

This article was downloaded by: [University of Rhodes]

On: 13 October 2014, At: 00:01

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



African Journal of Research in Mathematics, Science and Technology Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rmse20>

Primary Teachers' Experiences Relating to the Administration Processes of High-stakes Testing: The Case of Mathematics Annual National Assessments

Mellony Graven^a & Hamsa Venkat^b

^a Education Department, Rhodes University, Grahamstown, South Africa

^b Wits School of Education, University of the Witwatersrand, Johannesburg, South Africa

Published online: 07 Oct 2014.

To cite this article: Mellony Graven & Hamsa Venkat (2014): Primary Teachers' Experiences Relating to the Administration Processes of High-stakes Testing: The Case of Mathematics Annual National Assessments, African Journal of Research in Mathematics, Science and Technology Education, DOI: [10.1080/10288457.2014.965406](https://doi.org/10.1080/10288457.2014.965406)

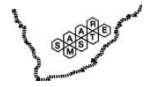
To link to this article: <http://dx.doi.org/10.1080/10288457.2014.965406>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever

or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>



Primary Teachers' Experiences Relating to the Administration Processes of High-stakes Testing: The Case of Mathematics Annual National Assessments

Mellony Graven^{1*} and Hamsa Venkat²

¹ *Education Department, Rhodes University, Grahamstown, South Africa*

² *Wits School of Education, University of the Witwatersrand, Johannesburg, South Africa*

*Corresponding author. Mellony Graven. Email: m.graven@ru.ac.za

In this paper we highlight teacher experiences of the administration of high-stakes testing, in particular, of the 2012 Annual National Assessments (ANAs). The exploration is based on data gathered across two primary numeracy teacher development projects in the Eastern Cape and Gauteng in the form of open-ended questionnaires designed to elicit teacher experiences of the 2012 Numeracy ANAs (at Grades 1–3) and Mathematics ANAs (Grades 4–6). Fifty-four teachers across 21 schools (including fee-paying and non-fee-paying schools) completed the questionnaire. Using a grounded approach to the analysis of data, we note that, while teachers state support for the purpose of the ANAs, several concerns emerge in relation to their administration. These concerns fall largely into two categories: concern for learner experiences and concern for the implications of the administration processes (including the use of exemplars and the marking process) for teacher practices. The primary purpose of the paper is to raise awareness of the need for further discussion and research into the way in which ANAs result in possible unintended consequences.

Keywords: Annual National Assessments; primary mathematics; numeracy

Introduction

The authors of this paper collaborate, respectively, with teachers in 12 primary schools in the broader Grahamstown area and 10 primary schools in the Johannesburg area. The schools include both township and suburban schools, and fee-paying and non-fee-paying schools. We both engage in in-school teacher development activity working alongside teachers, and run in-service primary mathematics teaching courses for the teachers in these projects. Recurring concerns have been raised in relation to the Mathematics Annual National Assessments (ANAs) by the teachers we work with. These systemic assessments, broadly implemented since 2011, are focused on providing system-wide information on learner performance for both formative and summative purposes (Department of Basic Education, 2011). Reporting of national results on achievement according to gender, poverty index quartiles and the language of learning and teaching is intended to stimulate follow-up action across all levels of the education system:

the immediate target is the various tiers of education with the intention of supplying credible information to assist teachers, principals and department officials to strengthen their existing and planned efforts of improving the quality of teaching and learning. (Department of Basic Education, 2012, p. 3)

The ANAs have been written by all public school learners across Grades 1–6 since 2011 and in 2013 were extended to Grade 9 learners. Aside from 2011, when 2010 tests were deferred to early 2011 owing to teacher industrial action and disruption to the normal school year because of the 2010 Soccer World Cup, the ANAs have been written in September, that is, 9 months into the school year rather than at the end. Given our focus on primary schools, our attention in this paper is on the Numeracy ANAs that are written in the Foundation Phase (Grades 1–3) and Mathematics ANAs in the Intermediate Phase (Grades 4–6).

We sought specifically to explore teachers' perceptions and experiences of the administration processes related to the Numeracy/Mathematics ANAs encompassing the use of external exemplars, and the administration, marking and collation of results. Our motivation for this focus relates to our work within the perspective of the curriculum as 'contextualised social process' (Cornbleth, 1990). This remit means that teachers' local contextual concerns about curriculum, pedagogy and assessment have to be taken seriously if we are to impact the ground in positive ways. Additionally, if patterned responses and regularities are seen in teachers' identification of elements of the policy terrain that are perceived as disruptive to their work as teachers of mathematics or to the learning of mathematics, these areas are particularly important to investigate and understand. We refer to a cluster of possible teacher concerns, which include learner stress, the length of the ANAs as an assessment process, the fairness of the assessment in relation to language and reading abilities of learners and the timing of the tests. Some of these issues have strong antecedents in the international literature on 'high-stakes' assessments and the consequences of this for curriculum coverage and breadth in classrooms. We also review alternative bodies of research relating to notions of 'care' as a more emergent theme that has been written about in the terrain of primary teachers' identities.

Related Literature

High-stakes Assessments – Nature, Purposes and Consequences

That high-stakes assessments have an influence on what happens in schools and in classrooms is widely acknowledged in the international literature (e.g. Elmore, Ablemann & Fuhrman, 1996). A range of purposes can underlie high-stakes assessments. Postlethwaite and Kelleghan (2008) note that gaining information at a system-wide level about the 'quality' of educational outcomes is one purpose. Kellaghan and Greaney (2001) point out that this information often includes information on outcomes at disaggregated levels within the system – for example, by gender, race, socio-economic background – with this kind of disaggregation featuring, as noted above, within the South African government's reporting of results. In some cases, systemic assessments may enable policy-makers to judge the success of intervention mechanisms such as curriculum reform or resource-level changes. In the South African context, Kanjee (2007) has noted that national assessments can provide tools for evaluating the impact of educational policies in the system.

Accountability measures often make differences between schools the subject of particular attention – with this leading to further policy-level intervention consequences for low-performing schools and positive praise and market-driven 'choice' consequences for high-performing schools (Fuhrman & Elmore, 2004). More mathematically orientated forms of disaggregation can also feature, within both the design of tests and the reporting of performance, through focus on specific content areas within the curriculum and cognitive demand for instance. While the numeracy/mathematics ANAs are reported to be designed with attention to the breadth and balance of topics in the national curriculum statement, reporting currently pays less attention to mathematical disaggregation by topic or cognitive demand, and focuses more on disaggregation by province, urban/rural school background and socio-economic background through attention to school quintile (Department of Basic Education, 2012).

Negative consequences of high-stakes assessment have also been noted in the international literature. Key amongst these consequences is concern relating to a narrowing of the enacted curriculum through the practice of 'teaching to the test'. Popham (2001) describes this phenomenon in a US context:

American teachers are feeling enormous pressure these days to raise their students' scores on high-stakes tests. As a consequence, some teachers are providing classroom instruction that incorporates, as practice activities, the actual items on the high-stakes tests. Other teachers are giving practice exercises featuring 'clone items' – items so similar to the test's actual items that it's tough to tell which is which. In either case, these teachers are teaching to the test. (p. 16)

Alongside this 'cloning' there is also evidence of reorganization of the timing of teaching around high-stakes assessments (Borko & Elliott, 1999). Thus Diamond (2007, p. 306) argues, following his study of high-stakes testing in Chicago:

What has been consistently demonstrated across multiple studies is that teachers reallocate instructional time to specific content that is likely to be tested and spend more time on preparing for tests. These types of instructional responses are likely to lead to the inflation of test scores.

In the South African assessment context, the language used in standardized international (Howie, 1997; Reddy, 2006) and national assessments such as the ANAs (Henning & Dampier, 2012) and learners' reading abilities (Bohlmann & Pretorius, 2008) have been raised as contributing factors to poor assessment performance. These issues were raised specifically in the context of African-home-language children having difficulties accessing item instructions given in English or Afrikaans on international tests. In the ANA context, in the Foundation Phase (where use of the home language in classrooms is encouraged in the national language policy context), test papers are provided in learners' home languages as requested by schools, and questions are read out to learners in Grades 1 and 2. However, in the Intermediate Phase (where language policy demands a switch to either English or Afrikaans), the ANA tests are only administered in English or Afrikaans.

The effects on performance of this shift in the language of assessment for the majority learners are, however, not clear. The political nature and power of language means that epistemological access to mathematics often comes secondary to concern for access to the power of learning English in schools in South Africa (Setati, 2008). It is thus uncertain as to whether learners, schools and parents would elect to write assessments in African languages if they were made available for learners in Grades 4 upwards. Broader assessment evidence in South Africa continues to point to lower mean performance for learners taking tests in African languages in comparison to those taking the tests in English or Afrikaans (Green, Parker, Deacon & Hall, 2011). This could be connected with Bohlman and Pretorius's (2008) findings that reading ability in English is more significant than English language proficiency as a predictor of mathematics achievement.

Teacher Identity of Care

An emergent finding from our data analysis related to teachers' concerns about learner anxiety and lack of 'care' in the administration processes of the tests. Anxiety fuelled by the nature of mathematics teaching and the culture of mathematics classrooms has a history across many countries and across phases (see, Boaler, 1997; Geist, 2010; Graven & Buytenhuys, 2011). A smaller body of writing has considered anxiety specifically relating to high-stakes assessments in primary schools. Walls (2008) conducted research on learner experiences of a year 5 standardized nationwide numeracy test in Queensland, Australia. Her study found that:

Participants' reports revealed a significant proportion of children experienced mild to severe stress and some suffered severe feelings of disappointment, loss of confidence and decline in mathematical self-belief as a result of the test process; participants also reported extra pressures on families created by the test, including managing children's stress and disparities between siblings and friends' test performances. (p. 476)

Walls concludes with a call for further research into how learning and assessment must find ways to serve our children's 'best interests'. Similarly, Reay and Wiliam (1999), working in the context of Year 6 Standard Assessment Tasks (or 'SATs') in England, note the 'strong currents of fear and anxiety permeating children's relationships to the SATs process' (p. 349). They pointed too, to the

shifting self-identifications of learners in relation to their expected levels of performance, poignantly picking up the learner comment that features in the title of the paper: 'I'm frightened I'll do the SATs and I'll be a nothing' (p. 345). At one level, there is an argument that the relationship between anxiety and 'high-stakes' assessment is perhaps an inevitable stress that learners need to learn to deal with as they progress through school. The small body of writing relating to high-stakes assessments in primary schools raises, though, specific concerns about the self-fulfilling nature of identifications with failure at a young age, and the negative consequences of these internalizations for learners' preparedness to put further effort into future learning.

Focusing more broadly on issues related to caring for learners in primary schools in the context of standardized curricula and assessments and the highly comparative orientations to performance that frequently characterize these contexts, Nias (1999) has discussed care in relation to 'affectivity' and holistic responsibility for learners. Her concern in the English National Curriculum context was that these standardizations often stood markedly against the construal of primary teaching as 'a "culture of care", whose underlying values emphasize the importance of making children feel secure, happy and cared for' (p. 68).

This overview presented four types of concerns, that is, those related to test content, fairness in relation to language access, issues related to test administration and care/anxiety. These issues were structured, in relatively open-ended ways, into the kinds of questions we asked in the teacher questionnaire. These issues also function as a conceptual framework of assessment issues that allow us to locate the findings in relation to the broader literature.

Methodology, Sample and Related Limitations

In order to gather teacher views on the ANAs we designed an open-ended questionnaire with seven questions, and a final request for 'any other issues/suggestions' with respect to the ANAs. Our questionnaires addressed a range of issues concerning teacher perceptions and experiences of the ANAs relating to: their purpose and value; the use of exemplar papers; the administration and marking of the ANAs; the correspondence with topics taught; and the extent to which performance reflected learners' numeracy competence. These foci were driven primarily by the issues that teachers raised with us in earlier conversations, but reflected too the four elements of our awareness of the literature base around high-stakes assessments. This paper is based on three clusters of questions relating to the administration of the ANAs that were phrased as follows:

Cluster 1: the use of exemplars

Did you use the exemplar papers given to you before the ANAs? If you did use them did you find them easy to use? Did you find them useful to use? If so, in what ways?

Did you feel that using them helped your learners? If so, in what ways?

Cluster 2: conducting the ANAs

Tell us about how you administered the ANAs in your school. What issues arose for you in administering these? Did you have to monitor behavior during the tests? Did you have to repeat questions more than once? Did your pupils understand the instructions provided in the tests?

Cluster 3: marking processes

How did you find the marking of the ANAs? Did any issues or concerns arise in the marking? Was the memorandum accurate and easy to follow?

Participation in the questionnaire was voluntary and questionnaires were given to teachers present in broader teacher workshop sessions in Grahamstown and Johannesburg, respectively. In this respect it is an opportunity sample. Fifty-four teachers from across 21 schools completed the questionnaires (24 teachers from the Eastern Cape and 30 from Gauteng). The relatively small sample of teachers across only two provinces means that what is illuminated here cannot be considered representative of the general population of teachers but rather is generative of aspects of ANA administration that appear

to require further research, consideration and discussion. While all 54 teachers completed the questionnaire, in each cluster one to two teachers left questions unanswered.

Teacher responses were coded and analysed in grounded ways according to themes that recurred across teacher responses. To find recurrent themes every sentence was coded within a category. In some cases sentences were coded in two categories. That is, a sentence referred to two emergent themes, for example, 'some learners were nervous and confused'. In some cases two sentences that related directly to one another were jointly allocated to one theme. For example: 'I did not like that we did not facilitate our own classes. We needed to shift classes'. If comments did not relate directly to the questions they were coded as 'other'. The themes that emerged within teacher responses to each cluster of questions were summarized in tables with counts of the number of teacher responses to each theme or subtheme. Below we discuss issues of concern that arose in relation to teachers' experiences of the administration of the ANAs (including use of exemplars, conducting the ANAs and the marking process). Since concerns across the clusters are interrelated, we present these as crosscutting themes, thus referring to data derived from multiple clusters.

Findings and Discussion

Purposes of the National Assessment Practice

Teacher views on the administration of the ANAs were located within relatively widespread acceptance of the purpose and validity of the ANAs. Thus concerns raised relating to the administration of the ANAs, discussed below, were juxtaposed with responses to other questions showing overall support for their value (see also Graven & Venkatakrishnan, 2013).

In summary the ANAs were perceived as 'good' for (a) standardizing content coverage; (b) making explicit expectations about what will be assessed; (c) providing information on learners' levels of understanding; and (d) providing guidance on required content coverage. Examples of the kinds of positive points summarized above are as follows:

The values and purpose of ANA are good because they help educators to do curriculum pacing very well and to cover the content prescribed for that class or grade.

I think the idea of the ANA is brilliant because it evaluates where our standards is.

Within this context of general support for the ANAs we discuss below the emergent themes of teacher concerns across our data summary tables of the three clusters of questions.

Concern for the Implications for Learners

Teacher concerns largely fell into two categories relating to the implications for learners and the implications for their mathematical teaching and assessment practices. The emerging theme *Concern for the implications for learners* includes the following subthemes, each of which will be presented below:

- processes of conducting the ANAs;
- language of the ANAs and/or language fluency of learners;
- learner negative experiences;
- time issues related to learner pace.

Processes of conducting the ANAs

While several teachers commented on the smoothness of the administration process of the ANAs and several others provided neutral descriptions of how they were administered according to the rules, there were 25 comments that raised concern over the rules of how the ANAs must be administered. Twelve of these comments related to teachers not being allowed to facilitate their own classes. For example:

I had to monitor the grade 3s of which I didn't like – to supervise somebody else's class, my mind was thinking about my own class as there was a 'stranger' in front of them.

We administered in a very rigid way. The learners are not used to the strict way of invigilation where someone else who does not teach them has to guard them.

Another six teacher comments related to not being allowed to provide any explanation for what was being asked or not being allowed to mediate the questions. For example:

We could not go through the questions with the learners (Not Allowed!). Only example questions which the learners understand.

Another seven teachers raised concern about the rule of not being allowed to read to learners in Grade 3 and above, which links with the concern for learners' poor language and reading levels that arose repeatedly in teacher comments across questions (discussed below). For example:

The challenge was with the learners who have a language barrier because in grade 3 we were not allowed to read questions for learners & that gave some learners a challenge as they cannot read.

Language of the ANAs and/or language fluency of learners

Thirteen comments related to the language of the 2012 ANAs being too difficult for learners. Five similar comments were noted in relation to the exemplar papers with another seven comments relating to questions being too challenging or difficult. (In this case it was difficult to tell whether this referred to mathematical challenge or the challenge of the language). For example:

Language used in ANA, learners are not familiar with it, our learners have problems with language. You have to repeat the question more than once but ANA does not allow that, not giving us an opportunity to guide learners through the questions.

Related to this issue of language, although connected here with learners' reading abilities, 10 comments (eight for the ANAs and two for the exemplars) noted learners' reading difficulties as a key concern. For example:

ANA question papers are full of language, learners have to read a lot before solving or doing numbers. Learners cannot read and others are slow in reading this resulted in their poor performance.

The papers were too long, some questions were not clear, the language they used or wording was not easily understood by the learners.

We note here that poor reading skills have been noted widely as a problem in the South African landscape (Fleisch, 2008; NEEDU, 2013), and acknowledge that there are good reasons for expecting learners to be able to read questions independently at Grade 3 level. These arguments notwithstanding, teacher comments suggest that lack of reading ability has blocked learners' access to problems that they may be able to solve mathematically. Spaul (2013) confirms the close connection between speaking English at home and reading benefits.

From the perspective of validity of assessment for all learners from different linguistic backgrounds, this situation has to be viewed as problematic. Other areas of validity and fairness, raised by several teachers, related to timing and content coverage and the issue of fairness of the marking memorandum. These issues are discussed later.

Learner negative experiences

Connected with the above 'rules' of administration and 'strictness' of exam conditions 18 comments related to negative learner experiences of the ANAs. Most of these related to learner anxiety and stress, with several including comments about learner confusion, for example:

Anxiety was a big factor. Children were nervous. Learners' behavior was different as when writing internal tests/exams. I did not like the fact that we did not facilitate our own classes. Children were confused.

Especially foundation phase learners. Foundation phase learners need their own educators.. I neglected my own assessment for third term ... Learners had 'exam fear'! Poor learners!

We were placed in classes we did not teach. (For example I went to Grade 4). A problem with this is some learners become nervous with a new teacher in their class.

Increasingly over the past 20 years there have been articles published on the notion of care, particularly in relation to primary and elementary teachers. Nias's (1999) work emerges from her experience within the UK context following the introduction of a National Curriculum and national assessments, resulting in comparative school league tables. This context is similar to South Africa with parallels in the recent curriculum restructuring in which the ANAs were carried through. While Nias's (1999) discussion of care as affectivity and responsibility for learners is explored within the curriculum rather than the assessment context, her notion of care can be usefully applied to understanding teachers' care-related responses to the ANAs. According to Nias (1999, p. 68):

if we accept as generalisations that teachers are affectively involved with children, and the proposition of the relational feminists that affectivity is the basis of an ethic of care, then it is easy to construe primary teaching as a 'culture of care' whose underlying values emphasise the importance of making children feel secure, happy and cared for.

Time issues related to learner pace

Concern for the length of ANAs in terms of the time learners take to complete questions was noted in seven comments relating to the administration of the ANAs and in nine comments relating to the use of exemplar papers. Some related this to learners' inability to concentrate and others to learners being slow to read the questions. Some examples are:

The grade 3's I had to facilitate were so agitated. They could not sit still for so long. Other learners did not finish their papers.

The time was very short with long questions with few marks.

Concern for Implications on Mathematics Teaching and Assessment Practice

The theme *Concern for implications for teaching and assessment practice* includes the following sub themes to be presented below:

- content coverage and timing of the ANAs;
- exemplars are too close to ANAs (teaching to the test);
- inaccurate or inappropriate marking memorandums.

Content coverage and the timing of the ANAs

Since the 2012 ANAs were written before the end of the academic year (with almost a term still to go), several comments noted that content covered in the ANAs had not been fully taught yet:

They were not very useful because they cover the whole year's work in September. I can't rush to finish everything in September, because in that way I will be teaching the syllabus, not the learners.

Exemplars similar to ANAs

Several teachers (11 comments) noted how the exemplars were helpful for getting learners used to the structure, style and test format. For example:

They [exemplars] were useful because they set a good example of the exact way in which questions were to be asked, so it trained my learners to comprehend

While many positive comments (five) were noted about the exemplars being useful because many of the questions are the same or similar to the ANAs, two teachers made negative comments pointing to this similarity being problematic because it encourages teaching to the test:

[The exemplars are] just a duplicate of the ANA-papers.

I don't think it really helped the learners I was just like drilling them. It only helped them for specifically the exams. This term I am not rushing for ANAs.

This directly reflects the concerns noted in the international literature on focusing teaching on the practice of 'cloning' tasks (close replication of test items) for test preparation, resulting in a narrowing of the learning curriculum to what will be assessed.

Inaccurate or inappropriate marking memoranda

Of the 52 teachers who responded to the cluster of questions on marking of the ANAs, 38 said they found the memos easy to use, six gave mixed responses (e.g. 'some easy some difficult'), and eight said they were not easy to use. Across this latter group there were comments relating to three key challenges in the memos. These included eight comments related to unfair or inappropriate mark allocation, for example:

It was not accurate. Marks were unfair to learners, e.g. 1 mark for 4 items in a question. If a child is wrong 1 piece, the whole sum is wrong.

Our grade 6 maths memo okay, however I feel that marks allocated for some questions were too many e.g. 4 or 5 marks for 1 question which meant if a student failed three questions it meant he would have failed.

Six comments related to not allowing for multiple methods, for example:

Yes, memorandum for grade 2 & 1 was easy to follow. Question on adding 3 digit nos. & multiplication in grade 3 was difficult to mark as method was given on memorandum and learners did use different methods taught.

We see the singularity of 'accepted methods' as particularly problematic in the light of research on problem-solving processes encouraging sense-making through a variety of modelling strategies (Schoenfeld, 1985). Additionally though, these responses show evidence of teacher engagement with mark allocations and their substantive entailments and with the mathematics that underlies multiple solution methods – aspects that have been noted as weaknesses in the South African terrain (Taylor & Vinjevold, 1999).

Other Concerns

An area of concern, not falling clearly into either of the above two categories, is that of the burden on school resources and teacher time. The burden and expense of photocopying appeared in six comments across the clusters of questions (e.g. 'we also have to use a lot of paper for photostating which our school had to supply from limited resources'). In some cases where schools could not afford photocopying of exemplars, teachers explained that they wrote exemplar questions out on the board. This raises issues of equity across our system of education where it is noted that two vastly different systems exist – one well resourced and the other poorly resourced which services the majority of learners (Fleisch, 2008; Spaul, 2013). The issues raised in terms of access to photocopying indicate that the provision of exemplars will inevitably benefit those schools that are able to photocopy them for each learner more than those who cannot.

Similarly, the resource on teacher time will probably be exacerbated for teachers in schools where learner teacher ratios are high, thus increasing the time required for marking and the required analysis of results (noted by several teachers). Several teachers also noted pressure on their teaching time. The

teachers noted that they spent between 1 and 8 weeks on the ANAs with the average time noted across these teachers being 3.97 weeks (Graven & Venkatakrishnan, 2013). For example, one teacher wrote:

It took nearly +4 weeks. The first week was spent photostatting. One exemplar was 26 pages! One paper took nearly 2 weeks to go through. I neglected my other assessment tasks.

Reflections and Concluding Remarks

The data above highlight several clusters of questions relating to possible unintended consequences of the ANAs, including challenges to the fairness and validity of assessments. We argue that, given the emphasis on the ANAs by the DBE, the time allocated to them by teachers and the extent to which they are reported on in the press, further research is needed in terms of the effects of ANAs on learner experiences, on teacher practices and on issues of equity in the schooling system. In relation to each of these areas we raise several questions below.

The Effect of ANAs on Learner Experiences

We suggest that the effects of the ANAs in terms of learners' affective experiences need further research. If we accept that children who feel safe and secure at school are better able to concentrate on learning (Nias, 1999), then how do we balance the purpose of the ANAs with learner experiences of anxiety? For instance, how might a reduction of the assessment burden by conducting the ANAs only at the end of each phase affect learner anxiety and the systemic evaluation purposes of the ANAs? Noting that foundation-phase learners tend to only have one teacher and do not move to different teachers for subjects, would invigilation by, or the presence of, their own class teacher reduce anxiety levels?

Equally, we feel that the data raise issues around the method of communication of the ANA tasks and equity in accessing these tasks, and thus the validity of the assessment strategies. Is the language level of the ANAs sufficiently accessible for the level of learners? Initial research by Henning and Dampier (2012) suggests that this is not the case. In this case we need to research a range of language-related questions: how should the language used in the Mathematics ANAs change so as to maximize learner access to what is being asked? To what extent do learners' reading levels impact on their performance in the mathematics ANAs? In this respect, would it be useful to read questions to Grade 3 learners and perhaps beyond, as in the case of Grade 1 and 2 learners? Would this provide increased access to the questions by learners and would this enable increase in performance for some learners?

The Effect of ANAs on Teaching Practice

Are the ANAs a fair assessment of learner progress expected at the time of writing? How is the timing of the ANAs influencing teacher practices in terms of teaching pace or content coverage? Do the ANAs fairly assess multiple methods taught and used by teachers and learners? If not, how is this influencing teacher practices? Graven, Venkat, Westway and Tshesane (2013) suggest that the ANAs overlook the assessment of the essential learner competence of mental strategies. If we accept that assessment influences practice, then how might the absence of this in the ANAs affect teaching practices relating to mental strategies emphasized within the CAPS documents for both foundation and intermediate phase?

The extent of the influence of national assessments on teaching time and the nature of teaching should not be underestimated. Given the widespread acceptance of the usefulness and purpose of the ANAs emerging from the data, special attention must be given to the influence of these assessments on classroom practice. Teacher utterances largely indicate acceptance of ANA questions as valid exemplars of 'the' appropriate standard, format, scope and coverage expected of teachers in relation to their teaching. As noted above, international studies have shown that teachers adapt the

content of what they teach and focus on the test even while pedagogy will not necessarily change. As one teacher noted, 'ANA is good as it upgrades both learners and teachers. Teachers will know what to teach and how to teach them'. In this respect 'how to teach them' may refer to the method indicated in the memorandum rather than pedagogy for mathematics more generally and narrow the range of examples that are taught.

Should ANA exemplars and ANAs over the years be too similar across style, scope, content and format each year, there is a danger of seeing improvements in performance that are not necessarily matched by improved mathematical learning and competence. As one teacher noted in relation to the use of the ANA exemplars, 'they set a good example of the exact way in which questions were to be asked so it trained my learners', and another noted that with the exemplars she felt she was just 'drilling' them for the test. Such statements indicate a focus on test preparation perhaps at the expense of learning. Thus we argue that extremely careful consideration must be given to the choice of questions over time, ensuring range in format, style, scope and content if we are to avoid a situation of narrowing coverage of the curriculum.

The Effect of ANAs on Equity in the School System

The ways and extent to which ANAs contribute to remedying the high levels of inequality in the schooling system also need further research. It appears that the ANAs exacerbated pressures on school physical resources and teacher time for poorer schools operating with limited financial resources, leading to questions about how these inequalities play out within learner access to support – do ANAs simply serve to confirm and even exacerbate the inequalities?

It is widely acknowledged that learning and being assessed in a language other than one's mother tongue is challenging. While translating ANAs for Grades 4–7 may not be viable or helpful given the language of instruction in these grades, how might the ANA administration process accommodate the language challenges of the majority of our learners so as to enable increased access to what the questions are asking?

As Spaul (2013, p. 436) notes, 'modeling a single schooling system when there are in fact two school systems can lead to spurious results and misleading policy conclusions'. He notes that South Africa should report performance results 'by wealth quartile, and not only by province, principally because South African averages are misleading: they are not an accurate representation of any "average" student, but rather overestimate the achievement of the majority of students' (p. 444). He concludes by saying that, without acknowledging inequalities in primary education, 'current patterns of poverty and privilege will remain unabated' (p. 444). Since the upper 25% of results from well functioning and resourced schools skew national results, the lack of a breakdown of ANA results by wealth quartile, as Spaul suggests, makes it difficult to know whether any national increase in ANA results is simply a result of wealthier schools doing better or whether learners were improving across various economic contexts. A challenge is therefore to understand to what extent and how the ANAs might take into consideration the two schooling systems. Can one set of numeracy and mathematics ANAs serve both schooling systems in terms of meeting the purposes intended, which include providing information for remediation planning? In this respect, if the majority of learners in the majority of schools get below 25% for the Grade 4 ANAs, then surely the ANAs tell teachers very little in terms of where to begin their remediation with learners as such results point to extensive gaps but little notion of where to begin. For example, if all the learners in a class cannot do 239×46 , this tells the teacher very little about whether the learners understand foundational multiplicative concepts such as multiplication as repeated addition.

On a final note, we must remain mindful that a 'problem with high-stakes testing policies (as currently implemented) is that while they get teachers' attention, they provide few resources for addressing issues of inequality in schools' (Diamond, 2007, p. 306). In this respect it is critical that we do not conflate spending on educational testing with spending on resources for quality learning and teaching that addresses our equity crisis in South Africa.

Acknowledgement

Our work as South African Numeracy Chairs is supported by the FirstRand Foundation (with the RMB), Anglo American Chairman's fund, the Department of Science and Technology and the National Research Foundation.

References

- Boaler, J. (1997). *Experiencing school mathematics: Teaching styles, sex and setting*. London: Open University Press.
- Bohlmann, C., & Pretorius, E. (2008). Relationships between mathematics and literacy: Exploring some underlying factors. *Pythagoras*, 69(special issue), 42–55.
- Borko, H., & Elliott, R. (1999). Hands-on pedagogy versus hands-off accountability: Tensions between competing commitments for exemplary mathematics teachers in Kentucky. *Phi Delta Kappa*, 80(5), 394–400.
- Combleth, C. (1990). *Curriculum in context*. London: Falmer Press.
- Department of Basic Education. (2011). *Report on the Annual National Assessments 2011*. Pretoria. Retrieved from <http://www.info.gov.za/view/DownloadFileAction?id=147629>
- Department of Basic Education. (2012). *Report on the Annual National Assessments 2012: Grades 1 to 6 & 9*. Pretoria: Department of Basic Education.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between High-Stakes Testing policy and classroom instruction. *Sociology of Education*, 80(4), 285–313.
- Elmore, R. F., Ablemann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: from process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65–98). Washington, DC: Brookings Institution.
- Fleisch, B. (2008). *Primary education in crisis: Why South African schoolchildren underachieve in reading and mathematics*. Johannesburg: Juta.
- Fuhrman, S. H., & Elmore, R. F. (2004). Introduction. In S. Fuhrman & R. Elmore (Eds), *Redesigning accountability systems for education* (pp. 3–14). New York: Teachers College Press.
- Geist, E. (2010). The anti-anxiety curriculum: Combating math anxiety in the classroom. *Journal of Instructional Psychology*, 37(1), 24–31.
- Graven, M., & Buytenhuys E. (2011). Mathematical literacy in South Africa: Increasing access and quality in learners' mathematical participation both in and beyond the classroom. In B. Atweh, M. Graven, W. Secada & P. Valero (Eds), *Mapping equity and quality in mathematics education* (pp. 493–509). Heidelberg: Springer.
- Graven, M., & Venkatakrishnan, H. (2013). ANAs: Possibilities and constraints for mathematical learning. *Learning and Teaching Mathematics*, 14, 12–16.
- Graven, M., Venkat, H., Westaway, L., & Tshesane, H. (2013) Place value without number sense: Exploring the need for mental mathematical skills assessment within the Annual National Assessments. *South African Journal of Childhood Education*, 3(2), 131–143.
- Green, W., Parker, D., Deacon, R., & Hall, G. (2011). Foundation phase teacher provision by public higher education institutions in South Africa. *South African Journal of Childhood Education*, 1(1), 109–121.
- Henning, E., & Dampier, G. (2012). Simple language is the answer to 'difficult' assessment tests. *Mail and Guardian*. Retrieved May 2, 2013 from <http://mg.co.za/article/2012-09-08-simple-language-is-the-answer-to-difficult-assessment-tests>
- Howe, S. (1997). *Mathematics and science performance in the middle school years in South Africa: A Summary report on the performance of South African students in the Third International Mathematics and Science Study (TIMSS)*. Pretoria: HSRC.
- Kanjee, A. (2007). Improving learner achievement in schools: Application of National Assessments in South Africa. In S. Buhlungu, J. Daniel, & R. Southall (Eds), *State of the Nation* (pp. 470–499). Pretoria: HSRC.
- Kelleghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*. Paris: International Institute for Educational Planning.
- NEEDU (2013). *National Report 2012: The state of literacy teaching and learning in the Foundation Phase*. National Education Evaluation and Development Unit.
- Nias, J. (1999). Primary teaching as a culture of care. In J. Prosser (Ed.), *School culture* (pp. 66–81). London: Paul Chapman.
- Popham, W. J. (2001). Teaching to the test? *Educational Leadership*, 58(6), 16–20.
- Postlethwaite, T. N., & Kelleghan, T. (2008). *National assessments of educational achievement*. Paris: The International Institute for Educational Planning.

- Reay, D., & Wiliam, D. (1999). I'll be a nothing: Structure, agency and the construction of identity through assessment. *British Educational Research Journal*, 25, 343–354.
- Reddy, V. (2006). *Mathematics and science achievement at South African schools in TIMSS 2003*. Cape Town: HSRC Press.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic Press.
- Setati, M. (2008). Access to mathematics versus access to the language of power: the struggle in multilingual mathematics classrooms. *South African Journal of Education*, 28(1), 103–116.
- Spaull, N. (2013). Poverty and privilege: Primary school inequality in South Africa. *International Journal of Educational Development*, 33(5), 436–447.
- Taylor, N., & Vinjevold, P. (Eds) (1999) *Getting learning right: Report of the President's Education Initiative Research Project*. Johannesburg: The Joint Education Trust.
- Walls, F. (2008). Children talk about Mathematics Assessment. In J. F. Matos, P. Valero & K. Yasukawa (Eds), *Proceedings of the fifth international mathematics education and society conference* (pp. 475–484). Albufeira, Portugal: Universidade de Lisboa and Aalborg University.