement?

Abstract
Two presidents, the National Governors Association and numerous blue ribbon panels have called for the development of state content standards for core subjects and examinations that assess the achievement of these standards. The Competitiveness Policy Council, for example, advocates that "external assessments be given to individual students at the secondary level and that the results should be a major but not exclusive factor qualifying for college and better jobs at better wages (1993, p. 30)." The American Federation of Teachers advocates a system in which:

> Students are periodically tested on whether they' reaching the standards, and if they are re not, the system responds with appropriate assistance and intervention. Until they meet the standards, they won' be able to graduate t from high school or enter college (AFT 1995 p. 1-2).

It is claimed that curriculum-based external exit exam systems (CBEEES), based on world class content standards will improve teaching and learning of core subjects. What evidence is there for this claim? New York' Regents Exams are an s example of such a system. Do New York students outperform students with similar socio-economic backgrounds from other states? Outside the United States such systems are the rule, not the exception. What impacts have such systems had on school policies, teaching and student learning?

Disciplines
Educational Assessment, Evaluation, and Research | Educational Methods | Education Policy

Comments
View on the CPRE website.

This report is available at ScholarlyCommons: http://repository.upenn.edu/cpre_researchreports/91

Do Curriculum-Based External Exit Exam Systems Enhance Student Achievement?

John Bishop

CPRE Research Report Series
RR-40

Consortium for Policy Research in Education
University of Pennsylvania
Graduate School of Education

8 **Copyright 1998 by the Consortium for Policy Research in Education**

Contents

Biography

John Bishop is an Associate Professor and Chair of the Department of Human Resources at Cornell University. His areas of research include education reform and work force preparedness, employer behavior and personnel, program evaluation and government training programs, income maintenance policy, and comparative human resource institutions and policies. He is a member of the Safety Net Study Group advising New York State's Commissioner of Education on how to minimize the number of students who fail to pass the New Regents examinations.

Author's Note

The preparation of this paper was made possible by support from the Center for Advanced Human Resource Studies and the Consortium for Policy Research in Education (CPRE), funded by a grant (No. OERI-R308A60003) from the National Institute on Educational Governance, Finance, Policymaking and Management, Office of Educational Research and Improvement, U.S. Department of Education. This paper has not undergone formal review or approval of the faculty of the New York State School of Industrial and Labor Relations (ILR). It is intended to make results of Center research, conferences, and projects available to others interested in human resource management in preliminary form to encourage discussion and suggestions. The findings and opinions expressed in this report are those of the author and do not reflect the position or policies of the U.S. Department of Education, ILR, CPRE, or its institutional partners.

Introduction

Two presidents, the National Governors Association and numerous blue ribbon panels have called for the development of state content standards for core subjects and examinations that assess the achievement of these standards. The Competitiveness Policy Council, for example, advocates that "external assessments be given to individual students at the secondary level and that the results should be a major but not exclusive factor qualifying for college and better jobs at better wages (1993, p. 30)." The American Federation of Teachers advocates a system in which:

> *Students are periodically tested on whether they're reaching the standards, and if they are not, the system responds with appropriate assistance and intervention. Until they meet the standards, they won't be able to graduate from high school or enter college* (AFT 1995 p. 1-2).

It is claimed that curriculum-based external exit exam systems (CBEEES), based on world class content standards will improve teaching and learning of core subjects. What evidence is there for this claim? New York's Regents Exams are an example of such a system. Do New York students outperform students with similar socio-economic backgrounds from other states? Outside the United States such systems are the rule, not the exception. What impacts have such systems had on school policies, teaching and student learning?

What's So Different About a Curriculum-Based External Exit Exam System?

Skeptics point out that American students already take lots of standardized tests. They ask, "Why should a curriculum-based external exit examination system significantly improve incentives and learning?" Advocates claim that CBEEES have uniquely powerful incentive effects because they have the following six characteristics.

1. **CBEEES produce signals of student accomplishment that have real consequences for the student.**

2. **CBEEES define achievement relative to an *external* standard, not relative to other students in the classroom or the school.** Fair comparisons of achievement across schools and across students at different schools are possible. Costrell's (1994) analysis of the optimal setting of educational standards concluded that more centralized standard-setting (state or national achievement exams) results in higher standards, higher achievement and higher social welfare than decentralized standard setting (such as teacher grading or school graduation requirements).

3. **CBEEES are organized by discipline and keyed to the content of specific course sequences.** This focuses responsibility for preparing students for particular exams on a small group of teachers.

4. **CBEEES signal multiple levels of achievement in the subject.** If an exam generates only a pass-fail signal, the standard will have to be set low enough to allow almost everyone to pass and this will not stimulate the great bulk of students to greater effort (Kang 1985; Costrell 1994).

5. **CBEEES cover almost all secondary school students.** Exams for a set of elite schools, advanced courses or college applicants will influence standards at the top of the vertical curriculum, but will probably have limited effects on the rest of the students. The school system as a whole must be made to accept responsibility for how students do on the

exams. A single exam taken by all is not essential. Many nations allow students to choose the subjects in which they will be examined and offer high and intermediate level exams in the same subject.

6. **CBEEES assess a major portion of what students studying a subject are expected to know or be able to do.** It is not essential, however, that the external exam assess every instructional objective. Teachers can be given responsibility for evaluating dimensions of performance that cannot be reliably assessed by external means.

Commercially prepared achievement tests such as the CAT, CTBS, ITBS, and ITED are not curriculum-based external exit exams because they fail requirement number one noted above (as well as several others). Students have no stake in doing well on these tests. Where stakes are attached to results, it is teachers and school administrators who experience the consequences, not individual students.

The minimum competency exams that many American states require students pass before graduation are not CBEEES because they fail requirements numbers three and four. Minimum competency exams are generally first taken in ninth and tenth grade and most students pass on the first sitting. High school transcripts indicate only whether the student eventually passes the test, not achievement levels above the minimum. For the great majority of students who pass on the first try, therefore, the tests no longer stimulate study. Incentive effects are focused on the small minority who fail on the first try and must repeat the test. Minimum competency exams can be a useful part of a CBEEES, but other more demanding curriculum-based exams that signal higher levels of performance are essential.

The requirement (number four) that a CBEEE signal different levels of achievement—not just whether the student has achieved a minimum—is essential because it has major effects on the incentive effects of exams. Students differ dramatically in their levels of achievement by age 13. On the National Assessment of Educational Progress, seven to nine percent of 13 year-olds are four or more grade-level equivalents behind their age mates and 15 to 17 percent are four or more grade-level equivalents ahead. When achievement differentials among students are this large, incentives for effort are stronger for most students if the full range of achievement is signaled rather than whether the individual has just passed some absolute standard. When a test generates only a pass-fail signal, many students pass without exertion and are not stimulated to greater effort by the reward for passing. Some of the least well-prepared students will judge the effort required to achieve the standard to be too great and the benefits too small to warrant the effort. They give up on the idea of meeting the standard. Few students will find the reward for exceeding a single absolute cutoff an incentive for greater effort (Kang 1985). Costrell agrees: “The case for perfect information [making scores on external examinations available rather than just whether the individual passed or failed] would appear to be strong, if not airtight: for most plausible degrees of heterogeneity, egalitarianism, and pooling under decentralization, perfect information not only raises Gross Domestic Product, but also social welfare.” (1994, p. 970)

The SAT-I reasoning tests are not curriculum-based external exit exams because they do not meet requirements numbered five and six. SAT-I tests fail to assess most of the material—history, science, economics, civics, literature, foreign languages and the ability to write an essay—that high school students are expected to learn. From the beginning the SAT was designed to minimize backwash effects on teaching and student study habits. Indeed, Richard Gummere, Harvard

College's admissions director when the machine scored multiple-choice Scholastic Aptitude Test (SAT) replaced the curriculum-based essay style College Board Examinations, was very candid about why the SAT had been adopted: "Learning in itself has ceased to be the main factor [in college admissions]. The aptitude of the pupil is now the leading consideration." (Gummere, 1943 p. 5)

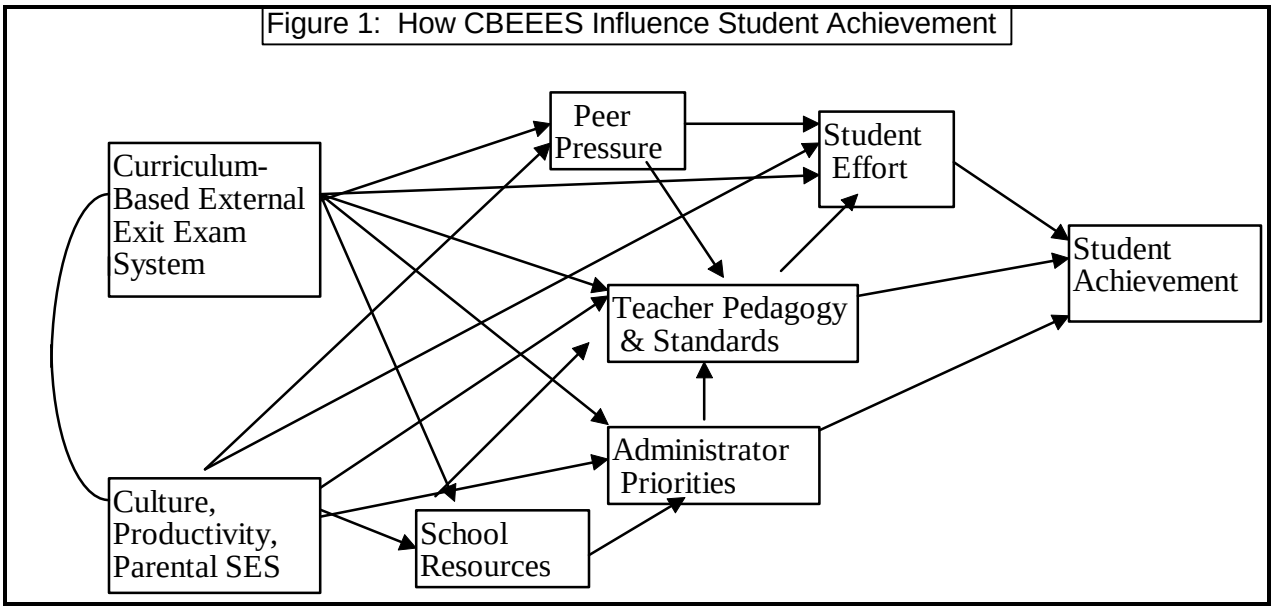
The subject specific SAT-II achievement tests fail requirements number one and five. Stakes are very low—few colleges consider SAT-II results in admissions decisions, and few students take them. In 1982-83 only six percent of SAT-I test takers took a science SAT-II and only three to four percent took a history or foreign language SAT-II test. Schools do not assume responsibility for preparing students for SAT-II tests.

The Advanced Placement (AP) examinations are the single exception to the generalization that the U.S. lacks national curriculum-based external exit examinations. Although growing rapidly, AP is still a very small program. In 1995 only 3.2 percent of juniors and seniors took AP english or AP history exams and only 2 percent took AP calculus or science exams (National Education Goals Panel, 1995). Low participation means that AP exams fail requirement number five and are, consequently, not a CBEEE system. They can, however, serve as a component of a larger system.

How are CBEEES Hypothesized to Increase Achievement?

Curriculum-based external exit exam systems (CBEEES) fundamentally change the signaling of student achievement, and by doing so they transform the incentives faced by students, parents, teachers and school administrators. Consequently, CBEEES hypothesized to influence the resources made available to schools and the priorities of school administrators, teacher pedagogy, parental encouragement and student effort. The many paths by which CBEEES influence student achievement are illustrated in Figure 1.

Figure 1: How CBEEES Influence Student Achievement



Impact on Students: CBEEES improve the signaling of academic achievement. As a result, colleges and employers are likely to give greater weight to academic achievement when they make admissions and hiring decisions, so the rewards for learning should grow and become more visible. CBEEES also shift attention toward measures of absolute achievement and away from measures of relative achievement such as class rank and teacher grades. By doing so, CBEEES ameliorate the problem of peer pressure against studying.

Interviews I conducted during 1996 and 1997 with middle school students in Collegeville, a small city dominated by two universities, indicate that most students (males especially) internalize a norm against "sucking up" to the teacher. How does a student avoid being thought a "suck-up?" He avoids making eye contact with the teacher; does not hand in homework early for extra credit; does not raise his hand in class too frequently; and talks or passes notes to friends during class (signaling that you value friends more than your reputation with the teacher). Steinberg, Brown and Dornbush similarly conclude, "The adolescent peer culture in America demeans academic success and scorns students who try to do well in school." (1996, p.19).

Why are the studious called **suck-ups**, **dorks** and **nerds** or accused of "**acting white**"? In part, because exams are graded on a curve, their study effort make it more difficult for others to get top grades. When exams are graded on a curve or when college admissions are based on class rank, joint welfare is maximized if no one puts in extra effort. Side payments (friendship and respect) and punishments (ridicule, harassment and ostracism) enforce the cooperative don't study solution. If, by contrast, students are evaluated relative to an outside standard, they no longer have a personal interest in getting teachers off track or persuading each other to refrain from studying. Peer pressure demeaning studiousness should diminish.

Impact on School Administrators: When there is no external assessment of academic achievement, students and their parents benefit little from administrative decisions that support higher standards, more qualified teachers or a heavier student work load. The immediate consequences of such decisions are all negative: higher taxes, more homework, having to repeat courses, lower grade point averages, complaining parents, a greater risk of being denied a diploma.

When student learning is not assessed externally, the positive effects of choosing academic rigor are negligible and postponed. If college admission decisions are based on class rank, GPA and aptitude tests, and not on externally assessed achievement in secondary school courses, upgraded standards will not improve the college admission prospects of next year's graduates. Graduates will probably do better in difficult college courses and be more likely to get a degree, but that benefit is uncertain and far in the future. Maybe, over time, the school's reputation and the college admission prospects of graduates will improve because the current graduates are more successful at local colleges. That, however, is even more uncertain and postponed. Publishing data on proportions of students that meet targets on standardized tests probably speeds the process by which real improvements in a school's performance influence its local reputation. However, other indicators—such as SAT test scores, proportions going to various types of colleges and the socioeconomic background of the students—tend to be more prominent. As a result, school reputations are largely determined by things that teachers and administrators have little control over: the socio-economic status of the student body and the proportion of graduates going to college.

American employers have historically paid little attention to student achievement in high school or to school reputation when selecting young workers (Bishop 1989, 1993; Hollenbeck and Smith 1984). Employers that do pay attention to student

achievement use as hiring criteria indicators of relative performance such as GPA and class rank rather than results on an external exam. Because higher standards do not benefit students as a group, parents as a group have little incentive to lobby strongly for higher teacher salaries, higher standards and higher school taxes.

External exams in secondary school subjects transform the signaling environment. Hiring better teachers and improving the science laboratories now yields a visible payoff—more students passing the external exams and being admitted to top colleges. This, in turn, is likely to lead to more spending on schools, more rigorous hiring standards for secondary school teachers and a higher priority assigned to student learning in the allocation of school budgets.

Additionally, reform-minded administrators have used results of CBEEES to shame and inspire teachers to raise standards for all students. The superintendent of a suburban New York district that has been nationally recognized for raising student achievement levels explained: "[External validators like Regents exams and International Baccalaureate] were the best and only way in which we could get teachers and staff to see themselves as others might see them and not just keep looking in the mirror and seeing themselves as they would like to see themselves." (Interview with superintendent of an All-Regents High school, August 1997)

Impact on Teachers: Thirty percent of American teachers say they "feel pressure to give higher grades than students' work deserves" and "feel pressure to reduce the difficulty and amount of work you assign" (Peter D. Hart Research Associates, 1994). Under a system of external exams, teachers and local school administrators lose the option of lowering standards to reduce failure rates and raise self-esteem. The available alternative is to demand more of their students in order to maximize their chances of being successful on the external exams.

An additional benefit of CBEEES is the professional development that teachers receive when they came together at centralized locations to grade the extended-answer portions of examinations. In May of 1996 I interviewed some teacher union activists about the examination system in the Canadian province of Alberta. Even though the union and these teachers opposed the exams, they universally reported that serving on grading committees was "…a wonderful professional development activity." (Bob, 1996) Coming to agree on what constituted excellent, good, poor, and failing responses to essay questions or open-ended math problems elicited a sharing of perspectives and teaching tips that most found very helpful.

Many, however, fear that external exams will negatively effect teaching. Opponents argue that "preparation for high stakes tests often emphasizes rote memorization and cramming of students and drill and practice teaching methods" and that "some kinds of teaching to the test permits students to do well in examinations without recourse to higher levels of cognitive activity." (Madeus 1991 p. 7-8)

CBEEES advocates challenge the assumption implicit in this argument that examinations developed by the committees of teachers working for state departments of education are or will be worse than tests developed by individual teachers. In fact, the tests teachers develop for themselves are generally of low quality. The 1983 Fleming and Chambers study of tests developed by high school teachers found that "over all grades, 80 percent of the items on teachers' tests were constructed to tap the lowest of [Bloom's] taxonomic categories, knowledge (of terms, facts or prin-

ciples)" (Thomas 1991, p. 14). Rowher and Thomas (1987) found that only 18 percent of history test items developed by junior high teachers and 14 percent of items developed by senior high teachers required the integration of ideas. College instructors, in contrast, required such integration in 99 percent of their test items. Secondary school teachers test low-level competencies because that is what they teach.

Carefully designed external exams can induce improvements in instructional practice. Sherman Tinkelman, New York State's Assistant Commissioner for Examinations and Scholarships, describes one such instance:

> *For years our foreign language specialists went up and down the State beating the drums for curriculum reform in modern language teaching, for change in emphasis from formal grammar to conversation skills and reading skills. There was not very great impact until we introduced, after notice and with numerous sample exercises, oral comprehension and reading comprehension into our Regents examinations. Promptly thereafter, most schools adopted the new curricular objectives* (Tinkelman, 1966 p. 12).

Do CBEEES Increase Achievement? A Look at the Evidence

The hypothesis that curriculum-based external exit examination systems (CBEEES) improve achievement will be tested by comparing nations, states and provinces that do and do not have such systems. Four different data sets will be examined: science and mathematics achievement of seventh and eighth graders in the 40-nation Third International Math and Science Study; science and math scores of 13-year-olds on the International Assessment of Educational Progress (IAEP) for 16 nations and nine Canadian provinces; and SAT test and NAEP math scores for New York State versus the rest of the United States. The theory predicts that CBEEES affect societal decisions about education spending, administrator decisions about school priorities, teachers decisions about standards and pedagogy and student decisions about studying. Much of the ultimate impact of CBEEES on student achievement derives from the changes they induce in spending, priorities and pedagogy. Most of the components of the full Figure 1 model have been estimated in data on Canadian schools and students in Bishop (1996). Educational systems are the units of observation in this paper, and in most analyses the objective is to assess the **total** effect of CBEEES on achievement (the sum of all the paths leading from CBEEES to student achievement in Figure 1). Total effects are estimated by a reduced form model that controls for parental socio-economic status (SES), productivity and national culture, but not the endogenous administrator, teacher and parent behaviors.

Third International Mathematics and Science Study

The recently released Third International Mathematics and Science Study (TIMSS) provides 1994-95 data for seventh and eighth graders for 39 countries. To determine which of the TIMSS nations have curriculum-based externally-set exit examinations in secondary school, [1] we reviewed comparative education studies, government documents and education encyclopedias and interviewed education ministry officials, embassy personnel and Cornell graduate students from the various countries. Twenty-two national school systems were classified as having CBEEES for both mathematics and science in all parts of the country: Austria, Bulgaria, Columbia, Czech Republic, Denmark, England, Hong Kong, Hungary, Ireland, Iran, Israel, Japan, Korea, Lithuania, the Netherlands, New Zealand, Russia, Scotland, Singapore, Slovak Republic, Slovenia and Thai-

land. Three countries—France, Iceland and Romania—had CBEEES in mathematics but not in science. Five countries—Australia, Canada, Germany, Switzerland and the United States—had CBEEES in some provinces but not in others. Norway has regular exit examinations in mathematics, but examines science only every few years. Latvia had an external examination system until very recently, so it was given a .5 on the CBEEES variable. The countries classified as having no CBEEES in either subject were Belgium (both Flemish and French speaking systems), Cyprus, Greece, Philippines, Portugal, Spain and Sweden. Based on the work of Madeus and Kelleghan (1991), the university entrance examinations in Greece, Portugal, Spain, and Cyprus, and the ACT and SAT in the U.S. were not considered to be CBEEES. University entrance exams should have much smaller incentive effects because students headed into work do not take them and teachers can avoid responsibility for their students' exam results by arguing that not everyone is college material or that examiners have set an unreasonably high standard to limit enrollment in higher education.

Figures 2 and 3 array the 40 TIMSS countries according to the science and mathematics achievement of their 13-year-olds. The U.S. ranks fifteenth in science and thirty-first in mathematics. The gaps between the vertical grid lines represent one U.S grade-level equivalent—the difference between seventh and eighth grade TIMSS test score means for the U.S. Achievement differentials across nations are very large. In science, Singapore, Korea, Bulgaria and Flemish Belgium are more than one grade-level equivalent ahead of the U.S.; Columbia, Phillipines, Lithuania, Romania and Portugal are more than three grade-level equivalents behind the U.S. In mathematics, Singapore, Korea, Japan and Hong Kong are four or more grade-level equivalents ahead of the U.S., while Columbia, Philippines and Iran are behind the U.S. by more than three grade-level equivalents. The countries represented by a solid black bar in the figures have a curriculum-based external exit exam in the subject; countries represented by white bars do not have CBEEES. The countries with a CBEEES in the subject tend to have higher TIMSS scores.

Regression Analysis: The mean seventh and eighth grade science and mathematics test scores were regressed on average per capita gross domestic product in 1987 and 1990 deflated by a purchasing power parity price index, a dummy for East Asian nation and a dummy for CBEEES. The results presented in Table 1 indicate that test scores are significantly higher in more developed nations, East Asian nations and in nations with a CBEEES in the subject.

The analysis of achievement at a particular grade level may be biased, however, by differing policies regarding grade retention, age of school entry and the grade chosen for assessment. CBEEES, for example, might be associated with high rates of grade retention. Therefore, a preferable dependent variable is a measure of student achievement at some fixed age. The third and fourth rows of each panel present estimated models predicting the median test score for each nation's 13-year-olds (Beaton et al., 1996a, b, Table 1.5). For countries not included in this table, the 13-year-old median was estimated by age adjusting the seventh and eighth grade means.[2] Switching to the age constant achievement somewhat reduces the estimated impact of the CBEEES but the effects remain statistically significant. Using two-tailed *t* tests, the CBEEES coefficient has a P = .08 in the mathematics model and a P = .01 in the science model. The estimated impacts are substantively important: 1.3 U.S. grade-level equivalents in science and 1.0 U.S. grade-level equivalents in mathematics.

One of the ways CBEEES may improve achievement is by inducing greater social investments in education. Row 4 presents results of regressions that add the share of GDP spent on education to

Figure 2. Math Achievement at Age 13

Figure 3. Science Achievement at Age 13

Table 1
The Effects of Curriculum-Based External Exams on Science and Mathematics Achievement

	External Exit Exam	LnGDP/Pop 1987 & 90	East Asia	$K-12/GDP	AdjR2/RMSE
TIMSS Science-1994					
Mean for 7th Graders	38.0***	33.8***	20.1		.317
	(2.93)	(3.44)	(1.24)		35.4
Mean for 8th Graders	42.4***	36.2***	14.4		.364
	(3.40)	(3.80)	(.92)		34.4
Median for 13 Yr Olds	34.9***	45.0***	21.5		.402
	(2.77)	(4.68)	(1.35)		34.7
Median for 13 Yr Olds	32.0***	38.0***	33.7*	13.6*	.442
	(2.57)	(3.71)	(2.01)	(1.86)	33.9
Diff-13 minus 9 Yr Olds	7.6	-32.3***	5.5		.258
	(.54)	(3.11)	(0.39)		26.4
TIMSS Mathematics-1994					
Mean for 7th Graders	29.6**	46.6***	66.0***		.469
	(2.09)	(4.60)	(4.01)		36.2
Mean for 8th Graders	36.0**	48.7***	62.0***		.476
	(2.54)	(4.81)	(3.75)		36.6
Median for 13 Yr Olds	24.7*	56.0***	9.4***		.537
	(1.82)	(5.77)	(4.37)		35.1
Median for 13 Yr Olds	21.5	53.9***	75.9***	5.7	.545
	(1.55)	(5.07)	(4.41)	(.75)	35.1
Diff-13 minus 9 Yr Olds	17.1**	-3.4	22.5***		.450
	(2.28)	(.66)	(3.28)		13.2
IAEP-1991					
Science % Correct (U.S. GLE = 6)	4.3	1.7	9.6**		.436
	(1.72)	(.61)	(2.81)		4.0
Math % Correct (U.S. GLE = 8)	15.7***	3.7	16.1**		.641
	(3.85)	(.25)	(2.81)		6.0

Note: Numbers in parentheses are t values. GLE = grade level equivalent.
* p < .10 on a two-tailed test
** p < .05 on a two-tailed test
*** p < .01 on a two-tailed test

the standard model. Coefficients on this variable are positive for both outcomes and significantly so for science. The estimated impact of spending is modest, however. A one percentage point increase in the share of GDP devoted to education increases the science achievement of 13-year-olds by one-half of a grade-level equivalent.

The bottom row of each panel assesses the impact of CBEEES on measures of science and math learning between ages nine and 13.[3] Coefficients on the CBEEES dummy are positive for both math and science, but statistically significant only for mathematics. The exams are taken during upper secondary school or at the end of lower secondary school, so CBEEES may have larger effects on learning during secondary school than during primary school. This prediction is supported for math but not for science. For mathematics the coefficients suggest that about two-thirds of the effect of CBEEES on achievement at age 13 was generated in the previous four years. Since exams are also likely to affect learning during upper secondary school, total effects at the end of twelfth grade are likely to be larger still.

Analysis of the 1991 International Assessment of Educational Progress

The 1991 International Assessment of Educational Progress (IAEP) is the second data set in which CBEEE effects can be tested. Data on fifteen nations are available for the analysis: England, France, Hungary, Ireland, Israel, Emilia Romagna/ Northern Italy, Korea, Portugal, Scotland, Slovenia, Soviet Union, Spain, Switzerland, Taiwan and the United States.

The average percent correct (adjusted for guessing) for 13-year-old students was regressed on the same set of variables used in the analysis of the TIMSS data. The results are presented in the second panel of Table 1. For mathematics, the effect of curriculum-based external exams is highly significant and quite large. The U.S. standard deviation was 26.8 percentage points in mathematics, so the CBEEE effect on math was more than one-half of a U.S. standard deviation or about two U.S. grade-level equivalents. CBEEES had a smaller non-significant effect on science achievement. East Asian students scored significantly higher than students in Europe and North America. Coefficients on per capita GDP were positive but not statistically significant.

These results are consistent with the causal hypotheses presented above. Causation is not proved, however, because other explanations can be proposed. Other sources of variation in curriculum-based exams need to be analyzed. Best of all would be studies that hold national culture constant. Two such studies follow: one comparing Canadian provinces, the other comparing U.S. states.

Comparing Canadian Provinces

In 1990-91, the year the IAEP data was collected, Alberta, British Columbia, Newfoundland, Quebec and Francophone New Brunswick had curriculum-based provincial examinations in English, French, mathematics, biology, chemistry, and physics during the senior year of high school. These exams accounted for 50 percent of that year's final grade in Alberta, Newfoundland and Quebec and 40 percent in British Columbia. The other provinces did not have curriculum-based provincial external exit examinations in 1990-91. Ontario eliminated them in 1967, Manitoba in 1970 and Nova Scotia in 1972. Anglophone New Brunswick had provincial exams in language arts and mathematics but exam grades were not reported on transcripts or counted in final course grades. Canadian provincial exams are medium-stakes, not high-stakes tests. They influence grades but passing the examination is not essential

for graduation. Employers appear uninterested in exam scores. Job application forms do not request that applicants report exam scores or grades.

The principals of schools sampled by IAEP completed questionnaires describing school policies, school resources and the qualifications of eighth grade mathematics and science teachers. Students were asked about books in the home; number of siblings; language spoken at home; hours spent watching television, doing homework, pleasure reading, and watching science programs on television; parental oversight of school work; and teaching methods of teachers.

The effects of curriculum-based provincial exit exams taken by twelfth graders on achievement and the behavior of Canadian 13-year-olds, their parents, teachers and school administrators were examined by estimating models predicting these behaviors using schools as observations. The data set comprises 1,338 Canadian schools. The model contained 11 variables: logarithm of the mean number of books in the home; the mean number of siblings; the proportion of the school's students whose home language was different from the language of instruction; logarithm of the number of students per grade in the school; and dummies for schools run by a locally elected Catholic (or Protestant) school board, independent secular and non-secular schools, schools with primary grades, schools that include all grades in one building, French speaking schools, and a dummy for EXAM provinces.

Table 2 presents regression results predicting four achievement outcomes, 12 measures of school administrator behavior, nine teacher behaviors and 11 student/parent attitudes and behaviors. The first column presents the hypothesized sign of the relationship between CBEEES and that variable. The means and standard deviations across schools of each dependent variable are presented in columns two and three. The R^2 corrected for degrees of freedom is reported in column 14. The coefficient for EXAM and its *t* statistic are presented in columns four and five. Provincial exit exams had large effects on achievement: 19 percent of a U.S. standard deviation (about four-fifths of a U.S. grade-level equivalent) in mathematics and 13 percent of a standard deviation (about half of a grade-level equivalent) in science.

Effect of CBEEES on Behavior of Students, Teachers and Administrators: Exit exams also affected the behavior of parents, teachers and school administrators. Schools in exit-exam provinces scheduled significantly more hours of math and science instruction, assigned more homework, had better science labs, were significantly more likely to use specialist teachers for math and science, and more likely to hire math and science teachers who had studied the subject in college. Eighth grade teachers in exam provinces gave tests and quizzes more frequently. Hours in the school year, library books per student, computers per student, class size and teacher preparation time were not significantly affected by CBEEES.

Opponents of externally set curriculum-based examinations predict that they will cause students to avoid learning activities that do not enhance exam scores. This hypothesis was operationalized by testing whether exam systems were associated with **less** reading for pleasure and **less** watching of science programs like NOVA and Nature. Neither of these hypotheses is supported. Indeed, students in exam provinces spent significantly more time reading for pleasure, more time watching science programs on television, while watching significantly less television overall. Parents in these provinces were more likely to talk to their children about their math and science classes and their children were more likely to report that their parents "are interested in science" or "want me to do well in math."

CBEEES do not seem to skew teaching in undesirable ways. Students did more (not fewer) experiments in science class; emphasis on computation using whole numbers—a skill that should be learned by the end of fifth grade—declined significantly. Teachers subject to the subtle pressure of a provincial exam four years in the future apparently adopt strategies that are conventionally viewed as "best practice," not strategies to maximize scores on multiple-choice tests.

Students responded to the improved teaching by becoming more likely to report that science was "useful in everyday life." The data provided no support for our hypothesis that CBEEES would induce employers to pay greater attention to high school achievement. Students in exam provinces were not more likely to believe that math was important in getting a good job and were less likely to believe that science was important in job hunting.

A skeptic might point out that the correlation between EXAM and other outcomes may not be causal. Perhaps, the people of Alberta, British Columbia, Newfoundland, Quebec and Francophone New Brunswick—the provinces with exam systems—just place higher priority on education than the rest of the nation. This trait may also result in greater political support for examination systems. If so, we would expect that schools in the exam provinces should be better than schools in other provinces along other dimensions, such as discipline and absenteeism, not just by academic criteria. Bishop (1996) predicts, to the contrary, that exam systems induce students and schools to **redirect** resources and attention to learning/teaching exam subjects and away from the achievement of other goals (such as low absenteeism, good discipline and lots of computers). These competing hypotheses are evaluated in the 3rd, 4th, and 12th rows of Table 2. Contrary to the "provincial taste for education" hypothesis, principals in exam provinces had not purchased additional computers, did not report significantly fewer discipline problems, were significantly more likely to report absenteeism problems.

The Impact on New York State Regents Examinations

In the early 1990s, New York State was the only state with a CBEEE System. It has been administering curriculum-based Regents Examinations to high school students since June 1878. As Sherman Tinkelman, Assistant Commissioner for Examinations and Scholarships, described in a 1966 report:

> *The Regents examinations are closely related to the curriculum in New York State. They are, as you can see, inseparably intertwined. One supports and reinforces the other.... These instruments presuppose and define standards.... They are a strong supervisory and instructional tool—and deliberately so. They are effective in stimulating good teaching and good learning practices* (Tinkelman, 1966 p. 12).

The Regents examinations are taken throughout one's high school career. A student taking a full schedule of college preparatory Regents courses would typically take Regents exams in mathematics and earth science at the end of 9th grade; mathematics, biology and global studies exams at the end of 10th grade; mathematics, chemistry, American history, English and foreign language exams at the end of 11th grade and physics exams at the end of 12th grade.

In 1993, about 56 percent of ninth graders took the Mathematics Course 1 exam and, of these, 24 percent scored below the 65 percent passing grade. Similar proportions of tenth and eleventh graders took the global studies, biology and English exams. Failure rates were 20 percent in global studies, 18 percent in biology and 13

percent in English. Those not taking Regents exams were typically in considerably less challenging courses than Regents level courses. A system of minimum competency tests in specific subjects set a minimum standard for those not taking Regents courses but, as in other states, the passing standard was low.

Impact on SAT Test Scores

New York students are more disadvantaged, more heavily minority and more likely to be foreign-born than students in most other states. Among northern states, only Maryland, Delaware and Illinois have a larger share of African-American pupils. Nationally, only California has a higher share of its population foreign-born; only California, Texas, Arizona, New Mexico and Colorado have larger Hispanic population shares. Literacy levels among adults in New York are substantially below the national average (NEGP 1993, Vol 2).

Consequently, when one compares student achievement levels, family background must be taken into account. Considering the high incidence of at-risk children, New York students do remarkably well. The proportions of students taking algebra, calculus, chemistry and physics is generally above national averages. A larger proportion of New York's eleventh and twelfth graders are taking and passing (9.4 percent) AP exams in English, science, math or history than any other state except Utah (NEGP 1993, Vol. 2).

Graham and Husted's (1993) analysis of SAT test scores in the 37 states with reasonably large test-taking populations found that New York State students performed better than comparable students in other states. Graham and Husted did not, however, test the statistical significance of the New York State effect and used an unusual log-log specification.

Table 3 presents the results of a linear regression predicting 1991 mean SAT-M + SAT-V test scores for the 37 states for which data are available. With the exception of the dummy variable for New York State, variables are proportions—generally the share of the test-taking population with the characteristic described. New Yorkers do significantly better on the SAT than students of the same race and social background living in other states. When this model is estimated without the NYS dummy variable, New York has the largest positive residual in the sample. Wisconsin has the next largest positive residual which is 87 percent of New York's residual. Illinois and Nevada have positive residuals that are about 58 percent of New York's value. Arizona, California, Colorado, Florida, New Mexico, Ohio, Rhode Island, Texas and Washington have negative residuals greater than 10 points. Many of these states have large populations of Hispanics and recent immigrants, a trait that was not controlled for in the analysis. When one considers that Hispanics and immigrants are a large share of New York children, this makes New York's achievement all the more remarkable.

For individuals the summed SAT-V + SAT-M has a standard deviation of approximately 200 points. Consequently, the differential between New York State's SAT mean and the prediction for New York based on outcomes in the other 36 states is about 20 percent of a standard deviation or about three-quarters of a grade-level equivalent.

Adding the teacher-pupil ratio and spending-per-pupil to the model reduces the NYS coefficient by 25 percent, although it remains significantly greater than zero. The significant coefficient on teacher-pupil ratio suggests that heavy investment in K-12 schooling in New York State (possibly stimulated in part by the Regents exam system) may be one of the reasons why New York State students perform better than comparable students in other states.

Table 3
Determinants of Mean Total SAT-1 Scores for States

	NYS	Partic Rate	Parents AA-BA+	Private School	Prop. Black	Large School	3+Math Courses	3+Eng. Courses	*ln*Teach/ stud	*ln*Expend/ stud	R2 RMSE
	46**	-68**	370**	60	-135**	-44*	85	-36			.926
	(2.7)	(2.6)	(6.4)	(1.6)	(3.2)	(1.8)	(1.3)	(.3)			14.8
	35*	-88***	367***	69*	-113	-36	45	-45	48*	13	.933
	(2.0)	(3.3)	(6.6)	(1.9)	(2.6)	(1.5)	(.7)	(.4)	(1.7)	(.8)	14.2
Mean	.027	.414	.581	.207	.078	.129	.617	.797	-2.822	1.648	SAT-I
SD	.164	.240	.097	.082	.064	.113	.067	.038	.113	.215	925
											55

*** p < .01 on a two-tailed test
** p < .05 on a two-tailed test
* p < .10 on a two-tailed test

Impact on Mathematics Achievement of Eighth Graders

The New-York-is-exceptional hypothesis can also be tested by analyzing data from the 1992 administration of the NAEP mathematics assessment to representative samples of eighth grade students in 41 states and the District of Columbia. As with the analysis of SAT scores, state test score means were regressed on variables controlling for the socio-economic characteristics of the state's population and a dummy for New York State. The five variables that controlled for student background were: the proportion of people under age 18 living in poverty; a schooling index for the adult population; percent foreign-born; percent public school students who are black; and percent public school students who are Hispanic. The results are presented in Table 4. Parents' education, the poverty rate, percent black and percent foreign-born all had significant effects on math achievement in the expected direction. New York State's mean NAEP math score was a statistically significant 9.6 points (or about one grade-level equivalent) above the level predicted by the regression model.

One of the ways Regents exams may improve performance is by inducing the public to hire extra teachers to reduce class size and provide special help. Models were estimated with pupil-teacher ratios on the right-hand side. Point estimates of the effect of pupil-teacher ratios were negative but not signigicantly different from zero for 8th grade math test scores.

Impact on High School Dropout Rates

Table 4 also presents the results of cross-state regressions predicting school enrollment rates at age 17 and high school graduation rates. New York State's high school dropout rate is not significantly different from that of other states with students from similarly disadvantaged backgrounds. Additional staff appears to facilitate higher graduation rates. A 10 percent reduction in the pupil-teacher ratio increases the high school graduation rate by 1.5 percentage points.

Table 4
The Impact of Regents Examinations on Achievement and High School Graduation

	Math NAEP Mean Test Score for 8th Grade [4]			*Percent of 17 Year Olds Enrolled in High School* [2]			*Secondary School Graduates per 100 persons 17 Years Old* [2]		
New York State	9.59**	8.86*	7.27	.44	.55	.0046	.36	-1.71	-3.50
	(2.05)	(1.86)	(1.49)	(.36)	(.44)	(.00)	(.074)	(.36)	(.72)
Parents Education Index[1]	.68**	.65**	.85***	.091	.10	.15**	.96***	.85***	1.028***
	(2.71)	(2.57)	(2.89)	(1.58)	(1.63)	(2.28)	(4.09)	(3.80)	(4.12)
Percent in Poverty (People 18 years or less)[2]	-.52**	-.50**	-.42*	-.034	-.035	-.017	-.017	-.035	-.092
	(2.49)	(2.38)	(1.91)	(.69)	(.71)	(.35)	(.085)	(.19)	(.48)
Percent Foreign Born[3]	-.66***	-.63***	-.50**	-.18***	-.18***	-.13**	-.51**	-.42**	-.26
	(3.21)	(3.04)	(2.17)	(3.39)	(3.39)	(2.23)	(2.41)	(2.06)	(1.14)
Percent of Public School Students Black[4]	-.32***	-.33***	-.36***	-.047***	-.046***	-.045***	-.14***	-.17***	-.19***
	(6.06)	(6.10)	(6.21)	(3.59)	(3.43)	(3.90)	(2.73)	(3.27)	(3.62)
Percent of Public School Students Hispanic[4]	-.0092	-.0067	-.057	-.012	-.012	-.027	-.067	-.055	-.10
	(.10)	(.070)	(.55)	(.48)	(.50)	(1.07)	(.68)	(.59)	(1.05)
Pupil Teacher Ratio[4]		-.29	-.27		.037	.042		-.74**	-.73**
		(.38)	(.86)		(.47)	(.56)		(2.51)	(2.48)
Hours of Instruction per Year[2]			.030			.0098*			.032
			(1.29)			(1.76)			(1.49)
Adj R Squared	.8313	.8303	.8336	.5713	.5636	.5840	.5475	.5961	.6071
RMSE	4.232	4.244	4.203	1.111	1.121	1.095	4.510	4.262	4.203
N Observations	42	42	42	51	51	51	51	51	51

* Statistically significant at 10% level **Statistically significant at 5% level ***Statistically significant at 1% level

[1] Average of the percent of parents obtaining a secondary school diploma and the percent of parents obtaining a university degree. *Education in States and Nation*s. National Center for Education Statistics. U.S. Department of Education. 1991. Page 139.
[2] *Education in States and Nations*. National Center for Education Statistics. U.S. Department of Education. 1991. Pages 49, 73, 119, 129 and 149.
[3] 1990 Census of Population. Social and Economic Characteristics United States. Pages 174-179.
[4] *Digest of Education Statistics*. U.S. Department of Education. 1993. Pages 61, 76, and 120.

Does New York State Invest More in K-12 Education?

The theory predicts that the existence of CBEEES will induce New York State to spend more on K-12 education and to focus that spending on instruction. Indeed, New York's ratio of K-12 teacher salaries to college faculty salaries is significantly above average. New York teachers are also more likely to have masters degrees than the teachers of any state except Connecticut and Indiana. New York ranks seventh in both the teacher-pupil ratios and the ratio of per pupil spending to gross state product per capita (Bishop 1996).

Clearly, New York invests a great deal in its K-12 education system. If the cause of the high spending were a strong general commitment to education or legislative profligacy, we would expect spending to be high on both K-12 and higher education. This is not the case. New York is first in the ratio of K-12 spending per pupil to higher education spending per college student.

The Regents exams are currently low-to-medium stakes tests, not high stakes tests. Exam grades count for less than one-quarter of the final grade in the course and influence only the type of diploma received. Employers ignore exams results when making hiring decisions. During the 1980s, scholarships sponsored by the Regents were based on aptitude test scores, not Regents exam results. A passing score on Regents exams is not necessary for admission to community colleges or out-of-state colleges. Students were aware that they could avoid Regents courses and still go to college. Some perceived an advantage to avoiding them; as a student explained:

> *My counselor wanted me to take Regents history and I did for a while. But it was pretty hard and the teacher moved fast. I switched to the other history and I'm getting better grades. So my average will be better for college. Unless you are going to a college in the state, it doesn't really matter whether you get a Regent's diploma.* (Ward, 1994)

Indeed, the small payoff to taking Regents exams may be one reason why so many students have not been taking Regents courses.

This is about to change. The Board of Regents has announced that students graduating in the year 2000 must take and pass (at the 55 percent level) a new Regents English examination. The class of 2001 faces the additional requirement of passing an examination in algebra and geometry. The class of 2002 must pass separate Regents examinations in global studies and American history. The phase-in of all five new required Regents exams will be completed, when laboratory science exam courses come on stream, with the graduating class of 2003. Once the system has adjusted to the new exams, the Regents intend to raise passing scores from 55 percent to 60 percent and then to 65 percent.

Conclusions

Our review of the evidence suggests that the claims by advocates of standards based-reform that curriculum-based external exit examinations significantly increase student achievement are probably correct. Students from countries with such systems outperform students from other countries at a comparable level of economic development. Not only did students from Canadian provinces with such systems know more science and mathematics than students in other provinces, they watched less TV and talked with their parents more about school work. Schools in Canadian provinces with external exams were more likely to:

- employ specialist teachers of mathematics and science;

- hire math and science teachers who had studied the subject in college;
- have high-quality science laboratories;
- schedule extra hours of math and science instruction;
- assign more homework in math, in science and in other subjects;
- have students do or watch experiments in science class; and
- schedule frequent tests in math and science class.

When student demography is held constant, New York State, the only state having a CBEEE system in the early 1990s, does significantly better than other states on the SAT test and the NAEP math assessments without experiencing a reduction in high school graduation rates.

CBEEES, however, are not the only important determinant of achievement levels. General productivity levels and standards of living and an East Asian culture appear to have even larger effects. CBEEES are common in developing nations where achievement levels are often quite low (for example, Columbia and Iran). Belgium, by contrast, has a top quality education system without a CBEEES. More research on the system-level determinants of average achievement levels is in order.

References

American Federation of Teachers. (1995) Setting Strong Standards: AFT's criteria for judging the quality and usefulness of student achievement standards. Washington, D.C.: American Federation of Teachers, 1-12.

Beaton, Albert et al. (1996). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. CSTEEP, Boston College, Boston MA. (http://wwwcsteep.bc.edu/TIMSS)

Beaton, Albert et al. (1996). *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. CSTEEP, Boston College, Boston MA. (http://wwwcsteep.bc.edu/TIMSS)

Bishop, John. (1996). "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning." *International Journal of Education Research*.

Bishop, John. (1993). "The Impact of Academic Competencies on Wages, Unemployment and Job Performance." *Carnegie/Rochester Conference Series on Public Policy*, edited by Burton Malkiel and Charles Plosser, Vol. 37.

Bishop, John. (1989). "The Productivity Consequences of What is Learned in High School." *Journal of Curriculum Studies*.

Bob. (May 1996). Interview conducted in Calgary, Alberta.

Competitiveness Policy Council. (1993) Reports of the Subcouncils, March, Washington, D.C.: Competitiveness Policy Council.

Costrell, Robert. (1994). "A Simple Model of Educational Standards." *The American Economic Review*. Vol. 84, # 4, 956-971.

Fleming, M. and Chambers, B. (1983). *Teacher-made Tests: Windows on the Classroom*. San Francisco: Jossey Bass.

Graham, Amy and Husted, Thomas. (1993) "Understanding State Variation in SAT Scores." *Economics of Education*, Vol 12, No. 3, 197-202.

Gummere, Richard. (1943). "The Independent School and the Post War World." *Independent School Bulletin*. Vol 4, April, (quoted in Arthur Powell, *Standards*, 1995 Chapter 6).

Hart (Peter D.) Research Associates. (1995). "Valuable Views: A public opinion research report on the views of AFT teachers on professional issues." Washington D.C.: American Federation of Teachers, 1995, 1-24.

Hollenbeck, K., and Smith B. (1984). *The Influence of Applicants' Education and Skills on Employability Assessments by Employers*. Columbus: National Center for Research in Vocational Education, Ohio State University.

Interview with superintendent of an All-Regents high school (1997).

Kang, Suk. (1985). "A Formal Model of School Reward Systems," in *Incentives, Learning and Employability*, edited by John Bishop, Columbus Ohio: National Center for Research in Vocational Education.

Madeus, George. (1991). "The Effects of Important Tests on Students: Implications for a National Examination or System of Examinations." American Educational Research Association Invitational Conference on Accountability as a State Reform Instrument, Washington, D.C., 1-19.

Madeus, George and Kellaghan, Thomas. (1991). "Examination Systems in the European Community: Implications for a National Examination System in the United States." Contractor Report for the Office of Technology Assessment, U.S. Congress, Washington, D.C, 1-100.

National Center for Educational Statistics. (1993). *The Digest of Education Statistics: 1993*. Washington D.C.: U.S. Department of Education.

National Education Goals Panel. (1995). *Data for the National Education Goals Report: 1995*. Vol. 1, Washington, D.C.: U.S. Government Printing Office, 1-199.

Rohwer, W.D. and Thomas, J.W. (1987). "Domain specific knowledge, cognitive strategies, and impediments to educational reform." In M. Pressley (Ed.), *Cognitive Strategy Research*. New York: Springer-Verlag.

Steinberg, Laurence, Brown, Bradford and Dornbusch, Sanford. (1996). *Beyond the Classroom*. New York: Simon and Schuster, 1-223.

Thomas, John W. (1991). "Expectations and Effort: Course Demands, Students Study Practices and Academic Achievement." Paper presented at a Conference on Student Motivation sponsored by the Office of Educational Research and Improvement.

Tinkelman, Sherman N. (1966). "Regents Examinations in New York State after 100 Years." Albany, N.Y: The University of the State of New York, The State Education Department, 1-15.

Ward. (1994). "A Day in the Life." *N.Y. Teacher*. Albany, N.Y.

Appendix A—Bibliography

Sources Used to Classify National Education Systems

General

American Federation of Teachers. *What College-Bound Students Abroad are Expected to Know about Biology*. Washington, D.C.: AFT, 1994, 1-120.

American Federation of Teachers. *What Secondary Students Abroad are Expected to Know: Gateway Exams Taken by Average-Achieving Students in France, Germany and Scotland*. (Dawn Krusemark), Washington, D.C.: AFT, 1995, 1-165

American Federation of Teachers. *What College-Bound Students Abroad are Expected to Know about Chemistry and Physics*. (Matthew Gandal), Washington, D.C.: AFT, 1996, 1-145.

Beaton, Albert et al. (1996) *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. CSTEEP, Boston College, Boston MA. (http://wwwcsteep.bc.edu/TIMSS)

Beaton, Albert et al. (1996) *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. CSTEEP, Boston College, Boston MA. (http://wwwcsteep.bc.edu/TIMSS)

Britton, Edward and Raizen, Senta. *Examining the Examinations*. Boston: Kluwer Academic Publishers, 1996.

Eckstein, Max and Noah, Harold. *Secondary School Examinations*. New Haven: Yale University Press, 1993.

Eckstein, Max and Noah, Harold. (Eds.) *Examinations: Comparative and International Studies.* New York: Pergamon, 1992.

Elley, Warwick, *How in the World Do Students Read?*, The Hague, The Netherlands: International Association for the Evaluation of Educational Achievement, 1992.

Heyneman, Stephen P. and Fagerlind, Ingemar. *University Examinations and Standardized Testing*. Washington D.C.: World Bank, 1988.

International Assessment of Educational Progress, *Learning Mathematics*. Princeton, NJ: ETS, 1992b.

International Assessment of Educational Progress, *Learning About the World*. Princeton, NJ: ETS, 1992c.

International Assessment of Educational Progress, *IAEP Technical Report*. Volume 1, Princeton, NJ: ETS, 1992c.

Kreeft, Henk P. J. (ed.) (1990) "Issues in Public Examinations." Paper prepared for the International Association for Educational Assessment, 16th International conference on Issues in Public Examinations, Masstricht, The Netherlands June 18-22, 1990.

Lundberg, Ingvar and Linnakyla, Pirjo. *Teaching Reading around the World*. The Hague: International Association for the Evaluation of Educational Achievement, 1993.

Madeus, George F. and Kellaghan, Thomas. *Student Examination Systems in the European community: Lessons for the United States.* Contractor report submitted to the Office of Technology Assessment. June 1991

National Center for Educational Statistics. *Survey of examination practices.* (Designed by James Guthrie, Marilyn Binkley and Gary Phillips), Conducted for OECD, 1991

Organization of Economic Co-operation and Development. *Education and Training after Basic Schooling*. Paris, France: OECD, 1985.

Organization of Economic Co-operation and Development. *Education at a Glance*. Paris, France: OECD, 1995.

Phelps, Richard. *Are U.S. Students the Most Heavily Tested on Earth?* Washington, D.C.: American Institutes for Research, 1996, 1-26.

Postlethwaite, T. Neville and Wiley, David E. *Science Achievement in Twenty-Three Countries*. London: Pergamon Press, 1992.

Robitaille, David and Garden, Robert. *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics*, New York: Pergamon Press. 1989

U.S. Congress, Office of Technology Assessment. *Testing in American schools: Asking the right questions.* OTA-SET-519 Washington, D.C.: U.S. Government Printing Office, Feb. 1992, 1-301.

Specific Countries

Australia

Fendley, G. C. "Australia." In Lee Deighton (ed.), *The Encyclopedia of Education*, vol. 1, Macmillan, 1971, 427-431.

Austria

Brockman, Erin. *Education in Austria*. Paper for ILR 698, Cornell University, 1995, 1-19 .

Lahey, Helen. "Austria." In Lee Deighton (ed.), *The Encyclopedia of Education*, vol. 1, Macmillan, 1971, 431-437.

Donnell, Julie, "Austria." In George Kurian (ed.), *World Education Encyclopedia*, Facts on File Publications, 1988, 96-105.

Federal Press Service, *The Austrian Education System.* Vienna: Federal Chancellery, 1994, 1-46.

Belgium

Plancke, R. L., "Belgium," Lee Deighton (ed.), *The Encyclopedia of Education*, vol. 5, Macmillan, 1971, 449-455.

Fletcher, Ann. *Belgium*. Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers, 1985, 1-183.

Ministerie Van de Vlaamse Gemeenschap. *Educational Developments in Belgium, 1992-94: The Flemish Community.* Brussels: Centrum voor Informatie en Documentatie, 1994, 1-42.

Ministerie van de Vlaamse Gemeenschap. *Education in Belgium: The Diverging Paths.* Review of National Policies of Education) Bussels: Ministerie van de Vlaamse Gemeenshap, 1991, 1-463.

Vanberrgen, P. "Belgium." in Postlethwaite, T. N. (ed.), *The Encyclopedia of Comparative Education and National Systems of Education*, Oxford: Pergamon Press, 1988,

Brazil

Interview with Romualdo Protela de Oliveira, Professor at University of Sao Paulo.

Canada

Alberta Education. *Diploma Examinations Report, Annual Report, 1993-94.* Edmonton, Alberta, Student Evaluation Branch, 1994, 1-56.

Alberta Education. *Back to School Information Kit: 1995/96.* Edmonton, Alberta, Student Evaluation Branch, 1995, 1-48.

Calder, Peter. *Impact of Diploma Examinations on the Teaching—Learning Process.* Edmonton: University of Alberta, Dec. 1990, 1-61.

McEwen, Nelly "Student Academic Outcomes in Canada." In APEC Symposium on performance Measurement Systems, October 23-26. 1995, Washington, D.C.: Asia-Pacific Economic Cooperation Education Forum, Pelavin Research Institute, 207-215.

Manitoba Education and Training. *Educational Assessment in Manitoba.* Winnipeg: Curriculum Services-Assessment Unit, Sept 1993, 1-27.

Minister of Education and Training. *Renewing Education: New Directions, The Action Plan.* Winnipeg: Manitoba Education and Training Jan. 1995, 1-71.

Evaluation Branch, Department of Education. *1993-94 Grade 11 Provincial Examinations Report and Analysis: English and Mathematics.* Fredericton, New Brunswick, Nov. 1994, 1-44.

Saskatchewan Education, *Saskatchewan Provincial Learning Assessment Program*. Regina, Saskachewan: 1993, 1-49.

Saskatchewan Education, *High School Review Advisory Committee Final Report*. Regina, Saskachewan: 1994, 1-104.

Saskatchewan Education, Training and Employment. *Secondary Level Statistics: Jan-Aug 1993-94 Reporting Periods.* Regina, March 1995, 1-11.

Schweitzer, Thomas Robert Crocker and Geraldine Gilliss. *The State of Education in Canada.* Montreal: Institute for Research in Public Policy, 1995, 1-141.

Statistics Canada. *Education in Canada, 1995*. Ottawa: Statistics Canada, 1996, 1-198.

U.S. General Accounting Office, *Educational Testing: The Canadian Experience with Standards, Examinations and Assessments*. (written by Kathleen D. White) GAO/PEMD-93-11, Washington, D.C., April 1993, 1-74.

Columbia

Mora, J., "Columbia," in Postlethwaite, T. N. (ed.), *The Encyclopedia of Comparative Education and National Systems of Education*, Oxford: Pergamon Press, 1988, 202-205.

Cyprus

_____ "Cyprus," in George Kurian (ed.), *World Education Encyclopedia*, Facts on File Publications, 1988, 276-285.

Denmark

Interviews with qjvind Brogger at a FOLKESKOLE (Main School)) near Arhus, and with Johanus Andersen and Dorte Bollerup of Katedral Gymnasium and principal, teachers and students at Crhus Krbmandsskole (Business College).

Ministry of Education and Research. *Facts and Figures, Education Indicators for Denmark.* Copenhagen: Undervisningsministeriers forlag 1993, 1-83.

Organization of Economic Cooperation and Development. *Denmark.* Paris: OECD, 1980, 1-161.

Finland

Interviews with Rita Asplund at ETLA and with principals, teachers and students at three secondary schools.

Kivinen, Osmo and Rinne, Risto. *Educational Strategies in Finland in the 1990s* Turku, University of Turku, 1992. 1-134.

National Board of Education. *The Education System of Finland.* Helsinki: National Board of Education, 1994, 1-38.

National Board of Education. *Senior Secondary School in Finland.* Helsinki: National Board of Education, 1994, 1-15.

France

Interviews of principals and teachers at two Lycee and officials at the Ministry of Education.

Ministere de l'Education Nationale et de la Culture, *Reperes and References Statistiques sur les enseignmements et la formation.* 1992 Edition, (Paris: 1992a)

Ministere de l'Education Nationale et de la Culture, *Education in France.* 4th edition, Vanves, France: Direction Del=Evaluation et de la Prospective, Oct. 1994, 1-41.

Warner, Raymond. *France*. Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers, 1975, 1-249.

Lewis, H. D. *The French Education System*, New York, St. Martin's Press, 1985.

Germany

Interviews of Antonio Ruiz-Quintanilla, and Martin Behrens)

Behrens, Martin. *Structure, Standards and Diversity of the German System of Secondary Education*, paper for ILR698, Spring 1997, 1-6.

Lehman, R. H., "Germany: System of Education," in Husen T. & Postelthwaite, T.N. (eds), *The International Encyclopedia of Education*, Pergamon, 1994.

Rust, Val D. "West Germany," in George Kurian (ed.), *World Education Encyclopedia*, Facts on File Publications, 1988, 450-457.

World Education News and Reviews. [abstracting newsletter Faits Nouveaux], Vol 6, No. 2, Spring 1993, p. 6.

Greece

Massialas, Byron, "Greece." Lee Deighton (ed.), *The Encyclopedia of Education*, vol. 1, Macmillan, 1971, 190-194.

Massialas, Byron G., Flouris, G., and Cassatakis, Michael, "Greece," in George Kurian (ed.), *World Education Encyclopedia*, Facts on File Publications, 1988, 488-495.

Iceland

Interview with chief of the Cultural Section Embassy of Iceland.

Fraser, Stewart, "Iceland," in Lee Deighton (ed.), *The Encyclopedia of Education*, Macmillan, 1971, 534-536.

Josepsson, Bragi, "Iceland," George Kurian (ed.), *World Education Encyclopedia,* Facts on File Publications, 1988, 560-567.

India

Interviews with Sukdeep Brar and Sarosh Kruvilla.

Bordia, A., "India: System of Education," in Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd*. Pergammon, 1994.

Iran

Kurian, George, "Iran," George Kurian (ed.), *World Education Encyclopedia,* Facts on File Publications, 1988.

Afzal, Manuchehr, "Iran," in Lee Deighton (ed.), *The Encyclopedia of Education,* vol. 5, The Macmillan Co. and Free Press, 1971.

Italy

Burton, Maria Concetta. *The Historic Development of Italian Education.* Paper for ILR698. Dec 1995, 1-21.

Kurian, George. "Italy," George Kurian (ed.), *World Education Encyclopedia,* Facts on File Publications, 1988, 650-657.

Organization of Economic Cooperation and Development. *Educational Reform in Italy.* Paris: OECD, 1985, 1-111

Japan

Dore, Ronald and Mari Sako. *How the Japanese Learn to Work.*

Kogirima, Atusushi. *Japanese Education in Early 1960s.* Cornell University, Dec. 1995, 1-14.

White, Merry, *The Japanese Educational Challenge.*

Korea

Quan, Ducchi. *The Korean Educational System: an Overview.* Cornell University, Dec. 1995, 1-22.

Shin, Se-Ho. "Republic of Korea: System of Education," in Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd.* 1994, 3161-3170.

Malaysia

Faan, Hew Seok. *Education in Malaysia.* Cornell University, Dec 1989, 1-18

The Netherlands

Interviews of officials at the Ministry of Education and Principals, teachers and students at a VWO and a LBO near Gronigen.

Schuler, Peter. *The Netherlands*. Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers, 1984, 1-191

Nijhof, W.J. and Streumer, J.N. "The Netherlands." Postlethwaite, T. N. (ed.), *The Encyclopedia of Comparative Education and National Systems of Education,* Oxford: Pergamon Press 1988.

Ministry of Education and Science. *The Dutch Education System.* Docinform No. 332E, Zoetermeer: Ministry of Education and Science, 1988, 1-50.

Organization of Economic Cooperation and Development. *Netherlands.* Paris: OECD, 1991, 1-148.

Social and Cultural Planning Office. *Social and Cultural Report: 1988, The Netherlands.* Rijswijk: Social and Cultural Planning Office, 1988, 1-427.

Central Bureau Voor De Statistiek. "Education Statistics," The Hague, Netherlands, 1993.

New Zealand

Kallenbach, Dolores. *New Zealand.* Paper for ILR698, Dec 1995, 1-17,

McMahon, Tim. "Performance Measurement Systems in New Zealand." In APEC Symposium on performance Measurement Systems, October 23-26. 1995, Washington, D.C.: Asia-Pacific Economic Cooperation Education Forum, Pelavin Research Institute, 39-62.

Norway

Interviews with Tove Hammer and Johan____)

Kurian, George. "Norway" in George Kurian (ed.), *World Education Encyclopedia,* Facts on File Publications, 1988, 952-961.

Bjorndal, I. R. "Norway: System of Education." In Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd.* 1994, 4188-4192.

Ministry of Education and Research, *Education in Norway,* Oslo: Royal Norwegian Ministry of Education and Research, 1990, 1-36.

Ministry of Education and Research, *The Development of Education 1990-92: Norway,* prepared for International Conference on Education, 43 session Geneva 1992, Oslo: Royal Norwegian Ministry of Education and Research, Aug. 1992, 1-53.

Ministry of Education and Research, *The Upper Secondary School,* Oslo: Royal Norwegian Ministry of Education and Research, 1992, 1-21.

Nordic Statistical Secretariat. *Educational indicators in Nordic Countries.* Stockholm, Norstedts Tryckeri AB, 1991, 1-62.

Philippines

Interviews with Cornell graduate students from the Philippines: Pia Gavino, Carol Hau, Lorna Acebedo and Noel Yap.

Arcelo, Adriano, *Higher Education and the Labor Market in the Philippines.* New Delhi: Wiley Eastern Limited, 1981.

Bascani, Ramon C. "Performance Measurement Systems in the Republic of the Philippines." In APEC Symposium on performance Measurement Systems, October 23-26. 1995, Washington, D.C.: Asia-Pacific Economic Cooperation Education Forum, Pelavin Research Institute, 105-117.

Gonzalez, Hernando, Josefina Cortes, "The Philippines." in George Kurian (ed.), *World Education Encyclopedia,* Facts on File Publications, 1988, 992-1002.

Guevara, Elizabeth. *Educational System in the Philippines.* Paper for ILR698, 1995.1-23.

_____, "Philippines: System of Education," in Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd.* 1994, 4435-4442.

Sutaria, M. C. "Philippines," in Postlethwaite, T. N. (ed.), *The Encyclopedia of Comparative Education and National Systems of Education,* Oxford: Pergamon Press, 1988, 554-559.

Wilcox, Lee. (ed.) *The Evaluation of Asian Educational Credentials.* New York: National Asociateon for Foreign Student Affairs, 1996, 34-52.

Poland

Wesolowski, Katherine. *Education in Poland.* Cornell University, December 1995, 1-16.

Portugal

Planchard, Emle, "Portugal," in Lee Deighton (ed.), *The Encyclopedia of Education,* Macmillan, 1971, 181-185.

Ministry of Education. *The Portugese System of Education: Facts and trends.* Prepared for Educational and Cultural Spaces, Lisbon March 1990.

Rau. J. M. "Portugal: System of Education," in Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd.* 1994, 4622-4631.

Romania

Neag, Marie, "Romania," in George Kurian (ed.), *The World Education Encyclopedia,* Facts on File Publications, 1988, 1046-1051.

Russia

DeWitt, Nicholas, "Union of Soviet Socialist Republics," in Lee Deighton (ed.), *The Encyclopedia of Education,* Macmillan, 197, 295-299.

Slovenia

Pozarnik, B. Marentic, "Slovenia: System of Education," in Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd.* 1994, 5490-5497.

Plut-Pregelj, Leopoldina. "Slovenia's Education in the Process of Change: Legal Aspects." *Slovene Studies,* vol 13. No. 2, 1991, 129-141.

Plut-Pregelj, Leopoldina. "Changes in the Slovene Educational System, 1990-1992." *East/West Education.* Vol. 14, No. 1, Spring 1993, 51-65.

Spain

Interview with Ferran Mane, lecturer at the University of Barcelona

Borruso, Chris. *The Spanish Educational System.* Cornell University, Dec. 1989, 1-15.

Hoz, Victor Garcia, "Spain," in Lee Deighton (ed.), *Encyclopedia of Education,* Macmillan, 1971

Gil, G.A., "Spain: System of Education," in Husen, T. & Postelethwaite T.N. (eds) *The International Encyclopedia of Education, 2nd.* 1994, 5650-5658.

Turner, Solveig and Colbert, Ava Goicoa, "Spain." In George Kurian (ed.) *The World Education Encyclopedia,* Facts on File Publications, 1988, 1120-1129.

Sweden

Organization of Economic Cooperation and Development. *Educational Reform in Sweden.* Paris: OECD, 1981, 1-105.

Rehn, Gosta and Petersen, K. Helveg. *Education and Employment in Sweden and Denmark.* New York: Carnegie Council on Policy Studies in Education, 1980, 1-191.

Ministry of Education and Finance. *The Swedish Way Towards a Learning Society*. Report to OECD, Stockholm: Ministry of Education and Finance, 1992.

Switzerland

Federal Statistical Office. *The Swiss Educational Mosaic: A Study in Diversity.* Berne: Federal statistical Office, 1991, 1-71.

Gretler, A., "Switzerland," in Postlethwaite, T. N. (ed.), *The Encyclopedia of Comparative Education and National Systems of Education*, Oxford: Pergamon Press, 1988, 641-646.

Gretler, A., "Switzerland: System of Education." in Husen, T. & Postelethwaite T.N. (eds) *The International Encyclopedia of Education, 2nd.* 1994, 5876-5883.

Haffner, Walter. *The Educational Systems of Switzerland.* (Speech given on May 5, 1995 by head of the Cultural Section of the Swiss Embassy).

Susskind, Jacob. "Switzerland," *The World Education Encyclopedia,* Facts on File Publications, 1988, 1184-1191.

Organization of Economic Cooperation and Development. *Switzerland.* Paris: OECD, 1991, 1-213.

Taiwan

Fang, Lih-Ling. *The Examination System in Taiwan.* Cornell University, Fall 1995, 1-20.

Kuo, Su-Feng. *Education in Taiwan.* Cornell University, Fall 1993, 1-17.

Su, Jin-Li "Performance Measurement Systems for Elementary and Secondary Education in Taiwan: Current status and Concerns." In APEC Symposium on Performance Measurement Systems, Washington, D.C.: Asia-Pacific Economic Cooperation Education Forum, Pelavin Research Institute, October 23-26. 1995, 105-117.

Thailand

Chantanrrich, S., & G. W. Fry, "Thailand," in Postlethwaite, T. N. (ed.), *The Encyclopedia of Comparative Education and National Systems of Education*, Oxford: Pergamon Press, 1988, 658-662.

_____, "Thailand: System of Education." In Husen, T. & Postelethwaite T.N.(eds) *The International Encyclopedia of Education, 2nd.* 1994, 6374-6378.

_____. "Thailand," in *The World Education Encyclopedia*, Facts on File Publications, 1988

United States

American Federation of Teachers. (1995) *Setting Strong Standards: AFT's criteria for judging the quality and usefulness of student achievement standards*. Washington, D.C.: American Federation of Teachers, 1-12.

U.S. General Accounting Office, *Student Testing: Current Extent and Expenditures with cost estimates for a National Examination*. GAO/PEMD-93-8, Washington, D.C., Jan. 1993, 1-86.

End Notes

1. Appendix A provides a bibliography of the documents and individuals consulted when making these classifications. The TIMSS report's information about examination systems does not distinguish between university admissions exams and curriculum-based exit exams, so its classifications are not useful for this exercise. The Philippines, for example, is classified as having external exams by the TIMSS report, but its exams are university admissions exams similar to the SAT. South Africa was excluded because its education system was disrupted for many years by boycotts that were part of the campaign to end apartheid. Kuwait was excluded because of the disruption of its education system by the Iraqi invasion and the Gulf War.

2. The Phillipines, for example, had a math score mean of 399 in eighth grade and a mean of 386 in seventh grade. The mean age of eighth graders was 14 and the mean age of seventh graders was 12.9. The math score for 13.5 year olds was estimated by interpolation between seventh and eighth grade means. Math13.5 = 386 + (399-386)*((13.5-12.9)/(14-12.9)).

3. This indicator of learning between age nine and 13 can only be constructed for the 25 countries that participated in both the primary school and middle school TIMSS studies. The small size of the sample lowers the power of our statistical tests.