Comparing Apples to Apples

There is a way to find out which schools are doing the best job.

By David Johnson

Tests measure what students have learned. In elementary schools, students are assessed by their classroom teacher on their performance in reading, writing, and math. These assessments identify the students' areas of strength and weakness.

Everyone ─ parents, students, and teachers alike ─ takes such assessments for granted. Students take home written summaries, in the form of report cards, several times a year. In many households, the arrival of the report card is a significant event.

Teacher-written assessments of their own students are, then, an accepted part of the education system. Much more controversial, however, is the concept of standardized testing, where almost all students in the province work on the same material, which is then marked outside the local school to a uniform standard.

Beginning in the 1996-97 academic year, Ontario introduced just such a standardized assessment of all its elementary students. The testing is carried out in May, graded over the summer, and returned to schools, students, and parents in the fall.

Opinions vary about the usefulness of these external assessments. The controversy really heats up, however, when the external test results are averaged across students at each school, made available to the public, then used as a measure of the school's effectiveness.

Is such a school report card useful? Are schools that contain students with higher average test results actually better schools? If better student test results are *not* an appropriate measure of a school's sues, what is? Is the Ontario province-wide elementary assessment program useful?

One important group in the elementary educational process, Ontario's elementary school teachers as represented by their federations ─ the Elementary Teachers' Federation of Ontario (EFTO) and the Ontario English Catholic Teachers' Association (OECTA) ─ is among those whose position is to oppose standardized testing in its current form.

Teachers dislike the publication of results averaged over students at individual schools, since such results are often used, implicitly or explicitly, to rank schools.

EFTO argues that "[t]he best way to judge schools is by visiting them and looking for evidence of learning and interest in learning", and claims that "average test scores are *more* indicative of the pooled characteristics of the students, such as socio-economic status".

In fact, my research found that *less* than half the variation in Ontario schools' results can be associated with the pooled socio-economic characteristics of their students. So a large part of the explanation for the variation must be, at least to some extent, the outcome of choices principals, teachers, boards, and parents have made for their schools.

Everyone involved in the education system knows that students vary greatly according to the social and economic background of the households in whch they live. Most people would agree that direct comparisons of the test results of schools with such disparate groups of students would be unfair to teachers, boards, students, and parents alike.

A comparison should not look simply at absolute test results and suggest that a school containing students with a high average mark is a "good" school; rather, it should focus on the relative ranking of schools after taking into account variations in the characteristics of the communities in which the schools are located.

In that way, one can compare schools whose students have similar socio-economic backgrounds and identify schools that have had substantially above-average assessment results over a number of years. This would be a useful and powerful exercise.

Happily, we know enough about the backgrounds of Ontario's students to undertake such an exercise. One can predict the average assessment result over a number of years of a particular school based on both a wide variety of known socio-economic factors concerning the school's student body and the estimated relationships between those factors and test results at all Ontario elementary schools.

Once these factors have been taken into account for all schools, one can make a valid comparison between two schools with the same *predicted* assessment results but widely varying *actual* results.

I was able to develop a method for making this comparison, and the resulting comparisons (for almost all Ontario publicly-funded elementary schools) can be found at www.cdhowe.org. There were a number of interesting findings.

For example, a substantial amount ─ approximately 75%─ of the variation in school assessment results is not associated with observable characteristics of school communities. Two school communities with identical measurable characteristics may have very different results in a given year or over a number of years.

Perhaps surprisingly, if a larger percentage of the school community consists of immigrants who have come to Canada within the past five years, achievement scores tend to be higher.

School size, on the other hand, appears to have no substantial effect.

It should go without saying that it is important to be able to identify schools where good practices are making a difference, so that we can in turn identify those good practices.

(Adapted with permission from* Signposts of Success. *See our review on page 3. Dr. Johnson is professor of economics at the School of Business and Economics at Wilfrid Laurier University.)