

Understanding PISA

Stephen DOWNES
Moncton-CANADA

INTRODUCTION

The headline was dramatic enough to cause a ripple in the reading public. "Students who use computers a lot at school have worse maths and reading performance," noted the BBC news article, citing a 2004 study by Ludger Woessmann and Thomas Fuchs (Fuchs and Woessman, 2004).

It was not long before the blogosphere took notice. Taking the theme and running with it, Alice and Bill ask, "Computers Make School Kids Dumber?" They theorize, "If you track the admitted decline of education, you'll probably notice that it follows along with the increase of technology in the classroom."

In a similar vein, James Bartholomew asks, "Do you think that the government will turn down the volume of its boasting about how it has spent billions introducing computers in schools (while keeping down the pay of teachers so much that there are shortages)? Do you think it will stop sending governors of state schools glossy pamphlets about insisting that computers are used in their schools as much as possible?"

Compounding the matter was the BBC's inclusion of statements by Prince Charles on computers and learning. "I simply do not believe that passion for subject or skill, combined with inspiring teaching, can be replaced by computer-driven modules, which seem to occupy a disproportionate amount of current practice."

While computers stole the headline, the Woessmann and Fuchs report contained numerous other bombshells for the educational sector. Small class sizes have no impact on educational outcome, they argued. Private schools have a positive impact. So do standardized exams.

Additionally, school autonomy (hiring of teachers, textbook choice and budget allocations) is related to superior student performance. And students in public schools perform worse than students in private schools. Better equipment with instructional material and better-educated teachers also improve student performance (See also the Subtext summary).

THE PISA PROCESS

The Woessmann and Fuchs report, along with many others by these and other authors, was derived from a set of data obtained by the by the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA), conducted in 2000.

PISA was not the first such study. It follows previous work such as IAEP, TIMSS and TIMSS-Repeat. However, the PISA study diverges from the previous work in several respects.

First, in addition to the study of achievements in mathematics and science, PISA adds a major focus on literacy. Two thirds of the test was based on literacy, and more students were tested on literacy - 174,227 in literacy, 96,855 in math and 96,758 in science. (Woessmann and Fuchs, 2004) The sample consisted of 15-year-olds in 32 countries (28 OECD countries), countries - the focus on age rather than grade "it captures students of the very same age in

each country independent of the structure of national school systems." (Woessmann and Fuchs, 2004)

More importantly, the outcomes test is not based on curriculum objectives. "PISA aims to define each domain not merely in terms of mastery of the school curriculum, but in terms of important knowledge and skills needed in adult life (OECD 2000, p. 8). That is, rather than being curriculum-based as the previous studies, PISA looked at young people's ability to use their knowledge and skills in order to meet real-life challenges (OECD 2001, p. 16)" (Woessmann and Fuchs, 2004).

PISA also looked well beyond educational attainment. It also included school demographics, such as whether it was a public or private school, had large or small classes, or had access or not to technological resources. It also measured student information-their family background, access to books and computers and parental support.

ANALYSING THE RESULTS

One might wonder why it would take four years for Woessmann and Fuchs to produce their report. The student results were available almost immediately, and as the authors point out, they created a stir in the press. "The Times (Dec. 6, 2001) in England titled, Are we not such dunces after all? , and Le Monde (Dec. 5, 2001) in France titled, France, the mediocre student of the OECD class . In Germany, the PISA results made headlines in all leading newspapers for several weeks (e.g., Abysmal marks for German students in the Frankfurter Allgemeine Zeitung, Dec. 4, 2001).

But such simple analysis, argue the authors, is misleading. For one thing, they consisted typically of comparisons between one country and another - hence Britain's joy and Germany's disappointment. And they are typically bivariate, that is, "presenting the simple correlation between student performance and a single potential determinant, such as educational spending."

In fact, note the authors, the various variables impact on each other, skewing the results. This sort of dispute has come up in other studies as well. For example, a study may show that charter schools produce poorer outcomes. However, it might be argued that charter schools attract students of a disadvantaged demographic, and when this disadvantage is taken into account, it may result that charter schools are better value for the investment.

That's what Woessmann and Fuchs do. Speaking very loosely - they estimate the weight of each measured variable on student performance. Then when assessing another variable, they subtract that weight from the results where that variable is present. For example, if parental influence is worth 0.4, then if they are measuring for the impact of computers, then for each student who has access to computers they subtract 0.4 if they are also benefiting from parental support. Thus, the impact of the computers, independent of parental support, is measured (please note that this is a gloss; readers should consult the actual paper for the precise model).

Thus we see the authors argue as follows: "once family background is controlled for, the relationship between student achievement and their having one or more computers at home turns around to be statistically significantly negative. That is, the bivariate positive correlation between computers and performance seems to capture other positive family-background effects that are not related to computers... Holding the other family-background characteristics constant, students perform significantly worse if they have computers at home." (Woessmann and Fuchs, 2004)

THE ECONOMIC MODEL

It is worth noting at this juncture that Woessmann and Fuchs are economists and that their methodology is informed by (what they believe to be) principles of their discipline. Indeed, it is clear from the report that to a large degree they approach the subject from the standpoint of what "economic theory says" (a phrase oft-repeated in the paper) and their intent is to a large degree to compare the results of the study to what economics says should be the case.

In approaching the study in this way, one assumes a stance very different than one that might be taken by an educator or a member of the public. For example, economists assume (as a matter of process, not prejudice) that economic theory applies to education. Thus, for example, it is taken for granted that "students choose their learning effort to maximize their net benefits, while the government chooses educational spending to maximize its net benefits." (Bishop and Woessmann, 2002)

The economic point of view, moreover, favours a depiction of the educational institution as a dominant influence in the production of educational outputs. "Economic theory suggests that one important set of determinants of educational performance are the institutions of the education system, because these set the incentives for the actors in the education process." (Woessmann and Fuchs, 2004)

Setting incentives is tantamount, on this view, with marketplace interference. "One reason why the institutional system plays such a crucial role especially in educational production may be that public schools dominate the production of basic education all over the world. As the Economist (1999, p. 21) put it, "[i]n most countries the business of running schools is as firmly in the grip of the state as was the economy of Brezhnev's Russia." (Bishop and Woessmann, 2002) This depiction puts the educational system at odds with marketplace theory, and thus the expectation (from economists, at least) is that a more efficient production will be obtained via more marketplace ideas.

Hence, the authors have a prior disposition to a market analysis of educational production. "It is argued that central examinations favor students' educational performance by increasing the rewards for learning, decreasing peer pressure against learning, and improving the monitoring of the education process." (Bishop and Woessmann, 2002) This disposition informs the manner in which data collected by OECD are assessed.

THE QUESTIONS

Without attributing motive to the designers, it is nonetheless plausible to assert that similar considerations led to the design and implementation of the PISA study. Certainly, there is room for criticism of the methodology, and therefore, for questioning the results obtained.

As noted above, the PISA survey departs from previous surveys in disregarding the stated curricula of the schools being measured. As Prais (2003) notes, "the stated focus was ostensibly distinct from details of the school curriculum, and was intended to elucidate how pupils might cope in real life with the help of what they have learnt." It is not clear, however, that the resulting set of questions is any more or less 'real life' than the school curricula. Moreover, the selection of an arbitrary set of "international" questions biased the results against countries which pursued different curricular objectives.

British students did well on the PISA tests. By contrast, in previous tests, which involved (for example) basic subtraction, they performed poorly. Prais (2003) argues (reasonably), "the kind of mathematics questions asked in PISA were deliberately different from those in earlier surveys, and were ostensibly not intended to test mastery of the school curriculum." And he suggests that the tests measured common sense rather than mathematical skill.

Despite the assertions of Prais along with Woessmann and Fuchs, it may be that the PISA test did not test "real life" applications at all. Adams (2003), for example, argues, "It is also quite explicitly stated that authentic settings are not primarily focused on day-to-day (everyday) applications of mathematics. Instead, the primary focus of PISA is on the ability to apply mathematical knowledge and thinking to a whole variety of situations." That would explain the question about seals. But the main criticism remains intact: insofar as the test ignores stated curricula, it ignores the intended output of the educational system, and can hardly thereby be said to be a measure of it.

THE SAMPLE

As mentioned previously, the sample surveyed students at a particular age, rather than students at a given grade level. Woessmann and Fuchs (2004) see this as a benefit. "It captures students of the very same age in each country independent of the structure of national school systems. By contrast, the somewhat artificial grade-related focus of other studies may be distorted by differing entry ages and grade-repetition rules in different countries."

Equally plausibly, however, it is a sample with a built-in bias. For one thing, as Prais (2003) notes, it impacted response rates. Where classes were tightly bound to age, such as in Britain, a larger percentage of students participated, as it resulted in less disruption of classes. Not so in Germany. "For countries where the date of entry is flexibly dependent on a child's maturity, etc., there is a clear difference between the population of pupils intended to be covered."

In addition to skewing participation rates, the measurement by age rather than grade also skews results. Again, the sampling methodology is independent of the intended product of the educational system, so much so that, according to Prais (2003), it creates "a kind of optical illusion without any underlying real change in pupils' educational attainments."

The increased age of the sample population (previous samples were taken at ages 14 and younger) may also skew results. In some nations, weaker students have dropped out of school by age 15. "Full coverage of academically weaker pupils is important if any reliance is to be placed on calculations of average attainments and of the proportion of under-achieving pupils," observes Prais, and it's hard to disagree.

Finally, there was an inconsistency in the school populations sampled. In Britain, students from 'special schools' were excluded. But in Germany, they were included. Adams suggests that Prais assumes without evidence that such students were "lower attaining" - one wonders, however, what else they could be when their own administrators decline the test on the ground that it would be "too challenging".

SMALL CLASSES AND COMPUTERS

One of the surprising contentions of the (Woessmann and Fuchs) study was that small classes did not improve performance. This runs contrary to the assertions of numerous educational groups. For example, Achilles (1997) observes, "4th graders in smaller-than-average classes are about half a year ahead of 4th graders in larger-than-average classes." This oft-cited Tennessee study notwithstanding, there is nonetheless considerable disagreement about the impact of small classes, with studies cited by people such as Kirk Johnson (2000) from the Heritage Foundation arguing that "class size has little or no effect on academic achievement."

The problem with class size is that it is itself subject to numerous determinates. As Woessmann and Fuchs (2004) observe, parents with lower achieving children may move to districts where smaller classes prevail. Moreover, not all small classes are the same: a class may or may not benefit from additional school resources. The influence of external activities may come to bear; Lindahl (2001) compensates for the effect of summer vacation to show that class sizes do have a positive impact.

A similar sort of effect is present with respect to the use of computers. As mentioned above, Woessmann and Fuchs (2004) argue that, "Holding the other family-background characteristics constant, students perform significantly worse if they have computers at home." But let's examine this.

The major variable eliminated in the normalization of the data is parent influence. Of course, this is the major variable - the one variable, it seems, common across all the studies - that is most predictive of outcome. The better off parents are, the more resources a student will have, the more encouragement and support the students will have, and the better the schools students will go to.

The provision of a computer for student use at home is, therefore, part and parcel of a supportive parental environment. Indeed, one wonders about the nature of, and the number of, students with access to numerous computers in poor non-supportive households (one suspects the number of instances is low enough to itself introduce a large degree of error).

That said, eliminating parental influence from the equation is tantamount to measuring the impact of a computer in households with non-supportive parents. No wonder they show no positive impact! Even Woessmann and Fuchs (2004) are willing to concede some ground here: "computers can be used for other aims than learning." Indeed, there appears to have been no effort made to distinguish between educational uses of computing (as in, "computers may not be the most efficient way of learning") and the non-educational use of learning. Given that the authors found a positive correlation between doing one's homework and positive outcomes, one would expect that playing Doom instead of doing one's homework - exactly what we would expect in an unsupportive environment - would have a detrimental impact on performance.

Indeed, in Fuchs and Woessmann (2004) they observe, "At home, the negative relationship of student performance with computer availability contrasts with positive relationships with the use of computers for emailing, webpage access and the use of educational software. Thus, the mere availability of computers at home seems to distract students from learning, presumably mainly serving as devices for playing computer games. Only by using computers in constructive ways can the negative effect of computer provision on student learning be partly compensated for."

What of the assertion that increased computer use at school decreases performance. Once again, we see the same sort of elimination of variables occurring. "Computerized instruction induces reallocations, substituting alternative, possibly more effective forms of instruction. Given a constant overall instruction time, this may decrease student achievement." (Fuchs and Woessmann, 2004)

Given the apparent result, Fuchs and Woessmann offer two hypotheses. First, computer usage may be determined by an ability deficit: teachers assign computers to the better students; this would explain why students who never use computers perform worse. But also, they suggest, "computerized instruction may substitute alternative, more effective forms of instruction, and it may also harm the creativity of children's learning."

But there is a third, equally plausible explanation. Remember, we are treating school resources as a constant. We are also compensating for student ability, such that good students and poor students are considered to be on a par academically prior to calculating the impact of computers in the classroom. Then, we note, that by comparison the students using computers progress less than the students not using computers.

How could this be? Not because students using computers are performing worse - before all the data compensation they appeared actually to be doing better. It must be because giving computers to some students - the good ones - helps the other students - the poor ones - perform better! The use of computers constitutes a reallocation of resources in a class, allowing teachers to concentrate more on the poor students, thus improving their performance substantially.

Only an economist would see this as a net loss. Which gets me to the point of this article.

ECONOMICS PRESUMPTIONS

It is important, when reading studies such as the one cited here, that education and economics are different disciplines. There are risks inherent in imposing the principles of one onto the other.

For example, one might ask why the OECD study focussed on literacy, science and math, and why Fuchs and Woessmann (2004) would limit their enquiry to the impact of computers on outcomes in these areas.

One need not wonder; the authors explain: "the ability to effectively use a computer has no substantial impact on wages. At the same time, they show that math and writing abilities do yield significant returns on the labor market. Thus, they suggest that math and writing can be regarded as basic productive skills, while computer skills cannot." (Fuchs and Woessmann, 2004)

There is not space in this article to review this new data, save to suggest that it is in its own way highly suspect. Neither Wayne Gretzky nor Paris Hilton required a computer education to obtain their substantial incomes, but it would be inappropriate to thereby conclude that computer literacy is not necessary (for the rest of us) in order to achieve a higher income.

More to the point, it is not clear that the maximization of income through work is the ultimate objective of an education, and it is clear (and has been stated above) that satisfaction of OECD's 'real life' competencies is not the stated purpose of various national education systems.

But the assumption of the economics stance produces an even deeper dissonance. The employment of the mathematical interpretation of statistics, as demonstrated in the Fuchs and Woessmann work, produces conclusions that are counterintuitive and in some instances factually incorrect.

To a large degree, economics functions as a science by smoothing differences that are assumed to (according to the principles of economics) make no difference. Discounting the non-economic motivations of an educational policy is but one example of this. Ignoring the stated objectives of the educational system is another. So is the process of compensating for extraneous variables.

But in the evaluation of educational policy, these differences do make a difference. And they do so not merely because education begins from different presumptions than economics, but because the nature of the entities being studied is also different.

Put simply: it is not possible to merely eliminate the influence of one variable or another from the calculation. The presence or absence of one variable has an impact on the nature and effect of the others. Having access to a computer is part and parcel that of parental support. Allowing students to use computers is the same things as freeing teacher time for other work.

It's like trying to describe the relation between the Moon and the Sun by factoring out the influence of the Earth. After the variations of the Moon's orbit around the Earth are smoothed, the path of the Moon appears to be a simple line around the Sun. Economists would conclude that the Moon orbits the Sun. But of course this is simply not so; it orbits the Earth - something that cannot even be considered when the earth is removed from the equation.

SOME FINAL REMARKS

So what can we conclude from the study?

Probably this: that a computer, all by itself, considered independently of any parental or teacher support, considered without reference to the software running on it, considered without reference to student attitudes and interests, does not positively impact an education.

Stated thus, the conclusion is not surprising, nor even wrong. It is like saying that, without the Earth, the Moon orbits the Sun. But it ignores the much more complex reality.

Unfortunately, such fine distinctions are missed in the reporting of results. Hence we read, "computers don't help people learn" and "computers make people dumb." even flawed and skewed as it is, the study reaches no such conclusion; and when the biases are taken into account, it is hard to draw any conclusions at all from the study.

The population as a whole - let alone legislators - is ill served by such studies and such reporting. It is indeed hard not to conclude that the conduct of such research is intended, not to assist, but to skew public understanding of such complex subjects. Insofar as it is the purpose of the press to correct misunderstandings in the public mind (and one wonders these days) a more thorough and critical analysis of such work would be strongly recommended.

BIODATA AND CONTACT ADDRESSES OF AUTHOR

Stephen Downes Born in Montreal, Quebec, Stephen Downes lived and worked across Canada before joining the National Research Council as a senior researcher in November, 2001. Currently based in Moncton, New Brunswick, at the Institute for Information Technology's e-Learning Research Group, Stephen has become a leading voice in the areas of learning objects and metadata as well as the emerging fields of weblogs in education and content syndication.

Stephen is perhaps best known for his daily research newsletter, OLDaily (short for Online Learning Daily), which reaches thousands of readers across Canada and around the world. His work also includes the development of educational content syndication systems such as Edu_RSS and DLORN along and the design of a digital rights management system for learning resources. Stephen is also frequently to be found the road giving seminars and

lectures on the field of online learning, including the notable Buntine Oration delivered in Perth, Australia, in October, 2004.

Prior to his employment with the National Research Council Stephen worked as an information architect at the University of Alberta, hired under contract to design and build a one-stop knowledge, learning and resource portal for the province's municipal sector, MuniMall. While with the University of Alberta he also contributed his unique approach to information architecture to projects such as the Alberta Motion Picture Industries Association (AMPIA) portal and to PEGGAsus, a learning and resource service for petroleum engineers, geologists and physicists. Stephen also worked as an instructional and new media designer at Assiniboine Community College in Brandon, Manitoba, designing the College's website and launching some of the first online courses in the country. Stephen lives in Moncton with his wife, Andrea, and three cats, Bart, Polly and Nadia.

Stephen reads a lot of science fiction and books in general. He enjoys movies, television and sports, especially hockey. He enjoys playing darts and NTN Trivia (though since Moncton has no NTN locations, it's a lot harder to play. He is an avid Civilization III player and an enthusiastic photographer.

For details see: <http://www.downes.ca/me/index.htm>

E-mail: stephen@downes.ca

REFERENCES

Achilles, Charles M. October, 1997. *Small Classes, Big Possibilities*. The School Administrator: American Association of School Administrators.

http://www.aasa.org/publications/sa/1997_10/achilles.htm

Adams, Raymond J. 2003. Response to 'Cautions on OECD's Recent Educational Survey (PISA)'. *Oxford Review of Education*, Vol. 29, No. 3, September 2003.

http://www.pisa.oecd.org/Docs/Download/adams_response_prais.pdf

Bartholomew, James. November 22, 2004. *Computers in schools damage student attainment*.

<http://www.tq-enterprises.com/bartholomew/2004/11/computers-in-schools-damage-student.html>

BBC News. Doubts about school computer use. November 24, 2004.

<http://news.bbc.co.uk/1/hi/education/4032737.stm>

Bishop, John H. and Woessmann, Ludger. 2002. *Institutional Effects in a Simple Model of Educational Production*. IZA Discussion Paper No. 484.

<http://www.iza.org/index.html?lang=en&mainframe=http%3A//www.iza.org/iza/en/webcontent/personnel/photos/index.html%3Fkey%3D621&topSelect=personnel>

Dillon, Sam and Schemo, Diana Jean. November 23, 2004. *New York Times*. Charter Schools Fall Short in Public Schools Matchup.

<http://www.nytimes.com/2004/11/23/education/23charter.html?adxnnl=1&oref=login&adxnnlx=1101834123-vijC3np3LH29aohFY8bitQ>

Fuchs, Thomas, and Woessmann, Ludger. 2004. *Computers and Student Learning: Bivariate and Multivariate Evidence on the availability and use of computers at home and at school*. CESIFO Working Paper number 1321.

http://www.ifo.de/pls/ifo_app/research_output.abstract?p_id=9359&p_base=DLCI

Johnson, Kirk A. June 9, 2000. Do Small Classes Influence Academic Achievement? What the National Assessment of Educational Progress Shows. Center for Data Analysis Report #00-07. The Heritage Foundation. <http://new.heritage.org/Research/Education/CDA00-07.cfm>

Lindahl, Mikael. 2001. Home versus School Learning: A New Approach to Estimating the Effect of Class Size on Achievement. IZA Discussion paper number 261. <http://netec.mcc.ac.uk/WoPEc/data/Papers/izaizadpsdp261.html>

Organisation for Economic Co-operation and Development (OECD). 2000. Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy. Paris: OECD.

Prais, S.J. 2003. Cautions on OECD S Recent Educational Survey (PISA). Oxford Review of Education, Vol. 29, No. 2, 2003. <http://www.pisa.oecd.org/Docs/Download/prais.pdf>
Subtext. October, 2004. What works in education - PISA revisited. http://www.educationforum.org.nz/documents/e_newsletter/10_04/Oct04_Pisa.htm

Woessmann, Ludger and Fuchs, Thomas. September, 2004. What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data. IZA Discussion Paper No. 1287; CESifo Working Paper Series No. 1235. <http://ideas.repec.org/p/ces/ceswps/1235.html>

BIODATA AND CONTACT ADDRESSES OF AUTHOR

Stephen Downes born in Montreal, Quebec, Stephen Downes lived and worked across Canada before joining the National Research Council as a senior researcher in November, 2001. Currently based in Moncton, New Brunswick, at the Institute for Information Technology's e-Learning Research Group, Stephen has become a leading voice in the areas of learning objects and metadata as well as the emerging fields of weblogs in education and content syndication.

Stephen is perhaps best known for his daily research newsletter, OLDaily (short for Online Learning Daily), which reaches thousands of readers across Canada and around the world. His work also includes the development of educational content syndication systems such as Edu_RSS and DLORN along and the design of a digital rights management system for learning resources. Stephen is also frequently to be found the road giving seminars and lectures on the field of online learning, including the notable Buntine Oration delivered in Perth, Australia, in October, 2004.

E-mail: stephen@downes.ca
<http://www.downes.ca/>