Version 16/02/2013

Uses and abuses of education assessment in Brazil[1]

Simon Schwartzman

The advantages and pitfalls of education assessment, or, more specifically, comparative and quantitative assessments, have been discussed endlessly, and my personal conclusion is that it is not a question of which is better, but of what world we prefer do live in: the world without education measurements and comparisons, with its issues of imprecise goals of lack of standards, or the world of explicit targets, indicators and statistical measurements, with its problems of unwarranted assumptions, bad measurements, misplaced goals and wrong incentives. If I lived in Scandinavia or, for that matter, Boston, places where the quality of education is beyond question, I would probably choose the first; but I live in Brazil, and have always chosen the second.

[1] 2013 Kneller Lecture, Comparative & International Education Society, Annual Conference, New Orleans, March 10-15, 2013. I am grateful to Cláudio de Moura Castro, João Batista Araújo Oliveira, José Francisco Soares, Maria Helena Magalhães Castro, Renato H. L. Pedrosa and Robert Verhine for corrections, comments and suggestions.

Even for Latin American standards, Brazil is a latecomer in education. Its first professional schools date from the late 19th century, the first universities from the 1930s, and no serious attempt to create a nationwide system of universal public education was tried before the 1950s. In 1970, a third of the Brazilian adult population was still illiterate, unable to read a simple sentence or write their own names. Today, this figure is still 8.6%, and the number of functional illiterates, unable to read fluently and to understand a written text, is much larger[2].

Another Brazilian peculiarity is that, instead of developing its education from the bottom up, starting with primary education from the grassroots, it has always tried to work from the top down, developing first professional and higher education in a limited number of national institutions, and then, gradually, trying to expand it to secondary and basic education and to more people. This is an oversimplified picture, since there has been also local efforts to create grassroots education institutions at all levels in some states, particularly in the South and among immigrants of European and Japanese origin. But, since the 1950s, as the country became more urban and the size of the public sector increased, the dependency of education institutions regarding the national and state governments also increased, in spite of the fact that it was much easier for governments to create education institutions by decree and to impose regulatory legislation on private and local initiatives than to make sure that they did what they were supposed to do.

The assessment of graduate education and research

Not surprisingly, the first experience of external, comparative and quantitative assessment in Brazil was for graduate education and research. In 1968 the Brazilian government, under military rule, tried to reform its higher education institutions by adopting the American system of academic departments and graduate schools. (Law 5540 of November 28, 1968). Starting in the mid seventies, in a period of economic expansion which became known as the "Brazilian miracle", new or renewed agencies such as the National Council for

[2] Data from IBGE, Séries Estatísticas / Séries Históricas, http://seriesestatisticas.ibge.gov.br

Scientific and Technological Development (CNPq), the Financing Agency for Studies and Projects (FINEP) and the Coordination Agency for Support and Evaluation of Graduate Education (CAPES), within the Ministry of Education, started to pour resources into graduate education, research and technology, creating in a few years an impressive network of research institutions in universities and other settings, and trying to redress the opposition and mistrust that existed between the government and the scientific and academic communities since the military coup of 1964. Not by chance, two of these agencies, CNPq and FINEP, were brought under the Ministry of Planning, and for them the goal of these investments was to modernize the economy and strengthen the government's military capabilities. They believed in the power of technology and in technology planning, and put out two successive ambitious National Plans for Scientific and Technological Development (Schwartzman 1991; Schwartzman 1994). But only the scientists knew how science had to be organized to flourish, and the model they had in mind was the combination research university and science support agencies, which they knew from their experience in the United States and some other places. Most of these scientists worked in universities, organizing the country's first graduate programs, and helped CAPES to establish Brazil's first system of quantitative and comparative assessment of education[3].

The incentives for graduate education, including higher salaries for faculty with master's and doctoral degrees, research grants, fellowships and institutional support, created strong incentives for institutions to create degree programs with low academic standards, particularly outside the mainstream scientific disciplines such as physics or the biological sciences. This alarmed the scientists and academics in the best institutions, who worried that their resources would be wasted and their standards lowered.

[3] At the same time, Brazil was investing in some ambitious and, ultimately, failed attempts at high technology development in areas such as nuclear power, space technology and computers, which remained mostly outside the university circles. Another significant development in those years was in agricultural research, led by a special agency, EMBRAPA.

The assessment system developed by CAPES was an ingenious combination of two elements, peer review and the collection of systematic data such as staff qualifications, publications, number of students and degrees granted by the graduate programs (Castro and Soares 1986). Advisory Committees were established for each main field of knowledge with invited prestigious scientists, and the Ministry provided them with the information from the graduate programs. To increase the quality of the information, the advisers also visited the programs to get first hand impressions. Periodically, the committees would meet to access the programs in their fields, comparing the information received with their own knowledge, and ranking them in a seven-point scale going from to unacceptable to world-class. Those programs considered best received fellowships for their students and additional resources, while those considered unacceptable did not receive such benefits. Since the universities were autonomous, CAPES did not have the formal authority to close the unacceptable programs, but could use incentives to induce the universities to act, either by closing themselves down, or by bringing themselves up to acceptable standards.

The experience was so successful that the rankings developed by CAPES became the standards for other science and technology agencies, and the system established in those years is still in place. One explanation of its success was the legitimacy gained by the system through the peer review process, thanks to the careful selection of its members, and the established policy of the Ministry of Education of not overruling their recommendations. The other explanation was that the number of graduate programs under evaluation – about one thousand in 1981 – was still manageable.

Over time, as the graduate education and research sector expanded, the assessment procedure became more bureaucratic and more dysfunctional. Today, there are about 3 thousand graduate degree programs, with 52 thousand professors and 190 thousand master's and doctoral students, divided among 48 areas or disciplines, from traditional fields such as physics, biology and law to subjects such as arts and music, environment sciences, nursing and "interdisciplinary" study, for new subjects that could not fit the traditional categories. The selection of peer reviewers used to be based on reputation, at the

discretion of CAPES's authorities. Now the coordinator of each committee is chosen through nominations by the departments and scientific associations, therefore excluding more controversial personalities, and he in turn selects his peers. Scientific productivity is now measured by a complicated ranking system of scientific periodicals, developed from inputs from the researchers, called "Qualis"[4] which leads to rigid bibliometric algorithms that cannot be challenged by the peer reviewers. The assumption is that the publication patterns in all areas are similar to those prevalent in the natural sciences, which is far from obvious. The whole system is supervised by a "council of councils" (the Scientific and Technical Council, CTC) with power to make recommendations, consider appeals and eventually reform the decisions of specific advisory groups.

The end result is that, along with its achievements, the CAPES system became, to a large extent, a self-serving mechanism to sanction and validate the Brazilian academic establishment as it stands, with little space for change and innovation (Schwartzman 2010b). With a few exceptions, most research and graduate education takes place in fully subsidized public institutions. The number of scientific papers published by Brazilian scientists has expanded, but their quality, as measured by citation statistics, remains low (Leta 2012). There is a strong stimulus for researchers to remain within their disciplinary boundaries, and there is no incentive for interdisciplinary work; applied research that does not lead to academic publications is penalized; there are no incentives for researchers to look for partnerships with the productive sector; and most of the doctors who get their degree end up working very often in the same universities from which they graduated (Galvão, Viotti, and Baessa 2008; Velloso 2002). Brazilian graduate education and university research are among the largest in the developing world, with some remarkable achievements, but their social and economic impact on society is probably more limited than it should be.

Assessment of Higher education

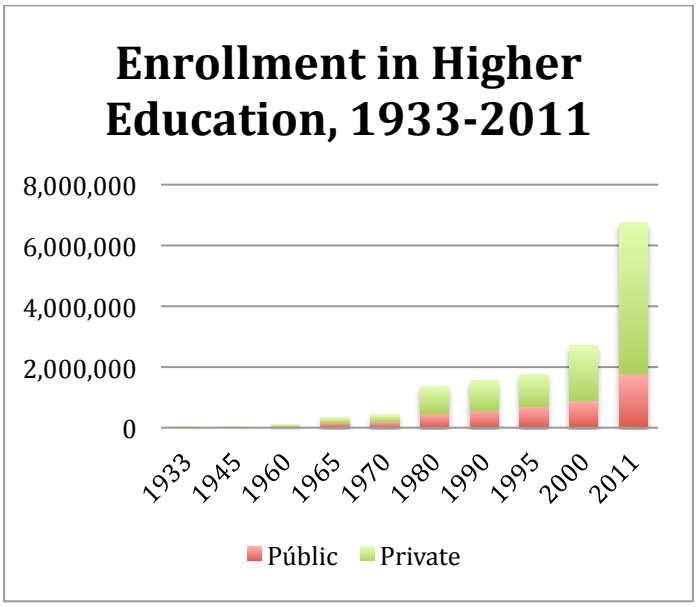
In 1985, after 20 years of military rule, the first civilian government created a broad based, high level Presidential Commission to take stock of what was

[4] http://www.capes.gov.br/avaliacao/qualis

happening with higher education and propose a new policy for the sector (Comissão Nacional para Reformulação da Educação Superior 1985). One of the main recommendations of the Commission was to establish a nationwide assessment system for higher education. Once published, the report drew strong resistance, the government decided to withdraw its recommendations, and it took another ten years for the first assessment mechanisms to start operating.

The report's recommendations were based on the perception that higher education in Brazil, although very small for the country's size, was spreading without any clear standards and at a growing cost for the public sector. Questions related to increasing access to higher education were discussed in the report in terms of the need to improve the quality of primary and secondary education, but did not include the issues of affirmative action that would become prevalent in the 2010s.

Higher education in Brazil started to expand in the 1950s with the creation of a federal network of public universities and the attempt to introduce the American research university model in 1968. Between 1945 and 1960, the number of students in higher education increased from 40 to a paltry 95 thousand, for a population of 70 million. In 1970, they were already 450 thousand, and, in 1980, 1.4 million. In the past, about half of the students enrolled in a few public institutions, particularly in the fields of Law, Medicine, Engineering and in social sciences and humanities, and another half in private universities and professional schools, mostly maintained by the Catholic Church. By 2011, there were 6.7 million students – still a small number for a population of 190 million - , 74% of which in private institutions.



The 1968 legislation required that all higher education institutions, dedicated so far mostly to teaching, should become research universities with full-time academic staff, laboratories, on-going research and graduate education. In practice, only public institutions could afford the high cost of fulltime contracts, with the staff paid by the government as part of the civil service, and they were also autonomous enough to limit the number of students they accepted every year. As the demand for higher education increased, the government allowed new private, teaching only institutions to get established and provide low cost, evening courses for students who needed to work and could not meet the admission standards of public universities. The result was that, contrary to the expectations of the 1968 legislation, the Brazilian higher education system did not converge to a single model based on the research university, but diverged in at least three very distinctive sectors: a small group of public, research intensive institutions providing good quality courses in the most prestigious careers; a larger group of public institutions that never reached high standards in professional and graduate education and research, but had similar costs because of their full-time staff and the limited the number of students they enrolled; and a much larger private sector providing low-cost, bare bone teaching courses

mostly in the social professions and admitting as many students as they could get (Balbachevsky and Schwartzman 2009; Durham 2004; Levy 1986; Sampaio 2000; Schwartzman 2001).

The rapid growth of the public and particularly the private sector led to the general notion that most Brazilian higher education was moving away from the quality standards it should have, even if no clear notion existed of what these standards should be like. This concern was particularly strong among the established professions of law, medicine and engineering, which worried about the growing number of persons with academic credentials provided by dubious institutions and threatening their job markets; by the small graduate and research academic community, who aspired to the high standards many of them had seen during their graduate studies abroad or in leading Brazilian institutions such as the University of São Paulo; and, more broadly, by the administrators in government and academic staff in the public universities, who mistrusted the values of private institutions, perceived as profiteering and not really caring about the values of science, culture and citizenship which supposedly prevailed in the public sector. There was also a general perception in the by the population in general and, increasingly, among politicians and the private sector leaders, that Brazilian higher education was not providing the country with the quantity and quality of highly skilled manpower the country needed to develop its economy.

The 1980s were the years of the "lost decade", when the grandiose projects of economic prowess and world power of the military regime had vanished, and the weak civilian governments that replaced it left the public administration to deteriorate, the economy to stagnate and the inflation to run out of control. The reason why the recommendations of the 1985 Presidential commission were not implemented was that the civilian government that succeeded the military regime, led by José Sarney, felt that there was no consensus on how the higher education sector should be reorganized, and did not dare to take a stand. It was the same weakness that made the government unable to deal with hyperinflation and to work for a new Constitution that could be more than the aggregate interests of all organized pressure groups.

This situation lasted until 1994, when the economy was at last stabilized thanks to the "Real Plan", which was followed by the election of Fernando Henrique Cardoso for the presidency, which also allowed for a higher degree of government stability and the implementation of long-range policies in different sectors. Between 1964 and 1984, the years of military government, Brazil had 12 ministers of education, and eight during the 1985-2004 period of civilian rule. In contrast, Paulo Renato de Souza, Cardoso's minister of education, remained in charge for eight years, and, among other initiatives, established a broad system of statistical information and assessment mechanisms both for higher and general education that are still in place in its main aspects.

The main idea, for higher education, was to use the same assessment model that had been working for so many years for graduate education: collection of systematic statistical data, peer review committees by subject area and careers, and in situ visits by reviewers (Castro 2004; Castro forthcoming). The National Institute of Education Statistics, INEP, a department within the Ministry of Education that had existed since the 1940s, was transformed into an agency for data collection and assessment, with a much-improved staff. Since then INEP implements two yearly censuses, one for general and another for higher education, with information provided by each school and teaching institution, which is used internally and also made available for researchers and the public.

The achievements in data gathering were unmistakable, but the peer review process was much less so, in part because of problems of scale: the 1998 higher education census identified almost 8 thousand course programs in about one thousand higher education institutions scattered throughout the country, which had to be assessed one by one to be accredited or reaccredited by the Ministry of Education according to existing legislation.

More serious than the problems of scale, which the Ministry of Education tried to solve by spending large amounts of money to fly hundreds of advisers to Brasilia and send them to visit the institutions around the country, was the lack of clear assessment parameters. The only existing criteria was to assume that all institutions should approach the university research model predicted by the

1968 legislation, with a significant number of full-time faculty with graduate degrees, as well as proper physical installations and library facilities. The reviewers could check whether the course programs followed the approved national curriculum for their field, if such curriculum existed, but not the quality of the education provided. The idea that the assessments should have a strong formative aspect, of providing the institutions with support and guidance to overcome their limitations, was never present.

In 1996 the Ministry introduced an innovative procedure, which consisted of an exam applied to all students concluding their university degrees in each career program (Schwartzman 2010a). The tests were prepared by advisory committees drawn mostly from public institutions; the mean results of each course program were standardized and distributed on a five-point scale, and widely publicized. Every year some fields and careers were selected for testing, to be repeated every year thereafter. Individual results were not published nor included in the student records, but students that did not participate could not get their degree.

This procedure, which became known as "Provão" (the big exam), had some obvious advantages: its interpretation was straightforward as an output-only measurement, uncontaminated with inputs that could or could not influence the results; and the final result was straightforward, like hotel stars. In spite of some initial opposition from students and institutions, the Provão was well received by the press, and became the main and most visible assessment mechanism for higher education in the country. In the growing higher education market, it became a distinction to be shown, or a problem to be confronted. A few years after being introduced, the Ministry was able to inform that the student demand for private courses with bad results was falling, as an evidence of the positive impact of the assessment (Ministério da Educação - INEP 2002)

This assessment had some well-known drawbacks, however: it required all course programs to follow the same curriculum content, limiting diversity; the results depended very much on the education levels of the students entering the courses – the most selective institutions would normally get better results

regardless of the quality of their work. And, more seriously, there were no standards of what was expected from the course programs – they were just ranked on a 5-point scale, with those at the bottom being considered bad and those at the top as very good (the public never learned, for instance, if it could trust a medical doctor coming from a "B" or "C" course). Not surprisingly, the most selective public institutions tended to appear at the top, with the non-selective institutions, mostly private, appearing at the bottom. The Ministry threatened and in some cases acted to limit admissions and even to close down courses appearing repeatedly at the bottom, but most of the effect of the rankings was to create a dispute for quality among the course programs in the private sector.

For the private sector, the assessment procedures created by the Ministry of Education were perceived as strongly biased against it, and complained that the lower rankings would become a stigma for its students. Some of private institutions reacted by getting external advisers and working to find out what was wrong and how to improve. But they also resisted by lobbying and pressuring the government to reduce the demands, by challenging the government in court and by dodging the system by, for instance, hiring consulting firms to prepare their paperwork to the satisfaction of the Ministry's requirements.

The pressure on the public institution was much lower, but the teacher unions and university administrators, as well as the student associations, resisted the idea that they should be compared with the most prestigious universities and eventually made responsible for the quality of the education they provided and for the public resources they used. The opposition to the assessment was quickly translated in ideological terms: for an influential group of academics, particularly in the schools of education, it was all part of a neoliberal plot of the government to privatize the public institutions and replace critical and humanistic education with the "productivist" requirements of the market (Gentili and Apple 1995; Sader, Aboites, and Gentili 2008; Sguissardi and al. 1997; Sobrinho 2000; Souza 2012).

This perception became dominant with the election of Luis Ignacio Lula da Silva as president in 2002. The new vision was that public policies should be based on the mobilization of "organized society" which, in the case of education, meant mostly the unions in the public sector. Qualitative self-evaluation was good, quantitative, external education was bad; public education was good, private education was bad. The problems of education were caused by social injustice and neoliberal, pro-market policies, and the main tasks for the government were to increase the resources for education, improve the teacher salaries and increase access to higher education.

To deal with Higher Education, a "Special Assessment Commission" was established with representatives of the unions, public universities, student associations and government appointees to propose a new approach, presented in 2003, and new legislation was introduced in 2004 creating a complex National System of Higher Education Assessment (SINAES) strongly based on the principles of self evaluation, working under a broad-based National Commission of Higher Education Assessment (CONAES) (Ministério da Educação - INEP 2004).

It took some years and three Ministers for the Lula government to get going in education, and with the nomination of Fernando Haddad in 2005 it became obvious that he had to reorganize and rely again on INEP, which, in practice, reintroduced most of the assessment procedures of the 1990s, including the student assessments (now under the name of "National Assessment of Student Performance", ENADE), and quantitative information on inputs. Going beyond that, INEP developed a "Provisional Index" to assess the course programs, combining data from ENADE and other sources making use of a complicated formula (INEP 2006), combining indicators which were in turn used to create an obscure assessment index for each higher education institution. The rationale was that, with this index, the Ministry could identify the most precarious institutions and focus its attention on them. In practice, the "provisional index"", in spite of its obvious shortcomings, was widely published in the press and became an official ranking of course programs and institutions, to the dismay both of the private sector, which continued to be the main target of the

assessments, and the original proponents of SINAES, who felt that the "neoliberal" and "productivist" approaches of the 1990s had returned with a vengeance (Pedrosa, Amaral, and Knobel 2012; Schwartzman 2008; Sobrinho 2008; Verhine, Dantas, and Soares 2006).

Between 1997 and 2012, the government issued about 1.5 thousand laws, decrees and regulations trying to make this evaluation system work, with questionable result. The private sector, instead of adjusting in response to the government assessments, retrenched by increased concentration of resources, developing a strong market orientation and fighting the government in court (Castro forthcoming). According to a publication from the association of the private sector, ABMES, quoted by Castro, "the evaluation system is nearing collapse. INEP holds approximately 5,000 assessment visits per year, or about 100 per week. The logistics to support an operation of this size, nationwide, and every day is overwhelming. For example, there are more than 400 flights per week to be scheduled, budgeted, accounted for, and issued by INEP. Yet, for a system with nearly 30,000 undergraduate programs and 3,000 institutions, not counting new authorization and accreditation procedures for courses and institutions, 5,000 visits are insufficient" (Garcia, Vianna, and Suñé. 2012).

The latest chapter of this saga is a proposal sent by the government to Congress to create a brand new National Institute for Supervision and Assessment of Higher Education, with more than 500 new civil servants and a huge budget to do what the Ministry of Education has been unable to accomplish. A detailed analysis of the project showed, among other problems, that, although it would have the characteristics of a regulatory agency for the higher education sector, it lacked a central component of such agencies, namely institutional autonomy and a clear legal authority (Nunes, Fernandes, and Albrecht 2012). According to an observer, it would be much better to create a non-governmental regulating agency that would require the higher education institutions to make public their information on goals, resources and achievements, allowing the students and their families to make informed choices (Mota 2012), instead of trying to force the whole higher education system into the uniform model of the 1968 legislation.

Basic education

In 1990 the Ministry of Education carried the first round of the National Assessment of Basic Education (SAEB), broadly inspired by the American National Assessment of Educational Progress (NAEP), which since 1995 has been the cornerstone of the Federal Government's policy for primary and secondary education (Crespo, Soares, and Mello e Souza 2000). K-12 education in Brazil is the responsibility of states and municipalities, with the national government providing broad orientations, managing education statistics and providing support for special programs such as school books and school meals, and also supplementing the local education budgets for the poorest states. SAEB consisted of tests on Portuguese language and mathematics applied to state-level samples of students at levels 5, 9 and 12, as well as socioeconomic questionnaires that could be analyzed to interpret the results. The Ministry did not establish the minimum acceptable standards for the tests, which were developed several years later by a private association, *Todos pela Educação* (Todos Pela Educação 2008). The results showed that the proficiency levels of Brazilian students were extremely low, a finding confirmed by Brazil's participation in two international comparative assessments, PISA, from OECD (INEP 2001; OECD 2009), and SERCE, from UNESCO (Ganimian 2009; SERCE 2008). At the same time, improvements in education statistics showed that Brazil had unacceptable levels of repetition in schools, leading to the introduction of automatic promotion in many parts of the country. These assessments and improved data helped to increase official and public awareness about the serious shortcomings of basic education, but did not contribute directly to its improvement: there was no explicit link between the contents of the tests and the curriculum adopted by the schools, and state governments and municipalities did not know how to deal with their shortcomings. The introduction of automatic promotion reduced repetition, but was often associated with the idea that students did not need to be assessed at all, which may have reduced even further the quality of the education they received.

In 2007-8 INEP introduced Prova Brasil, which was a version of SAEB assessments in language and mathematics to be answered in full by each student,

from which it would be possible to assess the standards of each school at levels 5 and 9[5]. The results of Prova Brazil, once standardized, are combined with data on student flow to produce an index of education development (IDEB) for each school. INEP compared the grades of students participating in SAEB with their results in PISA, the OECD international comparative education assessment, and set as the target for Brazil to reach PISA's average levels by 2022, the 200th anniversary of Brazil's independence.[6] This allowed the government to establish targets for each state, municipality and even schools. Most of the nearest targets could be easily met by just by improving student flow, but higher levels could only be reached by a combination of good quality education and high completion rates (Fernandes 2009; Fernandes and Gremaud 2009; Ministério da Educação 2007; Soares 2007).

IDEB is an ingenious creation, but has several pitfalls. Results can easily be improved by asking the less achieving students to stay home on test day (the Ministry considers the results valid if at least 50% of the students in each class participate). It is a low-stakes test for the students, since the individual results are not published nor added to the student records, leading to low motivation; and schools can game the system, at least at the lower stages, by increasing automatic promotion. IDEB has been criticized also because, by working with mean results, it does not take into account improvements among the worse students. Critics have also mentioned that, by emphasizing language and mathematics, it may induce the schools to just train the students for the test, neglecting other subjects in sciences and humanities. The assumption that the Prova Brasil scales can be linked to the PISA scales is open to question, not only because the tests are so different, but also because PISA applies only to 15 year olds. Finally, the combination of student flow and test achievements in the same

[5] In fact, only urban schools with 20 or more students at the designated levels participate.

[6] According to INEP, "a definição de uma meta nacional para o Ideb em 6,0 significa dizer que o país deve atingir em 2021, considerando os anos iniciais do ensino fundamental, o nível de qualidade educacional, em termos de proficiência e rendimento (taxa de aprovação), da média dos países desenvolvidos (média dos países membros da OCDE) observada atualmente. Essa comparação internacional foi possível devido a uma técnica de compatibilização entre a distribuição das proficiências observadas no PISA e no Saeb. http://portal.inep.gov.br/internacional-novo-pisa-opisaeideb

index makes it difficult to interpret what the scores of IDEB actually mean in each case.

To increase the awareness of IDEB results, there is a movement to require that each school should post its results in a large plaque at its entrance, and some states are already implementing this. One criticism is that, since the achievements are strongly dependent on the student's socioeconomic conditions, this would lead to the stigmatization of the schools in the poorer communities.

The low target set for IDEB for the first years, added to some improvements observed in the PISA results, led the government to proclaim that Brazilian education was finally improving its quality, after a long period of stagnation. The analysis carried out by *Todos pela Educação* showed that there was indeed some progress in proficiency levels at the 5th grade between 2005 and 2009, but very little or no improvement for students at the 9th year and in secondary schools (Todos Pela Educação 2012). Most of the advancements took place in a few states and municipalities that were able to improve the management of their school systems; and the overall improvement in mathematics seems to be more related to problems in the way the results were standardized than to an actual improvements in proficiency. Finally, there are good reasons to believe that the small improvements observed in the PISA results can be attributed to variations in the age groups of the PISA samples, rather than to higher proficiency (Klein 2011)

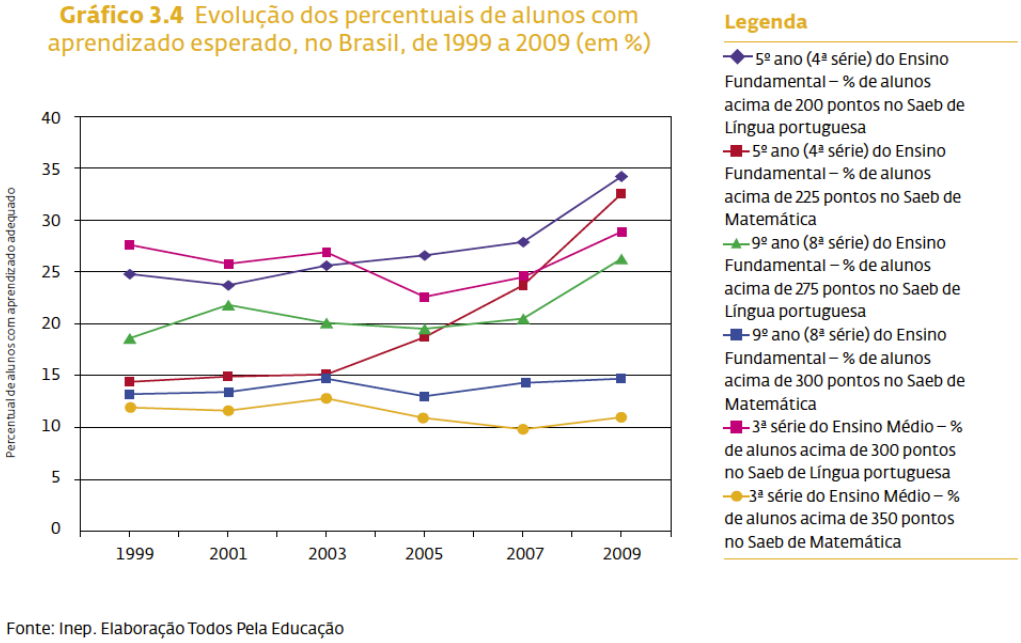


Figure 1 - Source: Todos Pela Educação

The assessment of Secondary Education

For the students concluding secondary education (which in Brazil lasts three years, for ages 15-17), the Ministry of Education decided to expand an existing voluntary exam for secondary education school graduates (ENEM), turning it into a major selection mechanism for higher education. Secondary education expanded very rapidly in the 1990s, to a large extent through the creation of evening schools, and there were no standards to assess what the students should know at the end of mandatory education, in spite of (and in part because of) the very detailed list of subjects required by the 1998 National Education Law. ENEM was devised at first as a voluntary test that could set a standard and become a reference for Brazil's secondary education. According to its proponents, ENEM was a single, multidisciplinary test consisting of an essay and 63 objective questions, based on a matrix of five competences and 21 abilities. It was supposed to measure fluency in Portuguese, and in the mathematical, artistic and scientific languages; the ability to use concepts to understand natural phenomena, historical-geographic processes, technological production and artistic manifestations; the ability to use of data and information to make decisions in problematic situations; the construction of consistent arguments; and the capacity to elaborate proposals to intervene in reality, respecting human

values and taking the socio-cultural diversity of the country into account" (Castro and Tiezzi 2004). The difficulty of measuring all these competencies and abilities in such a short test, and make it comparable from one year to the next, was not considered an issue, and several million students took the test in its first years.

In 2009 the government decided to turn the ENEM into a national selection mechanism for higher education in public universities, which so far had relied on their own exams (the so-called "vestibular"). To do that, the government expanded the contents to be assessed, to cover language, mathematics, the social and the natural sciences, plus an essay, in a marathon exam lasting two days; and persuaded the federal universities to admit all or part of their students according to their rankings on the ENEM exam (Ministério da Educação 2009). With that, the participants increased dramatically – there were 5.7 million applicants in 2012. The yearly implementation of ENEM, in the whole country and on the same days, became a logistic nightmare, subject to problems of fraud and leaking that led to the dismissal of two presidents of INEP.

One of the proclaimed advantages of ENEM was that the students could apply to any university in the country with a single exam, instead of having to travel to different places to participate in different selection procedures. However, in the absence of residential facilities and maintenance fellowships at the universities, only the richest students could make use of this possibility. Also, the state universities in São Paulo, among the most prestigious in the country, maintained their own selection mechanisms, making the ENEM irrelevant for the students applying to these universities. The government also used the average grades of students taking the exam to publish a ranking of the secondary schools, in spite of the fact that these students are not representatives of their student bodies and that students applying to private universities or to the São Paulo system would not take it.

ENEM is probably the most flawed assessment of Brazil's repertoire, but, because of its high visibility, its success became a point of honor for the government, making it resistant to recognizing its limitations. Instead, other

functions were added to it, using its results to screen students for PROUNI, the free tuition program for private institutions, and also as a certification of secondary education for students who did not complete the regular courses (Oliveira 2010; Soares 2012). The problems start with the way the tests are formulated, based on proposals from representatives of the different fields of knowledge, which tend to require the students to repeat what they learned in school, rather than to demonstrate their analytical competencies. This issue is considered of special importance in the social sciences and humanities, where many questions are ideologically biased and the students are supposed to answer whatever is considered politically correct. To make ENEM comparable from one year to another, in 2009 INEP announced that the multiple-choice tests would be developed making use of the Item Response Theory, at least in the sense that the weight of each item now depends on its relative difficulty; but it is not clear that they correspond to well-defined psychometric dimensions that are stable from one year to another. In addition, there are no clear parameters for the written essay, which are assessed manually by at least two examiners, and can weight 10% or more of the student's final score, depending on the course to which he or she is applying to (each university can give different weights to the different components of ENEM – mathematics, natural sciences, humanities, language and the essay).

In spite of its complicated structure and high ambitions, ENEM ends up reflecting the social stratification that brings the children from the richest and better educated families to the best, usually private schools, and then places them in the most prestigious universities, failing to contribute to the democratization of access to higher education that was supposed to be one of its main justifications. The graph bellow shows the standardized mean scores for all students taking ENEM according to their father's education. It should be noted that, to have access to prestigious professional faculties in Medicine, Engineering, Dentistry or Law, it was necessary to get between 750 and 800 points.

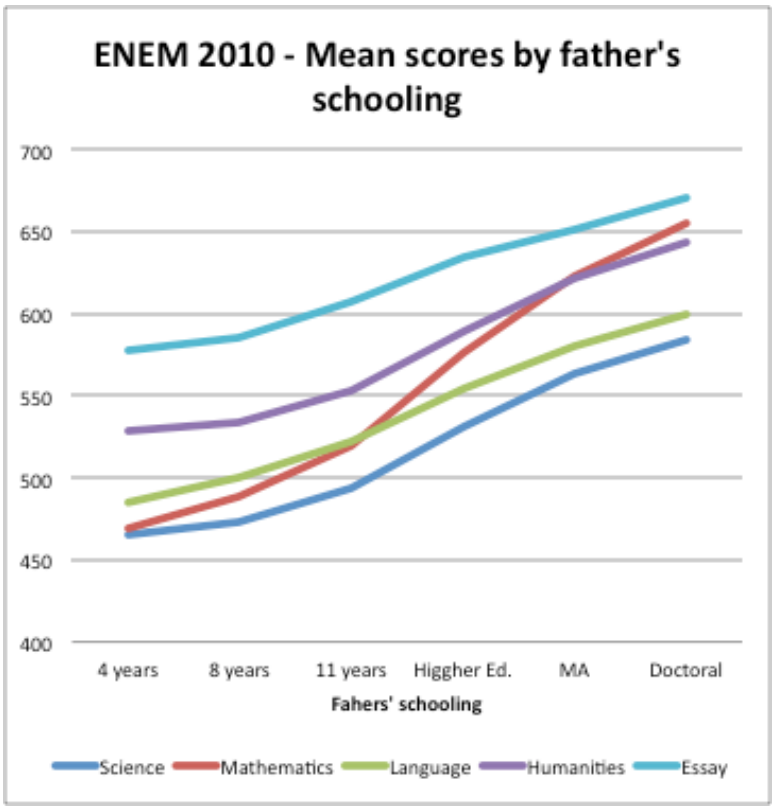


Figure 2 - Source: ENEM 2000 (microdata)

The most serious problem with ENEM, however, is that it placed a straightjacket on secondary education, requiring all the schools to teach the same extended curriculum for all the students, making it impossible to offer a variety of options, from academic to vocational or general education. The lack of choices would be a problem anywhere, but it is particularly serious in Brazil, where large numbers of young persons never finish secondary education, and only a small percentage of students reach the proficiency levels required for good quality university education (Schwartzman 2011).

Conclusions: whither the Goodhart Law?

In 1975 the English economist Charles Goodhart published a paper arguing that "as soon as the government attempts to regulate any particular set of financial assets, these become unreliable as indicators of economic trends" (Goodhart 1975). This notion was later expanded to other fields and became known as the "Goodhart Law", according to which, once a social or economic indicator or other surrogate measure is made a target for the purpose of conducting social or economic policy, then it will lose the information content that would qualify it to

play such a role. Another version of the same idea was stated by the American social scientist David Campbell, stating that " the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.[7]"

This "law" points to a serious risk, but should not be used as a justification to return to the situation that existed in Brazil before the 1990s, when there were no standards nor any information of what worked or did not in education. The debates on how to improve de quality of education remain, and the achievements so far have been less than what one would expect, but, thanks to the existing assessments, it is possible to identify with much more precision where the problems are, to compare experiences and to see what works and what does not (Menezes-Filho and Ribeiro 2009; Veloso 2011).

One of the gravest findings of the assessments of basic education is the high proportion of students that remain functionally illiterate even after several years of schooling. According to Prova Brasil, as analyzed by *Todos Pela Educação*, in 2011 only 35% of the students at the 4th grade had the minimum proficiency in Portuguese language expected at this level. This was further confirmed by of a sample test designed specifically to measure illiteracy (Prova ABC) which found that, among a sample of students at the 4th grade, only 56% were able to read as expected for that level (Todos Pela Educação 2012). This situation led the government to establish as a goal that all children should be fully literate at age 8, and to mobilize resources to support the schools to achieve that end. Some specialists believe that age 8 is too late, since in most parts of the world children become literate at 6, and also that the instrument used to assess literacy was not appropriate, since it did not include a direct measurement of reading fluency (Becskeházy and Oliveira 2012). But the fact that there is now such a target is by itself an improvement compared with the previous situation.

In about 20 years, Brazilian education moved from the world of ignorance or denial about its bad quality to a brave new world full of assessments, indexes

[7] http://en.wikipedia.org/wiki/Campbell's_law

and quantitative targets, with their own risks and pitfalls. The country knows more about its problems, education became much more important as a national priority, and there are many genuine efforts to improve it. These improvements, however, have been too slow and far apart, and there is a growing perception that this movement towards large-scale assessments at all levels may have gone too far, and promised much more than what it could deliver. It is time to look at the existing assessment instruments and to ask how they could be improved and better used to help, rather than to hinder or distort the country's education. To move ahead, some issues need to be addressed.

Reification

Reification occurs when quantitative indicators, which are in last resort tentative aggregations of subjective judgments, are taken to be "objective", and their original meaning is lost. This happens when they are presented as mathematical expressions pretending to be "scientific", or when the precarious nature of statistical estimations is lost in translation as they become official and sometimes legally binding figures (Schwartzman 1999). Number of publications, for instance, is an aggregation of decisions of journal editorial boards to publish articles, and rankings of journals are an aggregation of the decisions of scientists to read and quote each other's articles. This is a fair indicator of aggregate scientific quality in mathematics and the natural sciences, but cannot be applied mechanically to assess an individual researcher or a department, or fields in which other kinds of products prevail, as in engineering and the humanities. Reification also occurs with the development of indexes combining different dimensions, such as the IDEB or the "provisional" grading of higher education courses and institutions, with mathematical formulae hiding more or less arbitrary decisions on weights and scaling, leading to a simple number presented as a true objective measurement of "education development" or quality (and displayed at the schools' door as a trophy or a sign of their disgrace). Such combined indexes may be useful for analysis of broad trends, but should never be considered as an assessment of individual cases. Reification also occurs when low-stakes tests and student assessments, such as those of Prova Brasil, are used as high stake indicators to rank schools and their professors.

Technical quality

Brazil is well beyond the time when it was assumed that aggregate assessments were no different from traditional school tests, with results ranking from 0 to 10, or 100% of "right" answers. Today, most existing assessments take into account that items have different levels of difficulty, that they should be aligned with dimensions that are conceptually clear and statistically consistent, and that comparisons between results over time depend on large sets of interchangeable, well-tested items. In short, that they should obey the general cannons of what the literature calls the "item response theory", IRT (Baker 2001; Hambleton, Swaminathan, and Rogers 1991; Lord 1980). However, it is not clear that assessments such as ENEM or Prova Brasil, which are claimed to be based on the IRT methodology, work with dimensions which make sense from a pedagogical and psychological point of view, or that the item data banks are large and well-tested enough to warrant comparisons through time, or that the scaling used to adjust the results are robust. Brazil has competent economists and statisticians working on education, but little or no tradition in psychometrics, and the assessments suffer because of that. These technical questions are important because, if the tests do not measure what they intend, or measure things which are irrelevant or tautological, they could hardly be used for the implementation of proper policies.

Low-stake and high-stake tests

Most of the tests applied by the assessments in Brazil, with the exception of ENEM, are low-stake tests, in the sense that they bring no harm or benefit to the test-taker. For a student doing the ENADE, the only obligation is to participate in the exam, since the result will not appear in his personal record; the same applies to students taking Prova Brasil. For students taking the ENEM, however, the exam can determine his or her chances to study in the institution and course of his or her choice, and because of that the test needs to be of high quality. Low-stake tests are cheap and can be done quickly, and, if the instruments and the sample are correct, they can provide good estimations of general trends. As a rule, however, they should not be used as high-states assessments, which require more complex information and interpretation of its results. When the average

grades of students in Prova Brasil are used to rank a specific school and become part of the IDEB ranking, which is in turn pinned at the school's entrance, this rule is being violated. The same happens when the average results of ENADE, a low-stakes test of higher education, are combined with other estimations to produce a "provisional" ranking which the government then releases to the press and used them for accreditation purposes.. These estimations may be right on the aggregate, but may be extremely unfair in specific cases, creating a situation of mistrust that may affect the legitimacy of the whole procedure.

Transparency

Most of the data collected by the Ministry of Education in its assessments are made available both in official publications and as microdata, which can be freely downloaded and analyzed by independent researchers (with the proper protection of personal information). This is an important innovation that should be preserved. However, there is little or no information on the ways the scales were constructed, the reliability of the tests, the quality of the items and item banks and other technical aspects that external psychometrists and statisticians would need to assess the quality and relevance of the final results. Also, in some cases the Ministry decides which data should be published or not. For instance, the proportion of students taking the Prova Brasil in each school and grade, and some characteristics of these students in comparison with those who do not participate, would help to identify the schools that were trying to game the test. The only information available, however, is that only schools where 50% of more of the students participate are included in the rankings, which does not dispel the suspicion of widespread manipulation.

In-house vs independent assessments

The centralization of all assessments within the Ministry of Education leads to an undesirable situation in which the same institution is responsible for implementing the policies and evaluating their results. This conflict of interest is very clear regarding higher education, in which the Ministry runs its own network of universities and cannot possibly be too harsh against its own institutions. Pre-university education is the responsibility of states and municipalities, so here again the federal government needs to get consensus and

support among local governments for its policies. For instance, the new requirement that all students should be literate at age 8, instead of 6 or 7, was a clear concession to pressures from states that did not feel they could deliver better results in a short time. Also, in many cases the instruments developed for testing need to be adjusted to the prevailing pedagogical ideologies of vocal groups in the education community (as in the cases of the literacy test and the requirement of a written essay in ENEM), rather than be based on the best available international expertise.

Centralization vs. decentralization

There are good arguments in favor of the notion that Brazil needs a national curriculum for basic education in language, mathematics and science, which should set the basic core capabilities and knowledge that all students should have, before being able to go to different directions. So far, the country has only a set of broadly defined "curricular guidelines", established in the 1990s, together with a long list of general "transversal competencies". One consequence of that is that the contents of Prova Brasil are not linked with clearly identified competencies that could be identified by the schools and help them to deal with their shortcomings (Valverde 2009).

The argument in favor of a core curriculum gets weaker as one moves from elementary to secondary education, and then to higher and graduate education. There is an interesting paradox here, which is that the national government does not feel strong enough to implement a national curriculum for basic education, but, at the same time, does not dare to work towards increased diversification for the more advanced levels of schooling, in spite of heavy investments, in the last several years, to expand vocational education. There are practically no options in secondary education besides the traditional academic curriculum as assessed by ENEM (Schwartzman and Castro forthcoming) .

In higher education, all professional courses are supposed to provide the same diplomas, although the kind of education provided by a selective school of economics, for instance, is very different from what the students get under this name in most evening courses provided by private institutions throughout the

country. Instead of one, centralized straightjacket for all course programs, Brazil needs a large variety of certifications at the secondary and higher education levels, as well as many more options of post-university education than the existing CAPES system is able to recognize and assess.

When combined, centralization and in-house assessments create two additional problems, the tendency towards bureaucratic gigantism and the difficulty of turning the assessments into instruments for improving education, instead of imparting shame or punishment. Bureaucratic gigantism was obvious, for instance, in the complicated assessment procedures established for higher education, with armies of reviewers being sent to thousands institutions to collect information that could never be properly assessed. Instead of looking for a better approach, stimulating the disclosure of relevant information to the families and the creation of a plurality of competitive agencies, the government opted to create a new agency with hundreds of new employees, with very little assurance that they would do better work. With ENEM, an oversized and unified assessment system was put in place, stifling secondary education and restricting the ability of the universities to set their own criteria for student admissions. The huge size and the distance of these assessment systems from the institutions and persons being assessed created a situation in which the institutions did not have the chance for a fair consideration of their own peculiarities, and did not receive proper support to learn from their shortcomings and to improve.

The conclusion to be taken from Goodhart's Law and the Brazilian experience is not that one should give up the use of external assessments and quantitative measurements, but that that one must careful to not let the assessment procedures run away and gain independent life. Indicators are just that, precarious and simplified representation of much more complex realities. It is good, and indeed indispensible, to have good indicators, but the work of building good education institutions has to be carried on at each locality and at each school and institution, making use of the best available evidence of where one stands and what resources are appropriate in each case.

References

.

Baker, Frank. 2001. *The Basics of Item Response Theory*. University of Maryland, College Park, MD.: Clearinghouse on Assessment and Evaluation.

Balbachevsky, Elizabeth and Simon Schwartzman. 2009. "The academic profession in a diverse institutional environment: converging or diverging values and beliefs?" Pp. 145-164 in *The changing academic profession over 1992-2007: international, comparative and quantitative perspectives*, edited by University of Hiroshima's Research Institute for Higher Education. Hiroshima: Hiroshima University Press.

Becskeházy, Ilona and João Batista Araújo de Oliveira. 2012. "À espera de um currículo de padrão internacional." *O Estado de São Paulo*.

Castro, Cláudio de Moura and Gláucio A D. Soares. 1986. "As avaliações da Capes." in *Pesquisa universitária em questão*, edited by S. Schwartzman and C. d. M. Castro. São Paulo: Editora da UNICAMP.

Castro, Maria Helena Guimarães and Sertio Tiezzi. 2004. "The reform of secondary education and the implementation of ENEM in Brazil." Pp. 89-113 in *The Challenges of Education in Brazil*, edited by C. Brock and S. Schwartzman. Oxford, UK: Triangle Journals, Ltd.

Castro, Maria Helena Magalhães. 2004. "The State and the Market in the Regulation of Higher Education in Brazil." Pp. 179-– - 208. in *The Challenges of Education in Brazil*, *Studies in Comparative Educatio*, edited by C. Brock and S. Schwartzman. Oxford: Triangle Journals, Ltd.

—. forthcoming. "Higher Education policies in Brazil: failures in market regulation " in *Higher Education in the BRIC countries: investigating the pact between higher education and society1*, edited by P. Pilay, R. Pinheiro, and S. Schwartzman.

Comissão Nacional para Reformulação da Educação Superior. 1985. "Uma Nova Política para a Educação Superior." Ministério da Educação, Secretaria de Educação Superior.

Crespo, Manuel, José Francisco Soares, and Alberto Mello e Souza. 2000. "The Brazilian national evaluation system of basic education: context, process, and impact." *Studies in Educational Evaluation* 26:105-25.

Durham, Eunice Ribeiro. 2004. "Higher education in Brazil - public and private." Pp. 147-178 in *The Challenges of Education in Brazil*, edited by C. Brock and S. Schwartzman. Oxford, UK: Triangle Journals, Ltd.

Fernandes, Reynaldo. 2009. "Índice de Desenvolvimento da Educação Básica (IDEB): metas intermediárias para a sua trajetória no Brasil, estados, municípios e escolas." Brasília: Ministério da Educação, Instituto Nacional de Estudos e Pesquisas Educacionais 'Anísio Teixeira'.

Fernandes, Reynaldo and Amaury Gremaud. 2009. "Qualidade da Educação: Avaliação, Indicadores e Metas." Pp. 213-238 in *Educação básica no Brasil - construindo o país do futuro*, edited by F. Veloso, S. Pessoa, R. Henriques, and F. Giambiagi. Rio de Janeiro: Editora Campus - Elsevier.

Galvão, Antonio Carlos Filgueira, Eduardo B. Viotti, and Adriano R. Baessa. 2008. "Características do Emprego dos Doutores Brasileiros - Características do emprego formal no ano de 2004 das pessoas que obtiveram título de

doutorado no Brasil no período 1996-2003." Centro de Gestão e Estudos Estratégicos, Brasília.
Ganimian, Alejandro J. 2009. "¿Cuánto están aprendiendo los niños latinoamericanos? Hallazgos claves del Segundo Estudio Regional Comparativo y Explicativo (SERCE)."
Garcia, Maurício, Nadja Maria Valverde Vianna, and Letícia Soares de Vasconcelos Sampaio Suñé. 2012. *Diagnósticos e Propostas para a Avaliação da Educação Superior no Brasil*: ABMES Editora : Associação Brasileira de Mantenedoras de Ensino Superior.
Gentili, Pablo and Michael W. Apple. 1995. *Pedagogia da exclusão : crítica ao neoliberalismo em educação*. Petrópolis: Vozes.
Goodhart, C.A.E. 1975. "Problems of Monetary Management: The U.K. Experience." Reserve Bank of Australia.
Hambleton, R. K, H. Swaminathan, and H. J. Rogers. 1991. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Press.
INEP. 2001. "PISA 2000: Relatório Nacional." Brasília: Ministério da Educacão e Cultura, Instituto Nacional de Estudos e Pesquisas Educacionais.
—. 2006. "Sistema Nacional de Avaliação da Educação Superior (Sinaes), Resumo Técnico Enade 2005." Ministério da Educação, Brasília.
Klein, Ruben. 2011. "Uma re-análise dos resultados do PISA: Problemas de comparabilidade." *Ensaio - Avaliação e Políticas Públicas em Educação* 19:717-742.
Leta, Jacqueline. 2012. "Brazilian growth in the mainstream science: The role of human resources and national journals." *Journal of Scientometric Research* 1:44-52.
Levy, Daniel C. 1986. *Higher education and the state in Latin America : private challenges to public dominance*. Chicago: University of Chicago Press.
Lord, Frederick .M. 1980. *Applications of item response theory to practical testing problems*: ERIC.
Menezes-Filho, Naércio Aquino and Fernanda Ribeiro. 2009. "Os Determinantes da Melhoria do Rendimento Escolar." Pp. 171-188 in *Educação básica no Brasil - construindo o país do futuro*, edited by F. Veloso, S. Pessoa, R. Henriques, and F. Giambiagi. Rio de Janeiro: Editora Campus - Elsevier.
Ministério da Educação. 2007. *Aprova Brasil: O Direito de Aprender. Boas Práticas em Escolas Públicas Avaliadas pela Prova Brasil*. Brrasilia: Ministério da Educação.
—. 2009. "Termo de Referencia. Novo ENEM e Sistema de Seleção Unificada." Brasília.
Ministério da Educação - INEP. 2002. "Cai demanda por cursos com baixo desempenho no Provão." in *Notícias do Provão*, vol. November 5, 2003.
—. 2004. *SINAES – Sistema Nacional de Avaliação da Educação Superior: da concepção à regulamentação* Brasllia: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.
Mota, Ronaldo. 2012. "Regulação do Ensino Superior: É Hora de Inovar." *Folha Dirigida, Caderno de Educação*.
Nunes, Edson de Oliveira, Ivanildo Fernandes, and Julia Vogel de Albrecht. 2012. "O INSAES como agência reguladora atípica." Observatório Universitário, Universidade Cândido Mendes, Rio de Janeiro.

OECD. 2009. *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science*. Paris: OECD.

Oliveira, João Batista de Araújo e. 2010. *Os três senhores do ENEM*. Rio de Janeiro: http://www.schwartzman.org.br/sitesimon/?p=1929&lang=pt-br

Pedrosa, Renato H. L., Eliana Amaral, and Marcelo Knobel. 2012. "Assessing Higher Education Learning Outcomes: the Brazilian Experience." in *IMHE/OECD General Conference*. Paris.

Sader, Emir, Hugo Aboites, and Pablo Gentili. 2008. "La Reforma Universitaria - Desafíos y perspectivas noventa años después." Buenos Aires: Consejo Latinoamericano de Ciencias Sociales.

Sampaio, Helena. 2000. *Ensino superior no Brasil - o setor privado*. São Paulo: FAPESP / Hucitec.

Schwartzman, Simon. 1991. "The Great Leap Forward." Pp. 216-233 in *A space for science the development of the scientific community in Brazil*, edited by S. Schwartzman. University Park: Pennsylvania State University Press.

—. 1994. "Brazil: Scientists and the State--Evolving Models and the "Great Leap Forward."." Pp. 171-88 in *Domestic structures and the international context*. Solingen,-Etel, ed. Scientists and the state: Ann Arbor: University of Michigan Press.

—. 1999. "Legitimacy, Controversies and Translation in Public Statistics." *Science, Technology & Society* 4:1-34.

—. 2001. "A Revolução Silenciosa do Ensino Superior." in *o ensino superior .em transformação*, edited by E. R. Durham and H. Sampaio. São Paulo: Nucleo de Pesquisas sobre Educação Superior -USP / CEBRAP.

—. 2008. "O " conceito preliminar" e as boas práticas de avaliação do ensino superior." *Estudos - Revista da Associação Brasileira de Mantenedoras de Ensino Superior* 38:9-32.

—. 2010a. "The National Assessment of Courses in Brazil." Pp. 293-212 in *Public Policy for Academic Quality – Analyses of Innovative Policy Instruments*, edited by D. D. Dill and M. Beerkens. Dordrech Heidelberg London New York: Springer.

—. 2010b. "Nota sobre a transição necessária da pós--graduação brasileira." Pp. 34-52 in *Plano Nacional de Pós-Graduação 2011-2020*, vol. 2, edited by Brasil, Ministério da Educação, and CAPES. Brasilia: Ministério da Educação.

—. 2011. "Academic Drift in Brazilian Education." *Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana* 48:14-26.

Schwartzman, Simon and Cláudio de Moura Castro. forthcoming. "Estudo e Trabalho da Juventude Brasileira." Insiituto de Estudos do Trabalho e Sociedade / Instituto Teotônio Vilela.

SERCE. 2008. "Los aprendizajes de los estudiantes de América Latina y el Caribe - Primer reporte de los resultados del Segundo Estudio Regional Comparativo y Explicativo." Oficina Regional de Educación para América Latina y el Caribe, Santiago.

Sguissardi, Valdemar and et al. 1997. "Avaliação universitária em questão - reformas do Estado e educação superior." Campinas: Editora Autores Associados.

Soares, José Francisco. 2007. "Melhoria do desempenho cognitivo dos alunos do ensino fundamental." *Cadernos de Pesquisa (Fundação Carlos Chagas)* 37.

Soares, José Francisco. 2012. "Enem: seleção versus relevância educacional." GAME-FAE-UFMG.

Sobrinho, José Dias. 2000. *Avaliação da educação superior*: Vozes.

—. 2008. "Quality, evaluation: from SINAES to indexes." *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 13:817-825.

Souza, Carla Zottolo Villanova. 2012. "A emergência das políticas de avaliação no processo de reforma da universidade brasileira." Programa de Pós-Graduação em Políticas Públicas e Formação Humana, Universidade do Estado do Rio de Janeiro, Rio de Janeiro.

Todos Pela Educação. 2008. *De Olho nas Metas: Primeiro Relatório de Acompanhamento das 5 Metas do Movimento Todos Pela Educação*. São Paulo: Movimento Todos Pela Educação.

—. 2012. "De Olho nas Metas 2011 - Qiuarto Relatório de Monitoramento das 5 metas do Todos Pela Educação." edited by S. Paulo.

Valverde, Gilbert. 2009. "Padrões e avaliação." Pp. 49-80 in *Políticas Educacionais e Coesão Social – Uma Agenda latino-americana*, edited by S. Schwartzman and C. Cox. Rio de Janeiro, São Paulo: Elsevier; iFHC.

Velloso, Jacques. 2002. "A pós-graduação no Brasil formação e trabalho de mestres e doutores no país.". Brasilia: CAPES.

Veloso, Fernando. 2011. "A evolução Recente e Propostas para a Melhoria da educação no Brasil." in *Brasil: A Nova Agenda Social*, edited by E. E. Bacha and S. Schwartzman. Rio de Janeiro: LTC.

Verhine, Robert Evan, Lys Maria Vinhaes Dantas, and José Francisco Soares. 2006. "Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino Superior Brasileiro." *Ensaio - Avaliação e Políticas Públicas em Educação* 14:291-310.