C.D. Howe Institute
Commentary

www.cdhowe.org No. 215, August 2005 ISSN 0824-8001

The Education Papers

# School Class Size:

## *Smaller Isn't Better*

Yvan Guillemette

***In this issue...***

*Many provinces are spending millions of dollars on class-size reduction initiatives, with no solid evidence that they raise student achievement. The money could be better spent elsewhere.*

## *The Study in Brief*

Many provinces, most notably Ontario, are going ahead with class size reduction initiatives in the early primary grades. Many parents and teachers are strong advocates of reducing class sizes in the later primary and secondary grades, as well. However, in spite of the fact that such policies are often promoted as means to help students perform better, there is no solid base of empirical evidence to show that smaller classes improve student achievement beyond kindergarten and grade one, when pupils are being socialized into the classroom environment. Even in those very first school years, the gains in achievement observed are relatively small and do not carry through to later years. Recent standardized test scores from Canadian pupils aged 13 to 16 years old show no evidence that smaller classes are better, either for achievement or classroom atmosphere. One possible, though surely partial, explanation for the counter-intuitive effect is that school systems that have smaller classes have to employ more teachers than otherwise, forcing them to hire less-qualified teachers. There is some evidence that this effect is at play in Canada.

Because reducing class size is enormously expensive, it is very likely that the money being spent there could be better spent on other educational policies, such as continuous teacher training, which, unlike class size reduction, have been shown to improve student performance.

## *The Author of This Issue*

*Yvan Guillemette* is Policy Analyst at the C.D. Howe Institute.

* * * * * *

*C.D. Howe Institute Commentary*© is a periodic analysis of, and commentary on, current public policy issues. Kevin Doyle edited the manuscript; Diane King prepared it for publication. As with all Institute publications, the views expressed here are those of the author and do not necessarily reflect the opinions of the Institute's members or Board of Directors. Quotation with appropriate credit is permissible.

To order this publication, please contact: Renouf Publishing Co. Ltd., 5369 Canotek Rd., Unit 1, Ottawa K1J 9J3 (tel.: 613-745-2665; fax: 613-745-7660; e-mail: order.dept@renoufbooks.com), or the C.D. Howe Institute, 67 Yonge Street, Suite 300, Toronto M5E 1J8 (tel.: 416-865-1904; fax: 416-865-1866; e-mail: cdhowe@cdhowe.org).

$12.00; ISBN 0-88806-664-3
ISSN 0824-8001 (print); ISSN 1703-0765 (online)

Few policy issues in elementary and secondary education generate more heat than class size. Teachers and their unions are nearly unanimous in their support for smaller class sizes as a means of improving class behaviour and student performance. The public largely agrees: In each of five National Issues in Education polls commissioned by the Canadian Teachers' Federation (CTF) between 1997 and 2004, Canadians cited class size reduction as the most pressing educational spending priority. In the October 2004 poll, some 76 percent of those surveyed said that public school classes are too large.

On the other side of the debate, some economists, researchers and education specialists cite a lack of empirical support and considerable costs as their reasons for doubting that class size reduction benefits students.

The array of class size policies pursued by provinces in recent years reflect conflicting views: While some provinces like Ontario and Alberta are actively introducing policies to reduce class sizes in elementary grades, supported by extra funding, others, such as Manitoba and British Columbia, have decided against, or even removed, provisions that restrict class sizes. In the process, each province has pleased some interest groups and angered others, and the debate as to the effectiveness of restricting class size rages unabated.

Research on the issue shows that restricting class sizes to below 20 students in kindergarten and grade one improves student achievement, albeit modestly, with the strongest effects concentrated at the very beginning of a pupil's schooling. As a result, provincial initiatives that target class size reductions in the first few grades benefit from empirical support and partly for that reason generate less controversy. It is far from clear, however, that reducing class size is the most cost-effective strategy available to raise young pupils' achievement, particularly in the case of the later primary and secondary grades, where smaller classes have not been shown to produce tangible achievement gains. Those likely to gain most from smaller classes generally are teachers whose workloads are eased somewhat with fewer students.

After surveying the different strategies that provinces are following and reviewing some well-established research findings, I present Canadian-specific evidence on the link — or, more exactly, the lack of one — between class size and achievement in core subject areas at the high school level using recent test results from the School Achievement Indicators Program (SAIP) and the Programme for International Student Assessment (PISA). Using only simple descriptive statistics, I show that, in line with the bulk of empirical research on this question from many countries, smaller class sizes do not seem to produce better achievement results in Canadian schools for pupils aged 13 years and older. My conclusion is that even if Canadian efforts to limit class sizes in the early grades can be justified, other education reform strategies would likely achieve a much bigger bang for the buck. In short, class size reduction policies should not be extended beyond grade one.

* Without implicating them, I thank Pierre-Pascal Gendron, Steven Lehrer, Finn Poschmann, Tom Roberts, Bill Robson and one anonymous reviewer for comments on a previous draft.

## Different Strokes for Different Provinces: Same Debate

In recent years, the governments of P.E.I., Nova Scotia, New Brunswick, Quebec, Ontario and Alberta have announced strategies to reduce class sizes in some or all kindergarten-to-grade 12 classes. Some provinces have done it through legislation, others through collective agreements with their teachers' unions; others still have simply announced commitments for class size reductions, usually at budget time. For example, the British Columbia Liberals passed legislation in 2002 that stripped the firm class-size limits from their collective agreement with teachers, while also passing legislation that caps individual class sizes from kindergarten to grade three, as well as district-average class sizes in other grades up to 12. Similarly, in Ontario, the Liberal government has committed to reducing class sizes to 20 students from junior kindergarten to grade three by the end of its mandate. Spending increases typically accompany such commitments. In the case of Ontario, $90 million was earmarked in the 2004/2005 school year to implement the cap on class size, followed by $126 million for 2005/2006. The total amount is expected to grow to over $450 million by 2007/2008. In Ontario, the new money primarily goes to hiring additional teachers, 1,100 in 2004/2005 and 1,275 in 2005/2006.

One province has taken the opposite view. Manitoba, in the final report of the Commission on Class Size and Composition, tabled in April 2002, concluded that a blanket approach, with provincially legislated caps on class size, would fail to address local needs and that class composition is more important than class size.

Clearly, there is no uniformity among provincial policies. Within each province, however, there is a flourishing debate about the effects of class size on students' results.

Ideally, the substantial investments of some provinces in smaller classes would rest upon a solid base of empirical support. Such support should consist not only of a well-established relationship between smaller classes and improved student achievement, but one of sufficient magnitude to establish that the policy of reducing class size is more cost-effective than other feasible methods of improving student achievement (Addonizio and Phelps 2000). I first consider the empirical link between class size and achievement.[1]

### *Class Size and Achievement in the Early Grades*

Probably more studies on the class size issue have surfaced than on any other question in education.[2] Although there is no consensus, the empirical evidence generally points to limited positive effects from smaller classes on student achievement in kindergarten and grade one, with no significant effect on classes above that level.

---

1 In this paper, the term achievement refers to student performance on standardized tests of reading, writing, mathematics and science.

2 For surveys of the literature regarding class size, see Biddle and Berliner (2002), Hanushek (1997, 1999, 2003) and Edmonton Public Schools and University of Alberta Faculty of Education (2001).

Tennessee's STAR Project provides the most widely cited results on early-grade class size effects, but there are others with similar results.[3] Project STAR found that students in small classes (13 to 17 pupils) performed better on standardized tests than students in regular classes (22 to 25 pupils). As Hanushek (1999) and others have remarked, however, the small-class advantages observed in Project STAR were very modest (about one quarter of a standard deviation on achievement tests) and almost exclusively obtained in the first year of exposure to a small class. A recent analysis of Project STAR results by Ding and Lehrer (2004), which corrects for multiple problems in the experiment's implementation,[4] confirms and refines this now widely accepted conclusion. The authors found that Project STAR pupils benefited in all subject areas from attending a smaller class in either kindergarten or grade one, but without additional benefits from being in a small class in both years. By the time students finished grades two and three, however, those benefits had vanished; that is, whether students had been in small or regular classes since kindergarten made no noticeable difference in their performance on standardized tests at the end of grades two and three.

As Hanushek (2003) points out, this pattern of effects is at odds with the normal rhetoric about smaller classes permitting more individualized instruction, allowing improved classroom interactions, reducing disruptions and the like. Common sense leads to expectations of increased benefits in any grades where smaller classes are the rule. Instead, the Project STAR results and those of many other studies appear more consistent with smaller classes benefiting students during the phase of socialization or introduction into the behaviour of the classroom.

It is indeed in the early grades that children are first learning about the rules of standard classroom culture and forming ideas about whether they can cope with education (Biddle and Berliner 2002). Many children have difficulty with these tasks, and their efforts are greatly aided when they can interact with teachers on a one-on-one basis — a process more likely to take place when the class is small. If socialization into the classroom takes about one school year, then this theory is very consistent with the result that smaller classes make a difference if a pupil is exposed to a smaller class in kindergarten. If a larger class in kindergarten hinders the socialization process, it is also consistent with smaller classes yielding benefits in the second year of schooling (grade one) if pupils were not exposed to small classes in kindergarten, but no extra benefit if they were. The theory also squares well with the result that achievement improvements vanish after a few years because by then, children have been in school long enough to have adapted. It does not seem that the head start gains in achievement from early socialization are important enough to remain noticeable in achievement results after a few years of schooling.

---

3 Project STAR (Student/Teacher Achievement Ratio), initiated in the mid-1980s, covered 7,000 students in grades K-to-three in 80 Tennessee schools. Students were randomly placed in small and larger classes. Researchers measured differences in student scores on tests in mathematics, reading and writing, both while students were divided into small and large classes (grades K-to-three) and after they had been mixed back together (grades four-to-eight). For more discussion of Project STAR and its limitations, see the cited literature surveys.

4 The most important implementation problem, which may skew previously reported Project STAR results, was non-random attrition as pupils progressed from kindergarten to grade three.

By the same token, the socialization theory can explain why reducing class size may not generate significant advantages if introduced in the upper grades. Older students have long developed their own habits for coping with standard classrooms and these are not likely to change just because class size is reduced (Biddle and Berliner 2002).

Studies that look at later primary (three to six) and secondary grades indeed generally find no significant improvement in student performance from smaller classes. In his review of the evidence, Hanushek (1999) first points out that over the past quarter century, there has been a steady decline in pupil-teacher ratios in the U.S., without a corresponding increase in measures of student achievement. Turning to the international evidence, he observes that the wide discrepancies in pupil-teacher ratios across countries and time show little relationship to achievement. In science and math tests, for example, such countries as South Korea, with very large class sizes, routinely outperform richer countries, such as the U.S. and Canada, where classes are much smaller.

Comparisons based on aggregate country-level data suffer from a number of problems that cloud their interpretation, however. As a result, Hanushek next turns to summarizing the extensive econometric evidence about the effectiveness of reducing class size to raise student achievement. Econometric studies try to separate out the influences of families and other school factors that may bias the findings from aggregate data. His analysis, which incorporates almost 300 different estimates of the effect of class size on achievement, give no indication that general reductions in class size will yield any average improvement in student achievement. While some studies point in that direction, an almost equal number point to an opposite conclusion.[5] Moreover, for those studies that do find a positive effect for smaller classes, it is often very small.

Since Hanushek's review of the evidence, more studies have appeared without altering his overall conclusion. One of the most noteworthy published since his survey looked at fourth- and sixth-graders in Connecticut (Hoxby 2000). Using a clever strategy based on natural population variations to circumvent the usual statistical problems that arise when trying to study class-size effects, the study found no evidence at all that smaller classes improved achievement.

Canadian-specific evidence on class size is very sparse. Perhaps the best, albeit old, findings come from a study initiated in the 1970s by the Board of Education of the City of Toronto, together with the Toronto Teachers' Federation. The study was designed to examine experimentally the differences between four class sizes ranging from 16 to 37 pupils by randomly assigning students and teachers to a particular class size (16, 23, 30 or 37). Sixty-two classes of students in the fourth and fifth grades from 11 schools in Metropolitan Toronto participated. The study showed that class size makes a large difference to teachers' attitudes and expectations, while having little or no effect on students or on instructional methods used. There were no significant differences attributable to class size for achievement in art, composition, vocabulary, reading and mathematics problem

5 An earlier survey of the literature identified 112 studies of the effect of class size on achievement (Akerhielm 1995). Of these, nine identified a positive and statistically significant effect of smaller class size on achievement, 14 identified a statistically significant negative effect, and 89 studies showed no significant effect.

solving (Shapson et al. 1980). Only for mathematics concepts was there a measurable overall effect due to class size: Students in class size 16 had slightly higher mathematics-concepts scores than their peers in class sizes 30 and 37.

Canadian research on class size in the secondary grades is even sparser. But because teachers' efforts to limit the number of students in their classes are not limited to the early grades,[6] I now turn to an analysis of some preliminary Canadian evidence based on aggregate data for secondary-school grades.

## *Class Size and Achievement in Canadian Secondary Schools*

It is revealing to look at a simple correlation between class size and student achievement across provinces. One source of information we can use is the School Achievement Indicators Program (SAIP),[7] which every year tests 13- and 16-year-old Canadian students on a different subject and questions pupils, teachers and school principals on their backgrounds and on characteristics of their schools. Most useful for our purposes, SAIP asks principals about the average class size in their schools for courses concerning the specific subject tested. In addition, when a school comprises both 13- and 16-year-old students, a separate answer is given for the two age groups, allowing for a great degree of precision in linking class size for a given test subject and age group to test scores.

Figure 1 illustrates this link by jurisdiction for 13-year-olds using SAIP 2001, which tested mathematics content, and SAIP 2002, which tested writing abilities. It is clear that some provinces have a much higher percentage of schools reporting average class size of 25 students or more, but those provinces do not have a lower proportion of students achieving a level 2 or higher on the respective tests.[8] In fact, somewhat surprisingly, the relationship is positive; that is, provincial school systems that report a higher occurrence of large classes obtain better student achievement results for 13-year-olds on average.[9]

More detailed results from the OECD's Programme for International Student Assessment (PISA),[10] as well as SAIP 2002 and 2004 tell a similar story.

Table 1 shows the results for grade 10 students in Canada according to PISA's 2000 science, reading and mathematics tests and the 2003 mathematics test. Class sizes were reported directly by students, and specifically for the subject tested. In all tests, larger classes are associated with better achievement scores, although in

---

6 For example, in a *Vancouver Sun* column (Jan. 19, 2005, page A17) Jinny Sims, president of the B.C. Teachers Federation, argues with statistics on class sizes in different subjects from grade eight to 11 that B.C. high school classes are too large and vows "to continue to advocate and negotiate for firm class size limits." Similarly, in Quebec, 30,000 teachers marched on the legislature and threatened to strike in May 2005, demanding smaller classes in both elementary and secondary schools ("Quebec teachers threaten long strike if class size isn't reduced," *CP Wire*, May 6, 2005).

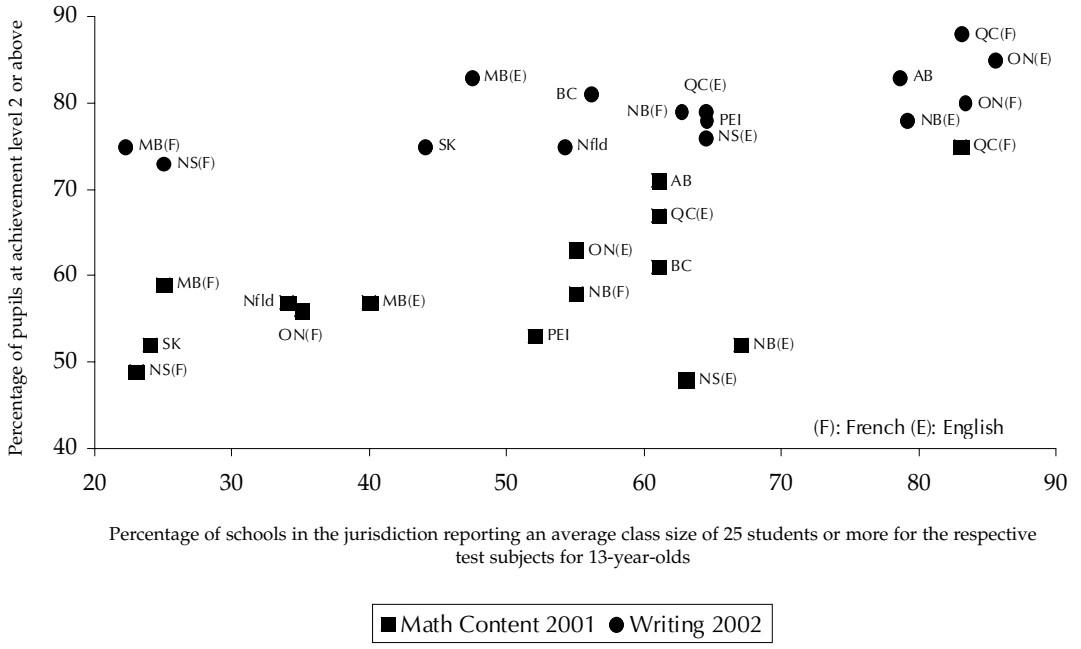
7 For details on SAIP, see www.cmec.ca/saip/indexe.stm.

8 Students are graded on a scale from 0 to 5, with 2 being the minimum satisfactory level for 13-year-olds.

9 Although not shown here, repeating the same exercise with the percentage of 16-year-olds who scored at level 3 or above (the minimum achievement level expected of 16-year-olds) in SAIP 2001 and 2002 also yields no significant correlation between class size and test scores.

10 For details on PISA, see www.pisa.gc.ca.

**Figure 1:** *Class Size versus Math/Writing Achievement for 13-year-old Canadians in SAIP 2001 and 2002*




Sources: Author's calculations using SAIP databases and public reports.

the 2000 reading and mathematics tests there was almost no difference in average scores for classes larger than 25 students. Results for classes with less than 20 students must not be given too much weight, because these classes are more likely to be remedial or special-needs classes, and thus more likely to produce lower average scores. But looking at classes with at least 20 students and those with 35 students or more, the normal range of variation in the country, the results give no reason to believe that student performance is negatively affected by larger classes.

Table 2 uses the SAIP 2002 writing test and the 2004 science test to look at average student performance for 13-year-old and 16-year-old Canadian students. Class size data are taken from school principals' responses about average class size in their schools for the given age group and test subject rather than directly from student responses. Results are therefore perhaps less reliable than those in Table 1.

Nevertheless, the SAIP results agree with those from PISA in that they show no evidence that large classes hinder achievement. In all four tests, the highest average score occurs in classes with 30 students or more.

## All Things Considered

While these results are interesting, and certainly challenge the commonly expressed view in Canada that smaller classes must produce better results, the true link between class size and achievement may be obscured by school characteristics and socio-economic factors that are correlated with class size. As

**Table 1:** Average Achievement Levels for 15-year-old Canadian Students in Grade 10 in PISA 2000/2003 by Class Size

| | PISA 2000 Science Test | | PISA 2000 Reading Test | | PISA 2000 Mathematics Test | | PISA 2003 Mathematics Test | |
|---|---|---|---|---|---|---|---|---|
| Class Size | Average Achievement Level | Number of Students in Sample | Average Achievement Level | Number of Students in Sample | Average Achievement Level | Number of Students in Sample | Average Achievement Level | Number of Students in Sample |
| Less than 10 | 470 | 347 | 464 | 625 | 475 | 357 | 497 | 396 |
| 10-14 | 515 | 398 | 514 | 712 | 507 | 379 | 512* | 836 |
| 15-19 | 524* | 1,125 | 533 | 1,959 | 526 | 1,058 | 532 | 1,439 |
| 20-24 | 529** | 2,545 | 540* | 4,635 | 526* | 2,606 | 535* | 2,780 |
| 25-29 | 546 | 3,615 | 556 | 6,542 | 545 | 3,623 | 558 | 2,849 |
| 30-34 | 553 | 3,636 | 561 | 6,522 | 553 | 3,621 | 575 | 2,061 |
| 35 or more | 557* | 514 | 558** | 897 | 551** | 474 | 595 | 245 |
| Mean/Total | 539 | 12,180 | 547 | 21,892 | 538 | 12,118 | 549 | 10,606 |

Sources and notes: PISA databases, author's calculations. Class size in the respective subjects are based on student responses. Means are predicted population means taking into account the stratified survey design. A * indicates that the estimate is not statistically different than the estimate immediately above using a 5-percent level Wald test. A ** indicates that the estimate is not statistically different from the two estimates immediately above it.

indicated, a more complete analysis needs to weed out other factors that vary among students, schools and jurisdictions and that may blur the true class size effect we are interested in.

Ascertaining the precise effect of class size on student achievement is not as straightforward as simply adding a few control variables. The central problem is that students are not assigned to classrooms randomly. For instance, schools often establish small remedial classes for lagging students or small enrichment classes for the so-called gifted and talented. In addition, school systems may direct students into schools with different average class sizes on the basis of their performance. Parents may also influence their children's class size. They may work hard to move their children to schools with smaller classes, where they believe they will receive more attention. Thus, variation in class size may be the result, rather than the cause, of differences in student ability and achievement.

Estimating the true effect of class size on student performance requires a strategy that looks only at variations in class size that are unrelated to any omitted factors that also affect students' current achievement. This condition can only truly be met through real-world trials that randomly assign students to classes of different sizes. Project STAR and the 1970s Toronto experiment cited earlier are examples of (imperfectly) randomized studies. Failing the availability of a randomized trial, econometric methods can be used to analyze data on class size and student achievement and abstract from the effect of ability and other factors on student achievement. Hoxby's (2000) study cited earlier, for example, used exogenous population variation to mimic the effect of randomization. Unfortunately, there are not many studies of that kind that look specifically at Canadian data.

Woessmann and West (2002) is an exception. The authors conduct a careful study spanning 18 countries, consider each nation separately and use a methodology that limits the potential biases due to the statistical problems just

**Table 2:** Average Achievement Level for 13- and 16-year-old Canadian Students on the SAIP 2002 Writing and SAIP 2004 Science tests by Class Size

| | SAIP 2002 Writing test | | | | SAIP 2004 Science test | | | |
|---|---|---|---|---|---|---|---|---|
| | 13-year-olds in Grade 8 | | 16-year-olds in Grade 11 | | 13-year-olds | | 16-year-olds | |
| Class Size | Average Achievement Level | Number of Students in Sample | Average Achievement Level | Number of Students in Sample | Average Achievement Level | Number of Students in Sample | Average Achievement Level | Number of Students in Sample |
| Less than 10 | 2.47 | 98 | 2.76 | 32 | 1.97 | 340 | 2.55 | 544 |
| 10-14 | 2.31 | 224 | 2.77 | 300 | 2.05 | 1,117 | 2.70 | 974 |
| 15-19 | 2.18 | 455 | 2.88 | 659 | 1.90 | 2,055 | 2.65 | 2,459 |
| 20-24 | 2.29 | 1,687 | 2.66 | 1,149 | 2.00 | 4,411 | 2.63 | 3,413 |
| 25-29 | 2.40 | 3,014 | 2.75 | 2,200 | 1.98 | 6,606 | 2.75 | 5,069 |
| 30-34 | 2.52 | 970 | 2.93 | 943 | 2.06 | 3,098 | 2.86 | 2,074 |
| 35 or more | 2.56 | 77 | 3.02 | 130 | 2.33 | 171 | 2.84 | 128 |
| Mean/Total | 2.40 | 6,525 | 2.77 | 5,413 | 2.01 | 17,798 | 2.72 | 14,661 |

Sources and notes: SAIP 2002 and 2004 databases, author's calculations. Student achievement is measured by assigning a level between 0 and 5 to a student's writing abilities (with 5 being the highest). Class size for the respective age groups are based on principals' responses. Means are predicted population means taking into account the stratified survey design. Statistical significance tests are not shown but, except for 13 year-olds in SAIP 2004, the mean score for the 30-34 class size category is always statistically significantly higher than for the 20-24 category.

mentioned.[11] The authors found that in six of the 18 countries, including Canada,[12] even a minimal relationship between class size and test scores in the middle grades could be ruled out. In an additional five school systems,[13] they could rule out large class size effects but not necessarily small ones, and in another five countries their strategy led to imprecise estimates that do not allow for any confident assertion about the effects of differences in class size.[14] There were only two countries where smaller classes did appear to elicit superior student performance, Greece and Ireland. In Canada's case, then, these results conform with the descriptive statistics presented in the previous section, in that class size has no noticeable impact on student performance in high school.

11 They use data from the Third International Mathematics and Science Study (TIMSS) for 13-year-olds. In a nutshell, their empirical strategy uses variation in class size within schools, instrumenting actual class size with average class size at the grade level. The difference in average class size between adjacent grades in a school should reflect natural cohort size fluctuations, which are presumably random.

12 The other countries are Belgium (Fl.), Slovenia, Portugal, Singapore and Japan.

13 Belgium (Fr.), Czech Republic, Romania, Spain and Korea.

14 Results were inconclusive for Scotland, France, Australia, the United States and Hong Kong.

**Table 3:** *Class Disruption by Class Size for 15-year-old Canadians in Grade 10*

| Class Size in Math | Percentage of Students who Report Noise/Disorder in their Math Classes in Every or Most Lessons | Number of Students in Sample |
|---|---|---|
| Less than 10 | 30.9 | 388 |
| 10-14 | 33.4 | 810 |
| 15-19 | 37.3 | 1,412 |
| 20-24 | 37.7 | 2,721 |
| 25-29 | 37.5 | 2,794 |
| 30-34 | 33.8 | 2,015 |
| 35 or more | 36.3 | 242 |
| Mean/Total | 36.6 | 10,382 |

Sources and notes: PISA 2003 database, author's calculations. Students are placed in class size categories based on their reports of the average number of students in their grade 10 math class. Percentages are predicted population parameters taking into account the stratified survey design. Wald tests reveal that all the estimates are not statistically different from each other at the 5-percent level.

## *Evaluating Some Claims on the Effects of Reducing Class Size*

According to the conventional wisdom, one way in which smaller classes would benefit students is by improving the conditions for learning, that is, by reducing the noise and disruptions believed to be associated with larger classes.[15]

### Learning conditions

PISA 2003 enables us to take a look at this hypothesis. On its questionnaire, students were asked how often "noise and disorder" occur in their mathematics classes. They had a choice between "every lesson", "most lessons", "some lessons" or "never or hardly ever". Table 3 shows the percentage of Canadian grade 10 students who answered either "every" or "most lessons" to that question according to the size of their mathematics class.

On average nationally, approximately 37 percent of students report a high frequency of class disruptions in their math classes due to noise or disorder, but according to standard statistical tests the proportions shown in the table are not statistically different from each other. In other words, there is no indication that disruptions are more frequent in larger classes.

### Teacher Quality

Why might lowering class size not yield improvements in student performance? What could be the reason that conventional wisdom fails in this matter? Part of the explanation may have to do with a consequence of class size reductions that

15 This conventional wisdom was formulated in theory by Lazear (2001).

does not receive proper attention in the debate: the dilution of teacher quality.[16] The point is easiest to grasp when comparing class size reductions on two different scales: the effect at the classroom level versus the average effect for all classes in the jurisdiction. In the first instance, pupils may well do better because teacher quality is held constant, though spread over a smaller number of children. At the provincial level, however, the number of teachers and the composition of the teaching force change as average class size declines. Unless enrolment is falling, reducing class size invariably means increasing the number of teachers, and it may be that hiring more teachers forces a school, board or jurisdiction to dip into the pool of less able teachers. This would certainly occur in the short-run because the supply of certified teachers is mostly fixed, and potentially over the long-run, as well, unless salary increases or smaller classes manage to attract more high-quality candidates to the profession. Somewhat ironically, if the current system does a good job of attracting the best candidates to education faculties and of selecting the best teachers for available positions, increasing the number of teachers to reduce class size logically leads to lower average teacher quality. Therefore, while it may be true that most students would do better in a smaller class with a given teacher, the same cannot be expected for the average student when increasing the number of teachers in a jurisdiction.

The available empirical evidence confirms the preponderance of the quality effect: "Variations in teacher quality have been shown to be extraordinarily important for student achievement, and econometric studies indicate that these variations completely dominate any effects of altered class size" (Hanushek 1999, p. 163).[17]

If this effect were at play, we might observe boards or individual schools with smaller classes on average making use of less qualified teachers. Table 4 provides evidence that this is what may be occurring in Canada. The table uses PISA 2003 mathematics results in Canada, using schools as the unit of observation rather than students as was the case in Tables 1 and 2. In each school, principals were asked how many mathematics teachers worked at their school and how many possessed a specialization in mathematics. Because many principals did not answer this question, the sample size is much smaller than in previous tables, which explains why only three class size categories are shown. The available data nevertheless show that schools with larger mathematics classes in grade 10 tend to employ a higher proportion of teachers specialized in mathematics.[18] As we saw, those same classes also do better on average in standardized tests.

16 Addonizio and Phelps (1995, p. 155) among others allude to this possibility.

17 See Jepsen and Rivkin (2002) for evidence that the California Class Size Reduction Program lowered teacher quality on the margin, and Rivkin et al. (1998) and Hanushek (1992, 1997) for other evidence and references.

18 One possible alternative explanation is that larger schools tend to have larger classes and tend to be located in larger cities, and larger cities have a relatively larger supply of qualified teachers.

**Table 4:** *Average Achievement Levels, Teacher Qualification and Class Size for 15-year-old Students in Grade 10*

| Average Class Size in Math | Average Achievement Level | Percentage of Math Teachers with Degree Specialization in Math | Number of Schools in Sample |
|---|---|---|---|
| Less than 15 | 519 | 47 | 63 |
| 15-24 | 533* | 63 | 219 |
| 25 or more | 566 | 70* | 116 |
| Mean/Total | 539 | 62 | 398 |

Sources and notes: PISA 2003 database, author's calculations. Class size categories are based on schools' average class sizes in grade 10 math. Average achievement levels are computed using only schools for which data to calculate teacher qualifications are available. Teacher qualifications use data for both part-time and full-time teachers, with full-time teachers given greater weight. Means and proportions are predicted population parameters taking into account the stratified survey design. A * indicates that the estimate is not statistically different than the estimate immediately above using a 5-percent-level Wald test.

## Cost-Effectiveness

Any expected benefits in achievement from class size reductions must ultimately be balanced against the costs of such policies. Benefits of raising achievement are hard to value in economic terms.

We do not need to consider the cost-benefit trade-off of a reduction in class size beyond grade one. Having already established that such a policy is likely to have no noticeable impact on average student achievement, the costs would obviously outweigh the benefits. What about grade one? A complete cost-benefit analysis would take us beyond the realm of this paper, but consider a simple back-of-the-envelope calculation: Assume there are 120 students in grade one at a school, and five teachers, resulting in 24 pupils per class. By employing an additional teacher, class size would fall to 20, a reduction of four. By employing a second additional teacher (making seven), class size would then fall to 17.1, a marginal reduction of 2.9 students per classroom. Assume costs of $60,000 per teacher,[19] ignore capital costs, and assume that lowering class-size by four students improves average student achievement by one tenth of a standard deviation (roughly in line with what Project STAR and other similar studies have measured). For example, if the average test score is 500 with a standard deviation of 100, lowering class size by four students improves the average on the test by 10 points to 510.

Table 5 shows the results of a simple simulation of the salary costs and student achievement benefits involved in reducing class size in the range from 30 to 12 in the above example. Evidently, such a strategy is very costly, and the expected gains in achievement quite small. The marginal costs of raising achievement by

19 This is a low estimate: According to Yvonne Mahoney, finance director of the Catholic school board in Peterborough, Ontario, hiring 14 extra teachers to reduce class size in the board from 22 to 21 pupils per teacher would cost $1,090,557 in salaries and benefits, or $77,897 per teacher ("Larger classes, more prep time," *The Peterborough Examiner*, 26 May 2005, page B1).

**Table 5:** Costs and Benefits of Class Size Reductions — Assuming 120 Students in a Grade

| Number of Teachers | Class Size | Class Size Reduction | Test Average | Total Teacher Cost ($) | Teacher Cost per Pupil ($) | Marginal Cost per 1-Point Increase in Achievement ($) |
|---|---|---|---|---|---|---|
| 4 | 30.0 | n/a | 500 | 240,000 | 2,000 | n/a |
| 5 | 24.0 | 6.0 | 515 | 300,000 | 2,500 | 4,000 |
| 6 | 20.0 | 4.0 | 525 | 360,000 | 3,000 | 6,000 |
| 7 | 17.1 | 2.9 | 532 | 420,000 | 3,500 | 8,400 |
| 8 | 15.0 | 2.1 | 538 | 480,000 | 4,000 | 11,200 |
| 9 | 13.3 | 1.7 | 542 | 540,000 | 4,500 | 14,400 |
| 10 | 12.0 | 1.3 | 545 | 600,000 | 5,000 | 18,000 |

Source: Adapted from Addonizio and Phelps (1995).

one hundredth of a standard deviation (one point) rises as average class size is reduced because an extra teacher always cuts average class size slightly less than the preceding teacher. In this fictional grade, raising average achievement by only one third of a standard deviation would involve spending twice as much in salary costs, with the last few extra points in the average score costing $18,000 each. Because I ignored capital costs, this estimate can be thought of as a lower-bound. If we believe, as the literature suggests, that performance improvements from class size reductions are modest, at best, then it is clearly not possible to achieve an educationally significant impact on average student performance in core subject areas at reasonable costs.[20] In contrast, it seems reasonable to believe that each $5,000 (approximately) necessary to raise average scores by one point in the range between 24 and 20 students would buy greater performance improvements if it were spent updating the skills of current teachers by offering them professional development courses and seminars instead. Other educational strategies targeting institutions such as greater school accountability, more competition, better teacher incentives and expanded school choice have also proven to be effective strategies in raising student achievement.[21]

## Conclusion

Canadians who have children in school obviously have a strong interest in the teaching conditions their children are exposed to, including class size. A majority of Canadians seem convinced that smaller is better — partly because the notion has intuitive appeal and perhaps because they have been convinced by relentless appeals from teachers who want the number of students per class lowered. Some

20 Other potential benefits, such as classroom atmosphere and teacher-pupil bonding, should not be overlooked, but those are not the benefits that class reduction policies are typically sold on.

21 See the collection of papers in a Forum entitled "Institutions for Better Education" in *CESifo DICE Report* 2(4): Ifo Institute for Economic Research.

provincial governments have responded by pursuing class size reduction goals supported by heavy investments. Other provinces, perhaps because of stringent budget constraints, have resisted imposing class size limits.

Empirical research has shown that smaller classes may benefit pupils for the short period of time during which they are being introduced and socialized into a classroom environment. For most pupils, this occurs in kindergarten or grade one. However, the best available empirical evidence shows that the small advantages observed then do not carry through to grade two or three when all pupils have been properly socialized. Also, the costs of reducing class sizes can be substantial.

Although few provinces currently have specific class size limits beyond grade three, the pressure is always there, especially from teachers' unions, to reduce class sizes in these grades as well. It is certainly convenient for some teachers to parade as student advocates and claim improvements in student achievement to demand class size limits, while the real reasons for such demands have more to do with improving their own working conditions. Simply put, although class size is an attractive policy target for politicians because it is easily understood by the electorate, easily measured, and reductions in it easily trumpeted as major achievements, no solid evidence exists to show that smaller classes improve student achievement in the later primary and secondary grades in Canada. Scarce public funds spent on class size reduction goals would therefore most probably be better spent elsewhere, for example on better teacher training and development. Increasing class sizes may even raise achievement scores by exposing more students to the best teachers.

In setting class size policy, as in many other policy areas, evidenced-based policymaking will prevent incurring large costs when no noticeable improvement in results can be expected.

## References

Addonizio, Michael F., and James L. Phelps. 2000. "Class Size and Student Performance: A Framework for Policy Analysis," *Journal of Education Finance* 26: 135-56.

Akerhielm, Karen. 1995. "Does Class Size Matter?" *Economics of Education Review* 14(3): 229- 41.

Biddle, Bruce J., and David C. Berliner. 2002. "What Research says about Small Classes and Their Effects," WestEd *Policy Perspectives*. San Fransisco: WestEd.

Ding, Weili, and Steven Lehrer. 2004. "Estimating Dynamic Treatment Effects from Project STAR," Mimeo. Kingston: School of Policy Studies, Queen's University.

Edmonton Public Schools and University of Alberta Faculty of Education. 2001. *Small Class Size Project — Final Report,* Appendix IV — Review of the Literature. Available at http://www.education.gov.ab.ca/ClassSize/ (March 28, 2005).

Hanushek, Eric A. 1992. "The Trade-Off Between Child Quantity and Quality," *Journal of Political Economy* 100(1): 84-117.

———. 1997. "Assessing the Effects of School Resources on Student Performance: An Update," *Education Evaluation and Policy Analysis* 19(2): 141-64.

———. 1999. "The Evidence on Class Size." In Susan E. Mayer and Paul Peterson (eds.) *Earning and Learning: How Schools Matter*. Washington, D.C.: Brookings Institution.

———. 2003. "The Failure of Input-Based Schooling Policies," *The Economic Journal* 113 (February): F64-F98.

Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement: New Evidence From Population Variation," *Quarterly Journal of Economics* 115(3) (November): 1239-85.

Jepsen, Christopher, and Steven Rivkin. 2002. "What is the Trade-off Between Smaller Classes and Teacher Quality?" *NBER Working Paper* 9205. Cambridge, Mass.: National Bureau of Economic Research.

Lazear, Edward. 2001. "Educational Production," *Quarterly Journal of Economics* 116(3): 777-803.

Rivkin, Steven G., Erik A. Hanushek, and John F. Kain. 1998. *Teachers, Schools and Academic Achievement.* NBER Working Paper 6691. Cambridge, Mass.: National Bureau of Economic Research.

Shapson, Stan M et al. 1980. "An Experimental Study of the Effects of Class Size," *American Educational Research Journal* 17(2) (Summer): 141-52.

Woessmann, Ludger, and Martin R. West. 2002. *Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS*. Institute for the Study of Labour (IZA) discussion paper No. 485.

## NOTES

## NOTES

## Recent C.D. Howe Institute Publications

| | |
|---|---|
| August 2005 | Kelleher, Maria, Janet Robins and John Dixie. *Taking Out the Trash: How To Allocate the Costs Fairly*. C.D. Howe Institute Commentary 213. |
| August 2005 | Goldfarb, Danielle. *U.S. Bilateral Free Trade Accords: Why Canada Should Be Cautious About Going the Same Route*. C.D. Howe Institute Commentary 214. |
| August 2005 | Goldfarb, Danielle. *The Canada-Mexico Conundrum: Finding Common Ground.* C.D. Howe Institute Backgrounder 91. |
| June 2005 | Robson, William B.P. *Reading the Currency Compass: Keeping Monetary Policy on Course Through a Choppy Exchange Rate.* C.D. Howe Institute e-brief. |
| June 2005 | Sikorski, Radek. *Cleaning Up the UN in an Age of U.S. Hegemony*. C.D. Howe Institute Commentary 212. |
| May 2005 | Harris, Richard. *Canada's R&D Deficit — And How To Fix It.* C.D. Howe Institute Commentary 211. |
| May 2005 | Goldfarb, Danielle and William Robson. *Canadian Workers Need the Tools to Do the Job And Keep Pace in the Global Investment Race.* C.D. Howe Institute e-brief. |
| April 2005 | Chen, Duanjie and Jack M. Mintz. *Federal Corporate Tax Cuts Would Lift Canada's Standard of Living.* C.D. Howe Institute e-brief. |
| March 2005 | Faruqui, Ahmad and Stephen S. George. *Preventing Electrical Shocks: What Ontario — And Other Provinces — Should Learn About Smart Metering*. C.D. Howe Institute Commentary 210. |
| March 2005 | Hart, Michael. *Great Wine, Better Cheese: How Canada Can Escape the Trap of Agricultural Supply Management.* C.D. Howe Institute Backgrounder 90. |
| March 2005 | Johnson, David. *Signposts of Success: Interpreting Ontario's Elementary School Test Scores.* Policy Study 40. |
| March 2005 | Boyer, Marcel. *Raise Electricity Prices in Quebec — and Benefit Everyone.* C.D. Howe Institute e-brief. |
| March 2005 | Boyer, Marcel. *Augmentons le prix de l'électricité au Québec — pour le bien de tous.* Institut C.D. Howe e-brief. |
| March 2005 | Gendron, Pierre-Pascal. *A Taxing Issue: Enhancing Quebec's Investment Attraction*. C.D. Howe Institute Backgrounder 89. |
| March 2005 | Gendron, Pierre-Pascal. *Un problème imposant : améliorer l'attrait du Québec pour l'investissement*. Bulletins de recherche de l'Institut C.D. Howe no 89. |
| February 2005 | Huberman, Michael. *Are Canada's Labour Standards Set in the Third World? Historical Trends and Future Prospects.* C.D. Howe Institute Commentary 209. |
| February 2005 | Payne, A. Abigail. *Firm Foundations: Putting Private and Public Foundations on Level Ground.* C.D. Howe Institute Backgrounder 88. |
| February 2005 | Robson, William B.P. and Finn Poschmann. *Restoring Purpose: A Shadow Federal Budget for 2005.* C.D. Howe Institute Commentary 208. |
| January 2005 | Chen, Duanjie and Jack M. Mintz. *How To Become Seductive: Make Canada More Investment-Friendly*. C.D. Howe Institute e-brief. |
| January 2005 | Guillemette, Yvan. *School Enrolment Is Down; Spending Is Up. What's Wrong With This Picture?* C.D. Howe Institute e-brief. |
| December 2004 | Wright, David S. *Managing Global Crises — and the U.S. Colossus*. C.D. Howe Institute Commentary 207. |
| December 2004 | Robson, William and Danielle Goldfarb. *Tools for Workers: How Canada is Faring in the Competition for Capital Investment*. C.D. Howe Institute Backgrounder 87. |
| December 2004 | Mintz, Jack M. and Tom Roberts. *Holes in the Road to Consensus: The Infrastructure Deficit — How Much and Why?* C.D. Howe Institute e-brief. |
| December 2004 | Guillemette, Yvan, David Laidler and William Robson. *The Real Reasons for the Canadian Dollar's Power Trip — And What Not To Do About It*. C.D. Howe Institute e-brief. |
| November 2004 | Goldfarb, Danielle. *How Canada Wins from Global Services Outsourcing.* C.D. Howe Institute Commentary 206. |

**C.D. Howe Institute**
67 Yonge Street Suite 300
Toronto, Ontario
M5E 1J8