

The Edumometer: The commodification of learning from Galton to the PISA

Michael Corbett

Acadia University, Wolfville, Nova Scotia, Canada

Abstract

This essay is a critical reflection on educational standardization, particularly the standardization of assessment flowing down from major international assessment projects like the OECD's Project for International Student Assessment to national and sub-national school systems. I argue that the establishment of an increasingly uniform metric to measure educational performance has developed out of the mental measurement movement from the 19th century and scientific management from the early 20th century. As this movement is articulated with commodification and neoliberalism, I use the image of an ongoing quest for a standardized educational measurement instrument similar to a thermometer or a speedometer. I call this quest the search for an edumometer and argue that it is by now well advanced and draw on Canadian examples to illustrate the argument.

There is a definite social relation between men that assumes, in their eyes, the fantastic form of a relation between things. In order, therefore, to find an analogy, we must have recourse to the mist-enveloped regions of the religious world. In that world the productions of the human brain appear as independent beings endowed with life, and entering into relation both with one another and the human race. So it is in the world of commodities with the products of men's hands. This I call the Fetishism which attaches itself to the products of labour... (Marx, Capital Volume 1, Part 1, Section 4)

Introduction: The quest for the edumometer

I am making a rather audacious claim here that education is hamstrung on a set of myths that were developed in the 19th century. These myths have in an important sense built modern educational systems and in many ways, they form the foundation of how we think about schools and learning. Specifically, I want to challenge the now

sacrosanct idea that learning and intellectual capacity can and should be measured like any other quantifiable commodity. For more than a century, educators and educational thinkers of all stripes have been searching for an elusive holy grail, the edumometer. The edumometer is the technology, and more recently the set of technologies, which will, once and for all, provide an objective measure of what a person has learned and what a person is capable of learning. For the most part, the public has bought into (or has been sold) the idea promoted educationalists, policy people and administrators, lobbyists, corporate test manufacturers and other interested parties, that test results reveal the truth about what children have learned. Those who resist tend to be people close to the working surface of life in diverse schools, i.e. some teachers, their organizations and unions, and in some cases parent groups representing schools and children that do not fit the norm.

Standardized tests that have been developed to measure and grade intellectual and academic performance are a multinational business that has its tentacles so far inside schools and education systems that there is probably no turning back in the foreseeable future. The development of these tests gives shape and meaning to another whole industry of university based and corporate research and entire fields of intellectual inquiry that have been busily producing knowledge about how to measure learning and how the brain works more generally, that to imagine the whole thing grinding to a halt is virtually unthinkable. There are legions of specialist working (school psychologists, learning disabilities specialists, psychometrists, etc.) in schools whose entire professional mandate is shaped by the assumption that learning and intellectual “capacity” can be measured. Finally, teachers themselves have been trained in and work in administrative structures which routinely and unproblematically grade and sort children and youth on the basis of technologies of mental measurement that have come to characterize what school is all about for everyone involved.

Recently, the search for an edumometer has moved out of the cloistered spaces of the universities and national and sub-national education systems and into realm of the private lobby group. Groups that favour the privatization of education like Canada's Fraser Institute have begun to compile and present analyse of educational “performance” essentially based on the principles of financial accounting. While there

is some debate about which forms of data ought to be used to constitute the line items on the educational balance sheet, there is no debate about the existence of an objectively measurable educational bottom line. So these think tanks follow the lead of Britain and the United States publishing league tables that single digit or letter rankings of all secondary schools in a given jurisdiction.

The argument that I will present in the following pages attempts to raise critical questions about the search for the edumometer and the largely unproblematic acceptance of educational measurement in public discourse.[1] I want to encourage us to imagine schools without tests, without grades, and that do not devote so many scarce resources (in the case of public schools) to finding out what parents, teachers, relatives and neighbors to a large extent already know and that is that there are differences between children. These differences can be measured with contrived instruments of assessment, but what do they tell us that was previously hidden? I want to argue that good educational assessments provide us with thick descriptions of how particular children learn. I want to argue that good educational assessments are predicated on deep knowledge of individual children and on how to support those children's growth. I want to argue that only by understanding the history of current assessment practices can we begin unravel the mess we have made of the public schools and the children and families who must endure them.

I do this because as a teacher I have been troubled by too many conversations with parents who are fixated on the notion that tests exist that could tell them the final truth about their children. These parents and the children themselves understand school as a series of measured high stakes hurdles or ordeals to be conquered within preordained time frames, and conquered at a particular standard so that the child may move up the ladder with his peers. I have also been troubled by conversations with colleagues, educational bureaucrats, support people, administrators and politicians who are also scrambling toward the same kind of standardizable educational measurement, even though many of them recognize the myriad problems, nonsense, and casualties associated with the whole enterprise. What results for all of this is tremendous frustration and very predictable failure for most of those children whose parents do not possess what Pierre Bourdieu (1984) called the “cultural capital” to support their children through the minefield of measurements that school represents. I understand

that educational assessment must take into consideration the life circumstances and social context of each individual child and that comparison of children who live in unique family and community circumstances are both inappropriate and damaging regardless of how entrenched and common-sensical it seems.

We now have a research base that extends back three generations demonstrating that educational performance is predicted by two key linked social attributes of children: parental income and parental education (Riordan, 2003, Berliner, 2006). I have spent my career as a professional teacher working in communities in which most people did not acquire very many formal educational credentials. These are also communities that contain a high proportion of low income residents. Parents and children themselves cannot critique the process, they are caught in it, and they are forced to perform or lose. Because they themselves lack the cultural capital represented by socially powerful professional middle class diction, grammar and literacy, they often feel inadequate and uneducated themselves. These parents are themselves victims of “objective” educational assessments that demonstrated to them, from a very early age and in quantitative term, their deficiencies.

So the system that dammed, miseducated and silenced them as children returns to them in adulthood with similar assessments of intelligence and academic performance for their own children. This is what Bourdieu called “symbolic violence” and it represents one of the most powerful ways the education system discriminates against identifiable populations in which school failure has become endemic. The sad irony is that the processes of education including edubabble and the pseudoscience of assessment appear to those farthest outside the process as a sacred mystery. Those middle class parents who understand systems of formal education from the inside understand both the educational discourse and the fact that it is just discourse and not the final truth about learning. They also understand as Annette Lareau showed in *Unequal Childhoods* (2003), that the “system” can be manipulated and controlled by strategic intervention and that people working in the education system can be influenced.

A look inside the head: Enter psychology

The history of mental measurement actually began with a literal measurement of the capacity of the head.[2] The foundational assumption of phrenology is that intelligence is directly correlated with the girth of the cranium. The theory went this way: those who have large heads also have large brains and the bigger the brain, the greater the intelligence. Through the investigations of the phrenologists White male European scientists learned that they were more intelligent than women, Blacks and Asians because their head size was larger. Thus, the science of phrenology supported existing social arrangements, explaining male domination of women, colonialism and the exploitation of half the world by a minority. The important part for the “scientists” was that by reporting (and often fudging) phrenological data, they could use the allegedly objective tools of rational inquiry to show how the exploitation of one group of people by another was not a question of power and domination, but rather a question of natural advantage (Lewontin, 1990).

It is easy to see how the losers in the phrenological game were always the powerless and it was easy for the scientists of the late 19th and early 20th centuries to shape data to support the prevailing structure of social power. For instance, the notoriously small heads of some members of the aristocracy and intelligentsia were conveniently ignored. And of course for a long time the work of the phrenologists appeared to be so obscure and technical that there was little question about what intelligence is in the first place. It probably seemed obvious: intelligence was the ability to get rich, run businesses, rise in the emerging bureaucracy, manage the empire, the military, the universities, the state, and generally, to hold power. Those who were rich and who held power were intelligent; their success proved it. So the task was essentially to show that these people have bigger heads. This, in the end, was what phrenology tried to do until it finally bogged down in its own contradictions and outright fabrications.

The phrenology farce was followed by the quintessentially sexist practice of a man sitting on park bench rating the beauty of women (Gould, 1982). The man was Francis Galton, eugenicist and founding father of intelligence testing. It was his particular aim to show that anything could be measured and if we kept careful records and took proper care to keep our data in order, we would be able to make predictions and essentially control our world better than we could otherwise. So Galton sat on park

benches and rated women using a punch card. The punch card allowed Galton to record the level of beauty of each woman who passed him quickly and accurately in columns. The result of an hour of this work provided him with a ready-made bar graph. Galton realized that if he simply speculated that the women in Paris were more beautiful than the women in London or Glasgow, he could be accused of making a subjective judgment. But, if he could show evidence of a mighty accumulation of judgments, he might be able to make the case that his is an objective study of the phenomenon at hand.

From these humble beginnings Galton launched the mental measurement movement and the field of social statistics. The core idea is that complex human characteristics, abilities, values, attitudes, capacities, inclinations, opinions, attitudes and a host of other non-standard perceptions and human states that we had previously imagined to be unmeasurable came to be laid on a grid for quantification, inference, and comparison. And one of the first questions the quantifiers of qualities asked themselves was: how can we measure learning and intellectual capacity? This simple question catapulted Westerners' sense of what learning is into the realm of the commodity the ongoing project of measuring the contents of the heads of the powerless began. Once begun, this process has changed, but it has never looked back.

The first major systematic attempt at mental measurement was perpetrated on soldiers using the instruments designed by another guy named Alfred Binet who advanced the rather bizarre pseudoscience of phrenology from physical measurement into the realm of what we currently understand as testing. Tests of one kind and another had been around at least since the 11th century when European monks first “graded” acolytes in terms of their suitability for admission to full-fledged brotherhood and teaching in the monastery (Ball, 1990). What Binet and his collaborator Theodore Simon[3] did was to once again take Galton's lead and carefully quantify things. In essence his intelligence test was the first intellectual thermometer that measured performance of particular tasks that were thought to demonstrate pure intellectual ability or what he called judgment.

It seems to us that in intelligence there is a fundamental faculty, the alteration or the lack of which, is of the utmost importance for practical life. This faculty is judgment,

otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. (Binet and Simon, 1916)

Binet's work was sponsored by a research institute established to develop a method for spotting and sorting children in need of alternative education. So what Binet and Simon did was develop an abstract test that could be done at some level by virtually all children. The test consisted of thirty tasks of increasing difficulty and by testing a large number of children Binet and Simon developed norms of performance for each age of child. They called these norms the “mental age” because they represent the level of ability of a statistically normal child falling in the middle of the range at a particular age-stage.

The development of the Binet-Simon test led quite naturally to other increasingly sophisticated tests of intellectual capacity and the industry that produces them is a powerfully entrenched lobby that services most contemporary educational systems public and private. Once the logic of the intellectual thermometer was established and legitimated, the educational testing industry was born. The only remaining task, and it is a massive ongoing task, was to continually improve, refine, and most importantly, to differentiate the edumometer to measure a wider variety of phenomena.

It is important to remember the context or the temper of the times in the early years of the 20th century. The latter decades of the 19th century were marked by the rise of positivist science and the application of experimental methods and the idea of objective observation from the natural world to the social world. Just as natural science was providing a non-magical systematic way of understanding natural processes (such as the chemical composition of matter, the principles that governed motion, natural history and geology, biology, etc.), nascent social sciences were in the process of attempting to develop similarly structured ways of understanding social “nature.” Social sciences such as sociology, psychology, political science, developed through this period and grew out of established forms of inquiry such as philosophy, philology, economics, political economy, and theology.

In 1909 F. W. Taylor published *The Principles of Scientific Management* establishing the field of scientific management. Scientific management was an attempt to use the positivistic logic of the emerging natural sciences to understand questions of industrial

productivity. Taylor studied and measured the micro-movements of efficient workers, those capable of consistently high levels of production and then sought to modify workplaces and teach less productive workers to be more like productive ones. Taylor effectively developed what he understood to be an objective and rational science of work measurement and workplace modification to enhance productivity. The story is well known, as is the story of the limitations of scientific management given that workers are not production machines and that by paying attention to their humanity, production is generally improved.

At this point it is appropriate to introduce another figure, Franklin Bobbit. Bobbit saw the potential for Taylor's ideas in an educational context and in his opus, *The Curriculum* (1918) he developed a framework for a rationally organized curriculum which would analyse the actual activities adults undertake in their ordinary affairs and break these behaviours down into sets of teachable skills. By atomizing and rationalising ordinary tasks, children could be taught complex tasks in lock-step fashion in the context of a rational, scientifically managed fashion. With this development we encounter the establishment of the modern school curriculum. While Bobbit's focus on the importance of training children in the mastery of practical life skills would shift in and out of fashion as the 20th century unfolded, what remained constant was the notion that a curriculum can and should be a rational mechanism for teaching relatively simple skills and measuring the results using some kind of standard calculus.

And the rest is history: educational history. From Bobbit's early work in curriculum studies and the subsequent developments in the psychology of learning assessment, we have never really looked back from the general idea of the edumometer. Notwithstanding historical and ongoing liberal experiments like the progressive schools dating from the early decades of the 20th century and the deschooling movement of the 1960s and 70s, there have been few large scale attempts to challenge the general hegemony of the technical-rational view of learning, efficiency, and the idea of quantifiable educational “performance”, particularly in the public schools where calls for public accountability are understandably more intense.

The big picture: enter sociology

Some policy reports on education begin and end with a macro view presenting and analysing large data sets that are supposed to describe educational systems and how they perform. These systems represent organized aggregations of disparate school sites, and they can indeed be viewed from the point of view of aggregated data that produce statistical aerial photographs that at one level show “everything” there is to see at a particular moment. At the same time they tell us nothing. Large scale snapshots say little about the quality of education anywhere. Each aggregated school contains a particular collection of students, teachers, administrators, and neighbourhoods.

These statistical aerial photographs can then be combined with similar images of other schools over a vast terrain. The statistical profile of schools and student performance allows the public to have access to collections of aerial photographs of varying degrees of resolution. These images give ordinary citizens the privileged administrative view of the corporate manager or the military commander. We can then look within our collection of photographs, comparing one shot with another, or we can take all of our pictures and piece them together to form what might be seen as a topographical image of all of the schools over a wide area. Indeed, organizations like the OECD produce images of schooling which are global in scale and represent effective satellite images of national systems of education. Furthermore, we are taught to believe that we are receiving some hidden truth about schools when we are given this aggregated information.

The “big picture” is also used to convince ordinary soldiers that their particular view of the world is inferior and partial. Soldiers are supposed to trust their commanders to make good decisions for two reasons: 1) because commanders have access to a broader scale of information, and, 2) because commanders are trained for systemic decision making rather than the actually doing the work at hand (in this case educating children). Applied to education, the idea here is that decision making cannot be left to educators on the ground working with real children and youth precisely because educational soldiers know children and communities in an intimate way. It is actually the intimate knowledge of local conditions which is problematic in

this vision because ordinary participants see human beings with different abilities living in different contexts.

The aggregated view is useful to the administrative perspective and in the new era of “accountability,” data is shared with the public in order to justify and support commanders' decision making. Armed with this information ordinary citizens can imagine that a given political jurisdiction contains within it a “school system” that can and should be understood as a unit of production. The large fiction that gets created in this vision is one of a very large manufacturing process whose workings can be measured and captured in quantitative terms simple enough that elementary school aged children can understand them.[4] And indeed the movement to create these kinds of systemic accounts of the work that school systems do is part of a larger umbrella movement that operates under the rubric of accountability. In the present era of globalization and neoliberalism, select slices of the educational lives of children in particular places are measured and compared with those of children across the world. In this matrix, previous accounts given by a parent or a teacher observing a child learning to read, write, ask questions, do mathematics and any of the other things modern children are expected to do as they grow to maturity are considered to be unscientific, unquantifiable and thus, unimportant.

Reports on education seldom describe particular schools, or the particular children who work within them, or the diverse communities in which the schools sit. In fact, most reports tell us about all of the things lacking in these communities, children, teachers and schools. As one Canadian provincial education minister commented in late February of 2004 when test results were released, there is nothing wrong with the curriculum. Apparently, the children or the teachers are the problem.

Indeed most reports on education present aggregated data taken largely from testing that is done on children and youth at particular points in their school careers, removed from the life context of their particular schools, and neighbourhoods. It is as though children in one place, given the same curriculum as children in another place, ought to produce the same results. These data are quite literally “crunched” into condensed pellets of systemic information and from these pellets we believe that we are learning things about schools and the children and teachers who work in them that could not otherwise be discerned.

These data also systematically obscure the differences between children which are normal. Anyone who has worked with children understands that a fourth grade classroom or an eighth grade classroom is full of individuals whose understanding and performance in academic work and in social behaviour is widely distributed. To use the crude and vague language of “grade level,” in a fourth grade classroom we find children who could easily operate in particular subject areas in the ninth grade and others who could easily operate with pre-schoolers. This variance is normal. The bizarre notion that each ten year old child in a given classroom or in a given nation, state, province, city or town ought to be reading the same sorts of material in the same way is a fiction promoted in part by statistical aggregation of large data sets. It is no surprise that the earliest forms of normative mental measurement were conducted not only on students to see which were “educable”, but also on soldiers in order to sort out those unfit for service (Gould, 1982; Postman, 1991). Grade-levelled children are like soldiers and labour forces, they are a mass of humanity constructed for deployment in campaigns. These non-existent normal phantom children are social constructions, fictional pseudo people generated by developmental psychology and statistical norms. They have no parents, interests, inclinations, abilities, wants, hopes, fears ... like all good soldiers they have no personal context and when they are trained properly, they obey and behave predictably according to grade level. When they do not obey, they should be disciplined and even shamed, as should the officers whose job it is to organize the soldiers’ work.

Ironically though what we have learned from several generations of educational research is that context is absolutely crucial to children’s educational success. Statistical aggregation and constructed data stories do not tell tales of school success or failure or of lives made or broken in the school system and in communities. These stories are told in the rich context of a child’s life and there are few surprises. We know which family and community contexts support school success. We know which kinds of school contexts support children’s learning. Sensible observers know that the children who tend to succeed in school are those children who grow up in middle class homes, who live in middle class neighbourhoods, whose parents have been successful in school and higher education themselves, and whose language patterns are middle class and similar to those of their teachers. We also know that those families that use their resources for high status, literacy-rich cultural products like

high quality educational toys, books, theatre and museum visits, after-school lessons, and travel to high cultural hot-spots (as opposed to pop-cultural hot-spots, e.g. Disneyland, resorts, etc.). Millions of dollars have been spent to document this common sense phenomenon all across the western world.

And yes, there are many exceptional individuals and even some exceptional schools like Deborah Meier's Central Park Elementary in New York that buck the trend (Meier, 2002). However, one only has to walk around an urban or suburban neighbourhood or spend a bit of time in a rural community talking to children to discover the power that context, money and culture hold over the intellectual growth of children. This is hard for us to face, because when schools were established in the 19th and early 20th centuries, there was indeed great faith in the potential of the application of efficient technologies of instruction and moral training to perform the alchemy of what Gramsci called turning everyone into bourgeoisie. This residue of 19th century moral training and 20th century scientific management, combined with the idea that schools are the crucible for the preparation of a democratic citizenry have created the great myth that schools can and ought to be fair and to actually generate the conditions for social equality and the unfettered growth and development of all children.

The truth is that schools have never plucked children out of their life circumstances placing them on an even playing field. A quick tour of a modern high school demonstrates how many children wear their social class position on their backs and how identity is embodied in studied performances that have little to do with academics. The tour will also reveal another cadre of youth whose manner and embodied performance is more similar to those of their teachers. It is no surprise that these children perform better on the tests designed to measure strictly academic kinds of performance. It is no surprise that these children can see a future in higher education, a future that represents a well-established path in their families. It is no surprise that the way these children use language is tacitly understood to represent some form of generic "intelligence" by teachers who measure others in terms of their own standards of articulation. It is no surprise that many of the best teachers eventually transfer into schools that contain large populations of already academically successful children because in these schools the work is typically more congenial,

easier and the results on standardized tests are more positive. The flip side of this is that schools that contain relatively large populations of academically unsuccessful students have relatively more young and inexperienced teachers as well as those teachers who are not marketable and mobile for one reason and another. And as increasingly meticulous measures of academic performance are developed and as urban schools become increasingly fragmented by specialized offerings and consumerist school choice schemes the process is likely to continue to polarize further.

It is no surprise that schools in particular neighbourhoods contain larger or smaller groups of academically successful children on a trajectory into higher education and professions. This is the story that the macro view misses. From the aerial view, all of the bodies on the ground look the same and this allows us to maintain the fiction that their lives and their life chances are all the same. In fact, in the Canadian context, most data show a difference that seldom amounts to more than 10% between the results of the worst performing and best performing provinces when the numbers are crunched into statistical pellets. These results actually hide significant and important differences between the educational worlds of children living in different communities.

It is also significant that post-secondary participation appears also to be related to the simple availability of service and perhaps even a lack of easily accessible employment opportunities in economically disadvantaged communities. For instance, in largely rural, “low performing” (in PISA and other standardized and commodified educational measures) Nova Scotia, university attendance is nearly double that of the Canadian average and nearly two thirds of Nova Scotian women between the ages of 18 and 21 were enrolled in university programs (Nova Scotia Advisory Council on the Status of Women, 2002: 18). We do not know what is going on here. It could be that this form of higher education is more easily accessible close to home in Nova Scotia with its large number of universities. We do have recent census evidence that shows massive outmigration from rural areas of Atlantic Canada. It could be that these universities are preparing a massive cadre of young Nova Scotians for outmigration. These are the kinds of complexity and ambiguity we need to understand better (Corbett, 2007). When we look closely, we can see the complexity of the common

sense worlds of teachers and students. We also see how and why we should expect to find enormous differences in standardized test results. It is this difference that is interesting and it can only be explained in a close look at the spaces in which children and youth produce it.

The commodification of learning

In large scale statistical reports, particular schools become what contemporary educational sociologists have called call “black boxes.” The black box is an amorphous school space into which children are inserted and out of which they are eventually extruded as products. From time to time, externally administered measurement instruments are inserted to gauge a particular aspect of children’s progress inside the black box, particularly at the end of what are considered to be key stages in children’s lives. The point is that the activity inside the black box of the school is not generally of concern in this kind of analysis. Black box studies tend to make the leap between the analysis of educational inputs and outputs and infer that better educational processes happen in those schools that produce better outputs. These sorts of studies have consistently demonstrated that the social and personal characteristics of the child who enters the black box not usually altered markedly from the youth who emerges at the end of the process. At their best, black box studies have pressured educational researchers, policy makers, bureaucrats and politicians to look inside the box to see what is going on to produce such predictable results. However, in recent years there has been a move away from qualitative analysis of educational sites in favour of what is constructed as scientific, results-based research that typically ignores and even dismisses context (, 2003).

The result is that schools tend to reproduce the social class structure because children who enter the box from middle class families, with middle class language and social patterns tend to succeed in school and eventually move into middle class and professional social positions similar to those of their parents. The same is true of most working class children. Most children of the wealthy attend private schools which effectively segregate them from children of other social classes, in part insuring their continued exclusive privilege and marriage patterns that protect family money and social privilege. One dangerous implication of this process is that the instruments used to measure the effects of the black box come to unproblematically signify

education. It is effectively assumed that closer monitoring, standardization of curriculum and assessment will improve the quality of the activity inside the box as well as the academic and social outcomes that emerge from it. Commodification is only truly developed when nobody any longer thinks to question the metric and it passes into use and into discourse as though it were the same as the process it measures. Just as money comes to take on a life of its own as a material force, so too does the standardized test score fetishize the measurement instrument conflating it with the process it seeks to measure.

To cite a Canadian example, in 2003, the Council of Ministers of Education Canada (CMEC) published results from their pan-Canadian indicators project. The report provides a considerable amount of data concerning the composition of school and other student populations, educational funding, and a variety of other types of statistical information. The most interesting material for most observers of the educational scene is the reporting of results around what are called learning outcomes or student achievement. The release of reports like this one are media events that often generate days and even weeks of coverage investigating the latest “crisis” in education. Things were quiet this time because this time there is no demonstrable crisis in student achievement. Canadian students are performing near the top internationally in all subject areas tested. It is difficult not to be cynical about this but it does appear that since this report does not show Canada’s educational performance in a bad light, it did not attract much attention. Much the same was the case when the 2006 PISA results were released in 2007. Rather, much of the analysis in the Canadian media shifted to regional differences in performance which vary somewhat, typically around 10% in terms of mean scores in most curriculum areas.

Two central features of educational policy are the increasing propensity to rank schools, allegedly to inform “consumer choice” in education and the measurement of learning outcomes (Portelli and Solomon, 2001), a movement the CMEC calls, “a hallmark feature of education policy in the last ten years or more” (2003: 53). What has emerged in educational policy through the 1990s and into the first years of the current decade is a strong focus on quantifying learning (the discourse of “outcomes”), and a subsequent movement to publish and disaggregate commodified outcome measurement data to the level of the school. The commodification occurs

not just in the transformation of standardized tests into data points, but also in the way that aggregations of these scores at the school level are used to compare schools as though they were educational production sites. It is well known for instance how in many jurisdictions funding is tied to standardized test performance as well as how real estate listings contain detailed information about school catchment areas and the performance of local schools. Schools are literally sold to prospective home buyers.



In many western jurisdictions school rankings are now commonplace. Initially schools were ranked in a relatively unsophisticated way, stacked up in league tables. In many places this is still the case, but in recent years it has become more fashionable to develop increasingly sophisticated statistical procedures to rank and compare schools in a given jurisdiction (Council of Ministers of Education Canada, 2003; Audas and Curtwill, 2003; Marceau and Cowley, 2003). For instance, in Canada, the Fraser Institute has been involved in this enterprise since 1998. In its 2003 “report card” on education the Fraser Institute argued that, “comparisons are the heart of improvement” (Marceau and Cowley, 2003: 4). In Canada this trend has spread to regional nodes. For example, in largely rural Atlantic Canada, a parallel school comparison movement has developed under the auspices of the Atlantic Institute for Market Studies.[5]

Pro-market lobby groups understand the importance of commodification in their ultimate quest to introduce elements of what they characterize as consumer “choice” and ultimately privatization public institutions. If school performance can be measured using a common metric, then is it quite simple to argue that underperforming public schools ought to be given over to private interests for improvement, the same kind of improvement elite private schools and private remedial clinics are alleged to demonstrate. The argument holds that where the public sector has failed, the private sector ought to be given a chance. Without an unproblematic edumometer, it is impossible to sustain the argument. Without a common metric, what we are left with are individual schools operating in diverse contexts educating groups of children. Where educational assessment is nuanced, contextualized and both quantitative and qualitative it becomes obvious that school

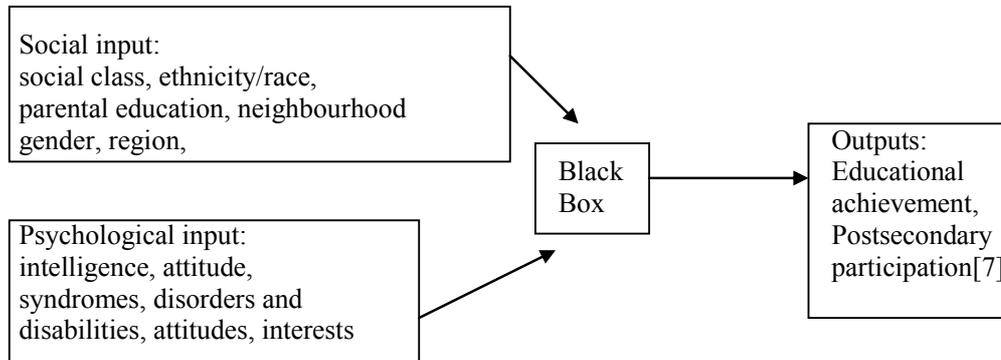
quality is much more complex than any single uniform measure can gauge. Within the context of this market logic, competing metrics, multiple measures, and qualitative investigations of diversity and context have to be shown by neoliberals to be irrelevant, unscientific, and obfuscatory and this is precisely how this form of research and educational assessment is presented by the proponents of the edumometer.

School ranking systems appear to present the performance of schools in a clear and unbiased light by generating objective data about comparative performance. More sophisticated versions of this process attempt to understand that not all schools are equal, however, like must be compared with like. This leads to more complex statistical procedures which allow analysts to differentiate between schools whose “inputs” are weighted to simulate conditions of equality. Here I wish to raise two key concerns with school rankings: 1) the problem of the “black box” which has dogged quantitative educational sociology from its inception, and 2) the related problem of measuring and comparing the performance of schools more generally.

Contemporary analysts of school performance quite rightly point out that it is inappropriate to compare schools that operate in different community, serving different populations in different quality facilities as though they can and should produce comparable results. There is nothing new in this approach. In 1961 James Coleman published the first major large scale quantitative study of schooling in the United States. Using statistical techniques, Coleman (and an army of graduate students) analysed the extent to which social variables such as race, ethnicity, social class, parental education and religion influenced educational outcomes. He discovered that social variables particularly those associated with race and social class were the strongest statistical predictors of student educational achievement (measured by standardized test scores).[6] This seminal work initiated a series of investigations in the sociology of education that sought to replicate, refute, reject, compliment or enhance Coleman’s findings. This project continues to the present day.

In sociological studies like Coleman’s the school is essentially an obscure container that accepts “inputs” and produces “outputs.” Individual children and youth are also erased from the picture, they become a collection of independent variables like race, social class, parental education and religion when they enter the box. They then emerge from a given period of “processing” inside the box as outputs or dependent

variables such as educational achievement or post-secondary participation. The inputs can be social characteristics or variables such as the level of income and education of parents, the race/ ethnicity, or they can be psychological characteristics such as intelligence, disability.



The assumptions in this analytical model are glaring. This is a simplistic industrial production or information processing model that assumes that what goes on in all school is essentially the same. The model also assumes that quantitative sociology and psychology are able to effectively measure and quantify complex human phenomena like intelligence, attitude, social class, the influence of neighbourhoods and how well a person has learned. The truth is that the definition and measurement of all of these “variables” is fundamentally contested and are likely to remain so indefinitely. To paraphrase Danish novelist Peter Hoeg in *Borderliners* (1994), we have no standard metric for distinguishing the value or quality of one sentence from another, one work of art from another, one chair from another. In other words any human activity more complex than, for instance, counting the number of times a ball is kicked into a net cannot be quantified to universal satisfaction. Referring to a lifetime of work in the field of mental measurement, Edward Thorndike pointed to a few key “defects” with intelligence testing: “just what they measure is not known; how far it is proper to add, subtract, multiply and divide and compute ratios with the measures obtained is not known; just what the measures signify concerning intelligence is not known” (cited in Postman, 1993: 131). Something similar could be said about virtually all variables used in quantitative educational sociology, in fact the difficulty of obtaining anything approximating experimental conditions in quantitative sociology makes the kind of methodological problem Thorndike addresses even more problematic.

The assumptions underlying the model used by the school comparison movement are seldom challenged, and when they are, the work is done in obscure journals by critical sociologists, academic educators, and allegedly self-interested groups like teacher organizations. Another ubiquitous yet “hidden” critique of such work is carried out by educators working in the system who understand that there are complexities on the working surface of school which are not at all understood by this kind of comparative analysis.

Recent work in the sociology of education has attempted to bridge the gap between large scale statistical “black box” analysis of schooling and the life world of the schools themselves. In order to understand school “performance” it seems entirely logical that we need to understand something of what is going on inside the black box. Such analysis would allow us to move between a look at what schools “produce” as quantifiable educational outcomes and what actually goes on inside the complex organizations that schools are.

The black box model is important but incomplete and problematic leading to serious misunderstandings, particularly when the analysis is used to compare individual schools. Quantitative educational sociology can provide us with some large scale trends and patterns of a very limited kind of educational achievement. The authors of many powerful and comprehensive analyses in the quantitative sociology of education like that of Coleman and others who have followed his lead understand this very well. They know that they are in the business of providing aerial photographs of very large “systems”. They understand that this system contains many constituent parts about which their results say very little. There is no comparison of individual schools in this work.

What is most disturbing is the way that black box studies have been co-opted by the neoliberal right. The very worst kinds of black box studies are those designed by corporate interests for the purpose of undermining trust in public education. These studies take techniques designed for structural-systemic analysis and inappropriately apply them to individual schools. In this way it is possible to construct rankings and effectively tell a story about the individual schools without ever having to darken the corridors or talk to a single student, teacher or parent. Using data they construct, these interests make the case that they know more about teachers and students and the

schools in which they work than do the players themselves. They can argue to parents, using allegedly “hard data,” that they have uncovered some essential truths about their children’s schools that are hidden from us by bureaucrats who want to keep the public in the dark. And softened up as we are by a barrage of polling and data we easily accept the maxim that the numbers reveal that which our eyes cannot see. Of course, the public school system like all public institutions is portrayed as an unmitigated disaster. Indeed, as one analyst found, the strategy is deliberately sensationalist. Quoting the head of Britain’s Adam Smith Institute, George Monbiot (2004) writes: “We propose things,” Madsen Pirie once boasted, “which people regard as being on the edge of lunacy. The next thing you know, they're on the edge of policy.”

All of this slides nicely into our daily lives where we as consumers are forced/schooled to become expert quantifiers looking for bargains and knowing exactly what everything is “worth.” Comparing prices and values is what we do in order to survive as consumers and as workers. When a group presents us with an analysis of our schools as simple to read as a McDonald’s menu, we thank them for having made it simple for us. We have a symbolic metric which is clear and unequivocal. We don’t really want to look very hard at our schools, or engage in complex conversations about how to improve them and support the learning of our children and those children down the road, never mind the children of the working poor and those on social assistance whose children are essentially doomed to educational failure. We want information that is simple, digestible and reducible to a common metric operating within a linear market. This is the ultimate effect of the process Marx described so long ago. All human relationships are now measurable and comparable. We rate everything from parenting skills, to love-making to happiness and caring, to education as though we were reporting batting averages or goal totals. We have arrived.

It seems to me that a simple menu is not appropriate in the analysis of schools. Schools are not black boxes, they sit in communities that are very different, they contain teachers and students who interpret and enact curriculum in very different ways, and they contain very different physical spaces and are equipped with different materials and resources. It is like imagining a series of very differently designed and

structured automobile factories that use different raw materials, each piece having a mind of its own, as do the production workers who each work alone with the raw materials for 1/30 of the production time, nine or so months of the year. There is and there can be, no instrument capable of measuring this. The instruments that we currently possess are woefully inadequate because they largely ignore the crucial components of educational performance, lived elements of social context and experience.

Then there is the question: Who are “we” anyway? Not all of the larger democratic “us” have the time, the leisure, the requisite education, the money and the inclination to pay careful attention to school options as a serious matter of choice.[8] In families where education has historically been an ordeal and a chore to be abandoned as soon as possible, parents tend to have little interest in shopping for educational options. A research base is now emerging that shows how what is commonly called “school choice” and more extreme neoliberal forms of education like voucher systems amount to allowing middle class parents to remove their children from schools that have high proportions of the children of the working class. The result is that increasingly fragmented schools are segregated by a collection of individual “choices” supplementing that defacto school segregation that is already accomplished by real estate markets.

So obviously, one of the things lacking in most statistical, documentary analyses, as well as in the psychological in-the-head studies of learning is the lived sense of what things are like in particular schools. If we agree that it is inappropriate to compare very different contexts, then questions of how to support quality education shift into a different register. No longer can we sit in the command post with data at our finger tips and imagine that we might know what is best for schools. Accountability operates at the level of the site and the players engaged in the process are principally accountable to each other. The argument here is to orient educational research on what Labaree described in metaphorical terms as a rural landscape. Rurality is micro, it is multifaceted, it resists homogenization of the blending and coordination of the commodified urban space. The metaphorical landscape introduced by large scale standardized assessments such as the PISA and other international and national projects is imagined as homogenous, flat, linear, standard and urban. One can hope

that the intractable differences within and between schools and the diverse people who work within them will challenge commodification, educational markets and resist the current trend toward and increasingly standardized edumometer.

Notes

[1] In the United States and in the Britain there has been an ongoing weak critique of educational standardization that calls into question such national initiatives as the British National Curriculum and No Child Left Behind in the United States. Much of this critique, I would argue tends to accommodate the standardizing structural changes introduced into national education systems as fait accompli. It is important though to note that there has been a persistent group of anti-standardization scholars and lobbyists who have maintained the position that such tests are at least deeply problematic and at worst fundamentally distort educational practices. Some of these critics include Alfie Kohn in the United States, Marita Moll in Canada, and Guy Claxton in the UK. The rise of home schooling, alternative schools and charter schools also represent ways that parents who are critical of educational standardization have responded to and resisted the development of linear educational commodification.

[2] This section sketches briefly what I consider to be some of the main currents in the development of the field of mental measurement and the subsequent development of standardized testing. This movement has been analyzed in a great deal more detail in a number of places. Ladson Billings (1994, 1998) and Gillborn (2006) work for instance details the historical and contemporary ways in which standardized testing has functioned and continues to function as a key legitimation tool protecting white privilege. Lemann's *The Big Test* (1999) also provides detailed analysis of the development of standardized testing in the American context.

[3] Simon was a teacher of the Swiss psychologist Jean Piaget who developed what became for most of the 20th century the leading theory of cognitive development in children and essentially shaped teacher training and curriculum in a psychological register

[4] And let us think about this one for a moment as well. What does it mean to publish school scores that show young children, year after year, that their school and their own performances are substandard on a provincial and national scale. In economically disadvantaged communities and neighborhoods, we can easily predict that this will continue to occur. There is absolutely no evidence to suggest that schools will ever be able to overturn economic inequality and there is massive evidence to suggest that economically privileged neighborhoods and communities do better on standardized tests.

[5] Which of course raises the key question about whether or not schools are markets. The answer is a political one. The work of this group has created the rather bizarre situation where dozens of small community-based rural schools are compared against one another in terms of standard performance indicators. In most of these communities school choice is not an option so the whole specious simply sows fear and loathing in many rural communities struggling to deliver educational services under difficult conditions.

[6] Large scale quantitative studies of educational achievement virtually all use standardized testing results as the dependent variable. The model works like this: independent variables like race, ethnicity, gender, intelligence, social class and other independent variables are correlated with variance in the single dependent variable academic achievement. As far as I know, nobody has treated standardized testing scores (which are simply assumed to be reasonable measures of academic achievement) as an independent variable that influences other kinds of outcomes. For example, it would be interesting to know if the implementation of standardized testing has had any impact on the social class composition of particular communities. Such a study might investigate the impact of the implementation of standardized testing in the Canadian provinces of Alberta or Ontario in the 1990s on particular economically disadvantaged communities, like First Nations reserves, rural communities or economically disadvantaged neighborhoods. A key question in such a study might be: Is there any evidence to show how an increased regimen of standardized has influenced variables such as: graduation rates, the proportion of students enrolling in advanced science courses which are most heavily tested, employment rates for these communities, average incomes in the communities, etc.

[7] Interestingly few studies investigate the actual life outcomes of students. It is as though the entire purpose of schooling is to proceed on to more schooling. Once the process is complete, educational researchers tend to lose interest or lose track of students. The challenges of doing such work are immense, but a qualitative/quantitative look at how students proceed through and beyond formal schooling and through the life course would be fascinating.

[8] The uses made of league tables and school choice information by differently placed families could generate a very interesting set of research questions. It is clear from US research that school choice is “exercised” predominantly by the middle class and the general result is that inner city and schools in economically disadvantaged communities are increasingly ghettoized as middle class children are removed from “underperforming” or “sink” schools. Obviously, school choice requires both mobility and the resources to afford both transportation and school fees. The real result of school choice is an increased level of segregation.

Bibliography

Audas, R. and Curtwill, C. (2003). *Grading our Future: Atlantic Canada's High Schools Accountability and Performance Context*. Halifax, NS: Atlantic Institute for Market Studies.

Ball, S. (Ed.). (1990). *Foucault and education: disciplines and knowledge*. London: Routledge.

Berliner, D. (2006). Our impoverished view of educational reform. *Teachers College Record*, 108, 6, 949–995.

Bobbit, F. (1918). *The curriculum*.

Conference Board of Canada. (1992). *Employment skills profile: The critical skills required of the Canadian workforce*. Ottawa: The Conference Board of Canada.

Corbett, M. (2007). *Learning to leave: The irony of schooling in a coastal community*. Black Point NS: Fernwood.

Corporate Higher Education Forum (1992). *Learning goals for K-12 education: Doing our best*. Montreal: Corporate Higher Education Forum.

Council of Ministers of Education Canada (2003). *Education indicators in Canada: Report of the Pan-Canadian Education Indicators Program 2003*. Ottawa: Canadian Educational Statistics Council.

Hoeg, P. (1994). *Borderliners*. Toronto: Seal Books.

Gillborn, D. (2006). Rethinking White Supremacy: Who Counts in 'WhiteWorld', *Ethnicities*, Vol 6(3): 318–340.

Gould, S. (1982). *The mismeasure of man*. London: Norton.

Labaree, D. (2004). *What's wrong with ed schools*. New Haven: Yale University Press.

Ladson-Billings, G. (1998) 'Just What is Critical Race Theory and What's it Doing in a Nice Field like Education?', *International Journal of Qualitative Studies in Education* 11: 7–24.

Lareau, A. (2003). *Unequal childhoods: Class, race and family life*. Berkeley: University of California Press.

Lather, P. (2003). *This IS Your Father's Paradigm: Government Intrusion and the Case of Qualitative Research in Education*. Guba Lecture, sponsored by AERA Special Interest Group: Qualitative Research, Chicago, April 2003 (www.coe.ohio-state.edu/plather/) Accessed November 2, 2005.

Lemann, (1999). *The big test: The secret history of American meritocracy*.

Lewontin, R. (1990). *Biology as ideology*. Toronto: Anansi.

Nova Scotia Advisory Council on the Status of Women (2002). *Learners and teachers: Women's education and training*. Halifax NS: Nova Scotia Advisory Council on the Status of Women. www.gov.ns.ca/staw/pubs2002-03/learners2002.pdf

Nova Scotia Department of Education (2002) *Learning for life: Planning for student success*. Halifax: Nova Scotia Department of Education.

MacKinnon, D. (2001). A wolf in sheep's clothing: a critique of the Atlantic Provinces Educational Foundation. In Portelli, J. and Solomon, R. P., *The erosion of democracy in education: From critique to possibilities*. Calgary: Detselig, p.p.117-144.

Marceau, R and Cowley, P. (2003). *Report card on Quebec's secondary schools*. Vancouver: The Fraser Institute.

Meier, D. (2002). *In schools we trust: Creating communities of learning in an era of testing and standardization*. Boston: Beacon Press.

Monbiot, G. (2004). On the edge of lunacy, *ZNet*, January 17.
<http://www.zmag.org/Zbios.htm>.

Portelli, J. & Solomon, R. P. (2001). *The erosion of democracy in education: From critique to possibilities*. Calgary: Detselig.

Postman, N. (1993) *Technopoly: The surrender of culture to technology*. New York: Vantage.

Riordan, C. (2003). *Equality and achievement*. Upper Saddle River, NJ: Prentice Hall.

Secretary's Commission on Achieving the Necessary Skills. (1991). *What work requires of schools: A SCANS report for America 2000*. Washington DC: United States Department of Labour.

Taylor, F. (1909). *The principles of scientific management*.

Whitty, G., Power, S., and Halpin, D. (1998). *Devolution and choice in education: The school, the state and the market*. Buckingham/Philadelphia: Open University Press.

Michael Corbett

Author's Details

Michael Corbett teaches in the School of Education, Acadia University, Wolfville,
Nova Scotia, Canada

Correspondence

michael.corbett@acadiau.ca