
Educational Accountability in Brazil: an overview¹

Nigel Brooke²

Abstract: This article charts the growth of large-scale educational assessment in Brazil starting with the implementation of the National System for the Evaluation of Basic Education (SAEB) and the creation of the first state system in Minas Gerais. The first purpose of the study is to determine whether these and any subsequent systems were created to promote accountability at school level. The second purpose is to determine the extent to which the country's first experiments in establishing consequences for teachers and school administrators on the basis of comparative school performance heralds the adoption of accountability policies on a wider scale.

Keywords: Assessment; SAEB; Accountability; Education policy; School management.

INTRODUCTION

It is frequently claimed that large-scale educational assessment has acquired its current strategic role in the management of educational services as the result of profound changes in the role of the State. In this view, widely-held in Brazil, the remodeling of State responsibilities through the privatization of State enterprises, the de-regulation of the economy, the downsizing of government bureaucracy and, above all, the decentralization of public

services, has provoked a transformation in the purpose and variety of controls that need to be exercised by government (CASTRO, 1998; MATTOS; PEREZ, 2001). As part of this scenario, the decentralization of public education has meant the abandonment of traditional methods for ensuring compliance with central government policy via direct intervention in the staffing and running of schools and the consequent development of new mechanisms of system control (BROAD FOOT, 1996). Better described as "remote control" (LIMA, 2000), these 'summative' methods focus more on the external assessment of the products of schooling, as measured at the end of different school grades, rather than on the 'formative' details of the educational process itself. As government gives greater autonomy to schools, or even steps back from the business of running them, so it looks for ways to remain in control of the curriculum and the broad objectives of educational policy by fixing the criteria by which educational success should be measured. Central to this view is the idea that the State

¹Paper presented in PREAL Seminar: *Accountability Educacional: Possibilidades y desafios para América Latina a partir de la experiencia internacional*, Santiago do Chile, 18-20 April, 2005.

no longer needs to be the sole direct provider of public services but more the evaluator of the quality of the decentralized services now also to be offered by others.

Critics see this conception of the role of the State to be heavily influenced by neo-liberal thinking insofar as it reflects the incorporation of a private sector management philosophy based on the logic of the market-place, the ethos of competition and the need to stipulate standards in order to control the quality of the final product (AFONSO, 1999). The need and the methods to determine whether central directives are being carried out, whether the curriculum is being followed and whether standards are being met is the result, therefore, of a change in the relationship between government and the system of public education where decentralization has become almost synonymous with privatization. However, it is still far from clear whether the adoption of methods for the large-scale evaluation of educational results is necessarily the product of educational decentralization or whether other factors play a determining role. In the case of Brazil, where the supply of basic education from pre-school to the end of secondary school has been the responsibility of state and municipal governments for more than fifty years, there is reason to doubt whether this long-established 'decentralization' can be taken as an important stimulus for the widespread adoption of systemic

assessment during the 1990's. Even when other definitions of decentralization are taken into consideration, including the policies to foster greater school autonomy that began to take shape in the transition from military to civilian rule in the mid-1980s, it is unlikely that the new systems of external assessment were designed to set limits on school autonomy or even to promote adherence to a central curriculum. In other words, although monitoring is clearly one of the functions of educational authorities, it is debatable whether the spread of assessment in Brazil can be seen as due to governmental demands for greater control of schools. For this reason, the first questions to be raised by this study concern the reasons behind the adoption of system assessment as an integral part of educational policy.

Further criticism casts the wide-scale adoption of educational testing in Brazil as part of a broader process that has seen the gradual replacement of traditional concerns for equity and equality with those of educational quality and control (GENTILI; SILVA, 1995). In this view, the expansion of assessment has not been to collect better information for the formulation of government policy or for the improvement of investment choices but to disseminate a concept of evaluation that can stimulate market-type competition between schools by making them publicly responsible for their results. (SOUZA; OLIVEIRA, 2003). In opposition to this view, it can be shown that the policy of

attributing responsibility for the products of schooling, now more commonly referred to as "accountability", can be achieved in a number of different ways and while it is always an attempt to attach consequences to the results of assessment, the idea of promoting competition rarely figures as an objective given the difficulty of making meaningful comparisons between schools.

Despite the critics of accountability policies, the logic for establishing an association between school activities and outcomes is clear: if professional members of the educational community can be held responsible for the quality of the educational experience directly or indirectly under their control, and the outcomes of this experience can be objectively measured through instruments that capture the academic performance of their pupils, then large-scale assessment of pupil performance can become part of accountability systems that, in the hands of either the local community or the educational authorities, can be used to apply sanctions and rewards for the purpose of stimulating necessary improvements. Nevertheless, there is a clear difference in interpretation between accountability as a way to promote market-type competition between schools and accountability as a legitimate device for the improvement of school-level results. This polemic creates the further need to determine whether the newly established mechanisms of assessment have indeed been used to

establish accountability for the results of schooling among those directly responsible for the learning process. Does the term accountability have this type of currency in Brazil? Do educational authorities at the different levels of the federation see value in attributing consequences to the results of their assessment systems and is this policy practicable given the still considerable resistance to the very idea of external evaluation within the educational community? If accountability systems have been instituted in Brazil, has the result been to generate more competition and less equity? If not, does this mean that the critics of large-scale assessment are mistaken regarding the corrosive effects of external assessment on the professional autonomy of teachers and on their collective responsibility for the different outcomes of schooling?

To answer both sets of questions requires an overview of the genesis of large-scale educational assessment in Brazil, at both national and state levels, and the more detailed study of the first state-level systems to associate pupil performance with pay incentives in an effort to institute forms of accountability. The discussion that follows seeks to determine the extent that these pioneer "high-stakes" accountability policies are likely to be copied in other parts of the country and whether "low-stakes" alternatives exist in the form of evaluations that can foster a wider distribution of information on

school quality without having pecuniary or career consequences for teachers.

THE NATIONAL SYSTEM FOR THE EVALUATION OF BASIC EDUCATION (SAEB)

The National System for the Evaluation of Basic Education (SAEB) was implemented for the first time by the Federal Ministry of Education in 1990 but since its creation has undergone a number of important methodological modifications to make it today an extremely detailed portrait of pupil abilities in Maths and Portuguese at the end of the 4th and 8th grades of elementary school and at the end of the third (and last) year of secondary school. In 1993, the structure of the national sample of schools was altered and in 1995 other innovations were incorporated, including a sample of private schools, the use of more appropriate methods for the measurement of performance and the use of a single performance scale for each subject for better communication of pupil results. In 1999 and 2001 the sampling procedures were further improved and a theoretical frame of reference established for the socio-environmental background questionnaires applied to students, teachers and school principals based on the concept of school effectiveness. Applied every two years and employing a version of Item Response Theory in the confection and analysis of items that permits the comparability of results over time, SAEB has provided an extremely

valuable tool for the monitoring of change in the level of results from one application to another, between the different states and between different levels and areas of study.

What this brief description should also make clear is that SAEB was designed as a research instrument rather than a tool for collecting school-level performance data. Based, first, on a sample of schools that permitted analysis only at the regional level and then on representative samples drawn from each state of the federation so as to permit state-level comparisons, SAEB was never intended as a mechanism of system control. For system control to have been its purpose, with or without a notion of accountability, it would have been necessary to apply tests to all schools throughout the federation even if the students tested in each school represented no more than a sample. This was manifestly not the case. In the first description of its purposes, SAEB was created in order to

[...]develop the assessment capabilities of the managerial units of the educational system; decentralize and regionalize the operation of the assessment process....so as to create the connections and a stimulus for the local development of research and educational assessment; propose a common methodology for the different research and assessment efforts already undertaken and in the process of implementation (WAISELFISZ, 1993, p.12).

The test instruments were seen as a way

to expose education managers at all levels of the federation to the fundamentals of education assessment and to establish the use of standardized measures of performance as the dependent variable in a variety of research models. This latter task of raising methodological issues among the research community was especially evident through the application of questionnaires to collect the responses of school principals and teachers on such topics as the level of school autonomy, the physical conditions of the school, school plans, involvement of the community, teacher work conditions and problems, levels of teacher education and teaching methods and materials.

More recent descriptions of the purposes of SAEB confirm its earlier characterization as a research instrument capable of monitoring the progress of the country's different systems of basic education and of providing a description of the "pedagogical dimension of schooling for those in a position to take steps to improving the ability of schools to satisfy the needs of pupils" (LOCATELLI, 2002). To this original purpose have been added the tasks of

[...]offering concrete assistance in the formulation and reformulation of government policy" and of "providing education authorities and society with a clear vision of the results of the education process and the conditions under which they are produced (LOCATELLI, 2002, p.18).

Overall, the national assessment of basic education has been motivated not by a need

for control but by the belief that policies to improve the quality of schooling should be based on reliable, objective information generated through wide-scale measures of performance, rather than hit-and-miss gambles based on received wisdom or guesswork.

Responsibility for the control of the nation's basic education systems constitutionally resides with the state and municipal education authorities. However, despite attempts by the 1996 Education Act to redefine the division of educational responsibilities between state and municipality, the ambiguities in the relationship between these two legally independent spheres of government means there is no overarching authority for the measurement of pupil performance nor for the overall control of the different systems of education within any given state. In some states of the federation, what exist are state-level assessment policies to which the municipal systems can adhere voluntarily. The first states to develop their own policies in the early 1990s were Minas Gerais, São Paulo and Paraná. In a second wave, the States of Ceará and Rio de Janeiro also developed state-level assessment procedures, followed by the State of Bahia that for the first time created an instrument to evaluate first grade reading abilities. The most recent states to implement state-level assessment policies include Pernambuco, Goiás, Mato Grosso do Sul, Acre, Maranhão and Tocantins.

THE MINAS GERAIS BASIC EDUCATION EVALUATION SYSTEM

The intention of those in charge of the development of SAEB, even more clearly stated after the adoption of educational evaluation as the corner-piece of federal government policy in 1994, was to create a single instrument that could be used by education planners and policy-makers throughout the nation. Although there was no policy to preempt the creation of similar, state-level instruments, the Ministry of Education offered no help to those states interested in the development of their own evaluation systems in the belief that these might lead to a fragmentation of effort and resources and undermine SAEB. Despite this lack of federal support, the government that took office in Minas Gerais in 1991 made the creation of a state-level evaluation system one of its top policy priorities, along with school autonomy, teacher training, the strengthening of school governance and state-municipal integration. As such, Minas Gerais was the first Brazilian state to adopt an education assessment policy and the first to formulate an assessment system based on the idea of repeated applications of census-type instruments in all schools. In this regard, the Minas Gerais assessment model was more radical than SAEB insofar as it held out a real offer of pedagogical information of potential use to all teachers as well as the means to determine differences between schools.

The principal argument in favour of greater school autonomy was the sheer size of the Minas Gerais state education system that the new state secretary of education had described as "unmanageable" in his first speech to the State Assembly (MARES GUIA NETO, 1992). The system at that time was comprised of 6,500 schools, 2.7 million students and 204,000 teachers and the decentralization of administrative, financial and pedagogical authority to the school was seen as a precondition for the reform of both school-level and central management procedures³. The question to be answered is whether at this early stage or in any of its later versions the Minas Gerais education assessment system was seen as a counterpoint to the school autonomy policy, designed to ensure that schools were exercising their new-found power of decision in a responsible fashion, and whether pupil results were seen to be indicative of school-level decision-making and amenable, therefore, to efforts to promote accountability.

The official purposes of the first version of the Minas Gerais evaluation system, involving the application of tests to third and eighth grade pupils in March 1992, were the following (ANTUNES; XAVIER; FREITAS, 1992):

1. measure the level of performance of pupils at the end of the Basic Literacy Cycle (an initial, ungraded period of literacy instruction equivalent to the first two years of elementary

³After more than a decade of a policy to decentralize responsibility for elementary education from State to Municipal governments, the State of Minas Gerais now has 3.925 schools and 121.659 teachers.

school, introduced experimentally in 1985 in the first attempt in Brazil to eliminate grading);

2. generate data for use in the formulation of a reading and writing instruction policy;

3. identify critical curriculum issues;

4. offer teachers an indication of how to overcome difficulties.

Together, the four objectives indicate the concern of the Secretariat of Education to gather information to help in the evaluation and formulation of policy, principally with regards the teaching of literacy and other components of the curriculum. Although it is not entirely clear how far the application of Portuguese language tests to third grade pupils would permit an evaluation of the Basic Literacy Cycle policy, given the absence of previous baseline data, the identification of the strengths and weaknesses of pupils at the end of this initial period of literacy training was an attempt to contribute constructively to the improvement of teaching. However, in his speech to the State Legislature, the Secretary of Education introduced his assessment policy as offering the chance "to see whether the system improves in comparison with itself. This will be a way to measure the results of our own work" (MARES GUIA NETO, 1992, p.26). In other words, the evaluation policy was also to be a measure of the effectiveness of the new state education reform policies, in a self-assessment that no previous government had ever proposed.

The Secretary of Education also raised the possibility of making comparisons between schools on the basis of test results:

With this test we will be able to compare the results of different schools, cities and regions. We will be able to help schools that need assistance. We will also identify more successful schools so as to extend their experience to other schools (MARES GUIA NETO, 1992, p.26).

The proposed comparisons between schools were never carried out nor were any concrete steps taken to analyse the characteristics of successful schools for the purpose of transferring their success to other institutions. Nevertheless, the language used and the threat of comparisons could well have been part of a policy to alert schools as to their responsibility for the results achieved in what would have been an early version of a policy to generate accountability. Poulson (1998) argues that changes in education can be achieved not only through legislation but also through the discursive practice of educational authorities and gives the example of how the continual use of the concept of "choice" and the representation of education as a commodity during the Thatcher years in England led to a change of attitude among both teachers and the population. The reiteration of statements and metaphors suggesting that parents should have "choice" with regards their children's schooling led parents to see themselves as education consumers and to accept the need for externally-imposed regulatory

mechanisms for schools in order to protect their consumer rights. Was the Secretary of Education in Minas Gerais embarked on a similar process to instill accountability by adopting an external measurement of the success of teachers and by emphasizing a policy of comparisons in order to show which schools performed better?

In an article written after the second evaluation of the Basic Literacy Cycle in 1994⁴, the Secretary describes the evaluation system as a "compass for both the school and the Secretariat in the task of promoting change and elaborating plans for improvement" and as a "reference so that parents can better understand the quality of their child's school" (MARES GUIA NETO, 1992, p.14). There is no mention, however, of comparisons and the official description of the evaluation process of that year, that had also included the second application of Maths and Portuguese tests to 5th grade pupils, repeats the objectives of previous applications, to help the Secretariat and teachers better identify curricular difficulties and better plan the development of education (ANTUNES; XAVIER; FREITAS, 1994). No mention is made, however

vaguely, of informing parents as to the quality of schools and much less of drawing up league tables from which comparisons might be drawn and schools in some way held accountable. On the contrary, not only is the language of comparisons and accountability missing, there is also the direct admission that teachers are not to be seen as responsible for the poor results of their schools even if the solution for poor results is somehow in their hands. In a telling but profoundly ambiguous phrase, the Secretary of Education states:

Recent research has shown that our school does not feel responsible for the failure of its pupils. We recognize that the teacher cannot be blamed for this situation. It is a vice of our culture that is prior to the teachers who are today in the classrooms. But only the teachers can alter this state of affairs. For this reason, even if they are not to blame, it is the teachers' responsibility to look for a solution to the problem (MARES GUIA NETO, 1992 p.10)

Although there was no clear notion as to how to proceed from the identification of problems to the necessary actions for improvement, the purpose of the Minas Gerais education evaluation system was

⁴ The different grades tested under the Minas Gerais evaluation system are shown in the following table:

Gread/Years	1992	93	94	95	96	97	98	99	2000	2001	2002	2003
3rd (BCL)												
4th Elem entary												
8th Elem entary												
5th Elem entary												
2th Secondary												
3rd/4th(Secondary/ Teacher Training)												

always diagnostic (SOUZA, 1999). Even when the responsibility of the school was specifically mentioned, the onus for the improvement of results was generally laid at the door of the Secretariat which, as the result of more accurate information, was expected to produce better policies for the investment of public funds and for the improvement of educational quality:

In synthesis, this program for the evaluation of the public school proposed to carry out a diagnosis and build up a school data base comprised of reliable information for the purpose of implementing a plan for the improvement of education (SOUZA, 1999, p.62).

In this vision, no attempt was to be made to enlist popular support for greater community pressure on schools nor to name and shame schools for their poor results by publishing league tables in accordance with levels of performance. On the contrary, in line with the state constitution which makes education evaluation a "cooperative" venture between secretariat and teachers, the idea of attaching consequences to the results of the evaluations was deliberately avoided. The secretariat was aware of the need to establish a bridgehead for the gradual dissemination of a culture of evaluation before any attempt could be made to even disseminate the language of accountability, let alone attach any real consequences to the results of the external tests.

There is likewise no evidence that the

creation and deployment of a census-type assessment policy was the necessary corollary of the decision to promote greater school autonomy. The key elements of the school autonomy policy, including the transferal of financial resources, the decentralization of authority for the local solution of a variety of administrative issues and the freedom to add locally-defined elements to the curriculum, were all seen as the way to unlock school-level creativity, democratize school governance, promote community involvement and foster a new sense of self-respect. As such, the promotion of greater school autonomy was seen as one more policy to improve educational quality rather than a real decentralization of power. Therefore, although occurring simultaneously with the policy to foster school autonomy, the new evaluation system was not seen to be the *quid pro quo* for a transferal of responsibilities. Nor indeed was the Secretariat in a position to determine whether the measured levels of pupil performance were in any degree associated with greater school autonomy. In this sense, not even the change in the method for selecting school principals from the traditional procedure of political nominations to the policy of instituting local elections was seen to require the adoption of a school monitoring process via the application of pupil performance tests. If the Secretariat had been concerned to establish new controls in the light of the loss of direct control over the schools then the more radical process of municipalization of state

schools would certainly have been accompanied by some form of evaluation. This was not the case nor was there any move to show autonomy to be a determining factor either in the improvement or drop in the levels of school performance.

If there was some unspoken reason for the creation of a state system of evaluation it was the accepted belief both inside and outside the Secretariat that the quality of education had suffered a serious decline and that hard evidence of this fact was required both to galvanize the efforts of teachers and to establish a base line for any future policy of educational standards. The results of the first round of SAEB had confirmed previous findings of the Carlos Chagas Foundation's research in a sample of Brazilian cities showing only a small minority of public school pupils to be achieving what teachers themselves defined as minimum standards (VIANNA; GATTI, 1988, VIANNA, 1989). The flight of the middle class in the direction of private education was another clear indication of the same fact, as was Brazil's shaming performance in the second International Assessment of Educational Progress⁵ (LA POINTE *et al*, 1992a, 1992b). With the demise of the traditional school inspection system, essentially disbanded after the return to democratic rule in the mid 1980s, the prolonged decline in the level of initial teacher education, the inability of the profession to attract the best candidates and the erosion of the ethos of professional accountability that

was once part of the teacher's contribution to the ideals of public education, the Minas Gerais education system was undoubtedly in need of a dramatic turnaround. It was the hope of those involved in the creation of state policy at the beginning of the 1990's that the generation of reliable evidence attesting not only to the general level of performance but also pinpointing the major learning difficulties of pupils would represent a significant contribution in the effort to improve quality and raise standards.

The new version of the Minas Gerais assessment system, created in 2000 under the name of the Minas System for the Evaluation of Public Education (SIMAVE), has continued the "low-stakes" tradition of its predecessor with regards the generation of information of use mainly to planners. The principal component of the system, the Program for the Evaluation of Public Basic Education, is described as being

for the purpose of evaluating schools belonging to the state system and to those municipal systems taking part in the assessments so as to produce a diagnosis capable of identifying the problems and contributing to the definition or reorientation of educational policies implemented by the public school systems of Minas Gerais (BARBOSA, 2004 *apud* BROOKE, 2005, p.9).

The idea of going further and using this information for the purpose of accountability

⁵ An assessment coordinated by ETS of mathematics and science performance among 9 and 13 year olds carried out in 20 countries in 1990-91. In Brazil, the study was restricted to the cities of São Paulo and Fortaleza (LAPOINTE, ASKEW and MEAD, 1992a and 1992b).

is expressly excluded by Article 126 of the legislation creating SIMAVE. This article prohibits the use of the results of the state system of assessment for the purpose of "classifying schools or other components of the Minas Gerais System of Education with a view to altering the educational process". In this explicit rejection of the use of test results for the purpose of control or accountability, the legislators obeyed the orientation of the teachers union and of other critics of external assessment. However, in other significant ways SIMAVE has gone further than its predecessor. By adopting Item Response Theory and by using some of the same items as in the national assessment, SIMAVE has been able to employ SAEB's national proficiency scale for each of the different subject areas in order to plot the performance of all pupils from the 4th grade of elementary school to the last year of secondary school. The use of this scale and the effort to characterize the proficiencies corresponding to each interval as well as specify the proficiencies expected at each different grade level is symptomatic of the endeavor of the Secretariat to feed the assessment scores back into the schools in such a way as to provide the basis for school-level plans to improve results. One of these reports, entitled the "Evaluation Bulletin" and sent to each school at the end of the assessment process, allows the school to make comparisons with the average performance for the municipality, the region and the state. The other reports, called the "Pedagogical Bulletins", cover the different

subject areas and contain detailed analysis of the results of the state assessment and how to overcome the difficulties encountered. Using graphs, tables and other devices, these documents enable the school to identify the average level of proficiency of its pupils and the proficiencies that are lacking in order to match the standards expected for each grade. At the same time, SIMAVE has involved a number of higher education establishments in the confection and application of the instruments in an effort to strengthen the dialogue between the school and the institutions responsible for teacher training and to reinforce the idea that the state's assessment activities are a learning exercise.

ACCOUNTABILITY

It has been argued that while wide-scale assessment was implanted in Minas Gerais in order to generate school-level information there is no evidence that the decentralization of authority to the schools was the motivation for the monitoring of pupil performance. Likewise, it would be incorrect to portray the state assessment policy as an accountability program given the absence of any component designed to influence school behaviour by making results public, typically by offering incentives for higher levels of school performance. However, before extending this conclusion to other states, it would be as well to review the notion of accountability and to restate the essential ingredients of an accountability program.

Accountability programs can take a variety of forms but they share the common characteristic of increasing the real or perceived stakes associated with test results for teachers and school administrators. This means that for an assessment program to become an accountability program the results of the tests have to be associated with consequences that, symbolically or materially, can affect the individual and are therefore perceived as important. Breaking this down into its component parts, what an accountability program requires in order to operate is 1. the desire of those in authority to make public the differences in the level of performance of the component parts of the educational system, 2. standardized tests that can effectively supply this type of information, 3. ways to analyse this information so as to determine which schools are performing adequately and which schools are not and, 4. the capacity to apply rewards or sanctions in accordance with the established standards. The reasons for governments to embark on this type of policy can vary. This is not the place to do a review of accountability programs worldwide but it is evident that while the movement is on the decline in the UK it is taking the US by storm. In state after state, governments have created new curriculum standards, established new tests aligned with these standards, created new rules for student promotion and graduation, implemented new rules for ranking schools and publicizing test results and generated

new systems for rewards and sanctions. Even President Bush, a Republican not otherwise expected to be legislating in favour of federal intervention, has created the "No Child Left Behind" policy that sets in place a national accountability system of annual testing and performance-based rewards (RAVITCH, 2000; MOE, 2002). The question remains as to whether this type of policy has begun to influence Brazilian assessment programs and what its long term prospects of success might be.

STATE-LEVEL HIGH-STAKES ACCOUNTABILITY IN BRAZIL

Despite the availability of school level performance data, the State of Bahia has avoided using this information in its recently instituted model for the evaluation of school principals⁶. Comprising seven different criteria and a broad range of indicators for the assessment of school principal performance and the payment of bonuses equivalent to up to 50% of regular salaries, this model employs both self-evaluation and the existence of school-based procedures for the evaluation of teachers but fails to utilize the results of the state-wide assessment of pupil learning. Likewise, in the State of São Paulo there have been no public attempts at instituting consequences for schools or teachers in the light of the annual measures of pupil performance carried out in that State.

⁶ Established through the State Secretary of Education's Instruction No. 7733 of May 15, 2003

In both these states, as indeed in the case of Minas Gerais, there was an explicit avoidance of developing accountability structures related to school-level results. This decision was based on the belief that it was necessary to first create a positive attitude towards educational evaluation. This is still described as the need to create an "evaluation culture", in recognition of the absence in Brazilian culture of a commitment to objective evaluation and, in particular, of the inherent resistance to any type of educational testing. In other words, in the case of Brazil, the creation of a state-wide assessment program involving all schools is not necessarily followed by a policy to establish some form of accountability.

1. Ceará

The State of Ceará, on the other hand, has legally created an incentive scheme for participating schools entitled The New Millennium Educational Prize that foresees the payment of rewards on the basis of external assessment as part of its Program for the Improvement of Basic Education⁷. This annual incentive scheme is based on the results achieved by "the school, its teachers and pupils". Taking the form of "financial incentives and public recognition", this scheme clearly fits the model of accountability that lays emphasis not just on the competition for students but on the belief that education professionals can cooperate at school level to improve collective results if offered group

financial incentives. As the law establishing the annual prize makes clear, the purpose is to promote public recognition for higher performing schools, improve the school environment by creating a climate of quality that can influence school results, raise the standards of public education and verify the proficiency of pupils regarding both their school performance and their use of computers.

This last objective is a reference to the fact that the evaluation of pupil performance in the State of Ceará is now carried out over the Internet. From 1992 up until 1998 the evaluation of pupil performance was carried out in traditional manner under the aegis of the Permanent System for the Evaluation of Basic Education in Ceará (SPAECE). Created to produce data of value to the Secretariat of Education in the formulation of policy, from 1992 to 1994 SPAECE involved the annual application of Portuguese and Maths tests to 4th and 8th grade pupils. In 1996 the spacing of test applications was increased to once every two years to coincide with the intervening years between the SAEB applications. However, in 2001, following a year in which the anticipated evaluation did not take place, the methodology of the testing system was radically altered to become a Computer Aided Testing (CAT) program, with instruments based on Item Response Theory, and the program renamed SPAECE-NET. Using the new system, three applications have now been undertaken (2001, 2002 and 2003).

⁷ Law 13.203 of February, 21, 2002.

Another of the changes to take place in 2001 was the creation of a connection between SPAECE and the Program for the Improvement of Basic Education. With this connection it became possible to establish average 4th and 8th grade performance in Portuguese and Maths as the principal indicator of school quality and the only criterion for the decision regarding which schools should receive the prize. In accordance with the law, the top 100 schools in terms of average performance in each subject and grade are eligible for a prize providing the school average is five or more (out of 10). For the top fifty schools, the prize is 100% of the specified value while for the next fifty the prize is worth 50% of this value. The one-time prize goes to each member of school staff and in 2002 was worth a maximum of approximately US\$ 270 for all full-time temporary and permanent members of teaching staff and US\$ 100 for administrative staff, there being a reduction in the case of part-time teachers proportional to the number of hours worked. For the top scoring pupils there are also prizes, the number and value of which are determined each year by an oversight committee comprised of representatives from the Secretariat, the University, the State Assembly and the State Council of Education.

The incentive program has yet to be subjected to an external evaluation in order to determine its effects and observe the practicalities of implementation but a

number of questions already spring to mind. The first of these concerns the comparability of schools from very different regions and with widely different student intakes. Given the single criterion of average pupil scores and the failure to establish reference groups in accordance with pupil socio-economic characteristics, one can predict that the winning schools will always be from the capital or other larger cities, rather than from the rural interior, regardless of the level of contribution of each school in a "value added" calculation of pupil performance. A further doubt concerns the impact of the prize in establishing a climate of collaborative concern for quality given the transitory nature of school-teacher relations, the absence of any school-level decision-making regarding the hiring and firing of teachers and the broader difficulty of creating an incentive scheme for groups of professionals rather than for the individual teacher.

2. Rio de Janeiro

The only state-level education accountability system to have produced documents that describe its operationalization is that of Rio de Janeiro. Created under the name of The New School Program⁸ by decree number 25.959 of January 12, 2000, this program is a structured attempt to influence the management of schools as well as the outcomes of classroom teaching for which that management is now held responsible.

⁸The full name is the State Program for the Restructuring of Public Education -The New School Program.

The New School was seen to be a way to improve the quality of education, establish more democratic methods of school governance, integrate governmental action and improve the working conditions of teachers by rationalizing the use of resources, universalizing attendance, improving teacher pay and training and strengthening the articulation of municipal and state-level policy. The principal mechanism for the operationalization of this program has been the System for the Permanent Evaluation of State Public Schools which was designed to evaluate the governance and educational processes of each school.

Following the creation of the New School program, the state government hired the Cesgranrio Foundation to supply the necessary expertise for the design of an appropriate model for the analysis of school performance data. One of the very few Brazilian organizations with the technical capacity to design and analyse tests using Item Response Theory, Cesgranrio saw the New School program as an opportunity to implement a longitudinal research model that would permit the collection of performance data from the same pupils as they progressed through the system. This would have allowed the control of non-school influences on pupil performance and a more reliable "value-added" measure of school effectiveness. In line with the demands of the State Secretariat of Education of Rio de Janeiro, the model would also contemplate the

external evaluation of eight different dimensions of school governance as well as indicators of school efficiency.

The eight dimensions of school governance for which data were collected in 2000 and 2001 were the following:

- Planning
- Management of Human Resources
- Management of Financial Resources
- School Buildings
- Participation
- School-Community Integration
- Management of the educational process
- Nutrition

For the same years the chosen indicators of school efficiency were rates of pupil progression and age/grade distortion.

In 2003, Cesgranrio presented a new proposal containing important modifications. The pupil performance dimension of the program was the most affected. Due to the absence of any data collection in 2002, as the result of a change of government and the decision of the incoming authorities to suspend the program, the longitudinal research model needed to be altered. In 2000, performance data had been collected from pupils in the 3rd and 6th grades and in the 1st grade of secondary school in all schools voluntarily adhering to the New School Program. In line with the original proposal, in 2001 the gra-

des sampled were the 4th and 7th grades of elementary school and the 2nd grade of secondary school. The data to have been collected in 2002, in the 5th and 8th grades of elementary and the 3rd grade of secondary school, would have completed the process of establishing the value-added standards by which the future performance of all schools was to be judged.

To provide an alternative, Cesgranrio proposed new instruments to test pupil performance that would use similar items to the SAEB and thereby permit that all Rio de Janeiro scores be plotted on the same scale as used in the national assessment. Providing that the New School program also tested the same 4th and 8th grades of elementary school and the 3rd grade of secondary school, this would make the results of Rio de Janeiro pupils comparable with SAEB's national averages.

To comply with this requisite, Cesgranrio set to work to produce the necessary instruments. This involved the production of an item bank based on the SAEB reference matrices for the assessment of Portuguese and Maths to ensure that the new instruments measured the same abilities as those measured on the national assessment. The items were pre-tested on 20,000 pupils from public schools in five different states in August 2003 in preparation for the application to 180,000 pupils from the 4th and 8th grades and

the 3rd grade of secondary school by external examiners in November. To fulfill the purpose of providing teachers with relevant information, Cesgranrio also agreed to produce a report for each grade and subject containing a discussion of Rio pupil performance levels in relation to the national scales as well as presenting a statistical and pedagogical commentary on each one of the items used in the tests.

The second component of the New School External Evaluation system, the school governance sub-project, was also modified in 2003. In order to ensure that the evaluation was "grounded in reality", the indicators for each dimension of school governance were reviewed with new emphasis going to questions of context. While in the first and second evaluations emphasis had been placed on generating a diagnosis and describing the schools objectives and priorities, the third evaluation was to pay more attention to the school's environment while giving weight to the dimension entitled "management of the educational process". In practical terms this meant that two education process management indicators were added and six of the new total often indicators were given more weight. These were: 1. Planning, 2. Participation, 3. School-Community Integration, 4. Management of the Educational Process (Priority), 5. Management of the Educational Process (Instrumentality) and 6. Management of the Educational Process (Teaching and Learning). The four remaining indicators,

Management of Human Resources, Management of Financial Resources, School Buildings and Nutrition were considered 'instrumental' rather than 'managerial' and given lesser importance.

The third component, comprising the indicators of school efficiency, dropped the calculation of age/grade distortion. The data collected by specially trained technical staff continued to cover a variety of measures of pupil flow, including dropout, pass rates and transferals but the actual index of school efficiency was to take only pass rates into consideration. The calculation of this index involved the aggregate pass rates of the 1st to 4th grades and 5th to 8th grades of elementary school, the 1st to 3rd grades of secondary school and the 1st to 4th and 5th to 8th phases of youth and adult education. The school's efficiency index was the average of the aggregate rates, standardized according to each school's reference group for each of the levels and modalities of education on offer.

The reference group, containing schools of a similar variety, represented the Secretariat of Education's effort to inhibit comparisons between schools with students from different social back-grounds. To this end, the Secretaria established five different school reference groups from A to E in accordance with a measure of average family income. Unlike in the case of Ceará, this meant that a school was classified exclusively within its own reference group

thus making it impossible to compare school indices across different reference groups.

The index of school academic performance was based on the average of pupil scores on the tests of Portuguese and Maths for the grades under study and, as with the efficiency index, took only the scores of the reference group into consideration. The management index was composed by calculating the weighted average of the different scores with the more important managerial dimensions given three times the weight of the instrumental dimensions. The final index for each school within its own reference group was given by averaging the standardized scores for each of the three dimensions of evaluation and expressing it on a scale which had a mean of 60 and a standard deviation of 10.

In 2004, the contract for the program's external technical support was awarded to the Centre for Educational Assessment (CAEd) attached to the Federal University of Juiz de Fora. What ensued was a simpler version of the program that eliminated the reference groups but in their place created a measure of the schools progress over time. The dimensions of school performance were reduced to 1) Performance standards, with up to 10 points for schools with 80% or more students with satisfactory performance according to the national SAEB scale, 2) Pupil Flow, with up to 10 points for schools retaining 90% or more of its students until the end of the school

year, and 3) School Management, worth a further 5 points and based on four different indicators: Transparency; Integration with Community; Staff Frequency; and Enrolment Management. Schools getting top marks were classified in the first of five levels. At the same time, schools were classified according to their progress since 2003 and a new reward created to benefit those schools that had progressed most (RIO DE JANEIRO, 2004a, 2004b).

According to the law creating the New School Program, participating schools can be classified according to five different levels of performance on the basis of their final overall scores. Each one of these levels implies a monthly gratuity of a different value with the top level worth approximately US\$ 170 for full-time teachers and US\$ 125 for school principals. Although low by international comparisons, this is not an insignificant bonus given that the minimum wage in Brazil is less than \$100 and teacher salaries average less than \$400.

Apart from the continuing complexities of the appraisal system, that have to be renegotiated every year and require a legion of trained data-gatherers, the New School Program has faced a number of criticisms. The first of these is that the classification of schools is purely normative given the absence of any benchmarks or standards that might indicate whether the better placed schools have indeed achieved something in terms of educational quality.

However, the procedure for judging schools according to the average performance of the reference groups was not devoid of logic providing the program was understood primarily as a incentive scheme--designed to motivate teachers and other members of school staff to produce above average scores--rather than an as fully-fledged accountability program where the motivation for improvement is supplied by the pressure to reach adequate standards of performance. In this latter case, professional pride and the ethos of a commitment to quality would also have a role to play, along with the disincentive of being seen as a "failing" school for not achieving acceptable standards.

Another criticism relates to the use of assessment data for the previous year in order to reward staff the following year. Given the high turn-over of staff and the consequent volatility of school performance, the ideal would be to reward staff in the same school year as the measurements are made. However, this would require the application of tests near the beginning of the school year that would then measure the results of teaching and learning from the previous year and preserve the same inequity of rewarding teachers who were not necessarily responsible for the school's level of classification. Other questions concern the difficulty of explaining the meaning of the brute scores for each of the tests and the purpose of using the SAEB scales as a reference for school performance. And although the reference

groups have made the comparisons between limited numbers of schools more meaningful, the lack of any "value added" calculations continues to hamper broader comparisons and restrict the identification of school level factors associated with different levels of performance. Without this information, the Secretariat loses the ability to initiate the process of school-wide transformation in those cases where schools are proving ineffective.

THE FUTURE OF HIGH-STAKES ACCOUNTABILITY IN BRAZIL

Of the different states to have created system-wide assessment procedures, few have progressed to the point of establishing high-stakes accountability mechanisms tied to measures of pupil performance. At least three of these States - Minas Gerais, São Paulo and Bahia - actively considered and then discarded this possibility on the understanding that it was important to establish first the "culture of evaluation" before attaching consequences to the outcomes of schooling. Underlying this view was the certainty that teachers would not take kindly to assessment procedures being used to make comparisons that could alert the community to differences between schools and, worse still, between the teachers themselves.

The argument that it was the right of the community to know how well its schools were performing, and of the society in general to know what sort of system it was

paying for, fell on deaf ears. The irony is that the teachers could well have argued that they were already in tune with the community, aware of its rights and beholden to its needs, as the result of long years of policies to democratize the structures of school governance. Starting in the mid - 80s with the policy of school councils, state governments have gradually reduced the powers of local political bosses - and of the school principals that these bosses had traditionally been allowed to nominate - by giving increasing authority to school-level collegiate bodies and by establishing procedures for the election of school principals that included the vote of parents. With both these mechanisms, the community was invited to establish a much closer relationship with the school and its teachers and, to a limited extent, to exercise control over school decision-making. When the school councils were given the opportunity to decide on spending priorities, as the result of the policy to decentralize financial resources direct to schools, the community was also offered inside knowledge of the procedures of school management and given the opportunity to make suggestions regarding even such matters as teacher training. With the creation of municipal school councils and such other bodies as the state and municipal oversight commissions established to keep an eye on the education funds distributed under the national equalization policy called FUNDEF, the involvement of the community in school management became a matter even of national policy.

In this scenario it is hard to argue that the school is out of touch with the community or that teachers are unaware of the needs and wishes of their pupils' families. The struggle to remove "the rubble of dictatorship", as the process of Brazilian re-democratization was known, was accompanied by parallel procedures to reform the structures of school governance that have given teachers the argument that their schools are now under "social control" and no longer require government-sponsored accountability programs designed to subject them to community pressures for improvement. If the decentralization of power to the schools has meant that, in principle, school activities now respond to the demands of the population (BATE, 1998) and school objectives now reflect the goals of its community, who has the authority to demand more accountability? In this discussion, it is not a question of whether the evidence supports the belief of policy-makers that the community has indeed secured a bridgehead inside the school or whether the community is as distant from school decision-making as ever. As one of the political banners shared by both left and right, the principle of community participation is so central to policy that it would be almost impossible to admit that it has yet to become reality.

There are further reasons to doubt the widespread adoption of high-stakes accountability programs in Brazil. The

first of these has to do with the still-hostile attitude towards testing of a significant number of teachers and of an important segment of the education establishment in general. Despite more than a decade of large-scale testing, and of numerous attempts to make test results available and useful to teachers, it is still widely held that as tests are designed to measure solely cognitive performance in a limited number of subject areas, and thereby ignore the whole range of non-cognitive school objectives, they are inappropriate for the task of classifying school performance and of little practical use to teachers. Further arguments to justify opposition to tests include their excessive costs, the supposition that all schools are comparable in terms of intake and physical conditions, the emphasis on products rather than process and the lack of any independent evaluation of their validity. However, what underlies many of these criticisms is the mistrust of official explanations and the suspicion that in the hands of an unfriendly government even those tests designed for low-stakes purposes can be used against the interests of teachers. In the case of high-stakes accountability programs, this possibility becomes explicit and can be countered by the threat of strike action by the teacher's union, one of the most active components of the Brazilian labour movement. After almost 15 years of large-scale testing, the culture of assessment in Brazil is still incipient.

A LOW-STAKES ALTERNATIVE?

In the hope that the state assessment results might be of use to schools, the Minas Gerais Education Secretariat has produced individual school reports containing data that permit comparisons between the grade/subject averages for the school, the municipality and the state, in order to supply benchmarks against which schools might judge their level of performance. However, these school-level reports have never been used in any systematic attempt to make school performance data available to the general public (RAVELA, 2003) nor to promote community interventions in favour of school improvement on the basis of published information. So, although the necessary data is available to turn the Minas Gerais assessment systems into what has been described as the "decentralization" model of accountability (LEITHWOOD, 2001), where the community can use its voice to exert pressure for improvements on the basis of knowledge regarding school performance differences, this option has not been adopted.

The State of Paraná, on the other hand, went much further than producing just a simple report containing average performance scores for each school. In its concern to promote greater transparency regarding the quality of government services and greater understanding among parents as to the different types of

information required in order to make a reasoned evaluation of their children's schools, Paraná's Secretariat of Education fully embraced the decentralization model of accountability. Instead of appealing to the economic interests of teachers by creating an accountability program based on pecuniary rewards, Paraná opted for the production of school reports that offered a range of relevant information while giving members of the school community the necessary instruction on how to interpret this information and on how to use it for the purpose of promoting school improvements. In what it labeled a "report card accountability program", the Paraná education secretariat adopted an apparently low-stakes approach to the use of school performance data that did not offer teachers any positive or negative incentives but which, in the long run, offered possibly greater consequences for schools than the programs of Ceará and Rio de Janeiro.

The School Report Card (SRC) was first introduced in 2001. It represented a further component of a broader strategy to support and strengthen parent associations all across the state of Paraná. Designed to be a positive influence on the quality of education as well as a countervailing force capable of offsetting the excessive power of the Paraná teacher's union, the parents associations were both the reason for establishing the SRCs as well as a crucial element in the strategy for their dissemination and use. In

the end, however, the political power of the teachers was the stronger and when the new government took office in 2003 the SRC was totally abandoned.

The SRC was comprised of three blocks of information. The first block contained the results of the assessment program in Maths and Portuguese in the 4th and 8th grades involving all state government elementary schools. Alongside the average score for the school in each subject and grade, the SRC also showed the average scores for all other schools in the same municipality and for the whole State of Paraná. In order to make these means more readily intelligible, given that there was no predetermined pass or fail mark, the scores were plotted on a scale that had 250 as the average for the State and then classified in accordance with four levels of performance, from level I to level IV. In an annex to the SRC these four performance levels were interpreted in accordance with the curricular tasks that pupils reaching these scores are commonly capable of performing. With this key to the four performance levels, parents and others were ostensibly able to draw conclusions as to their school's average level of performance in relation to other schools and in terms of what tasks the children are effectively able to perform.

In the second version of the SRC, distributed in 2002, the Secretariat incorporated an important innovation by

calculating the expected level of performance of each school on the basis of the socio-economic level of the students. By determining, first, the state-wide relationship between parental education, family income and pupil performance, it was possible to calculate the level of performance that could be expected from each school, given the socio-economic characteristics of the students, and to show whether the school was performing at above or below this level of expectation. This measure of the "school effect" permitted comparisons that were not previously possible due to the large differences between schools regarding the profile of their intake. The parents were shown at which level schools were performing through the use of symbols, as demonstrated in the example in Figure 1 (p.29). What the 2002 SRC did not incorporate were new performance data. The performance tables were based on the same round of assessment carried out in 2000 (AVA 2000) as had been used for the 2001 SRC.

The second block of information (see example in Figure 2, p.29) was based on the School Census carried out every year by the Ministry of Education and contained key indicators regarding pupil flow as well as school and teacher characteristics. For each of these indicators the SRC also gave the equivalent data for the municipality and the State of Paraná to permit instant comparisons for each level of education.

FIGURE 1 - Example of the table included in the 2002 Paraná State School Report Card containing average performance scores for the school, municipality and state, by subject and grade.

2002 SCHOOL PERFORMANCE ASSESSMENT

		THIS SCHOOL						OTHERS SCHOOLS IN YOUR MUNICIPALITY		PARANÁ		
		MEAN SCORE	STUDENTS TEST	STUDENTS LEVEL I	STUDENTS LEVEL II	STUDENTS LEVEL III	STUDENTS LEVEL IV	MEAN SCORE	STUDENTS TEST	MEAN SCORE	STUDENTS TEST	
PORTUGUESE	4 ^o	257	◆	33	15%	27%	30%	27%	262	263	250	39.239
	8 ^o	241	●	29	38%	17%	17%	28%	252	335	250	31.125
MATHEMATICS	4 ^o	265	▲	33	18%	12%	30%	39%	266	337	250	38.441
	8 ^o	240	●	25	33%	16%	16%	24%	253	310	250	31.007
SCIENCE	4 ^o	255	◆	30	17%	17%	40%	27%	261	330	250	38.033
	8 ^o	235	●	26	35%	27%	23%	15%	252	334	250	31.125

Source: SEED/NIE - AVA 2000

Note: 1. See attachment for a description of the performance levels.

Note: 2. --- to few students tested for a reliable estimate.

* school (or municipality) did not offer (or did not test) this grade level.

** school (or municipality) did not participate in the AVA 2000.

*** no other school in the municipality participate in the AVA 2000.

■ to few cases for estimating the adjusted mean.

▲ adjusted mean higher than the expected given student profile.

● adjusted mean lower than the expected given student profile.

◆ adjusted mean as expected given student profile.

FIGURE 2 - Examples of tables containing school and teacher data included in the 2002 Paraná State School Report Card.

PROMOTION - RETENTION AND DROPOUT

	THIS SCHOOL			MUNICIPALITY			PARANÁ		
	GRADES 1 ^a a 4 ^a	GRADES 5 ^a a 8 ^a	SEC. LEVEL	GRADES 1 ^a a 4 ^a	GRADES 5 ^a a 8 ^a	SEC. LEVEL	GRADES 1 ^a a 4 ^a	GRADES 5 ^a a 8 ^a	SEC. LEVEL
PROMOTION	93%	63%	75%	90%	79%	70%	89%	81%	75%
RETENTION	6%	29%	13%	9%	16%	16%	9%	13%	10%
DROPOUT	1%	8%	13%	1%	4%	14%	2%	6%	15%

Source: MEC/INEP/SEEC - Censo Escolar 2002 (Preliminary Results)

Note: 1. Data refer to the 2001 school year of regular education public school.

2. * School (or municipality) did not offer this grade level.

TEACHERS AND STUDENTS 1

	TOTAL ²	THIS SCHOOL			MUNICIPALITY			PARANÁ		
		GRADES 1 ^a a 4 ^a	GRADES 5 ^a a 8 ^a	SEC. LEVEL	GRADES 1 ^a a 4 ^a	GRADES 5 ^a a 8 ^a	SEC. LEVEL	GRADES 1 ^a a 4 ^a	GRADES 5 ^a a 8 ^a	SEC. LEVEL
ENROLLMENT	981	232	480	269	6.123	6.102	3.692	825.850	737.602	408.020
AVERAGE CLASS SIZE		29	34	34	26	34	37	27	34	37
TEACHERS		9	27	21	270	367	210	39.255	37464	22.938
TEACHERS WITH POST-SEC. EDUCATION		89%	100%	100%	66%	100%	100%	46%	97%	97%

Source: MEC/INEP/SEEC - Censo Escolar 2002 (Preliminary Results)

Note: 1. Data for the public schools of regular education.

2. Total enrollment in elementary and secondary grade levels.

3. * schools (or municipality) does not offer this grade level.

4. State standards for class size is a minimum of 25 and a maximum of 30 students in grades 1 through 4; 30 to 40 students in grades 5 through 8, and from 30 to 45 students at the secondary level.

The third block of information was derived from different sources but dealt with the subjective evaluation of different facets of school and family life that were seen as relevant for an overall assessment of the school. The choice of which pupil, parent or school director opinions to include in this part of the SRC obeyed the principal that the information should connect with other SRC indicators and help, therefore, in the interpretation of school results rather

than supplying just further, unrelated data (AYRES, 2003). The opinions of pupils regarding the teaching of Portuguese and Maths, taken from the student questionnaire administered at the same time as the AVA 2000, were an example of this principle insofar as they facilitated the interpretation of the pupil performance scores for the same school (see example in Figure 3).

FIGURE 3- Examples of tables containing pupil and school director opinions included in the 2002 Paraná State School Report Card.

STUDENTS FROM THIS SCHOOL REPORT							
	4TH GRADE			8TH GRADE			
a parent always reading their report cards	69%			73%			
a parent always going to school when requested	55%			70%			
a parent always or almost always attending school events	48%			24%			
never being held back in grade	67%			59%			
at least one parent having secondary degree	43%			20%			
having computer at home	7%			11%			
having car at home	44%			49%			
		PORTUGUESE	MATH	SCIENCE	PORTUGUESE	MATH	SCIENCE
always or almost always needing help with homework at home	47%	44%	40%	36%	32%	28%	
enjoying the way their teacher teach	93%	97%	93%	55%	58%	72%	
liking the adopted school book	57%	68%	55%	52%	54%	60%	
having no difficulty learning	50%	55%	57%	38%	17%	48%	

Source: SEED/NIE - AVA 2000: Students Questionnaire

Note: 1. --- too few students tested for a reliable estimate.

* school did not offer or did not test this grade level.

** schools did not participate in the AVA 2000.

STUDENTS FROM THIS SCHOOL REPORT

- that during his tenure the larger of his/her time was dedicate to the analysis and discussion with teachers about student performance.
- that at school meetings where all parents were asked to partipate, the most frequent topic was the ways parent could be involved in trying to solve school problems. On average, 75% of parents attended such meetings.
- that at school meetings where all parents were asked to partipate, the most frequent topic was the improvement of student performance assessment procedures. In geral, all or almost al teachers attended such meetings.
- hat the School Concil met every other month with focus on matters related to the educacional process.

Source: SEED/NIE - AVA 2000: Students Questionnaire

Note: 1. ** schools did not participate in the AVA 2000.

The school director's opinions regarding such matters as parent participation, student performance and the work of the school council were likewise taken from the AVA 2000 questionnaire. The parent opinions, on the other hand, required the establishment of a whole new data collection procedure. In June 2001, parents were invited to go to their children's school to fill in a questionnaire and attribute an overall grade to the school from 1 to 10. Although some 53,000 parents responded to this first call, many more than had been anticipated, no attempt was made to ensure a representative sample of parents at the school level. When the same procedure was repeated in 2002, still with no pretense of statistical reliability, the number of questionnaires sent to the schools was increased and the number of respondents rose to 79,000. As well as teaching quality, the instrument covered other issues of importance to parents such as school safety and communications, as Figure 4 (p.32) shows.

Intended for widely different populations in terms of education and the level of sophistication in the interpretation of tables and percentages, the SRC needed to be adequately disseminated to insure impact. To this end, the Secretariat of Education developed a training strategy involving conferences and workshops at the Faxinal do Céu, a well-appointed training centre owned by the State and capable of taking residential groups of as

many as 500 at a time. In addition, those members of the parents associations given training were invited to take on the role of multipliers so as to further expand the radius of the Secretariat's efforts to explain the correct interpretation and application of the SRC. In 2001, more than 1.3 million copies of the SRCs were printed to ensure that every parent and teacher of the 1,963 state schools got a copy. Two copies were also sent to each of the 3,647 municipal schools.

Independently of any positive impact of the SRC on the understanding of teachers regarding the quality of their school, the assumption underlying the Secretariat's efforts to disseminate the use of the SRC among the parent associations was that parents could effectively bring beneficial pressure to bear on their schools. This thesis has yet to be established empirically but what seems almost certain is that the existence of a structured parent's association movement was a pre-condition for the creation of the SRC strategy and that in any other Brazilian state the strategy would almost certainly have fallen short of its mark. It is also clear that for parents associations to have any impact they need a channel for their communication with the school. In Paraná, this channel is the school council, a collegiate body of varying degrees of effectiveness depending on the managerial style of the school principal, the participation of parents and the degree of democratic governance that professional school staff are willing to tolerate.

FIGURE 4 - Example of table containing parent opinions included in the 2002 Paraná State School Report Card.**PARENT OPINIONS ABOUT****TEACHING QUALITY**

62 of 65 parents (95%) are satisfied with the quality of education their childrens receive at this school.

63 of 65 parents (93%) believe school building and grounds are well kept.

60 of 63 parents (95%) believe school building and grounds are proper for theaching activities.

61 of 62 parents (98%) agree that teachers in this school are dedicated to their work.

62 of 64 parents (97%) are satisfied with teachers assiduity.

PARENTS INVOLVEMENT

61 of 63 parents (97%) would like to be more involved in school activities.

63 of 63 parents (100%) believe the school principal fosters the participation of all in the community o

46 Of 66 parents (70%) indicate the school promotes activities for parent involvement in teaching matt

35 deles participated in such activities

48 of 64 parents (75%) indicate the school promotes activities for parent involvement in school admini

28 deles participated in such activities

EXCHANGE OF INFORMATION

48 of 68 parents (77%) indicate the school promotes regular mettings between parents and teachers

41 of 59 paarents (69%) indicate the existence of a regular system of communications with the paren

61 of 54 (95%) indicate teachers develop homework activities promoting the interest of parents intheir

51 of 53 (81%) indicate being informed about homework activities to help parents monitor their child's

SAFETY

62 of 63 parents (98%) indicate they are confortable sending their children to this school.

32 of 59 (54%) indicate the school hyas discipline related problems.

19 of 62 parents (31%) indicate has safety-related (internal) problems.

32 of 62 parents (52%) indicate the school is subjected safety-related rpblems of it's neighborhood

SCHOOL GRADE

93 average grade based on the rating of 64 parents

PARENT PROFILE

6 of 66 parents (9%) indicate being members of the School Concl.

12 of 66 (18%) indicate being directors of the PTA.

8 of 61 paretnts (13%) indicate being volunteers at the school.

24 of 67 (36%) have more then one child in this school.

Source: Survey questionnaire given to parents during Parents at thr School Week, from april 20 through 28, 2002.

Note:

1. Survey limited to school with more than 160 students.
2. Percentages do not represent the opinion fo all parents in this school. Only of those who responded the questionnaire.
3. Percentages are based on the number valid answers, not on the number of returned questionnaires.
4. Percentages based on less than five responses are indicate by ---.
5. The symbol * indicates the school did not returne any questionnaire.

Although parents are allowed to vote in school principal elections they have little direct control of school decision-making and it is uncommon to hear of schools being heavily influenced by a parent lobby.

This means that the effective impact of the SRC depended fundamentally on the type of relationship between school and community and on the degree to which teachers were sensitive to parent opinions.

The demise of the SRC would suggest that this was not the case.

Other questions concern the degree of reliability of the data and their impact at the school level. In the description of the implementation process, Ayers (2003) admits that school directors believed the report card would be used in the process of director selection and that this had introduced a "bias" into the results. It is not hard to imagine ways in which school directors might have influenced the results of the parent questionnaire so as to portray the school in a more favourable light. It is also a pity that the second round of the SRC was only able to renew the school census and the parent questionnaire data. All other sources, including the pupil performance information and the pupil and school director information remained the same, regardless of any modifications that might have taken place due to the publication of the 2001 report card.

If it is the case that a bias was introduced into the parent opinion data, including the overall school grade⁹, this then indicates that at least some members of school staff attributed importance to the instrument and saw it as having possible consequences for their professional lives. This brings us back to the original definition of accountability and throws doubt on whether there really can be such a thing as a low-stakes

accountability mechanism. The assumption was that while accountability mechanisms generally require that the authorities have the capacity to apply rewards or sanctions in accordance with approved standards, there could perhaps be an alternative in which this element was replaced by pressures from both within and outside the school that, beyond the control of central authorities, could be motivated by a common demand for improvements in the quality of education. In this definition, accountability would begin to look remarkably like the ideal of a school run by autonomous professionals who are committed by their ethic to producing the best possible results and who take pride in seeing their schools improve. However, the suggestion that some teachers in Paraná viewed the dissemination of school information as potentially threatening and as raising, therefore, the level of the stakes involved, means that no accountability mechanism can be completely low-stakes. What the Paraná example also shows is that any type of assessment involving school comparisons can generate a huge amount of resistance amongst the teaching profession and that if put to the vote, even a modest low-stakes method can be expected to fail. It is really only for as long as there is an explicit political will to create and sustain an accountability mechanism, such as in Rio where successive Garotinho governments have failed to show interest

⁹ Ayers (2002) indicates that the average school grade is generally high with only 10% of schools receiving less than 7.

in fostering the political support of teachers, that the methods and practice of accountability can be expected to survive. While it is important to promote a sense of responsibility for the outcomes of schooling among teachers and other members of the educational community, it is hard to believe that accountability mechanisms that do not meet with approval within the teaching profession can become permanent structures for the dissemination of information on school quality.

Artigo recebido em: 12/03/2006.

Aprovado para publicação em: 18/03/2006.

Avaliação educacional no Brasil: um panorama.

Resumo: O artigo detalha o crescimento da avaliação em larga escala no Brasil, a começar pela implementação do Sistema de Avaliação do Ensino Básico - SAEB, e a criação do primeiro sistema estadual de avaliação em Minas Gerais. O primeiro objetivo é determinar se estes e outros sistemas subsequentes foram criados para promover a responsabilização a nível de escola. O segundo objetivo é determinar em que medida as primeiras experiências no Ceará, Rio de Janeiro e Paraná em estabelecer consequências para os professores e gestores escolares com base no desempenho comparativo da escola anuncia a adoção de políticas de responsabilização em escala maior.

Palavras-chave: Avaliação; SAEB; Responsabilização; Política educacional; Gestão Escolar.

Evaluación educativa en Brasil: un panorama

Resumen: El artículo detalla el crecimiento de la evaluación a gran escala en Brasil, comenzando por la implementación del Sistema de Evaluación de la Enseñanza Básica - SAEB, y la creación del primer sistema estatal de evaluación en Minas Gerais. El primer objetivo es determinar si este

y otros sistemas subsiguientes fueron creados para promover la *responsabilización* a nivel de escuela. El segundo objetivo es determinar en qué medida las primeras experiencias en Ceará, Rio de Janeiro y Paraná en establecer consecuencias para los profesores y gestores escolares con base en el desempeño comparativo de la escuela anuncia la adopción de políticas de *responsabilización* en escala mayor.

Palabras-clave: Evaluación; SAEB; *Responsabilización*; Política Educacional; Gestión Escolar.

REFERENCIAS

ANTUNES, Ana Lucia; XAVIER, Edir P. C.; FREITAS, Maria Auxiliadora de. Avaliação do Rendimento do Aluno da Escola Estadual de Minas Gerais: O Ciclo Básico de Alfabetização. **Estudos em Avaliação Educacional**, São Paulo, n. 6, p.29-41, 1992.

_____. Avaliação da 8ª Série do Ensino Fundamental em Minas Gerais. **Estudos em Avaliação Educacional**, São Paulo, n. 9, p.17-23, 1994.

AYERS, Sandra. Boletim da Escola: Diagnóstico Multidimensional das Escolas do Paraná. **Ensaio**, Rio de Janeiro, v.11, n. 38, p. 33-44, 2003

AFONSO, Almerindo J. **Políticas Educativas e Avaliação Educacional: para uma análise sociológica da reforma educativa em Portugal (1985-1995)**. IEP/CEEP/Universidade do Minho, Portugal, 1999.

BATE, Peter. Education: the Gordian Knot. Shortfalls in schooling are at the root of inequality. **IDBAmérica**, nov./dez., 1998. Disponível em <<http://www.iadb.org/idbamerica/Archive/stories/1998/eng/e1198e4.htm>> Acessado em 17/04/2004.

BROADFOOT, Patrícia M. **Education, assessment and society: a sociological analysis**. Buckingham: Open University Press, 1996.

BROOKE, Nigel. Programa de Promoción de la Reforma Educativa en América Latina y el Caribe Partnership for Educational Revitalization in the Americas - PREAL. **Accountability educacional en Brasil, una visión gener**. Santiago: UNESCO/OREALC, 2005 (PREAL. n.34). Disponível

em:< <http://www.lpp-uerj.net/olped/documentos/1599.pdf> >

CASTRO, Maria Helena G. Palestra inaugural. In: Seminário Internacional de Avaliação Educacional, 1998, Rio de Janeiro. **Anais... INEP**, Brasília, D.F., p. 9-11, 1998.

GENTILI, Pablo; Silva, T.T. **Neoliberalismo, Qualidade Total e Educação: visões críticas**. Petrópolis: Vozes, 1995.

LAPOINTE, Archie; ASKEW, Janice M.; MEAD, Nancy A. **Learning Science**. Report n. 22, CAEP-02, Educational Testing Service, Princeton, NJ, Feb. 1992a.

_____. **Learning Mathematics**. Report. n.22, CAEP-01, Educational Testing Service, Princeton, NJ, Feb.1992b.

LEITHWOOD, Kenneth. **Educational accountability and school leadership**. National College of School Leadership. 2001. Disponível em: <<http://www.ncsl.org.uk/>> 13/05/2004 media-store/image2/kpool-evidence-leithwood.pdf)

LIMA, L. Administração escolar em Portugal: da revolução, da reforma e das decisões políticas pós-reformistas. In: CATANI, A. M; OLIVEIRA, R. P. **Reformas Educacionais em Portugal e no Brasil**. Belo Horizonte: Autêntica, p. 41-76, 2000.

LOCATELLI, Iza. O Sistema de Avaliação da Educação Brasileira. In: Simpósio Iberoamericano de Investigación sobre Eficácia Escolar e Mejora de la Escuela. La Habana, n.1, 2002, Cuba. **Anais...** Cuba, 28 de janeiro a 1º de fevereiro, 2002.

MARES GUIA NETO, Walfrido. S. A realidade da educação em Minas Gerais. **Estudos em Avaliação Educacional**, São Paulo, n.6, 1992.

MATTOS, Maria José V. M.de; PEREZ, José Roberto R. A Reforma Educacional de Minas Gerais: Promoção da Autonomia Escolar. **Revista Brasileira de Política e Administração da Educação**, v. 17, n.2, p.187-200, 2001.

MOE, Terry M. **Politics, control, and the future of school accountability**. Paper presented at the Conference on Taking Account of Accountability, Kennedy School of Government, Harvard University, jun. 10-11, 2002. Disponível

em: <www.ksg.harvard.edu/pepg/pdf/TAConfPDF/PEPG02-14%20TA%20Moe.pdf> Acesso em 09/03/04.

POULSON, Louise. Accountability, Teacher Professionalism and Education Reform in England. **Teacher Development**, v. 2, n.3, 1998.

RAVELA, Pedro. Como os Sistemas Nacionais de Avaliação Educativa da América Latina Apresentam Seus Resultados. Programa de Promoção da Reforma Educativa na América Latina e no Caribe (PREAL) **Documentos**, n.22, 2003.

RAVITCH, Diane. Testing and Accountability, Historically Considered. In: Williamson M. Evers and Herbert J. Walberg, (Eds) **School Accountability**. Stanford, CA: Hoover Press, 2000.

RIO DE JANEIRO; SECRETARIA DE EDUCAÇÃO. **Programa Nova Escola: Avaliação 2004**. Secretaria de Estado da Educação/CAEd, Setembro 2004a.

_____. **Programa Nova Escola: Informativo 1- Avaliação do fluxo e da gestão escolar**. CAEd/Secretaria de Estado da Educação, 2004b.

SOUZA, Sandra Zákia L de.; OLIVEIRA, Romualdo P. de. Políticas de Avaliação da Educação e Quase Mercado no Brasil. **Educação e Sociedade**, Campinas, v. 24, n.84, p.873-895, 2003.

SOUZA, Maria Alba de. A Experiência de Avaliação Educacional em Minas Gerais: 1992-1998. **Estudos em Avaliação Educacional**, São Paulo, n.19, p.57-76, 1999.

VIANNA, Haroldo M. Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: um estudo em 39 cidades. **Educação e Seleção**, São Paulo, n. 20, v.2, 1989.

VIANNA, Haroldo M.; GATTI, Bernardette A. Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: uma aplicação experimental em 10 cidades. **Educação e Seleção**, São Paulo, n.17,1988.

WAISELFISZ, Jacobo. Sistemas de Avaliação do Desempenho Escolar e Políticas Públicas. **Ensaio: Avaliação e Políticas Públicas em Educação**, v.1, n.1, p.5-22.1993.

Sobre o autor:

¹ *Nigel Brooke*

D.Phil. em Estudos do Desenvolvimento da Universidade de Sussex, Inglaterra. Professor Convidado da Faculdade de Educação (Grupo de Avaliação e Medidas Educacionais – GAME), Uni-

versidade Federal de Minas Gerais, Brasil.

E-mail; n.brooke@terra.com.br

Endereço postal: Universidade Federal de Minas Gerais, Faculdade de Educação, Grupo de Avaliação e Medidas Educacionais. Av. Antonio Carlos, n. 6627, Pampulha-Belo Horizonte/MG-Brasil. CEP:31270-901.