



Cornell University
ILR School

Cornell University ILR School
DigitalCommons@ILR

CAHRS Working Paper Series

Center for Advanced Human Resource Studies
(CAHRS)

1-1-1997

The Effect of Curriculum-Based Exit Exam Systems on Student Achievement

John H. Bishop
Cornell University

Follow this and additional works at: <http://digitalcommons.ilr.cornell.edu/cahrswp>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Thank you for downloading an article from DigitalCommons@ILR.

Support this valuable resource today!

This Article is brought to you for free and open access by the Center for Advanced Human Resource Studies (CAHRS) at DigitalCommons@ILR. It has been accepted for inclusion in CAHRS Working Paper Series by an authorized administrator of DigitalCommons@ILR. For more information, please contact hlmdigital@cornell.edu.

The Effect of Curriculum-Based Exit Exam Systems on Student Achievement

Abstract

[Excerpt] Two presidents, the National Governors Association and numerous blue ribbon panels have called for the development of state or national 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)." 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? 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?

Keywords

national, standard, student, school, teach, achievement, exam, CBEEES

Disciplines

Educational Assessment, Evaluation, and Research

Comments

Suggested Citation

Bishop, J. H. (1997). *The effect of curriculum-based exit exam systems on student achievement* (CAHRS Working Paper #97-15). Ithaca, NY: Cornell University, School of Industrial and Labor Relations, Center for Advanced Human Resource Studies.

<http://digitalcommons.ilr.cornell.edu/cahrswp/156>

WORKING PAPER SERIES

The Effect of Curriculum-Based Exit Exam Systems on Student Achievement

John H. Bishop

Working Paper 97 - 15



The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement

John H. Bishop
School of Industrial and Labor Relations
Center for Advanced Human Resource Studies
Cornell University
Ithaca, NY 14853-3901

Working Paper #: 97 - 15

This paper is to appear in the *Journal of Economic Education*. Its preparation was made possible by support from Project B-4b of the Consortium for Policy Research in Education, a center funded by the Office of Education Research and Improvement, U.S. Department of Education. The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

This paper has not undergone formal review or approval of the faculty of the ILR School. It is intended to make results of research, conferences, and projects available to others interested in human resource management in preliminary form to encourage discussion and suggestions.

<http://www.ilr.cornell.edu/cahrs>

The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement

Two presidents, the National Governors Association and numerous blue ribbon panels have called for the development of state or national 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)."* 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? 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?

1. What is a Curriculum-based External Exit Examination System?

Let us begin by defining what a curriculum-based external exit examination system is. It has the following traits. It:

1. Produces signals of student accomplishment that have real consequences for the student.
2. Defines 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 now possible. Costrell's (1994a) 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 (i.e. teacher grading or schools graduation requirements).
3. Is organized by discipline and keyed to the content of specific course sequences. This focuses responsibility for preparing the student for particular exams on one (or a small group of) teacher/s.
4. Signals multiple levels of achievement in the subject. If only a pass-fail signal is generated by an exam, 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 1994a).
5. Covers 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 which subjects to be examined in and offer high and intermediate level exams in the same subject.

Commercially prepared achievement tests such as the CAT, CTBS, ITBS, ITED are not curriculum-based external exit exams, CBEEE, because they fail requirement # 1. Students have no stake in doing well on these tests. Where stakes attached to student performance, it is teachers and school administrators who experience the consequences, not individual students.

The minimum competency exams that many American states require students pass to graduate from secondary school are not CBEEEs because they fail requirements # 3 and # 4. These tests focus on basic skills taught in primary school and lower secondary school. While minimum competency exams have apparently reduced the numbers of students with very low basic skills levels (Lerner 1991), the passing standard is quite low. The tests are typically first taken in 9th or 10th 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. Thus, for the great majority of students who pass them on the first try, 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 much more demanding curriculum-based exams signaling much higher levels of performance are essential.

The requirement (# 4) 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. By age 13 students differ dramatically in their levels of achievement. On the National Assessment of Educational Progress, 7-9 percent of 13 year olds are four or more grade level equivalents behind their age mates and 15-17 percent are four or more grade level equivalents ahead of their age mates. When achievement differentials among students are as large as this, incentives for effort are stronger for most students if the full range of achievement is signaled rather than just whether the individual has passed some absolute standard. When only a pass-fail signal is generated by a test, many students pass without exertion and are, thus, 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 Wang 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 GDP, but also social welfare (1994, p. 970)."

The SAT-I reasoning tests are not curriculum based external exit exams because they do not meet tests # 3 and # 5. They are not organized around school subjects and 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. The SAT was designed from the beginning to minimize backwash effects on teaching and student study habits. Indeed, Richard Gummere, Harvard College's admissions director at the time 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 meet some of the requirements of a CBEEE. However, since colleges admit on the basis of the SAT-I, not SAT-IIs, the stakes are low and few students take them. In 1982-83 only 6 percent of SAT-I test takers took a science SAT-II and only 3 to 4 percent took one in history or a foreign language. Schools do not assume responsibility for preparing students for SAT-II tests.

The Advanced Placement (AP) examinations are the one exception to the generalization that the U.S. lacks national curriculum-based external exit examinations. While 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 # 5 and are, consequently, not a CBEEE system. They can, however, serve as a component of a larger system.

2. Why are CBEEES Hypothesized to Increase Achievement?

Curriculum-based external exit exams (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 admission and hiring decisions, so the rewards for learning

should grow and become more visible. CBEEEs also shift attention towards measures of absolute achievement and away from measures of relative achievement such as rank in class and teacher grades.

Grading on a curve or basing college admissions on class rank give students a personal interest in persuading each other not to study. The studious are called nerds or accused of "acting white," in part, because they are making it more difficult for others to get top grades. When exams are graded on a curve, joint welfare is maximized when no one studies. In the repeated game that results, side payments—friendship and respect—and punishments—ridicule and harassment—enforce the cooperative "don't study" solution. When, by contrast, learning is assessed relative to an outside standard, students no longer have a personal interest in getting teachers off track or persuading each other to refrain from studying.

In the absence of a CBEEES, school reputations are largely outside the control of school staff; determined instead by the social class of the student body, mean SAT scores and by numbers attending prestigious colleges. When a CBEEES is in place, exam results displace social class as the primary determinant of school reputations and this in turn should induce school staff to give enhanced learning higher priority. Teachers will upgrade curricula and assign more homework and parents will demand better science labs and more rigorous teaching. School administrators will be pressured to increase the time devoted to examination subjects and hire more qualified teachers. The competition for teachers is likely to bid up salaries.

A further benefit of CBEEES is the professional development that teachers receive when they are brought to centralized locations to grade the extended answer portions of the examinations. In May 1996 I interviewed a number of Alberta teachers about their experience serving on grading committees. They all said that having to discuss and agree with their colleagues about what constituted an excellent, good, adequate, poor, and failing response to essay questions had been the best professional development experience they had since entering the profession.

3. Do CBEEES Increase Achievement? A look at the Evidence.

The hypothesis that curriculum-based external exit examination systems 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 in the 40 nation Third international Math and Science Study, science and math scores on the

International Assessment of Educational Progress (IAEP) for 16 nations and 9 Canadian provinces and SAT test scores for New York State vs. the rest of the United States.

3.1 Third international Mathematics and Science Study–TIMSS:

The just released TIMSS provides 1994-95 data for 7th and 8th graders for 40 countries. Comparative education studies and education encyclopedias were reviewed and embassy personnel interviewed to determine which of the TIMSS nations have curriculum-based externally-set exit examinations in secondary school. Twenty-one national school systems were classified as having CBEEEs for both subjects 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 Thailand. 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. A list of the interviews conducted and books consulted is available from CAHRS. The countries classified as not having a CBEEE in either subject were Belgium (both Flemish and French speaking systems), Cyprus, Greece, Latvia, Philippines, Portugal, Spain and Sweden. Following Madeus and Kelleghan (1991), the university entrance examinations in Greece, Portugal Spain, 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 the job market 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.

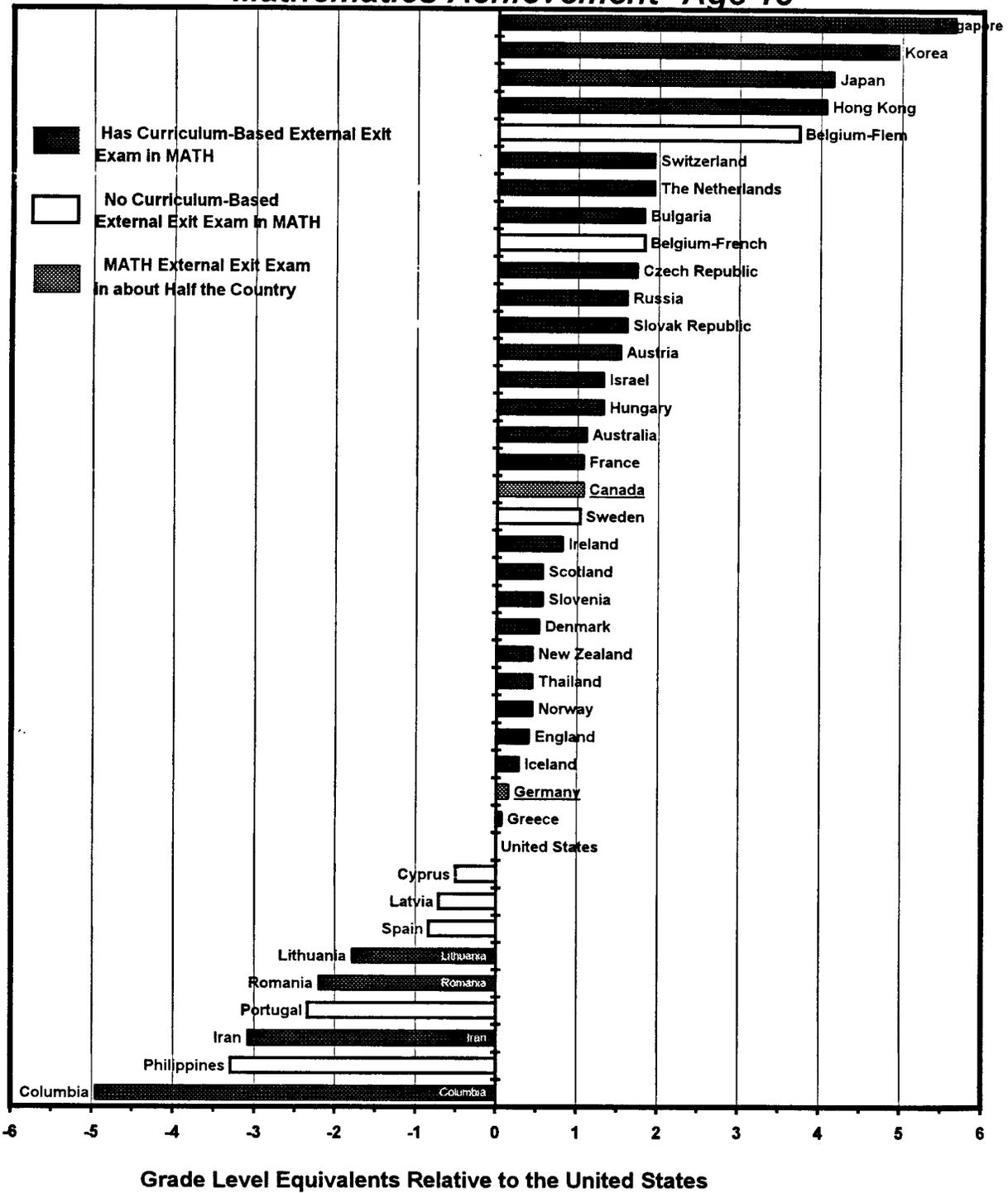
Figures 1 and 2 array the 40 TIMSS countries by the science and mathematics achievement of their 13 year olds. The US ranks # 15 in science and # 31 in mathematics. The gaps between the vertical grid lines represents one U.S grade level equivalent—the difference between 7th and 8th grade TIMSS test score means for the U.S. Achievement differentials across nations are very large. In science Singapore, Korea, Bulgaria and Flemish Belgium are more than 1 GLE ahead of the US and Columbia, Philippines, Lithuania, Romania and Portugal are more than 3 GLEs behind. In mathematics Singapore, Korea, Japan and Hong Kong are 4 or more grade level equivalents ahead of the U.S., while Columbia, Philippines and Iran are more than 3 GLEs behind. 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. Note that the countries with a CBEEE in the subject tend to have higher TIMSS scores.

Regression Analysis: The mean 7th and 8th grade science and mathematics test scores were regressed on per capita gross domestic product for 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 CBEEE.

The analysis of achievement at a particular grade level may be biased, however, by differing policies regarding grade retention, age of school entry and which grade was 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 fourth and fifth 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 7th and 8th grade means. Switching to the age constant achievement reduces the estimated impact of the CBEEES. Using two tail *t* tests, the CBEEES coefficient has a $P = .11$ on a 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 .9 U.S. grade level equivalents in mathematics.

Mathematics Achievement--Age 13



Science Achievement at Age 13

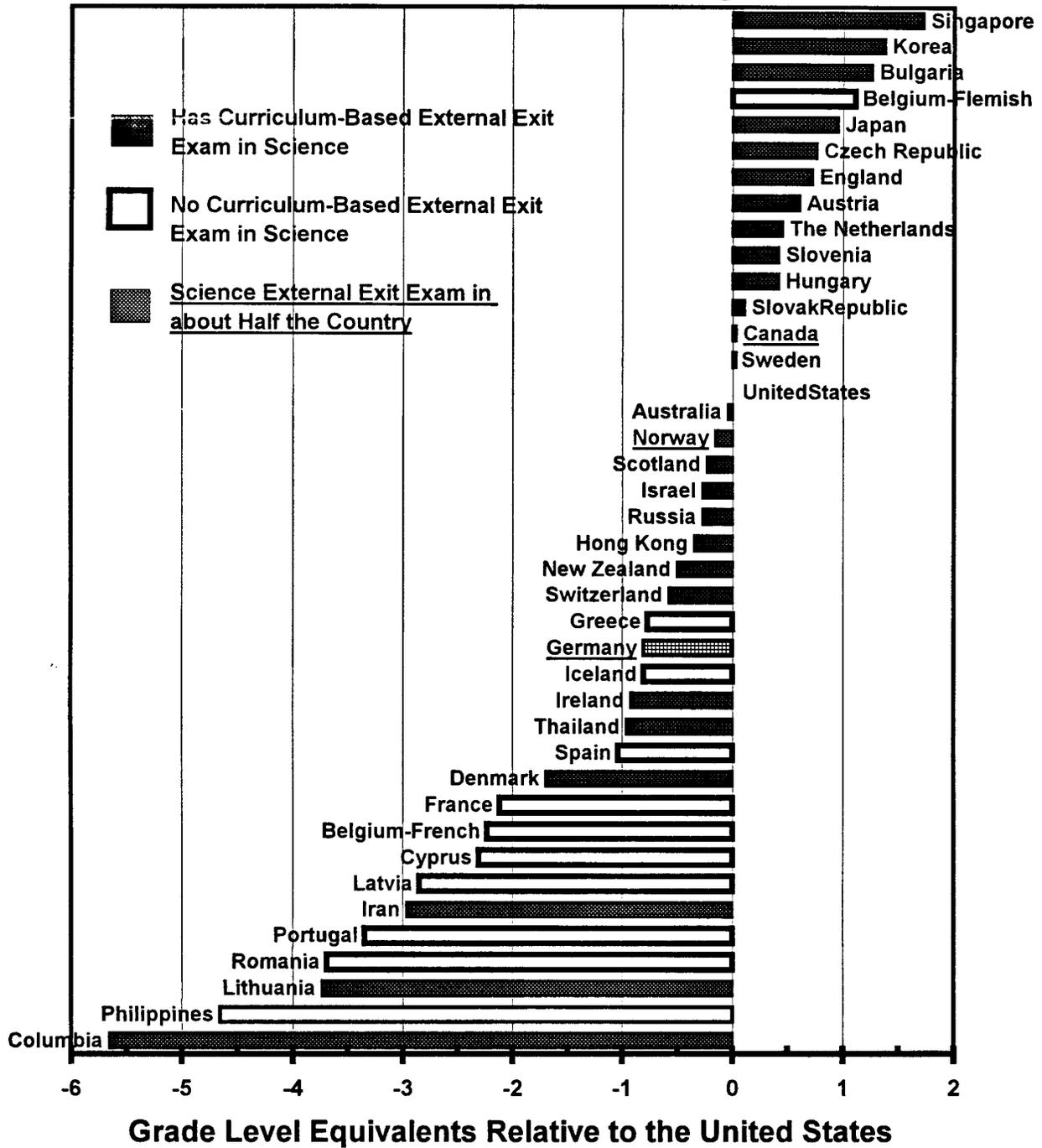


Table 1

The Effect of Curriculum-Based External Exams
on Science and Mathematics Achievement

(Analysis of 1994-95 TIMSS data)

	External Exit Exam	LnGDP/Pop 1987 & 90	East Asia	\$K- 12 1GDP	AdjR2/ RMSE
<u>TIMSS Science-1994</u> (U.S. GLE = 26)					
Mean for 7 th Graders	37.3*** (2.95)	45.2*** (3.33)	19.6 (1.19)		.320 35.4
Mean for 8 th Graders	40.8*** (3.33)	34.9*** (3.65)	13.8 (.87)		.357 34.6
Mean for 8 th Graders	38.7*** (3.08)	30.7*** (2.93)	21.5 (1.25)	8.3 (1.10)	.360 34.9
Median for 13 Yr Olds	34.3*** (2.79)	44.0*** (4.58)	20.8 (1.31)		.403 34.7
Median for 13 Yr Olds	31.2*** (2.56)	37.1*** (3.63)	32.9* (1.95)	13.5* (1.83)	.441 34.0
<u>TIMSS Mathematics-1994</u> (U.S. GLE = 24)					
Mean for 7 th Graders	26.1* (1.89)	45.2*** (4.47)	65.7*** (3.95)		.458 36.6
Mean for 8 th Graders	32.1** (2.31)	47.6*** (4.65)	61.7*** (3.67)		.462 37.1
Mean for 8 th Graders	29.9** (2.08)	46.7*** (4.11)	65.6*** (3.55)	3.1 (.38)	.456 37.7
Median for 13 Yr Olds	21.8 (1.64)	55.3*** (5.65)	69.2*** (4.32)		.529 35.4
Median for 13 Yr Olds	18.5 (1.37)	53.1*** (4.98)	75.8*** (4.36)	5.8 (.76)	.534 35.6
<u>IAEP-1991</u>					
Science % Correct (U.S. GLE = 6)	4.3 (1.72)	1.7 (.61)	9.6** (2.81)		.436 4.0
Math % Correct (U.S. GLE = 8)	15.7*** (3.85)	3.7 (.25)	16.1 ** (2.81)		.641 6.0

3.2 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. 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 as 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. Since the U.S. standard deviation was 26.8 percentage points in mathematics, the CBEEE effect on math was more than one-half of a U.S. standard deviation or about 2 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 no doubt 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 will be presented: one comparing Canadian provinces, the other comparing U.S. states.

3.3 Comparing Canadian Provinces

In 1990-91, the year the IAEP data was being 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 8th grade mathematics and science teachers.

Students were asked about books in the home, number of siblings, language spoken at home, hours of TV, hours doing homework, pleasure reading, watching science programs on TV, parental oversight of school work and teaching methods of teachers.

The effects of curriculum-based provincial exit exams taken by 12th 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 1354 Canadian and 106 U.S. schools. The model contained 9 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 religiously controlled school, secular non-public schools, French speaking schools, a dummy for U.S. school and EXAM province.

Table 2 presents regression results predicting 4 achievement outcomes, 12 measures of school administrator behavior, 9 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 2 and 3. The coefficient for EXAM and its T statistic are presented in columns 4 and 5. The R^2 corrected for degrees of freedom is reported in column 14. Provincial exit exams had very large effects on achievement: 24 percent of a U.S. standard deviation (about four-fifths of a U.S. grade level equivalent) in mathematics and 17.6 percent of a standard deviation (about three-fifths of a grade level equivalent) in science.

Table 2-LONG--Effects of Curriculum-Based External Exams in Canada

	Hyp.	Mean	StdDev	Curric Exam	U.S.	French	Relig	Indep	LnNumb	LnBook	Avg Num.	Different	Adj.
			Schl	Coefstat		Speakg	School	School	inGrade	in Home	Siblings	Hm. Lang.	R2
Achievement													
Mathematics	+	.464	.135	.061(9.4)	-.023*	.074***	-.045***	.074***	.006*	.156***	-.025***	.017*	.381
Science	+	.541	.096	.035(7.1)	.020**	.017**	-.039***	-.016	-.006**	.119***	-.023***	-.062***	.353
Discipline Prob	0/+	.765	.720	-.009 (.2)	.034	.18***	-.159***	-.236**	.117***	-.299	.067	-.308***	.090
Absenteeism Prob	0/+	.822	.766	.179(4.1)	.070	-.19***	-.068	-.288**	.153***	-.411***	.165***	-.004	.122
School Administrator Behavior													
Math Specialist Tchrs	+	.48	.50	.23(8.8)	.47***	.07*	-.201***	.057	.128***	.105***	-.041	-.056	.279
Science Specialist Tchrs	+	.49	.50	.19(7.2)	.35***	-.05	-.134***	.056	.167***	.152***	-.028	.068	.284
Took Math Courses in Univ	+	.66	.39	.19(8.3)	.23***	-.07**	-.127***	.011	-.007	.075**	.048**	-.060	.121
Took Science Courses in Univ	+	.70	.38	.22(9.9)	.15***	-.22***	-.182***	.037	.005	.064**	.013	.133***	.183
Math Class Hours	+	3.97	.89	.37(7.0)	.19*	.26***	-.028	-.095	-.088***	-.282***	-.104*	-.749***	.102
Science Class Hours	+	3.00	.82	.23(4.9)	.97***	-.11	-.415***	-.268**	.009	.009	.080	-.103	.171
Library Books per Student	?	21	21	2.48(1.9)	5.36**	8.00***	-.182	7.04**	-6.99***	3.73**	1.15	5.19*	.120
Computers per Student	?	.052	.043	.006(2.5)	.008*	-.009*	-.013***	-.006	-.012***	.004	.004	.016**	.086
Specialized Science Labs	+	1.95	.95	.43(8.0)	.06	.000	-.251***	.256**	.216***	.110	.008	.032	.158
Hours in School Year	+	949	89	1.0 (0.2)	42.6***	-10.0	-16.2***	3.7	-5.4	5.3	5.2	45.5***	.029
Class Size	-	24.8	6.1	-.38(1.2)	-2.3***	-.27	3.3***	4.4***	3.6***	.42	-.45	-2.27***	.347
Teacher Preparation Time	+	.31	.27	.01(1.5)	-.01	-.03**	.000	.080***	.063***	-.012	-.025**	-.042**	.192
Teacher Behavior													
Total Homework--Hrs/wk	+	4.41	1.62	.66(7.1)	1.33***	-.33	.821***	1.90***	.131***	.110	-.299***	1.024***	.168
Math Homework--Hrs/wk	+	1.66	.64	.20(5.0)	.18*	-.02	.165***	.219**	-.015	.115**	-.131***	.346***	.051
Science Homework--Hrs/wk	+	1.04	.47	.19(6.3)	.12***	-.06	.125***	.051	.016	.091**	-.007	.211**	.054
Emph Whole Number Compute	-	1.68	.49	-.11(3.7)	-.09	.10**	-.009	-.149**	-.010	-.038	-.010	.029	.026
Math Quiz Index	+	1.62	.52	.12(4.5)	.37***	.67***	-.077***	.173**	.107***	-.040	-.006	-.113*	.394
Science Quiz Index	+	.89	.38	.11(5.3)	.66***	.32***	-.089***	.026	.024**	-.044**	.032	.147***	.336
Science Do Experiments Ind.	+	1.52	.63	.26(7.3)	-.18***	.35***	.137***	-.088	.059***	-.019	-.028	.138	.165
Science Watch Experiments	+	2.42	.47	.15(5.3)	-.12**	.21***	.086**	-.012	.042***	-.100***	-.006	-.024	.116
Science Watch Films Index	-	.94	.48	-.05(1.7)	.31***	-.05*	-.054	.004	.061***	.002	.038	.001	.070
Home Behavior & Attitudes													
TV-Sch. Avg.-Hrs/wk	-	14.7	2.85	-.73(5.1)	.31	-2.0***	.50***	-2.53***	-3.48**	-.23	-.85**	-.276	.276
Read for Fun Index	?	1.85	.28	.04(2.8)	-.09***	.09***	.003	-.006	.265***	.033*	.230***	.143	.143
Watch Science programs on TV?		.97	.38	.06(2.7)	.05	.24***	.071***	.028	-.094***	-.032	-.178***	.113	.113
Parent Talk about Math Class	+	.62	.17	.05(5.1)	.08***	.03*	.030***	.042*	.028**	-.029***	-.029	.043	.043
P. Talk about Science Class	+	.47	.17	.06(6.4)	.06***	-.02	.004	.046*	.053***	-.007	.032	.050	.050
P. want me do well in Math	+	3.54	.22	.06(4.9)	.05**	-.03	.088***	.120***	.035*	-.077***	-.058***	.084	.084
P. interest in Science(0-4)	+	2.18	.34	.07(3.6)	.02	.08***	.088***	.017	.179***	-.060***	.073	.065	.065
Math Imp. to get Job(0-4)	+	3.57	.21	-.01(.7)	-.02	-.05**	.106***	.049	.019	-.043***	-.099***	.054	.054
Science Imp. to get Job(0-4)	+	2.93	.33	-.05(2.5)	-.13***	-.20***	.183**	-.019	.039	-.015	-.125***	.126	.126
Math Useful Solving Prob.	+	3.03	.31	.01(0.6)	-.05	.21***	.108***	.037	.103***	-.043**	-.146***	.095	.095
Sci. Useful Everyday Life	+	2.46	.31	.06(2.8)	-.02	.17***	.141***	.075	-.140***	.013	-.179***	.114	.114

Source: Regressions predicting the characteristics of 1366 to 1460 Canadian and American secondary schools. Provinces with external exams included in final course grade were Alberta, British Columbia, Newfoundland, Quebec and the Francophone schools in New Brunswick. Mean school char. based on n gt 8.

Effect of CBEEES on Behavior of Students, Teachers and Administrators: Exit exams also apparently affected the behavior of parents, teachers and school administrators. 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." Schools in exit exam provinces scheduled 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, class size and teacher prep 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 TV, while watching significantly less TV overall.

Attitudes Toward Science

Do CBEEEs skew teaching in undesirable ways? Madeus has pointed out 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 permit students to do well in examinations without recourse to higher levels of cognitive activity (1991 p. 7-8)."* Contrary to this hypothesis, however, students did more (not fewer) experiments in science class and emphasis on computation using whole numbers—a skill that should be learned by the end of 5th grade—declined significantly. Apparently, teachers subject to the subtle pressure of a provincial exam four years in the future adopt strategies that are conventionally viewed as "best practice," not strategies designed 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.

One possible skeptical response to these findings is to point out that the correlation between EXAM and other outcomes may not be causal. Maybe the people of Alberta, British Columbia, Newfoundland, Quebec and Francophone New Brunswick—the provinces with exam systems—place higher priority on education than the rest of the nation. Maybe this trait also results 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 and good discipline. These competing hypotheses are evaluated in the 3rd and 4th row of Table 2. Contrary to the "provincial taste for education" hypothesis, principals in exam provinces did not report significantly fewer discipline problems and were significantly more likely to report absenteeism problems.

Effectiveness of Non-Public Schools: In their influential 1990 book, John Chubb and Terry Moe argued that the constraints placed on public schools by bureaucracy and democratic government make them inherently less effective than non-public schools that must compete for students and that are, thus, required to survive a market test. Clearly, however, their theory does not apply to Canada. The students at the small number of secular nonpublic schools did better than public school students in mathematics but not in science. However, most of Canada's nonpublic schools were started by religious denominations. When background characteristics of the students were controlled, students at religiously controlled schools knew considerably less math and science at age 13 than public school students. This occurred despite that fact that the students at parochial schools were more likely to believe that math and science was useful in everyday life or important in getting a good job and had parents who were more interested in their school performance. The difference appears to result from the quality and amount of instruction students at parochial schools received. Students got 13 percent less science instruction time and worked in laboratories that were clearly inferior to those in public middle schools. Their teachers were less likely to be specializing in teaching the subject and less likely to have studied the subject in college.

3.4 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 ever 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 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, English and foreign language exams at the end of 11th grade and mathematics and physics exams at the end of 12th grade.

In 1993, about 56 percent of 9th graders took the Mathematics Course 1 exam and, of these, 24 percent scored below the 65% passing grade. Similar proportions of 10th and 11th 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 courses that were considerably less challenging 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. Of the 50 states, New York was the only one with such a system in the early 1990s.

Impacts on Achievement

New York's 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. Nationwide, only California has a higher share of its population foreign born and only California, Texas, Arizona, New Mexico and Colorado have larger Hispanic population shares. Literacy levels among adults 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 proportion of students taking algebra, calculus, chemistry and physics is generally above national averages. A larger proportion of New York's 11th and 12th graders are taking and passing (9.4 percent) AP exams in English, science, math or history than any other state except Utah (NGEP 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 did better than comparable students in other states. They 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, all right hand side variables are proportions—generally the share of the test taking population with the characteristic described. Clearly, 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. The next largest (Wisconsin's) positive residual 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. This makes New York's achievement all the more remarkable when one considers that Hispanics and immigrants are a large share of its school children.

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. It remains significantly greater than zero, however. 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 student perform better than comparable students in other states.

Table 3

Determinants of Mean Total SAT-I Scores for States

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

*** p < .01 on a two tail test

** p < .05 on a two tail test

* p < .10 on a two tail test

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 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 master's degrees than the teachers of any state except Connecticut and Indiana. New York ranks number 7 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 number 1 in the ratio of K-12 spending per pupil to higher education spending per college student.

The Regents exams are currently "medium" stakes tests, not "high" stakes tests. Exam grades count for less than a third of the final grade in the course and influence only the type of diploma received. Employers ignore exam results when they make 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. Indeed some perceived an advantage to avoiding them:

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 state, it doesn't really matter whether you get a Regent's diploma. (Ward, 1994)

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

Conclusions

Our review of the evidence suggests that the claims of the advocates of standards and examination-based reform of American secondary education may be right. The countries and Canadian provinces with such systems outperform other countries at a comparable level of development. New York State, the only state with a CBEEES, does remarkably well on the SAT test when student demography is held constant. CBEEES are, however, probably not the most important determinant of achievement levels. CBEEES are common in developing nations where achievement levels are often quite low [e.g. Columbia and Iran]. Belgium, by contrast, has a top quality education system without having a CBEEES. More research on the effects of CBEEES is clearly in order.

REFERENCES

- Association of Secondary Teachers of Ireland, Information Sheet opposing changes in Examination Systems, 1990.
- Beaton, Albert et al. (1996) Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study. CSTEED, Boston College, Boston MA. <http://www.csteep.bc.edu/TIMSS>.
- Beaton, Albert et al. (1996) Science Achievement in the Middle School Years: IEA's-Third International Mathematics and Science Study. CSTEED, Boston College, Boston MA. <http://www.csteep.bc.edu/TL4SS>.
- Bishop, John. (1996) "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning." International Journal of Education Research.
- Chubb, John and Moe, Terry. Politics and Markets and America's Schools. Washington, D.C.: The Brookings Institution, 1990.
- Costrell, Robert. (1994) "A Simple Model of Educational Standards." The American Economic Review. Vol. 84, # 4, Sept., 956-971.
- Graham, Amy and Husted, Thomas. "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 Chap 6).
- International Assessment of Educational Progress, IAEP Technical Report. Volume 1, Princeton, NJ: ETS, 1992c.
- Kang, Suk. "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, 1985.
- Lerner, Barbara. "Good News about American Education." Commentary, Vol. 91, No. 3, March 1990, 19-25.
- Madeus, George. "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., June 1991, 1-19.
- Madeus, George and Kellaghan, Thomas. "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, March 27, 1991, 1100.
- National Center for Educational Statistics. The Digest of Education Statistics: 1992. Wash. D.C.: US Department of Education, 1992.

National Education Goals Panel. Data for the National Education Goals Report: 1995. Vol. 1, Washington, DC: U.S. Government Printing Office, 1995, 1-199.

Steinberg, Laurence. Beyond the Classroom. New York: Simon and Schuster, 1996, 1-223.

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.

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.

Ward. "A Day in the Life." N.Y. Teacher. Jan. 1994, Albany, N.Y.

**Previous CAHRS Discussion Papers
on Similar Topics by John Bishop**

- "Is An Oversupply of College Graduates Coming?" (97-04)
- "The Effect of National Standard and Curriculum-Based Exams on Achievement" (97-01)
- "Do Some Employers Share the Costs and Benefits of General Training?" (96-19)
- "Nerd Harassment, Incentives, School Priorities and Learning" (96-10)
- "What We Know About Employer-Provided Training: A Review of the Literature" (96-09)
- "The Effect of Curriculum-Based Examinations on School Priorities and Student Learning."
(95-30)
- "Is the College Graduate Labor Market Headed for a Bust?—a Critique of Job Requirements
Projections"
- "Improving Education: How Large are the Benefits? How can it be done Efficiently?" (95-21)
- "Expertise and Excellence." (95-13)
- "The Effect of Curriculum-Based Examinations on Learning in Canadian Secondary Schools."
(94-30)
- "Signaling, Incentives and School Organization in France, the Netherlands, Britain and the
United States: Lessons for Education Economics." (94-25)
- "Signaling the Competencies of High School Students to Employers" (94-18)
- "The Incidence and Payoff to Employer Training: A Review of the Literature with
Recommendations for Policy" (94-17)
- "Impacts of School Organization and Signaling on Incentives to Learn in France, the
Netherlands, England, Scotland and the United States." (93-21)
- "Incentives to Study and the Organization of Secondary Instruction" (93-08)
- "Technical Education and Education Reform" (93-07)
- "Overeducation" (93-06)
- "The French Mandate to Spend: Is It Appropriate for the United States?" (93-05)
- "Underinvestment in Employer Training: Is a Mandate to Spend on Training the Answer?"
(93-02[E])
- "Improving Job-Worker Matching in the US Labor Market: What is the Role of the Employment
Service? (92-40)

- "Is a Skills Shortage Coming? A Review of BLS Occupational Projections" (92-04)
- "Workforce Preparedness" (92-03)
- "The Economic Consequences of Schooling and Learning" (92-22)
- "Employer Training and Skill Shortages: A Review of the State of Knowledge with Recommendations for Future Research by the Labor Department, (91-32)
- "The Impact of Previous Training in Schools and on Jobs on Productivity, Required OJT and Turnover of New Hires." (91-27)
- "Does the Targeted Jobs Tax Credit Create Jobs at Subsidized Firms" (91-25)
- "A Strategy for Achieving Excellence in Secondary Education: The Role of State Government," (91-24)
- "Department of Labor Testing: Seizing an Opportunity to Increase the Competitiveness of American Industry and to Raise the Earnings of American Workers" (91-14)
- "Academic Learning and National Productivity" (91-07)
- "Underinvestment in On-the-Job Training?" (91-03)
- "The Explosion of Female College Attendance" (90-19)
- "What's Wrong with American High Schools? Can State Governments Fix It?" (90-17)
- "The Worsening Shortage of College Graduate Workers" (with Shani Carter) (90-15)
- "The Deskilling vs. Upskilling Debate: The Role of BLS Projections" (with Shani Carter) (90-14)
- "The Skills Shortage and the Payoff to Vocational Education" (90-08)
- "The Federal Role in Education Reform" (90-05)
- "Achievement, Test Scores and Relative Wages" (89-22)
- "Why are Wage Profiles So Flat During the First Year on a Job?" (with Suk Kang) (89-18)
- "Why Students Don't Study. How You Can Make Studying Pay Off For Them" (89-15)
- "Scientific Illiteracy: Causes, Costs and Cures" (89-12)
- "On-the-Job Training of New Hires" (89-10)
- "Incentives to Study: Why American High School Students Compare so Poorly to their Counterparts Overseas" (89-09)
- "Occupational Competency as a Predictor of Labor Market Performance" (88-20)

- "The Productivity Consequences of What is Learned in High School" (88-18)
- "Productivity Growth and Tenure: A Test of the On-the-Job Training Theories of Wage and Productivity Growth" (with Stan Stephenson) (88-15)
- "The Economics of Employment Testing" (88-14)
- "The Motivation Problem in American High Schools" (88-13)
- "Employment Testing and Incentives to Learn" (88-12)
- "Vocational Education for At-Risk Youth: How Can It Be Made More Effective" (88-11)
- "Vocational and Academic Education in High School: Complements or Substitutes" (with Suk Kang) (88-10)
- "Occupational Training in High School: When Does It Work?" (88-09)
- "Do Employers Share the Costs and Benefits of General Training?" (88-08)
- "A Signaling/Bonding Model of Employer Finance of General Training" (with Suk Kang). (88-07)
- "Applying for Entitlements: Employers and the Targeted Jobs Tax Credit" (with Suk Kang) (88-04)
- "Match Quality, Turnover and Wage Growth" (88-03)
- "Why High School Students Learn So Little and What Can Be Done About It" (88-01)
- "Toward More Valid Evaluations of Federal Training Programs" (87-07)
- "Information Externalities and the Social Payoff to Academic Achievement" (87-06)
- "Is the Test Score Decline Responsible for the Productivity Growth Decline" (87-05)

To a order a copy of one of the papers, please contact the Center for Advanced Human Resource Studies at (607) 255-9358, 187 Ives Hall, Cornell University, Ithaca, NY, 14853-3901. Ask for the paper by title and paper number, which is listed in parentheses.