Research Notes/Notes de recherche

Comparative Analysis of Ontario Senior Division Mathematics Achievement 1968–82

Merlin Wahlstrom
Dennis Raphael
Les McLean
the ontario institute for studies in education[1]

During the school year 1981–82, Ontario joined 23 other nations participating in the Second International Study of Mathematics (SIMS). Although the prime focus of SIMS was a cross-national study of educational curriculum, teacher instructional approach, and student achievement in mathematics, an important aspect of the Ontario study was a comparative analysis of students' mathematics achievement from 1968 to 1982. This comparison was accomplished through the administration, as part of SIMS, of the same instrument, the Ontario Mathematics Achievement Test (OMAT), completed by students as part of their college and university entrance examination in May 1968 and by a sample of Grade 13 students in May 1976 studied in the Interface Study (Traub, Wolfe, Wolfe, Evans, & Russell, 1976). The issue addressed in the present report is the stability of Ontario students' achievement levels over the 1968–82 time period.

Declining Achievement in the United States and England

Few Canadian students since 1967, the year during which most provinces terminated end-of-schooling examinations, have written standardized post-secondary admission examinations. The scarcity of these assessments and concern about maintaining educational standards led some to claim that while expenditure on education has increased, the educational system has not provided value for money, and standards have been falling. This was the message of the *Black Papers*, a publication series in England by Cox and Dyson (1977).

Due to a scarcity of available Canadian data on secondary school achievement levels, the context for the comparative study of student achievement over the last 15 years must be largely developed from U.S. research evidence. In the United States, a consistent downward trend in the Scholastic Aptitude Test (SAT) over the 1965–80 period has been a cause of some concern.[2] Ebel (1976) summarized the situation:

> The decline in the scores of applications for college admission on the scholastic Aptitude Test of the College Entrance Examination Board has been well advertised. In the 1962–63 academic year, the mean score on the verbal portion of the test was 478. In 1974–75, 12 years later, the mean verbal score was 437. This is a drop of 41 score points. Applicant scores on the mathematical part of the test showed a similar though somewhat less rapid decline. The mean mathematics score was 502 in 1962–63. In 1974–75 it was 473, a drop of 29 score units. A rather similar decline is apparent in scores of college bound students on the American College Testing Program. (p.306)

In a monograph entitled *Achievement Test Score Decline: Do We Need to Worry?* Harnischfeger and Wiley concluded:

> Beyond doubt, beyond differences among assessment instruments, alterations in tests, or in pupil composition, achievement test scores have been declining for about a decade in all grades from five upwards. Score declines are more pronounced in higher grades and in recent years. They are more severe for tests probing verbal than mathematics achievements. Those are the facts and they describe a National phenomenon. (1976, pp. 115–116).

More recently, conclusions of the U.S. government-sponsored study, *A Nation at Risk: The Report of the National Commission on Excellence in Education*, summarized in a recent discussion (Goldberg & Harvey, 1983), suggested that the apparent failure of American schools to perform their traditional roles adequately may have disastrous consequences for the U.S. Among the data cited by the Commission as evidence for the view that mediocrity, not excellence, is the norm in education were findings that from 1963 to 1980 a virtually unbroken decline took place on average scores on the SAT and that a similar dramatic decline took place in the number of students who demonstrated superior achievement on the SAT. Can these conclusions be applied to aspects of Canadian education?

Since descriptions of the design and general findings of the Ontario SIMS study have been published elsewhere (McLean, Raphael, & Wahlstrom, 1983; Raphael, Wahlstrom, & McLean, 1983), the present report focuses upon the administration of the achievement measures providing data bearing upon the stability of student achievement over the past 15 years.

THE ONTARIO COMPARATIVE STUDY

The present analysis examines student achievement on the Ontario Mathematics Achievement Test (OMAT) from 1968 to 1982. Descriptions of the OMAT and the three student samples of 1968, 1976, and 1982 are followed by study findings.

The Ontario Mathematics Achievement Test (OMAT)

The OMAT was first administered as Part I, Form OB, in May, 1968 as part of the admissions procedure to Ontario Colleges and Universities. The 36 multiple choice items were constructed by Ontario Grade 13 mathematics teachers and members of the Ontario Mathematics Achievement Test Committee. Items were reviewed by the Committee and revised and selected for try-out in the fall of 1967 and administration as Form OB in May 1968.

Ontario Mathematics Curricula 1968–82

The structure of Ontario Senior Division Mathematics was modified in 1972. Prior to 1972, two courses entitled Math A and Math B were offered. These two courses were then expanded to three: Relations and Functions, Calculus, and Algebra. While for the most part course content remained unchanged, it must be kept in mind that the OMAT was designed to measure achievement in the Math A course.

Analysis of the 1976 Interface Study attempted to ascertain the validity of items for the new curriculum. The test appeared to provide a comprehensive coverage of Relations and Functions, and parts of Calculus, but did not sample differential equations, polar coordinators, and complex numbers. There was no coverage of Algebra topics. In 1976, 74 Calculus and 98 Relations and Functions teachers rated each item as to whether it measured knowledge prerequisite to the course, new knowledge that some or all students would learn in the course, or new knowledge that no student would learn. The combined ratings of Relations and Functions, and Calculus teachers found no items falling into the last category.

A similar procedure was carried out in the SIMS study and provided evidence for the validity of the OMAT in 1982. Sixty-four teachers of Grade 12 advanced-level mathematics, 69 of Relations and Functions, and 63 Calculus teachers indicated whether they had covered the mathematics required to answer the OMAT items. The mathematics required to answer each of the 36 items of the OMAT was covered in at least one of the three courses.

Students Completing the OMAT

The May 1968 administration of OMAT involved a volunteer sample of some 24,232 Ontario students seeking admission to university who had taken the Mathematics A course. The 1976 Interface Study sampled 1,404 students from 52 Ontario secondary schools enrolled in both Relations and Functions, and Calculus.

Approximately one out of nine of the SIMS sample of 5,687 students in classes of Grade 12 Advanced level, Grade 13 Relations and Functions, Grade 13 Calculus, and Grade 13 Algebra from each of 85 secondary schools completed the OMAT at the end of their courses.[3] The scores of the 329 students enrolled in Relations and Functions and Calculus were computed for the comparative analysis.

The main difference in study samples appears to be the more select nature of the 1968 group. The samples of the latter two studies, while certainly containing most university-bound students, did not involve self-selection of a volunteer group attempting to gain admittance to university.

Calculation of Difficulty Indices

Calculation of difficulty levels for items in the three administrations differed slightly. The 1968 results were based upon the proportion of students choosing the correct response to each item. The 1976 scores were computed through a two-step process: first, school item means were calculated, then the mean of school means was computed. The 1982 scores were calculated in a similar manner to that for the 1968 study. There seems to be little likelihood that these differences would have much practical effect on item scores.

RESULTS OF THE COMPARATIVE ANALYSIS

Reliability of the OMAT

The reliability of the 1968 administration of the OMAT, using an adaptation of the Kuder-Richardson formula 20 of internal consistency, was 0.78. Hoyt estimate of internal consistency reliability for the 1976 administration was 0.79 and for 1982, 0.81.

Item Difficulties and Total Scores

Item difficulties from the 1968, 1976, and 1982 administrations were examined. A striking similarity between the 36 items' difficulties for the

three sets of data was seen. The correlation coefficients of item difficulties were: 1968 and 1976, $r = .84$; 1968 and 1982, $r = .82$; and 1976 and 1982, $r = .97$

The average difference in item difficulty between 1968 and 1976 was 2.1%, and between 1968 and 1982, 2.5%. Between 1976 and 1982, the change was a negligible 0.4%. The average score on the entire OMAT (36 items) in 1968 was 19.01, the 1976 Interface Study, 18.25 and the 1982 SIMS study, 18.11. An analysis of variance of the item difficulties for the three administrations yielded a clearly nonsignificant finding ($F(2,105) = 0.13$, ns).

Although the sample in 1968 consisted of self-selected college- and university-bound students, the results are remarkably consistent from 1968 to 1982. A consistent standard of student mathematics achievement from 1968 to 1982 has apparently been maintained.

DISCUSSION

Although the mathematics curriculum was modified between the 1968 administration of the OMAT and the 1976 and 1982 studies, and slightly differing samples of students were studied, the levels of mathematics achievement remained remarkably constant over the time period. Conceivably, deletion of inappropriate items from the OMAT and their replacement by items reflecting the slight 1976–82 curriculum changes may have revealed some improvement in achievement levels in an absolute sense. Clearly, the problem of declining achievement seen in the United States was not apparent in the present analysis of student achievement over the past 15 years in Ontario.

NOTES

[1] The research was funded by the Ministry of Education, Ontario. The views expressed are those of the authors and not necessarily those of the Ministry.

[2] The SAT is a widely used university admission aptitude test consisting of verbal and quantitative components.

[3] In semestered courses, testing took place for some in December 1981 and for others in May 1982. Full-year courses were tested in May.

REFERENCES

Cox, C., & Dyson, A. (1977). *Black Paper V*. London: Temple Smith.

Goldberg, M., & Harvey, J. (1983, September). A nation at risk: The report of the National Commission on excellence in education. *Phi Delta Kappan, 65*(1), 14-18.

Ebel. R. (1976, December). Declining scores: A conservative explanation. *Phi Delta Kappan, 58*(4), 306-310.

Harnischfeger, A., & Wiley, D. (1976). *Achievement test score decline: Do we need to worry?* St. Louis, MO: Cemrel.

McLean, L., Raphael, D., & Wahlstrom, M. (1983, October). The Second International Study of Mathematics: An overview of the Ontario Grade 8 study. *Orbit, 67*, 4–7.

Raphael, D., Wahlstrom, M., & McLean, L. (1983, December). The Second International Study of Mathematics: An overview of the Grade 12/13 study. *Orbit, 68*, 7–10.

Traub, R., Wolfe, R., Wolfe, C., Evans P., & Russell, H. (1976). *Secondary-postsecondary interface project: Nature of students* (Vol. 2). Toronto: Ministry of Education and the Ministry of Colleges and Universities.