



High-stakes accountability using teacher salary incentives in Brazil: An update

Nigel Brooke

Universidade Federal de Minas Gerais – Game/FUFMG

n.brooke@terra.com.br

RESUMEN

En Brasil, las políticas de *accountability*, llamadas políticas de responsabilización, se basan fundamentalmente en la atribución de incentivos salariales y otros pagos de bonificaciones a maestros y escuelas. Para muchos autores, la ecuación entre la *accountability* educativa y los bonos salariales ha reducido la discusión del significado y las consecuencias de la responsabilización a una valoración del simbolismo y las injusticias de los incentivos salariales. Por otro lado, el concepto de *accountability* en Brasil se ha asociado con una sola medida de control gubernamental y la más fácilmente asociada con la gestión de recursos humanos. Si se van a continuar los incentivos salariales, como una entre varias opciones de política, sus costos deben ser

calculados y comparados con los beneficios medidos por métodos apropiados de evaluación de políticas públicas.

Palabras clave: Responsabilización, políticas de accountability, Incentivos salariales.

ABSTRACT

In Brazil, accountability policies, called *políticas de responsabilização*, are fundamentally based on the attribution of salary incentives and other bonus payments to teachers and schools. For many authors, the equation between educational accountability and salary bonuses has reduced the discussion of the meaning and consequences of the accountability to an appraisal of the symbolism and injustices of salary incentives. And the other hand, the concept of accountability in Brazil has become associated with a single measure of government control and the one most easily associated with human resource management. If salary incentives are to be continued, as one amongst a number of policy options, their costs need to be calculated and compared to the benefits as measured by appropriate methods of public policy evaluation.

Keywords: Responsabilização, accountability policies, salary incentives.

1. INTRODUCTION

In Brazil, authority for the supply and development of public basic education, from pre-school to the end of secondary, resides with state, Federal District and municipal governments, with the latter obliged to concentrate on pre-school and elementary schooling and the state governments on secondary education. The role of the federal government is to plan the national system, supply curriculum guidelines, offer financial and technical assistance and monitor results. This means that the overwhelming majority of public schools and teachers are under the charge of subnational educational authorities and that high-stakes accountability policies designed to influence school results, where they exist, are the purview of these same authorities.

According to a recent count, since the year 2000 a total of 10 state governments in Brazil have implemented accountability policies in the form of salary incentives for members of school staff on the basis of school performance scores while a further three have deployed salary bonus policies using other criteria, including teacher assiduousness and participation in in-service training (Scorzafave, Ferreira and Dorrigan, 2015). The first state governments to use salary incentives based on external assessment measures of pupil performance, in policies designed to impact school results, were those of Rio de Janeiro and Ceará in the year 2000 and 2001 and the latest, the State of Tocantins, in 2012. The number of municipal governments to adopt similar incentive payments is unknown but the pioneer was the municipality of Sobral which, starting in 2001 (MEC/INEP, 2005), created a bonus scheme for teachers of reading and writing that later served as inspiration for the State government of Ceará.

The operationalization of these salary incentives through one-off bonus payments, and the fact that the bonus payments are virtually the only form of high-

stakes accountability¹² to take root in Brazil, has heavily influenced the broader debate on *responsabilização*, as educational accountability is known. For many authors, the equation between educational accountability and salary bonuses has reduced the discussion of the meaning and consequences of *responsabilização* to an appraisal of the symbolism and injustices of salary incentives. In the absence of local legislation such as the No Child Left Behind act with its variety of possible consequences, often negative, for annual school performance figures, the concept of accountability in Brazil has become associated with a single measure of government control and the one most easily associated with human resource management.

The absence of a full range of accountability measures, and of a clear mandate for government to act on the behalf of citizens in the oversight of school results, have lead Brooke (2011) to affirm that the state salary incentive policies are first and foremost an employer's strategy to increase school productivity. Indeed, with the lack of widespread publication of school level results and the absence of any consequences for those schools that miss out on the staff-wide bonus payments by failing to reach their goals, it can be questioned if the model of "*responsabilização*" as used by the 10 states is really a version of educational accountability or not. While the bonus payments may serve the purpose of adding weight and focus to school level performance goals, the elements of accountability associated with transparency, the correct use of public funds and, above all, the achievement of a minimum levels of quality, are largely missing.

¹ Understood as a type of accountability that over and above the symbolic effects of publicly relating school quality to pupil results can have important financial or career consequences for schools and those involved.

² Another form of high-stakes accountability, of smaller scope, is the distribution of state funds to municipal governments in the State of Ceará, and possibly other states, in accordance with a municipal education quality indicator based on pupil performance. Some forms of low-stakes accountability, including signs on the outside of schools publicizing the school's overall quality score (IDEB), also exist.

The peculiarities of Brazilian educational accountability policies, their dissociation from questions of graduation rates, school-choice and standards, and their almost exclusive use as an instrument of public sector human resource management, has not protected them from high levels of criticism and grass-roots opposition from within the teacher unions. One of the purposes of the present article is to show that some of the criticism is inappropriate given the use of mainly US research carried out in a decidedly different cultural and educational context. In an attempt to make the discussion more relevant, the article will also look at recent research carried out in Brazil to determine whether the weight of the evidence supports the belief of educational administrators that bonus payments can promote gains in pupil performance without provoking damaging side-effects. The underlying goal of the article is to enable education policy-makers and managers to go beyond the ideological debate and take a dispassionate view regarding the benefits or otherwise of the model of salary incentive accountability currently in vogue in Brazil.

2. ONSET OF EDUCATIONAL ACCOUNTABILITY

Contrary to the chronology offered by some authors (Becker, 2014), the establishment of accountability policies in Brazil does not coincide with the start of external assessment at the beginning of the 90s. In the early phases of testing, starting with the federal government's sample-based Basic Education Assessment System (SAEB) in 1990 and universal state-level assessment systems in the following years in such states as Minas Gerais, Ceará, Paraná and São Paulo, the responsible government departments expressed the purpose of testing in exclusively diagnostic terms. SAEB, for example, was designed to "understand patterns of student learning", "promote discussion of curricular proposals", "construct a solid data base on the public

school teaching-learning process” (Maluf, 1996; Pilati, 1994). The state systems followed the same model, seeking, as in the case of Minas Gerais in 1992, to

carry out diagnoses and create a data base comprising reliable information on state schools ...with the purpose of designing a plan for the improvement of teaching capable of altering the current scenario of low productivity and high levels of grade repetition (Antunes, 1994).

For this reason, the language of accountability was entirely absent from the early documents and the purpose of testing restricted to the use of performance data to understand the curricular difficulties of both teachers and pupils. If the authorities thought to use the results of the tests to guide management policy, this was restricted to the design of in-service training based on the teaching-learning difficulties identified by the tests. Indeed, there is evidence that the idea of attaching accountability type consequences to the results of the tests was actively spurned. In a statement at the time, the Secretary of Education of Minas Gerais recognised that teachers could not be blamed for the failure of pupils even if it were their responsibility to look for a solution to the problem (Mares Guia Neto, 1992). The education authorities of the time saw public education as a cooperative venture between government and teachers and while committed to the need to disseminate a “culture of assessment”, were wary of making comparisons and of attaching moral or material consequences to the results of the tests (Brooke, 2006).

The first phase of test expansion in Brazil is therefore dissimilar to the expansion of state-level testing that took place in the United States in the mid-1970s. In that country, state-level authorities reacted to the widely discussed evidence that public schooling was in decline by adopting so-called minimum competency tests designed to assess students’ basic skills at elementary and secondary levels. Between

1975 and 1978, 33 states had already mandated the setting of minimum competency standards and the remaining states had some form of legislation under study (PIPHO, 1978). Simultaneously, a number of states fixed standards that defined the minimum test scores necessary for high school graduation, thereby making schools responsible for their students' progress beyond high school. With the expansion of testing in the following decade, provoked largely by the report *A Nation at Risk* and its emphasis on higher expectations and standards for high school graduates, course graduation requirements were tightened, further raising the stakes attached to high school test results (Dee, 2003). By the time the next phase of test driven reforms came into play, prompted by accountability initiatives in Texas and North Carolina in the 1990s and the ensuing No Child Left Behind (NCLB) legislation in 2002, the association between testing and high stakes accountability was already well established (West and Peterson, 2003). As far as teachers were concerned, testing by the state government had long been associated with performance standards and the way parents and authorities judged the work of schools.

The first salary incentive policies in Brazil, supported by a still incipient accountability discourse, took a decade to materialize. It seems likely that some of the thinking was supplied by the NCLB model. However, the use of pupil scores as at least one of the criteria for bonus payments in the States of Rio de Janeiro and Ceará was also the product of greater confidence in the reliability of the state testing systems and the ability to compare results over time. These technical advances had been achieved by the incorporation of Item Response Theory into the test construction process and the creation and interpretation of proficiency scales for Portuguese and Maths. Following the lead of the SAEB tests, that undertook its own technical overhaul in 1995, the state testing systems began to adopt TRI technology and incorporate the

national proficiency scales in 2000, starting with the Minas Gerais system of educational assessment.

Taking the first examples of salary incentives as indicative of government thinking, it is clear that the association between pupil results and bonus payments was part of a policy designed to improve school management and efficiency. In the case of the New School program in the State of Rio de Janeiro, for example, the point system, that determined whether the school had either a satisfactory overall evaluation or had “progressed” enough from one year to the next to merit the staff-wide bonus payment, took into consideration a number of school governance indicators as well as the measures of school efficiency (retention rates) and school academic performance. In the 2004 version of the policy, the school management indicators included the correct presentation of school accounts, staff assiduity, integration with the community and enrolment management (Rio de Janeiro, 2004a; 2004b). There being no externally provided parameters or standards to determine the satisfactoriness of school performance, schools were grouped according to average pupil socioeconomic status and compared with each other to determine which had progressed most. By creating a bonus scheme that offered all members of staff the chance to benefit from improved school management and results, the authorities laid bare their beliefs regarding the possible reasons for poor school performance and the benefits of promoting collaboration between members of school staff.

The general question of government efficiency and performance had grown as a political objective with the adoption of a Brazilian version of government reform, loosely modelled on the UK’s New Public Management model (Abrucio, 2006), and the end of inflation due to the introduction of the new currency, the Real. With the 1995 financial crisis, the privatization of state banks, the need to renegotiate state debt and a number of other regulations to control subnational finances, the State governments

were forced to abandon the predatory financial practices that had used inflation as a way to solve the fiscal deficit and seek new means to increase efficiency (ABRUCIO; GAETANI, 2006). At the same time, the idea of strategic government planning and the use of indicators gained momentum with the launch of two Federal Government initiatives: the 1997 Study of National Integration and Development and the 2000-2003 Multi-Year Plan, called Advance Brazil, which referenced all government programs in terms of objectives, target populations, goals, indicators and standards and procedures for the evaluation of quantifiable results (Muniz, Silveira and Bechelaine, 2010).

In the case of Ceará, the program to pay bonuses to teachers was much simpler. In a straightforward ranking of schools, that eschewed any attempt to take pupil socioeconomic characteristics into consideration, the New Millennium Educational Prize offered payment of an annual bonus to all staff members of the top 100 schools according to average performance in 4th and 8th grade Portuguese and Maths. For the top 50 schools the prize was 100% of the specified value while for the next 50 the prize was worth half this value. For the top scoring pupils there were also prizes. The purpose of the program was to promote public recognition for higher performing schools, to improve the school environment by creating a “climate of quality” in order to influence school results and to raise the standards of public education (Law 13.203 of February 21, 2002). While avoiding the use of the word accountability or its Portuguese equivalent, the authorities repeatedly expressed the belief that education professionals can cooperate at school level to improve collective results. The Ceará incentive program has undergone numerous modifications over the years and remains the only example with a track record of more than a decade.

3. RECENT DEVELOPMENTS

More recent state-level salary incentive programs have sought to avoid some of the pitfalls of their predecessors. For example, the new São Paulo program³, introduced in 2008, gets round the problem of a lack of standards by relating bonus program payments to the degree to which the school has achieved its targets. The setting of school level performance targets is based on a composite indicator of school quality comprised of both pupil survival rates and performance averages, called IDESP. To create a relationship between the bonus payment and the degree to which the IDESP targets are met, the Secretariat of Education has created an index based on the degree of target achievement for each cycle at each school (Ferraz, 2009). In schools that have an achievement rate of 100% of the target, for example, all teachers and school staff receive a bonus equivalent to 20% of their annual salary, making the bonus scheme very much like a private sector performance pay policy. However, given the other goal of the accountability system, to reduce absenteeism, all individual payments depend on employee assiduousness. To receive the bonus, employees must have worked at least two thirds of school days over the previous year.

To avoid the risk of teachers concentrating on pupils most likely to show improvement on test results, thereby giving less attention to those at the higher and lower extremes of attainment, the program has innovated by creating four levels of pupil performance, going from Below Basic to Advanced, and calculating the proportion of pupils at each level. In this way it is possible to frame the performance component of the composite IDESP index in terms of the discrepancy between the measured position of the pupils and the ideal of having every pupil at the advanced level (Soares, 2009). The use of this discrepancy indicator as the measure of school performance signals the

³ A previous bonus program, introduced in 2001, was based on average staff assiduousness and average pupil performance as measured by the state assessment system – SARESP (BROOKE, 2008).

desirability of improving the scores of the lower performing students, thereby favouring greater equality of results.

The bonus program of the State of Espírito Santo, introduced in 2010, is also based on a locally generated indicator of school quality. In this case, however, the Indicator of School Development is a function of both school results and school effort. The logic of the effort component is that the higher the socioeconomic level of students the lower the school's efforts to produce good results. The component uses a calculation of the difference between the socioeconomic level of the school, on a scale from 0 to 5, and a maximum value arbitrated at 10. The school with the highest socioeconomic level will have an effort index of 5, while the school with the lowest socioeconomic status will have an index of 10. Using the same logic in relation to the different education levels offered by the school, and assuming an increase in the effort with the advance of the student through the first grades of elementary school to high school, total school effort is estimated in accordance with the composition of the body student by educational level. In this way, the School Effort Indicator is found to take into account both the socioeconomic level of the school and the diversity of educational levels and to allow comparisons between schools that are not vitiated by differences in student background and conditions.

Further innovations include the creation of an additional level of pupil performance entitled "The Excluded", given zero weighting, comprising all pupils who drop out or who fail to appear for the end of year assessments. Overall, the performance levels send out the same message as in São Paulo, benefiting the school that manages to promote its students to higher levels while reinforcing school efforts to avoid dropout and discourage the "hiding" of lower performing students before the tests.

The Minas Gerais bonus program is somewhat different from the São Paulo/Espirito Santo model. Called the Results Agreement, the methodology implemented in 2008 and repeated annually since then⁴ extends to all government departments. In the case of the Secretariat of Education, the agreement relates to the results of the central Secretariat, the regional educational superintendents (SREs) and all schools. For a particular school, the final performance score is a composite, deriving 65% from the Secretariat's score, 10% from the local SRE scores and 25% from the school's own "results indicators". This composition clearly shows that the responsibility for the school's final score is shared between the three administrative levels and mitigates somewhat the criticism that accountability unilaterally shifts responsibility onto the shoulders of teachers.

The school's results indicators relate to a set of 10 targets, typically expressed in terms of average student proficiency per subject and year. As an example, the sixth target, weighing 2.14% of the total school note, concerns average 9th grade student proficiency in Portuguese. The school gets top marks if the observed value for the target is greater than or equal to the recommended level for the year, in accordance with the classification system published by the State's assessment system. If the school result is lower than the recommended level, an equation is used to calculate the difference between the proficiency achieved and the target. The percentage of the productivity award to which the school is entitled is given by the addition of its own indicators to those of the Secretariat and the regional office. This maths is intended as a way to ensure that the school receives some benefit even if it fails to reach 100% of the targets.

One of the interesting aspects of the Results Agreement is the process for setting targets. This process culminates in the signing of a term of commitment on the

⁴ Discontinued in 2015.

part of SREs and schools after a period in which the targets are listed school by school on the internet. If the school does not agree with its targets, there is a procedure for the suggestion and evaluation of new goals. Although a centralized process, the possibility of school counterproposals gives legitimacy to the targets and justifies the signature of the term of commitment. Not that the Secretariat abides by all the counterproposals: of 3,988 schools in 2010, 10% sought to change their goal, and of these, only 10% succeeded (Brooke, 2011).

State-level accountability programs were given new impetus by the creation of a composite 'indicator of basic education development' (IDEB) by the federal government in 2007. By creating a single indicator of both pupil performance and pupil flow, the Ministry of Education made it possible to express the quality of the school on a scale from 1 to 10 and for the first time establish a national discussion on the direction and speed of educational improvement. More importantly, by combining the results of the biennial Prova Brasil tests of 5th and 9th grade Maths and Portuguese, launched in 2005, with average rates of grade completion using data from the annual education census, and then calculating the necessary growth in school performance in order to reach the target of average OECD performance by the year 2021, the Ministry created what has become the main instrument of educational monitoring and management for both municipal and state governments (Vidal and Vieira, 2011).

At the time the IDEB was launched, the head of the Ministry's educational assessment institute had no doubts as to its potential as an instrument of accountability. In an article two years later, he explained:

When analysing the potential and risks of an educational accountability program, we concluded that such programs can help raise student achievement, but are not without risks. In Brazil, the main risk with an

accountability system focusing only on student performance in external examinations is to aggravate the grade repetition and dropout problems. That was the main reason for creating the IDEB. By combining the Prova Brazil scores with approval rates, the IDEB aims to curb both the indiscriminate failing of students and the practice of passing students who have learned nothing. The IDEB was also used to set targets for schools and school systems and thus provoke the greater commitment of state education departments and municipalities so that, by 2021, Brazil can reach the current educational stage of developed countries. As long as there is dissemination of results and target-setting for schools, the emphasis of the accountability system will belong to the states and municipalities (Fernandes and Gremaud, 2009. p.238)

The only deficiency of the Ministry's system of targets is that it depends on the Prova Brasil whose biennial frequency is less than ideal for education authorities seeking a more rigorous monitoring based on annual targets. Local authorities also emphasise the need for control over the production and dissemination of school results to ensure deadlines and continuity. No surprise, therefore, that many States with their own assessment systems, like São Paulo, Espírito Santo, Amazonas, Rio de Janeiro and Pernambuco, have created their own purpose-built state-level indicators of basic education quality rather than adopt IDEB. Other states, Paraíba, Tocantins and Sergipe, have chosen not to establish their own assessment systems and are currently using either Prova Brasil results or the IDEB indicator as the principal criterion of school performance and the payment of salary incentives.

4. OPPOSITION

Given that Brazil has 26 states and one Federal District, the fact that only 10 states have created salary bonus accountability programs, 7 on the basis of their own

assessment systems and 3 on the basis of biennial federal government data on school performance, indicates that the policy is considerably less than unanimous (Scorzafave and Dorigan, 2015). There is no equivalent data for the municipalities but it is telling that of the six state capitals with proprietary assessment systems, only one, Rio de Janeiro, has instituted bonus payments for teachers. There is evidence that some authorities created plans to establish accountability policies that were then abandoned for political reasons. This was the case of the Federal District in 2011 and the state capitals of both Belo Horizonte and São Paulo. The demise of the first salary incentive policy in Rio Janeiro, the Nova Escola program, brought to an end by the state governor elected in 2006 on a campaign promise to the teachers' union to do away with the policy, is also documented (Brooke, 2008).

The strength of opposition to the policy varies from state to state, depending on the nature of the relationship between the state government and the local public teachers' union but in no case have teachers declared support for the policy. The overwhelming opinion in academic circles is also contrary to the use of salary incentives. The arguments against the policy were laid out in systematic form at the 33rd meeting of the National Association of Postgraduate Studies and Research in Education (ANPED) in 2010 when faculty and students met to create the Movement Against High-Stakes Testing in Education (MOVIMENTO, 2010). The same concerns were also aired in a National Education Council document concerning basic education evaluation (CNE, 2012). Among the document's conclusions are the need to avoid high-stakes tests, because of their potential to "corrupt education quality indicators and generate fraud" (p.24), and to shun pupil results as the basis for the calculation of salary incentives because of the lack of empirical evidence as to the possible consequences of such a policy. A special edition of the Education and Society journal in 2012 on Public Education Accountability Policies also serves as a source for the

arguments against high-stakes accountability based on salary incentives (FREITAS, 2012).

The arguments against the policy can be roughly divided into two categories (Brooke, 2013b; Furtado and Magrone, 2015). The larger of the two contains the reflections of those who reject accountability for doctrinaire reasons, often based on beliefs regarding the structural determinants of accountability, as demonstrated by developments in the United States. In this view, Brazil is just the latest to succumb to changes in the role of the State, the move towards privatization and the expansion of external assessment to better adapt schooling to the demands of capitalist production (Freitas, 2012). Accountability, in this perspective, is part of the strategy to advance the privatization of public education through the advance of meritocracy as a philosophical and organizational ideal.

Further doctrinaire arguments include the rejection of accountability not primarily for its connection with capitalist development but for the impropriety of neo-liberal models of educational management adopted over the last decades by conservative governments. In this view, the problem of accountability resides in its authoritarian, private sector origins that denote a competitive, market-oriented philosophy entirely inappropriate for the public sector and alien to the ideals of an education system premised on democratic freedoms. The key words of the criticism are conservatism and managerialism, both used to describe a method of undemocratic school control based on the production of academic results, as measured by tests, that leaves out the “political quality” of education in terms of the teaching and exercise of citizenship (DEMO, 1998) and substitutes concerns for equity and equality with those of control (Gentili and Silva, 1995). In these terms, high-stakes accountability distorts the true purpose of education by constraining teachers and creating inappropriate goals.

These criticisms are not easily countered. Their authors are unlikely to be convinced of the possibility of gains for the Brazilian education system as the result of the borrowing and use of neo-liberal methods. This difficulty is revealed by the refusal even to see the relevance of such concepts as efficiency and efficacy in school management because of their prior association with Taylorism (Lima and Afonso, 2002). Arguments citing the advances in education access and better test results as indicative of democratization and the improvement of education quality would not be accepted as evidence of the advantages of neo-liberal thinking as these gains are exclusively individual. As such they do not capture the collective, societal impact of education nor include school processes as part of the definition of quality (Ximenes, 2012).

The difficulty of establishing a dialogue between those responsible for the framing and implementation of accountability policy and those who reject neo-liberal philosophy and methods on principle, does not mean that the only problem of salary bonus accountability is a question of ideology. There is a considerable amount of evidence that the policy is prone to numerous other problems, both regarding the outcomes desired by its proponents and the undesired side-effects it is able to provoke. However, it is only worth studying these practical questions of public policy if we are allowed to suspend ideological judgement and assess the policy in the same terms in which it is framed by state and municipal governments. Given the purpose of the article to inform public sector education administrators, we need to admit the legitimacy of accountability as an instrument of educational management, albeit unproven, and evaluate the pertinence of the criticisms regarding its practical consequences once implemented.

5. INEQUALITY

Practical considerations regarding the undesirable side-effects of salary incentive accountability fall into the second category of arguments used to combat the policy. Of these, the most urgent, given Brazil's struggle to reduce regional and social differences in pupil performance levels, is that salary incentives lead to an increase in inequality. In newspaper articles published in 2010 and 2012, Maria Alice Setubal, President of the Centre for Studies and Research in Education, Culture and Community Action (Cenpec), raised the thesis that the payment of bonuses could create an unhealthy competition between schools and cause even greater inequalities. The reasoning was that the teachers with permanent civil servant status and freer, therefore, to choose their place of work, would seek schools with better results to increase their chances of receiving the bonus. These schools are likely to be centrally located, with students of higher sociocultural level. Consequently, schools located in the urban periphery, with highly vulnerable clienteles, would be left to the temporary contract teachers, with less experience and probably less expertise.

Although Setubal failed to produce evidence based on research, a recent focus group study in Rio de Janeiro has shown that at least in some cases teachers have been known to migrate to schools that have reached their targets and received bonus payments in precisely the way predicted by Setubal (Cerdeiran and Almeida, 2013). The research does not indicate, however, if the movement of teachers could exacerbate inequality as it simultaneously suggests that the schools with greatest difficulty in reaching their targets are precisely those with higher levels of performance.

A further criticism, raised by Freitas (2012), cites research on the Chicago school system taken to show that academically disadvantaged students were left behind because the NCLB proficiency counts did not provide incentives for schools to

direct more attention towards them (Neal and Schanzenbach, 2010). Because these US results also showed significant increases in achievement for students near the proficiency standard, the results were taken as confirmation of the tendency of teachers to concentrate their efforts on students who are near the average so as to maximize the proportion of students who pass. This tendency, if real, can obviously harm students of lower performance levels and lead to increasing inequality. To the extent that the worst performers tend to be of lower socioeconomic level, this strategy of concentrating effort on average performance students can lead to socioeconomic segregation within the school.

However, both Setubal and Freitas may have overgeneralized by extending their criticism to all models of salary incentives. In the case of Setubal, concerned with between school inequality, the model of salary incentives under observation was that of São Paulo which at the time had yet to make any allowance for differences in pupil SES levels in the fixing of targets and the calculation of bonus payments. In other states, including Espírito Santo and Rio de Janeiro, this criterion had already been adopted, thereby eliminating the advantage of teachers moving to a higher SES school. As Brooke (2013a) suggests, the Espírito Santo model of “school effort” could perfectly well be extended to include such factors as teacher/pupil ratios, school location, school violence etc. and thereby incorporate an “affirmative action” component to attract higher qualified teachers to poorer neighbourhoods.

In the case of Freitas, concerned with within school inequality, the model was the criterion of Average Yearly Progress (AYP) based on pass rates, as per NCLB guidelines. None of the Brazilian states use a pass rate based on standards as the criterion for their incentive policies. On the other hand, five states use criteria that could be construed as similar to AYP, including Ceará, and Amazonas which employ average school performance levels as measured by their own assessment systems, and

Tocantins, Sergipe and Paraíba which determine salary incentives on the basis of the school average indicators supplied by the ministry. The other states with their own assessment systems have incorporated the idea of school targets that, in increasing number, take into consideration the proportion of students at each level of performance, precisely as a way to combat any tendency of teachers to concentrate their efforts on pupils near the average and neglect students of lower levels of attainment. Therefore, in the case of within school inequality, it would seem that the criticism is pertinent for one particular model of salary incentive accountability but does not indicate an inherent defect of all policies. In Brazil, these policies vary according to type, methodology, periodicity, coverage, bonus and other criteria. The degree to which the chosen policy is able to promote equality between and within schools, and counter the other undesirable side-effects still to be discussed, could be taken as a measure of its quality, which would then suggest that the overall problem of salary incentive accountability is one of appropriate design rather than of fatal flaws.

Notwithstanding, recent research in Brazil raises questions regarding the thesis that the impact on equality depends on whether the policy is designed for this purpose or not. In a national study of urban 5th and 9th grade pupils, the authors used the Prova Brasil results from 2007, 2009 and 2011 to compare the average performance of all states divided into three categories: those without any bonus scheme since 2007; those with bonus schemes that in their design seek to reduce inequality by creating targets for different levels of performance; and those with bonus schemes that do not seek to reduce inequality (Scorzafave and Dorigan, 2015). Different measures of system equality were employed, including the standard deviation for pupil scores and the Gini coefficient, in a model designed to relate change in system inequality to a range of factors including the type of bonus. Unexpectedly, findings show that states with policies designed to reduce inequality showed an increase in the inequality of

results between 2007 and 2011 when compared to others. However, as the researchers were unable to track the effects of the bonus policies for more than one or two years, given that many were introduced in 2009 or later, it is possible that the impact had still to appear. This seems likely given that the state systems with the highest levels of inequality were precisely those that implemented bonus schemes designed to reduce inequality. More important for the appreciation of the research, however, is the fact that the type of inequality that the bonus schemes seek to attack by promoting improvement at all levels of pupil performance is the within-school variety while the inequality measured by the researchers is the range of results across the system. If the research were to be repeated using the standard deviation of within school scores it would be possible to offer an answer on this crucial question.

A further source of inequality, that could possibly be related to the existence of salary incentives, is the informal, unseen selection of students by school directors. Although illegal, given that schools have mandatory catchment areas and are obliged to enrol all pupils on a first come first served basis, researchers in São Paulo and Rio de Janeiro see the operation of a hidden quasi-market, in which some schools seek to encourage or discourage pupils from enrolling in accordance with their area or residence, colour or socioeconomic status, leading to informal patterns of school specialization (Érnica and Batista, 2011; Costa and Koslinski, 2011). Although still not clear whether the preference for higher performing pupils is due to the incentive policy, to broader definitions of school prestige and quality held by parents and/or staff or just the demands by teachers to reduce pupil heterogeneity, the work of Cerdeira and Almeida confirms that there are indeed some school principals in Rio “who do not accept low performing pupils” (Cerdeiran and Almeida, 2013. p.9).

6. CURRICULUM

The most common criticisms regarding the side-effects of Brazil's high-stakes accountability policies concern their undesirable impacts on the curriculum. The types of curriculum impact can be classified as a. the definition of school purpose, b. the curriculum that is effectively taught and c. the style of teaching.

- a) It is held that the payment of salary incentives contingent on external assessments can lead not only to a reduction in the diversity of school activities but also the suppression of social, cultural and political objectives that contribute to the integral development of the pupils (XIMENES, 2012). In other words, high-stakes accountability in Brazil can reduce the purpose of schooling to a small number of largely academic goals due to the importance of good test results to the chances of receiving a salary bonus.
- b) The consequences of high-stakes testing on the choice of specific academic contents is also seen as a problem, commonly referred to as curriculum narrowing, when schools and teachers focus their resources on the teaching of the subject areas which are covered by the tests while other subjects get insufficient attention or are eliminated. As Freitas (2012) states:

When the tests include certain disciplines and leave others out, teachers tend to teach the subjects addressed in the tests [...] Assessments generate traditions. They direct the attention of teachers, administrators and students. If what is valued in a test are reading and mathematics, then special attention will be devoted to these, leaving other formative aspects out [...]. (Freitas, 2012. p.389)

- c) In a further criticism, it is held that "teaching to the test" becomes more of a problem given the increased importance of test results. The choice of subject content according to the probability of its inclusion in the tests and the

exaggerated use of test preparation are seen as typical of this style of teaching, more concerned with test results than the permanent acquisition of abilities.

In the absence of a solid body of local research on the impact of salary incentives on the way schools and teachers define their objectives and organize their practice, most critics rely on evidence from the United States. For this purpose, there is a vast array of research, some of it summarized by Pedulla (2003), Madaus, Russell and Higgins, (2009) and Newberg-Long (2010). By and large, this literature confirms the existence of curriculum narrowing and relates this phenomenon to the growth in test-based reform and high-stakes testing in that country.

The assumption of many of those who transpose the findings of US research to the Brazilian context is that as high-stakes accountability affecting teachers is largely a US invention then all we need to determine this policy's local impact is to study its history in its country of origin. However, as is well-documented, in the dynamics of policy borrowing the original proposal is subjected to a process of internalization and 'indigenization' that can lead to policy formats that are entirely new (Steiner-Khamsi, 2012). The political pressures to which the policy is subjected, as the result of the local cultural and educational environment, requires the use of the local policy context as the analytical unit.

One of the relevant differences between the two countries regarding the forms and consequences of accountability is related to the culture of government employment and the legislation governing public teacher careers. In Brazil, teaching jobs are acquired through public exams, promotions awarded on the basis of academic qualifications or years of service and teachers protected from being evaluated, transferred or losing their jobs in an environment that is free of risk of school intervention and closure. This scenario makes it impossible for the authorities to relate

teacher career movements to any measure of teacher performance, including pupil results, or to suggest any type of negative consequence for teachers or schools that fail to reach their targets. This would explain why the only type of high stakes accountability in use in Brazil is a collective award given to the school, rather than the individual, and expressly excludes any threat of loss of pupils, diminished funding, school intervention, school closure or loss of career opportunities, as in the US. As the Secretariat of Education of Rio de Janeiro states on its web page:

Meritocracy is ... above all, the appreciation of the most competent. The... appreciation of the good professional. It also includes the appreciation of professionals in schools that present the best performance. And in this case all professionals in the school unit receive a bonus. There is no punishment. Only the reward. What is more, there is no individual bonus payment. The winner is the school team. All for the sake of teamwork. Everyone wins at the school: teacher, director, kitchen staff, guard. The goal is not individual; it is the school unit. There is no punishment, even if the school does not reach its goal (Seeduc, 2016).

The absence of any negative consequences and the fact that the salary incentives are the only example of accountability in those states to have implemented the policy, makes for a fairly mild accountability climate. This in contrast to the US situation in which different types of accountability coexist and are compounded in what might be described as a more saturated accountability environment. Marsh et al (2011) allude to this type of environment when ascribing the lack of results for the bonus scheme for teachers in the city of New York to what they call “other accountability pressures”. For these researchers, the lack of clear correlation between the bonuses received by teachers and student outcomes can be attributed, at least in part, to the numerous other high-stakes consequences linked to student results of greater

significance to teachers, including the NCLB goals, parental demands and the city's own teacher assessment and promotion policies. This would make for the considerably more anxiety producing accountability environment described by Ravitch (2010). The way in which the bonus policy is received and impacts teacher behaviour could therefore be more a function of local conditions, including the co-existence with different types and levels of accountability, than of the nature of the policy itself.

Brazilian research on curriculum narrowing has been slow to get off the ground even in states such as Ceará with more than 15 years of experience of teacher salary incentives. Indeed, the lack of academic production in that State might even be taken as an indication that curriculum narrowing is not the burning issue the critics suggest. In Minas Gerais, one piece of research has broached the issue using a number of the same questions used by the team of Boston College researchers who in 2003 carried out a national survey of teachers in the United States to determine the perceived effects of state-mandated testing programs on teaching and learning (Pedulla et al, 2003). In the Minas Gerais version of this research, relevant to the present discussion because of the salary incentive program in this State, the questionnaire was answered by 978 school principals from all regions and with more than three years of job experience (Borges, 2016).

The responses of 80% of the school principals indicate that the external assessments to which the school is subjected have led to greater articulation in the management of the curriculum between teachers and between the different phases of basic education. But this articulation has not been achieved as the result of a slavish attention to the details of the tests. Only 15% responded that they recommend their Portuguese and Maths teachers to follow the test matrices when doing their planning and only 20% admit to a reduction in the time dedicated to artistic or cultural activities as the result of the external assessments. When asked the direct question as to

whether the pressure to obtain good results leads teachers to narrow the curriculum, only 17% agreed.

One reason for school principals to believe that the tests themselves have had little effect on the choice of subject matter is the high level of alignment between the tests and the official curriculum as described in detail on the Secretariat of Education's website and in the material distributed to teachers. This material, called the Common Basic Contents, act as common core standards and are followed by teachers according to 94% of school principals.

The high degree of observance of the common core raises the question as to whether the narrowing to which the critics refer is more the product of official decisions regarding the curriculum, reflecting the desire of the authorities to focus attention on certain contents rather than others. This, however, would be a separate discussion, related to the perceived need to guarantee that all students acquire the basic skills related to Portuguese and Maths in the lower years, and Maths, Portuguese and Science in the final years of primary and secondary schooling. In this case, the authorities are using the official curriculum and its related tests to signal priorities. If the salary bonus contributes to this end, it would be favouring compliance with official curriculum policy rather than distorting or narrowing it.

On the positive side, between 76% and 88% of school principals in Minas Gerais feel that external assessments have promoted greater attention to teaching-learning objectives, have increased the rigor of teachers in the design of their own assessments, have increased the monitoring and oversight of teaching, have increased the accompaniment of pupils and the adoption of strategies to increase school success and have led teachers to be more attentive in adapting their practice to the needs of the pupils. While only 14% feel that the pupils are under intense pressure to improve

their test results and only 13% that teachers are pressured to give less attention to pupils with learning difficulties, 79% see external assessments as a good strategy to measure what pupils learn and 84% that they have led to greater exchange of knowledge and pedagogical experience between teachers.

In other states with salary incentives, research suggests that while the practice of school principals has changed little, the bonus policy has increased the level of interest regarding school targets (Medeiros, 2014). In Rio de Janeiro, for example:

.. It was noticed that the bonus policy implemented by SEEDUC/RJ did not cause major changes in the management practices of state schools in the Valencia municipality. It only generated greater concern regarding the achievement of school goals. As a result, school managers began to focus more on making parents, students and teachers more aware of the importance of achieving the goals rather than on appropriating the results of evaluations to improve learning (Medeiros, 2014. p.87).

While the lack of concern with the use of test results for the purpose of guiding future teaching-learning is worrisome, the conclusion that the bonus policy has increased attention to school goals would be deemed proof of its success by the Rio de Janeiro education authorities. The assumption behind the policy, especially when based on school-level performance targets, is to give significance to the targets and thereby encourage teachers and others to seek gains in pupil performance. If this can be achieved without any apparent disruption to teaching or management, the policy can be seen as contributing to the improvement of education quality.

Other research in Rio de Janeiro, already mentioned in relation to the question of inequality, also describes the contradictory opinions of teachers and school principals for and against external evaluation (Cerqueira and Almeida, 2013). It is telling,

however, that the authors believe school principals and teachers to be living a “process of legitimation” of IDEB and other performance indicators as a consequence of the ability of these to express the reality of the schools. By and large, the authors believe the “resistance of teachers to be lessening in tandem with the visible improvement in school results” (p.13). School principals agree that it was necessary to implement external assessment and the bonus policy in order for teachers to demonstrate greater effort, that the external assessments have brought more dedication and collective effort and that the impact on the curriculum has been positive.

Although it is still early days for Brazilian research into the curricular consequences of high-stakes assessment, some articles have attempted to summarize the situation. One such article, by Bonamino and Sousa (2012), concentrates on the States of São Paulo and Pernambuco. The first observation concerns the efforts of these states to improve the school-level appropriation of external assessment results by establishing an official “unified” curriculum and corresponding curriculum materials. Even if the natural order of events would have put the assessment system as a successor to the definition of curriculum guidelines rather than its predecessor, the alignment of curriculum and testing is to be applauded. And although it is unclear to what degree this alignment can be attributed to the bonus payment policies, it can be presumed that, in the absence of clear curricular guidelines, a bonus policy based on school-level performance targets would represent an even greater challenge to teachers.

The second observation is that the assessment systems have gradually adopted the role of orienting the way teachers plan and teach their classes, especially in the way they evaluate their students. According to Bonamino and Sousa (2012) this indicates the appropriation of the methods of external assessment through an emphasis on the application of different types of tests, including mock exams. The

authors agree that it is early days yet but believe that the assessments might be “exacerbating the concern of school principals and teachers to prepare their pupils for the tests” (p.396). On the other hand, it is taken as proven that high stakes testing has led to a more informed discussion on the school curriculum in terms of the basic skills in reading and Maths that need to be guaranteed for all pupils (Bonamino and Sousa, 2012). The ambiguity of this conclusion is due to the fact that the concern with testing, and even the signs of “teaching to the test”, can, under certain circumstances, be an indication that teachers who were without guidelines have now adopted a way to improve their teaching.

7. EFFICACY

The most scathing of the criticisms levelled at the salary incentive policy is that it doesn't work. This would mean that despite the organizational effort and expense, the gains in pupil performance taken to be the measure of policy impact, are either non-existent or insufficient to justify the investment. Given this lack of proof of the policy's efficacy, one author has not only decried the waste of resources but accused education administrators of a lack of ethics for subjecting schools to untried remedies (Freitas, 2012, 2013).

Looking at US research, the results are wildly contradictory, due to the variety of contexts, policy formats, research designs, researcher biases and other factors. This creates the need for an independent meta-analysis, happily met by the 2011 report on research carried out in the US and elsewhere by the National Research Council Committee on Incentives and Test-Based Accountability in Public Education published by the National Academies Press (Hout and Elliott, 2011). After applying strict criteria for the selection of the best research methodologies, the National Research Council study looks closely at just 11 experiments, five from the United States and six from

other countries. The US research shows that salary incentives programs for teachers, also called payment by results, performance incentives or variable remuneration, have not influenced the performance levels of students in a consistent and meaningful way. On the various scales used by researchers to measure learning gains, the average effect ranges from just 0.02 to 0.06 standard deviations. In the US context, therefore, the argument in favour of monetary incentives is not proven.

Outside of the United States, however, the research tells a different story. Studies meeting the meta-analysis criteria conducted in India and Israel, for example, show an average 0.08 s.d. impact. In the case of India, Muralidharan and Sundararaman (2009) conducted a two year experiment on the impact of a 3% salary bonus on the learning of mathematics and the mother tongue, granted either individually or collectively to teachers. The average gain was 0.19 s.d., but with extremes of 0.22 s.d. in Maths in the case of schools with individual teacher incentives and 0.15 s.d. in schools with collective incentives (Muralidharan and Sundararaman, 2009). In the case of Israel, the incentives were paid to teachers of English and Maths in 49 high schools. There was a significant increase in both the number of students who passed the year and in average test scores, mainly of lower performing students. The average effect of the salary incentives on student scores was 0.12 s.d. (Lavy, 2002; 2009).

Incipient Brazilian research begins to show results that look similar to those of India and Israel. In São Paulo, researchers used a matching technique to create a control group composed of either municipal schools in São Paulo or state schools from elsewhere in the country and a “differences in differences” methodology to compare the control group with the state school treatment group. The matched, “shadow” schools, with the same observable characteristics as the São Paulo state schools, including Prova Brasil results from before the start of the bonus policy, showed lower results

after the introduction of the bonus policy. The authors found positive and significant gains for 4th grade results in the São Paulo state schools. On the national SAEB scale, the effect was equivalent to 6.4 points for Maths and 3.7 for Portuguese (Oshiro and Scorzafave, 2011).

More recent research in the State of Pernambuco also shows positive results. To get round the problem of the lack of control group, Furtado (2015) also used a matching technique, based on propensity scores, to compare the performance of schools included in that state's Educational Performance Bonus policy with municipal schools untouched by the policy. To this end, the research created 5th and 9th grade pairs of state and municipal schools from the State of Pernambuco that were equal on the municipal human development index, SES of pupils, total number of pupils and pre-bonus (2008) performance in Maths and Portuguese. The results of the comparison are broken down by grade. In the 5th grade the bonus policy schools are shown to have a low but significant advantage for the early years of the policy, between 2008 and 2011, with an average gain of 4 points in comparison with the control group in Maths and 3 points in Portuguese. The 9th grade results, on the other hand, show a consistent and significant advantage of, on average, 5 points for both Math and Portuguese for each of the periods studied (Furtado, 2015; 2016). However, the fact that the gains for the full period of the study (2008-2012) are the same as for the early period (2008-2010) leads to the conclusion that the bonus-induced performance improvement is not constant over time but a phenomenon concentrated in the first years of the policy, that permits the school to reach and then maintain a new level of attainment.

Incidentally, the research was able to show that the increase in performance as measured by the Pernambuco state assessment system (SAEPE), used to determine whether the schools have reached their targets, is the same as the increase measured by the national SAEB tests that have no bearing on the bonus policy. This finding is

relevant insofar as it shows that SAEPE has not been subjected to inflationary pressures, brought on by the “gaming” sometimes provoked by high-stakes tests, and continues, therefore, a reliable measure of performance. The loss of reliability is a problem raised by a number of authors (Hout and Elliott, 2011) and has been used, in the Brazilian case, as a further argument for the abandonment of accountability policies (Freitas, 2013).

8. OTHER SIDE-EFFECTS

The Brazilian literature identifies a number of other possible side-effects of the bonus policy. These include the increase in undesirable competition between teachers and schools (Nogueira et al, 2013), a loss of teacher autonomy in the definition of curriculum objectives, an undue increase in pressure on teachers (Ximenes, 2012), the generalization of the myth of teacher irresponsibility (Ximenes, 2012) and the promotion of dishonesty and fraud (Movimento, 2010; CNE, 2012).

Because of the tendency to assume a correspondence between the United States and Brazil in the forms and consequences of accountability, most criticisms that seek to sustain arguments regarding the side-effects of the policy have drawn on US literature. This is clear in the claim that accountability provokes undue competition. In the Brazilian case, the goals and the bonus payments are school-wide, paid on a collective basis and in accordance with levels of target attainment by the school. In this scenario, a school is effectively competing with its own past performance and any competition between schools and teachers would be unexpected and probably extraneous to the policy.

The perceived loss of teacher autonomy in the definition of curriculum objectives also seems misplaced in the Brazilian context. For those who study

classroom practice, there is no memory of a golden era in which the teacher was able to express his or her competence and creativity in the choice of content or the construction of a curricular sequence. The curriculum, as taught, has rarely been the subject of individual teacher choice. Prior to the advent of tests and their matrices, curriculum guidelines were provided by the text book, supplied free for teacher and pupils by the Ministry. After the introduction of testing, and the more recent phase of “basic content” guidelines by states concerned to create an alignment between the curriculum and its assessment, such as São Paulo, Minas Gerais, Pernambuco, Ceará and Rio de Janeiro, teachers have had access for the first time to a common core. It has been argued that the definition of curricular content by year and subject and the correspondence between the curriculum and the tests in these states has finally made it possible for the authorities to explain and for teachers to understand the results of assessments (Brooke and Cunha, 2011). Considerably overdue, the state common cores and the progress in the definition of curricular expectations for each grade are steps in the right direction that, to some extent, can be attributed to both low and high-stakes assessment systems having taken the lead.

If the local research existed, other criticisms might also be proven inappropriate given the nature of accountability in Brazil. In its absence, further discussion will have to wait.

9. CONCLUSIONS

The fact that more than half of the Brazilian states have chosen not to use their test results for high-stakes accountability purposes does not sit well with the belief that this country is inevitably set to repeat the experience of the United States. This over-mechanistic belief in the inexorable demands of capitalist development, that underestimates the importance of local history and culture in the decisions of

administrators to take or leave the use of salary incentives as a management tool, may even have contributed to the virulence of the opposition to the policy in Brazil on the grounds of the exaggerated use of all types of educational accountability in the US.

The history of the adoption and rapid development of accountability in the US shows the existence of a logic concerning the supply and oversight of public services that is closely related to the American citizen's identity as a taxpayer and the belief in the right to information regarding the correct and constructive use of all government funds. With the addition of middle-class parental pressure for good school results, deep confidence in the idea of individual merit and government conviction regarding the need for standards to drive educational improvements, US culture has been able to generate such a stifling education accountability environment as to provoke a popular revolt and the end of the NCLB legislation.

The onset and development of high-stakes accountability in Brazil obeys a different logic. The use of pupil attainment to calculate salary incentives is first and foremost about gains in efficiency and employee productivity, in a reaction to demonstrably poor results and the absence of any alternative method to promote better performance among public sector employees. As evidence, it should be remembered that all the salary incentive policies are contingent on indicators of teacher assiduousness, including those states which do not include pupil performance criteria. This means that if the teacher is absent in excess of the annual limit for days off, he or she is not awarded the school-wide bonus. In this situation it is moot whether the salary incentive is a policy to promote gains in pupil performance or an attempt to reduce widespread teacher absence. Whichever the answer, it is clear that the policy has little to do with public information and transparency, the correct use of resources, the satisfaction of curriculum standards or the promotion of students.

In reality, the discussion should be more about the use and consequences of performance pay policies, as used in the private sector, than of accountability. Do they work in their sector of origin and are they likely to work in schools if the public sector doesn't sell products for profit in a competitive market? A partial answer is supplied by Adams, Heywood and Rothstein (2009) who show that in the rare cases of collective bonuses in the US, the size of the group is a crucial factor and that in the case of small groups, the increase in collective productivity is generally the result of an improvement on the part of the least productive members. However, where the indicators of productivity are quantitative, as in the case of pupil results, the risk of their suffering distortion are considerable. Nevertheless, one of the study's conclusions is that, depending on the circumstances, the positive benefits of performance pay can outweigh the negative side-effects:

Several analyses by economists, management experts, and sociologists have concluded that narrowly quantitative incentive schemes have, at times, somewhat improved the average performance of medical care, job training, welfare, and private sector agents. The documentation of perverse consequences does not indicate that, in any particular case, the harm outweighed the benefits of such narrow quantitative accountability (Rothstein, 2009, p.97).

What the present article has argued is that given the enormous differences between the US and Brazil, including the social and cultural characteristics of that part of the population attended by public education and the way the population relates to government, there is actually little reason to imagine that the accountability policies of the two countries would be similar in their meaning and consequences. While in the US the accountability climate in schools is seen as multi-layered and often threatening, in Brazil the climate is described as tame and largely free of negative consequences,

given that the accountability policies are essentially an adaptation of private sector performance pay schemes.

It is these differences that makes it problematic to imagine that Brazil will follow the same path as the US and inappropriate to use research from that country as evidence for the evaluation of Brazilian salary incentives to promote pupil performance gains. What, then, does the Brazilian research show?

The evidence would seem to confirm that the nature and impact of Brazilian accountability policies are indeed the product of the local context and conditions. When the impacts of salary incentives are measured in terms of performance gains, Brazil behaves like other developing or middle-income countries rather than like the US. The gains are significant and if not a solution to Brazil's educational problems, look potentially positive in cost-benefit terms. If the case of Pernambuco is an example, the average 5 point gain on the national SAEB scale would not be easy to achieve with other types of policy such as in-service teacher training.

A possible explanation for the positive results, in contrast to the near zero scores in the US, is related to the recency of both performance targets and salary incentive policies and the fact that these policies are effectively alone in expressing the message of the importance of pupil results as a school objective and as a measure of school quality. In a more saturated accountability environment, this message is transmitted in numerous different ways, making it more difficult to show gains in relation to any specific accountability measure. In other words, where pupil results are used to determine school choice, school finance, school interventions, teacher evaluations and promotions etc. the likelihood of finding an association between salary incentives and performance is likely to be lower.

The research also makes it clear that there are significant differences in the operation of the incentive policies, making it difficult to assume they are all equally defective or likely to have equal impact. In other words, the design of the policies can be expected to affect their influence and consequences. This is especially important for the discussion of the side-effects of the incentive policies.

The capacity of salary incentives to exacerbate inequality is the more worrying of the possible side-effects. It seems that incentive policies can contribute to between school inequality even when their design acts explicitly to inhibit discrimination or the movement of teachers by incorporating a measure of pupil SES or “school effort” in the fixing of school targets. This could be the result of the contribution of the salary bonus to the school’s prestige and subsequent illegal selection of pupils, but more research is needed. An increase in within school inequality on the other hand, due to teachers paying less attention to pupils with lower performance, is less likely to occur where the incentive policy is not based on pass marks or a single school average. Any growth in inequality of this type, even in cases where different performance levels have been incorporated, might be due to the wider tendency for the benefits of school quality improvement to accrue to pupils of higher SES, as shown by research demonstrating that the increase in the school average is generally accompanied by increasing inequality (Franco et al, 2007).

Research on the effect of incentive policies on the curriculum has not shown the purely negative consequences foreseen by its critics nor that the quality of teaching is suffering unduly for this reason. The signs are that high-stakes assessments can be having a positive effect on the fulfilment of official curriculum objectives, on the level of collaboration between teachers, on the importance attributed to achieving school targets and even on the attention paid to adapting teaching practice to the needs of pupils. Where there is evidence of narrowing and “teaching to the test”, the discussion

as to how far these effects can be considered negative is far from conclusive. It is probable that some of the diversity has gone from the curriculum, as the price paid for the increase in performance by virtue of the acquisition of basic skills by a larger proportion of pupils. However, this conclusion is tentative and should be the subject of further research.

What the research also says is that education policy-makers can no longer ignore a growing body of evidence regarding the differences between alternative salary bonus models and the probability of different impacts. The dispassionate appraisal of the policy must not only weigh the evidence regarding side-effects but also observe the relationship between these and the type of incentive scheme under consideration. If salary incentives are to be continued, as one amongst a number of policy options, their costs need to be calculated and compared to the benefits as measured by appropriate methods of public policy evaluation.

REFERENCES

Adams, S. J., Heywood, J. S. and Rothstein, R. (Eds.) (2009). *Teachers, performance pay and accountability: What education should learn from other sectors*. Washington, D.C.: Economic Policy Institute.

Becker and Fernanda da Rosa (2014). Large-Scale Assessments and Quality in Basic Education. In D.P.Napier (ed) *Qualities of Education in a Globalised World* (pp. 63-81). Rotterdam: Sense Publishers. Retrieved from http://s3.amazonaws.com/academia.edu.documents/40804231/Napier_2014_Qualities_of_Ed_Diversity_1.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1466896676&Signature=H3RqT%2F0e363kKqXklu%2FoFyTDyKs

[%3D&responsecontentdisposition=inline%3B%20filename%3Dquality_of_education.pdf#page=72](#)

Bonamino, A., Sousa, S. Z. (2012). Três gerações de avaliação da educação básica no Brasil: Interfaces com o currículo da/na escola. *Educação e Pesquisa*, 38 (2), 373-388.

Borges, E.M. (2016). Avaliações Externas em Larga Escala no Contexto Escolar: Opiniões de Diretores Escolares da Rede Estadual de Ensino de Minas Gerais. (PhD Thesis). Faculty of Education. Federal University of Minas Gerais.

Brooke, N. (2006). Educational Accountability in Brazil: An overview. *Gestão em Educação*, 9 (1), 7-36

Brooke, N. (2008). Responsabilização Educacional No Brasil. *Revista Iberoamericana de Evaluación Educativa*, 1(1) Retrieved from <http://www.rinace.net/riee/numeros/vol1-num1/art7port.pdf> Acesso 01/07/2016

Brooke, N. (2011). As novas políticas de incentivo salarial para professores: uma avaliação. In: Fontoura, H. A. (Org.) *Políticas Públicas, Movimentos Sociais: Desafios à Pós-graduação em Educação em suas múltiplas dimensões*. Livro 3. Rio de Janeiro: ANPEd Nacional. Coleção ANPED SUDESTE. Retrieved from <http://www.fe.ufrj.br/anpedinha2011/ebook3.pdf> p. 2011

Brooke, N. (2013a). Sobre a Equidade e Outros Impactos dos Incentivos Monetários para Professores. *Est. Aval. Educ.*, 24(55), 34-62

Brooke, N. (2013b). Políticas Estaduais de Responsabilização: Buscando o Diálogo. In Adriana Bauer e Bernardete A. Gatti (orgs.) *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: Implicações nas redes de ensino, no currículo e na formação de professores*. V.2. Florianópolis: Insular.

- Brooke, N. and Cunha, M.A.A. A. (2011). Avaliação externa como instrumento da gestão educacional nos estados. In Fundação Victor Civita (Ed.), *Estudos e Pesquisas Educacionais - Volume 2* (pp. 17-79). São Paulo: Fundação Victor Civita.
- Cerdeira, Diana Gomes da Silva and Almeida, Andréa Baptista de. (2013). *Os Efeitos da Política de Avaliação e Responsabilização Educacional na Rede Pública do Rio de Janeiro*. 36ª Reunião Nacional da ANPEd. 29 de setembro a 02 de outubro. Goiânia-GO. Retrieved from http://36reuniao.anped.org.br/pdfs_trabalhos_aprovados/gt14_trabalhos_pdfs/gt14_3062_texto.pdf
- CNE. (2012). *Diretrizes Conceituais e Operacionais para a Avaliação na Educação Básica (Subsídios)*. Brasília: Conselho Nacional de Educação.
- Costa, M.C., Koslinski, M.C. (2011). Quase-Mercado Oculto: Disputa por Escolas “Comuns” no Rio de Janeiro. *Cadernos de Pesquisa*, 41(142)
- Fernandes, R., Gremaud, A.P. (2009). Qualidade da educação: avaliação, indicadores e metas. In: Veloso, F. et al. (Orgs.). *Educação básica no Brasil: construindo o país do futuro* (pp. 213-238). Rio de Janeiro: Elsevier
- Ferraz, C. (2009). Sistemas educacionais baseados em desempenho, metas de qualidade e a remuneração de professores: Os casos de Pernambuco e São Paulo. In: Veloso, F. et al (Eds.). *Educação Básica no Brasil: Construindo o país do futuro*. Rio de Janeiro: Elsevier,
- Franco, C. et al. (2007). Qualidade e equidade em educação: reconsiderando o significado de “fatores intra-escolares”. *Ensaio: Avaliação e Políticas Públicas em Educação*, 15(55) 277-298

- Freitas, L. C. D. (2012). Apresentação. *Educ. Soc.*, 33(119), 345-351 Retrieved from http://www.scielo.br/scielo.php?script=sci_issuetoc&pid=0101-733020120002&lng=pt&nrm=iso
- Furtado, C. (2016). Impacto da bonificação educacional em Pernambuco. *Estudos em Avaliação Educacional*, São Paulo. – no prelo.
- Furtado, C. (2015). *Responsabilização Educacional em Pernambuco*. (PhD Thesis). Programa de Pós-graduação em Educação. Faculdade de Educação. Universidade Federal de Juiz de Fora (UFJF).
- Furtado, C. and Magrone, E. (2015). Responsabilização em Educação: Mapeando Elementos. *Rendimento, Desempenho e Desigualdades Educacionais*, 5(1)
- Gentili, P. and Silva, T.T. (1995). *Neoliberalismo, Qualidade Total e Educação: visões críticas*. Petrópolis: Vozes.
- Hout, M. and Elliott, S. W. (Eds.). (2011). *Incentives and Test-Based Accountability in Education*. National Research Council Committee on Incentives and Test-Based Accountability in Public Education. Washington, D.C.: The National Academies Press.
- Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy*, 110(6),1286-1317.
- Lavy, V. (2009) Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. *American Economic Review*, 99(5), 1979-2011. Retrieved from https://www2.warwick.ac.uk/fac/soc/economics/staff/vlavy/lavy_aer_teachers_bonuses_99.5.pdf

- Lima, L. C.; Afonso, A. J. (2002). *Reformas da Educação Pública: Democratização, Modernização, Neoliberalismo*. Porto: Afrontamento,
- Madaus, G., Russell, M., Higgins, J. (2009). *The Paradoxes of High Stakes Testing*. Charlotte, N.C.: Information Age Publishing Inc.
- Mares Guia Neto, W. A (1992). Realidade da Educação em Minas Gerais. *Estudos em Avaliação Educacional*, 6.
- Mec/Inep. (2005). *Vencendo o Desafio da Aprendizagem nas Séries Iniciais: A Experiência de Sobral/CE. Série Projeto Boas Práticas na Educação, No. 1*. Brasília: MEC/INEP.
- Medeiros, Denise Barra. (2014). *Os Efeitos da Política de Bonificação do Estado do Rio de Janeiro nas Ações Gestoras de Escolas Estaduais do Município de Valença*. Dissertação Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública CAEd/UFJF. Juiz De Fora,. Retrieved from <http://www.mestrado.caedufjf.net/wp-content/uploads/2015/04/DENISE-BARRA-MEDEIROS.pdf>
- Movimento Contra Testes De Alto Impacto Em Educação. (2010) Retrieved from <http://sites.google.com/site/movimentocontratestes/home> .
- Muniz, R. M., Silveira, M. C. D., Bechelaine, C. O (2010). Caminho em Direção à Gestão por Resultados em Minas Gerais: Uma análise do Plano Mineiros de Desenvolvimento Integrado. In: Guimarães, T. B.; Pordeus, I. A., et al (Ed.). *Estado para Resultados*. Belo Horizonte: Editora UFMG.
- Muralidharan, K. and Sundararaman, V. (2009). Teacher performance pay: experimental evidence from India. *National Bureau of Economic Research: 119*(1), 39 – 77.

- Neal, D.; Schanzenbach, D.W. (2010). Left behind by design: proficiency counts and test-based accountability. *Review of Economics and Statistics*, 92(2), 263-283.
- Newberg-Long, Donna Kay. (2010). *Narrowing of Curriculum: Teaching in an Age of Accountability*. PhD Dissertation. University of Denver
- Oshiro, C.H.; Scorzafave, L.G. (2011). *Efeito do Pagamento de Bônus aos Professores sobre a Proficiência Escolar no Estado de São Paulo*. ANPEC - 39° Encontro Nacional de Economia. Foz de Iguaçu: ANPEC. Retrieved from <https://www.anpec.org.br/encontro/2011/inscricao/arquivos/000-e95d077296d03faa27e429dc66192da8.pdf>
- Pedulla, J. et al. (2003). *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers*. Boston: The National Board on Educational Testing and Public Policy. Retrieved from <http://www.bc.edu/research/nbetpp/reports.html>
- Pipho, Chris. (1978). Minimum Competence Testing in 1978: A Look at State Standards. *Phi Delta Kappan*, 59(9), 585-88. Retrieved from http://www.jstor.org/stable/20299107?seq=1#page_scan_tab_contents
- Ravitch, D. (2010). *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic Books.
- Rio De Janeiro. (2004b). *Programa Nova Escola: Informativo 1 – Avaliação do fluxo e da gestão escolar*. CAEd/ Secretaria de Estado da Educação,
- RIO DE JANEIRO. (2004a). *Programa Nova Escola: Avaliação 2004*. CAEd/ Secretaria de Estado da Educação.

- Rothstein, Richard. (2009). The Perils of Quantitative Performance Accountability. In Adams, S. J., Heywood, J. S. and Rothstein, R. (eds) *Teachers, performance pay and accountability: What education should learn from other sectors*. Washington, D.C.: Economic Policy Institute,
- Scorzafave, L. G.; Dorigan, T. A. (2015). *Efeito das políticas de bonificação aos professores sobre a desigualdade de proficiência no ensino fundamental*. VIII Reunião da ABAVE - Avaliação de Larga Escala no Brasil: Ensinaamentos, Aprendizagens e Tendências. 18 a 21 agosto, Florianópolis. Retrieved from http://abave.com.br/ojs/index.php/Reunioes_da_Abave/article/view/344/141
- Seeduc- Rio de Janeiro. (2011). *Educação. Notícias. Entenda Para Que Servem o Saerj e o Saerjinho*. Retrieved from <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=616581> .
- Soares, J. F. (2009). Índice de desenvolvimento da Educação de São Paulo (IDESP): bases metodológicas. *Revista São Paulo em Perspectiva*, 23, (1), 29-41.
- Steiner-Khamsi, G. (2012). Understanding Policy Borrowing and Lending: Building Comparative Policy Studies. In Gita Steiner-Khamsi and Florian Waldow (eds.) *World Yearbook of Education 2012: Policy Borrowing and Lending in Education*. Routledge..
- Vidal, E. M.; Vieira, S. L. (2011). *Gestão Educacional e Resultados no Ideb: Um Estudo de Caso em 10 Municípios Cearenses*. VI Reunião da Abave. Fortaleza, Ceará.
- Ximenes, S. B. (2012). Responsabilidade Educacional: Concepções Diferentes e Riscos Iminentes ao Direito à Educação. *Educ. Soc.* 33(119), 353-37