CAHRS
Center for Advanced Human Resource Studies

CAHRS / Cornell University
187 Ives Hall
Ithaca, NY 14853-3901 USA
Tel. 607 255-9358
www.ilr.cornell.edu/CAHRS/

WORKING PAPER SERIES

The Effect of National Standard and Curriculum-Based Exams on Achievement

John H. Bishop

Working Paper 97-01

School of Industrial and Labor Relations
CORNELL UNIVERSITY
Advancing the World of Work

The Effect of National Standard and Curriculum-Based Exams on Achievement

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

Working Paper #97-01

This paper was presented at the American Economic Association meetings, New Orleans, Louisiana, January 4, 1997. It's preparation was made possible by support from Project B-4b of the Consortium for Policy Research in Education (CPRE), Grant No. OERI-R308A60003), a center funded by the Office of Educational 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, the U.S. Department of Education, or the institutional partners of CPRE.

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

The Effect of National Standards and Curriculum-Based Exams on 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?

Let us begin by defining what is meant by a system of curriculum-based external exit examinations (CBEEE). A CBEEE 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, b) 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 (ie. 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, b).
5. **Covers almost all secondary school students.** Exams for a set of elite schools or for those specializing in a particular field will influence standards in that segment but may have limited effects on the bulk of students. A single exam taken by all is not essential. Some nations offer high and intermediate level exams in the same subject.

Curriculum-based external exit exams (CBEEEs) improve the signalling 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 gives students a personal interest in persuading each other not to study. The studious are called nerds, 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 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 CBEEE 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; school administrators will hire more qualified teachers and increase the time devoted to examination subjects; parents will demand better science labs and more rigorous teaching.

Do CBEEEs Increase Achievement of 13 year olds? A look at the Evidence.

The hypothesis that curriculum-based external exit examination systems improve achievement will be tested by comparing nations (and provinces) that do and do not have such systems.

Third International Mathematics and Science Study--TIMSS: The just released TIMSS provides 1994-95 data for 7th and 8th graders for 39 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. Nineteen national school systems were classified as having CBEEEs for both subjects in all parts of the country: 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. Four countries--France, Iceland, Norway 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. The countries classified as not having a CBEEE in either subject were Belgium, Cyprus, Greece, Latvia, Philippines, Portugal, Spain and Sweden. Following Madeus and Kelleghan (1991), the university entrance examinations in Greece, Portugal Spain and Cyprus were not considered to be CBEEEs. University entrance exams should have much smaller effects because students headed into the job market do not take them and teachers can avoid responsibility for their student's success. Disappointing results can be blamed on shortages of university places or the recondite standards of the exam.

Policies regarding age of entry into school and grade retention vary across countries, so comparisons must hold student age constant, not grade in school. Consequently, the dependent variable for this analysis is the median test score for the 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 mean for 7th graders. The median math and science 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 CBEEE. 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 impact of a CBEEE is about one U.S. grade level equivalent in mathematics and 1.2 grade level equivalents in science.

International Assessment of Educational Progress-1991: 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.

Canada--IAEP Data: 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 had no curriculum-based provincial 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.

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 backwash effects of curriculum-based provincial exit exams taken by 12th graders on 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 model contained 8 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, USA and EXAM province.

The estimated impacts of exit exams are presented in Table 2. Each row represents a separate regression on data from 1366 to 1460 schools. The means and standard deviations across schools of each dependent variable are presented in columns 1 and 2. The coefficient for EXAM and it's T statistic are presented in columns 3 and 4. Complete results are reported in Bishop (1996). 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.

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 "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.

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.

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 permits 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.

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 EXIT EXAM provinces did not report significantly fewer discipline problems and were significantly more likely to report absenteeism problems.

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 comparable levels of development. In addition, New York State, the only state with a CBEEE, does remarkably well on the SAT test when student demography is held constant (Bishop 1996). CBEEEs are not,

however, the most important determinant of achievement levels. CBEEEs are common in developing nations where achievement levels are often quite low [eg. Columbia and Iran]. Belgium, by contrast, has a top quality education system without having a CBEEE. More research on the effects of CBEEEs is clearly in order.

REFERENCES

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://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. CSTEEP, Boston College, Boston MA. http://www.csteep.bc.edu/TIMSS.

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

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

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

Costrell, Robert. (1994b) "Centralized vs Decentralized Educational Standards Under Pooling." July 1994, Dept of Economics, U. of Massachusetts, Amherst, Mass., 1-35.

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.

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 F. and Kellaghan, Thomas. (1991) "Student Examination Systems in the European Community." Report Submitted to the Office of Technology Assessment, 1-40.

Table 1
The Effect of Curriculum-Based External Exams
on Mathematics and Science Achievement of 13 Year Olds
(Analysis of TIMSS and IAEP data)

	External Exit Exam	LnGDP/Pop 1987 & 90	East Asia	AdjR2/ RMSE
Third International Math and Science Study-1994/95				
Mathematics-level	**23.0***	**55.6*****	**68.8*****	**.534**
(U.S. GLE = 24)	**(1.74)**	**(5.61)**	**(4.25)**	**35.6**
Science-level	**34.3*****	**44.0*****	**20.4**	**.398**
(U.S. GLE = 26)	**(2.77)**	**(4.56)**	**(1.27)**	**35.0**
International Assessment of Educational Progress-1991				
Mathematics % Correct	**15.7*****	**3.7**	**16.1****	**.641**
(U.S. stdev=	**(3.85)**	**(.25)**	**(2.81)**	**6.0**
Science % Correct	**4.3**	**1.7**	**9.6****	**.436**
(U.S. stdev=19.9)	**(1.72)**	**(.61)**	**(2.81)**	**4.0**

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

	Dep Var Mean	Std Dev	School Coef	Exit Exams Tstat
Achievement				
Mathematics	.464	.135	.061	(9.4)
Science	.541	.096	.035	(7.1)
Discipline Problems	.765	.720	-.009	(.2)
Absenteeism Problems	.822	.766	.179	(4.1)
School Administrator Behavior				
Math Specialist Teachers	.48	.50	.23	(8.8)
Science Specialist Teachers	.49	.50	.19	(7.2)
Math Teacher Studied Math in Univ.	.66	.39	.19	(8.3)
Sci. Teacher Studied Science in Univ.	.70	.38	.22	(9.9)
Math Class Hours	3.97	.89	.37	(7.0)
Sci. Class Hours	3.00	.82	.23	(4.9)
Good Science Labs	1.95	.95	.43	(8.0)
Teacher Behavior				
Total Homework Hrs/wk	4.41	1.62	.66	(7.1)
Emphasizes Whole Number Computation	1.68	.49	-.11	(3.7)
Student Do Experiments	1.52	.63	.26	(7.3)
Home Behavior & Attitudes				
TV Hours/wk	14.7	2.85	-.73	(5.1)
Read for Fun Index	1.85	.28	.04	(2.8)
Watch NOVA, Nature	.97	.38	.06	(2.7)
Parents Talk about Math Class	.62	.17	.05	(5.1)
Parents Talk about Science Class	.47	.17	.06	(6.4)
Parents want me to do well in Math	3.54	.22	.06	(4.9)

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

	Hyp.	Mean	StdDev Schl	Curric Exam Coef	Curric Exam Tstat	U.S.	French Speakg	Relig School	Indep School	LnNumb inGrade	LnBook in Home	Avg.Num. Siblings	Different Adj. 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
Principal repts Discipline Prob	0/+	.765	.720	-.009	(.2)	.034	.18***	-.159***	-.236**	.117***	-.299	.067	-.308***	.090
Principal repts Absenteeism Prob	0/+	.822	.766	.179	(4.1)	.070	-.19***	-.068	-.288**	.153***	-.411***	.165***	-.004	.122
School Administrator Behavior														
Math Specialist Teachers	+	.48	.50	.23	(8.8)	.47***	.07*	-.201***	.057	.128***	.105***	-.041	-.056	.279
Science Specialist Teachers	+	.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***	-1.82	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 Computation	-	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
Read for Fun Index	?	1.85	.28	.04	(2.8)	-.09***	.09***	.003	-.006		.265***	.033*	.230***	.143
Watch Science programs on TV	?	.97	.38	.06	(2.7)	.05	.24***	.071***	.028		-.094***	-.032	-.178***	.113
Parent Talk about Math Class	+	.62	.17	.05	(5.1)	.08***	.03*	.030***	.042*		.028**	-.029***	-.029	.043
P. Talk about Science Class	+	.47	.17	.06	(6.4)	.06***	-.02	.004	.046*		.053***	-.007	.032	.050
P. want me to do well in Math	+	3.54	.22	.06	(4.9)	.05**	-.03	.088***	.120***		.035*	-.077***	-.058***	.084
P. interested in Science(0-4)	+	2.18	.34	.07	(3.6)	.02	.08***	.088***	.017		.179***	-.060***	.073	.065
Math Important to get Job(0-4)	+	3.57	.21	-.01	(.7)	-.02	-.05**	.106***	.049		.019	-.043***	-.099***	.054
Sci. Important to get Job(0-4)	+	2.93	.33	-.05	(2.5)	-.13***	-.20***	.183**	-.019		.039	-.015	-.125***	.126
Math Useful Solving Everyday Pb.	+	3.03	.31	.01	(0.6)	-.05	.21***	.108***	.037		.103***	-.043**	-.146***	.095
Sci. Useful Everyday Life(0-4)	+	2.46	.31	.06	(2.8)	-.02	.17***	.141***	.075		-.140***	.013	-.179***	.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.