Next Big Thing in Education: Small Data

Semester 1 2017

www.cpl.asn.au/journal

Pasi Sahlberg and Jonathan Hasak

First published in Washington Post, 9 May 2016

One thing that distinguishes schools in the U.S. from schools around the world is how data walls, which typically reflect standardized test results, decorate hallways and teacher lounges. Green, yellow, and red colors indicate levels of performance of students and classrooms. For serious reformers, this is the type of transparency that reveals more data about schools and is seen as part of the solution to how to conduct effective school improvement. These data sets, however, often don't spark insight about teaching and learning in classrooms; they are based on analytics and statistics, not on emotions and relationships that drive learning in schools. They also report outputs and outcomes, not the impacts of learning on the lives and minds of learners.

After The No Child Left Behind Act became law in 2001, education legislation in the U.S. required all students in grades 3 to 8 each year and once in high school to be tested in reading and mathematics using external standardized tests. On top of that states had their own testing requirements to hold schools and teachers accountable. As a result, various teacher evaluation procedures emerged in response to data from these tests. Yet for all of these good intentions, there is now more data available than can reasonably be consumed and yet there has been no significant improvement in outcomes.

If you are a leader of any modern education system, you probably care a lot about collecting, analyzing, storing, and communicating massive amounts of information about your schools, teachers, and students based on these data sets. This information is "Big Data," a term that first appeared around 2000, which refers to data sets that are so large and complex that processing them by conventional data processing applications isn't possible. Two decades ago the type of data education management systems processed were input factors of education system, such as student enrolments, teacher characteristics, or education expenditures handled by education department's statistical officer. Today, however, Big Data covers a range of indicators about teaching and learning processes, and increasingly reports on student achievement trends over time.

Despite the outpouring of data, international organizations continue to build regional and global data banks. Whether it's the United Nations, the World Bank, the European Commission, or the OECD today's international reformers are collecting and handling more data about human development than before. Beyond government agencies, there are global education and consulting enterprises like Pearson and McKinsey that see business opportunities in Big Data markets.

Among the best known today is the OECD's Programme for International Student Assessment (PISA) which measures reading, mathematical, and scientific literacy of 15-year-olds around the world. OECD now also administers an Education GPS, or a global positioning system, that aims to tell policymakers

where their education systems place in a global grid and how to move to desired destinations. OECD has clearly become a world leader in the Big Data movement in education.

Despite all this new information and benefits that come with it, there are clear handicaps in how Big Data has been used in education reforms. In fact, pundits and policymakers often forget that Big Data, at best, only reveals correlations between variables in education, not causality. As any introduction to statistics course will tell you, correlation does not imply causation. Data from PISA, for example, suggests that the "highest performing education systems are those that combine quality with equity." What we need to keep in mind is that this statement expresses that student achievement (quality) and equity (strength of the relationship between student achievement and family background) of these outcomes in education systems happens at the same time. It doesn't mean, however, that one variable would cause the other. Correlation is a valuable part of evidence in education policymaking but it must be proved to be real and then all possible causative relationships must be carefully explored.

The problem is that education policymakers around the world are now reforming their education systems through correlations based on Big Data from their own national student assessments systems and international education data bases without adequately understanding the details that make a difference in schools. A doctoral thesis in the University of Cambridge, for example, recently concluded that most OECD countries that take part in the PISA survey have made changes in their education policies based primarily on PISA data in order to improve their performance in future PISA tests. But are changes based on Big Data really well suited for improving teaching and learning in schools and classrooms?

We believe that it is becoming evident that Big Data alone won't be able to fix education systems. Decision-makers need to gain a better understanding of what good teaching is and how it leads to better learning in schools. This is where information about details, relationships and narratives in schools become important. These are what Martin Lindstrom calls Small Data: small clues that uncover huge trends. In education, these small clues are often hidden in the invisible fabric of schools. Understanding this fabric must become a priority for improving education.

To be sure, there is not one right way to gather Small Data in education. Perhaps the most important next step is to realize the limitations of current big data-driven policies and practices. Too strong reliance on externally collected data may be misleading in policy-making. This is an example of what small data look like in practice:

1. Reduced census-based national student assessments to the necessary minimum and transfer saved resources to enhance the quality of formative assessments in schools and teacher education on other alternative assessment methods. Evidence shows that formative and other school-based assessments are much more likely to improve quality of education than conventional standardized tests.
2. Strengthened collective autonomy of schools by giving teachers more independence from bureaucracy and investing in teamwork in schools. This would enhance social capital that is proved to be critical aspects of building trust within education and enhancing student learning.
3. Empowered students by involving them in assessing and reflecting their own learning and then incorporating that information into collective human judgment about teaching and learning (supported by national big data). Because there are different ways students can be smart in schools, no one way of measuring student achievement will reveal success. Students' voices about their own growth may be those tiny clues that can uncover important trends of improving learning.

Edwards Deming once said that, "without data you are another person with an opinion." But Deming couldn't have imagined the size and speed of data systems we have today. Automation that relies on continuously gathered data is now changing our daily lives. Drivers today don't need to know how to use maps anymore when they can use smart navigators that find them the best routes: airline pilots spend more time flying on autopilot than by hand. Similar trends are happening in education systems with countless reformers trying to "disrupt" schools as they are.

Big Data has certainly proved useful for global education reform by informing us about correlations that occurred in the past. But to improve teaching and learning, it behooves reformers to pay more attention to small data – to the diversity and beauty that exists in every classroom – and the causation they reveal in the present. If we don't start leading through small data we might find out soon enough that we are being led by big data and spurious correlations.

Jonathan Hasak, based in Boston, is working to change public policies to better support youth who are disconnected from the labor market and disengaged from school. Follow him on twitter @JonathanHasak

Available online at Pasi Sahlberg Blog: http://pasisahlberg.com/next-big-thing-education-small-data/