Can Common Tasks of Assessment in the Foundation Phase make a difference? Evidence from a provincial project on the strengths weaknesses of common task assessments.

Du Toit, Roelien, Mrs., JET Education Services, rdutoit@jet.org.za
Periera, Carla, Mrs., JET Education Services, carla@jet.org.za

Recently there has been a move towards setting teacher administrated common assessment tasks as one aspect of the national movement towards raising learner performance in Literacy and Numeracy. One such an example is the Annual Assessment Tasks within the Foundations for Learning Campaign, which was launched by the Minister of Education, Naledi Pandor, in March 2008. Although common tasks of assessment in the form of provincial projects have been around for some time, their benefits to teaching and learning have largely been ignored by the wider educational community. One such multi-year provincial project is examined in terms of process, alignment to the NCS, cognitive levels addressed within the school level assessments, learner results, perceptions of teachers and district officials in regards to the benefits of the project. The biggest benefit, as perceived by teachers and district facilitators alike, is providing teachers with grade-level appropriate assessment items and immediate feedback regarding the level of teaching and learning. As a whole common assessment tasks seem to impact indirectly and positively on learners' performance through improving teaching and assessment practices. However, should one want more reliable learner performance data, certain process related issues need to be addressed. Furthermore, the amount of standardization seems to depend on the outcomes the project wants to achieve.

Introduction

In the National Gazette of 14 March 2008 the Minister of Education, Naledi Pandor, launched the Foundations for Learning Campaign, which aims to provide clear directives regarding the minimum expectations for schooling in the primary grades and ultimately aims to improve learner performance. One aspect of the Foundations for Learning Campaign is the annual assessments to be written by all learners in grades 1 to 6. These annual assessments can be classified as common task assessments, which aim to access the level of learning and teaching as well as providing the teacher with examples of good quality, grade-appropriate assessment activities. Can common task assessments work in the Foundation Phase? Can common tasks of assessment access the level of learner achievement and provide teachers with examples of good quality assessment items at the same time?

In this paper, evidence from a province wide project will be discussed in terms of factors that contributed to the level of success achieved in one specific provincial project. The success of the project will be measured in terms of the goals set in the project. The discussed includes National Curriculum Statement (NCS) alignment, cognitive levels of the items, perceptions of the district officials and teachers and project management issues.

Project history

Since 2002 (with the exception of 2006) a provincial department of education has conducted an annual Numeracy project in which all public primary schools participated.

At the first level of the Numeracy project (i.e. the School Level) teachers conducted the provided assessment tasks, or tests, with the classes on a specified date for each grade. The question papers, or tasks, were scored by teachers and the summary statistics were recorded and submitted to the respective district office. In 2007, data for a total of

240 207 learners (i.e. 77 032 Grade 1 learners, 76 767 Grade 2 learners and 86 408 Grade 3 learners), who participated in the School Level was received by JET.

The project aimed to achieve two broad objectives, namely:

a. To provide Foundation Phase teachers with a broad variety of level-appropriate Mathematical activities that demonstrates the complexity and range of the Mathematics Learning Outcomes in the National Curriculum Statement (NCS); and

b. To provide information regarding the achievement of learners, schools, districts and the province as a whole in the Mathematics Learning Area in Foundation Phase.

Can both these objectives be obtained to a satisfactory level simultaneously? We will refer back to this issue throughout the paper.

The process and project management

The project required service providers, provincial and district officials, schools and teachers working closely together, each having a vital role in the successful completion of the project.

In short, a service provider together with the provincial and district officials developed the assessment task items for each grade level as well as a teacher's guide for the administration of the task on a specified date. All the papers were available in nine official languages as well as Braille. All learners thus had the opportunity to write the task in the LoLT of their class. This was, according to all involved, one of the factors that contributed to the success of the project. A few translation inaccuracies were identified, which highlighted the need for not just translation but also a need to translate back into the original language. This could involve a second or even third translation process.

The district officials trained the Heads of Department of each school in the administration of the assessment tasks and completion of the data sheets. In most cases the assessment papers were delivered to schools approximately one to three days before the administration date to allow all the teachers to familiarize themselves with the content of the assessment tasks. This may have been a risky move in terms of giving teachers the opportunity to "teach the items" before the administration date, but was considered by teachers and district officials alike as beneficial in terms of teachers being prepared before administrating the task. The opportunity to clarify queries regarding items or the language used before administrating the task was therefore provided. In standardized tests this is an unheard of procedure, but the acceptability of the procedure largely depends on the purpose of the project as well as how the data is to be used. If you want to use the data for making comparisons the opportunities for "teaching the items" should be limited. If you want to provide teachers with good quality items, limiting the teacher's interactions with the items before the test administration becomes less important.

Teachers of a specific school administrated the tasks at their school, but not in their own class. This was an attempt to limit the amount of "cheating" or "teaching" during the administration of the task. Most teachers interviewed admitted that in regard to limiting the amount of teaching it was a good strategy, but also pointed out that it limited the amount of knowledge the teacher gained regarding the performance of her own learners. Under ideal circumstances the teacher would receive feedback regarding her learners' performance and behaviour during the administration from her colleague who administered the task in her class, but this seldom happened to a satisfactory degree. Is this strategy beneficial to the success of common task assessments? That depends on the outcome that you want to achieve: Providing teachers with good quality assessment items and feedback regarding the performance of her learners or reliable and valid learner performance data. When the aim is to provide quality assessment items and feedback to the teacher, having the teacher administer the test in her own class, at the risk of her "teaching to the items", would be acceptable. If the aim is to gather reliable learner performance data, then limiting the amount of "teaching to the items" would be the route to go.

Grade 1 teachers were advised to read the items one by one allowing learners to complete the item before proceeding to the next item. Grade 2 and 3 teacher could choose from three options:

- Read through all the questions and let learners complete the paper independently; **or**
- Read through half of the questions and let learners complete these questions before continuing; **or**
- Read the questions out loud to learners one by one, and let learners answer each question before continuing.

As indicated in Table 1, the third option was the most popular in both Grade 2 and Grade 3.

Table 1: Preferred procedure in Grades 2 and 3

Challenge methodology	Gr2	Gr3
Read all	4%	10%
Read half	3%	9%
Read one by one	86%	70%
Combination of above	0%	1%
No response	6%	10%

Most educators were happy with option 3, as evidenced by the following quotes:

"It made learners understand and finish one question at a time."

"Yes, children could keep up, especially those who can't read."

"All the learners can easily follow and it gives them a chance to work out answers easily."

One very important aspect is highlighted by the second quote: learners with reading barriers were also granted the opportunity to complete the Numeracy assessment task. The assumption is that by allowing teachers to read the items, the influence of reading ability on the learners' performance is limited.

From the above it is clear that the administration of the tests was not standardized, and therefore comparisons between schools and districts or between the previous and current provincial results were not possible. However, the strengths and weaknesses in teaching and learning could be qualitatively identified, even if they could not be quantified.

After learners completed the task the teachers scored and completed data sheets summarizing the learner performance of her specific class. When interviewed teachers could identify the areas in which their learners struggled or performed well, but they struggled to identify the reasons why learners struggled and often expressed their amazement since the concept was taught in class. Seen in conjunction with teachers admitting to finding certain items difficult themselves, or not knowing how to teach certain items, it seems very important for teachers to receive in-dept training on common task assessment items. Such training should focus on the concept addressed in the item, how it relates to the curriculum as well as strategies to teach the item. It seems not enough to just provide teachers with good quality items, when we strive to improve teaching and learning. Once again if the focus is on gathering learner performance data, providing basic training on the administration of the items should be enough. The following perceptions of teachers seem to support this point.

Table 2: Teacher perceptions

	Strongly agree	Agree	Disagree	Strongly disagree	No response
We would like more information regarding the project before the tasks are conducted.	44%	41%	11%	2%	2%
We would like more feedback regarding the results after it is conducted.	46%	46%	5%	1%	1%

As far as was possible the district officials monitored the administration of the assessment tasks. This was done at randomly selected schools and classes. However, due to manpower constraints only a small number of schools and classes could be monitored. This, once again, raises the question to the amount of "teaching" that took place during the administration of the tasks. Since these are teacher administered tasks this factor would be very difficult to quantify, but the lack of standardized procedure during classroom administration severely impacts on the statistical reliability and validity of the learner performance data. Does this matter? It depends on your intentions for the use of the data. If comparisons between the performance of learners, classes, schools, districts or provinces are to be made, standardized procedure is necessary. When the intention is to provide teachers with examples of grade specific good quality assessment items or activities or general feedback on how learners perform in specific items then standardized procedure is not so important.
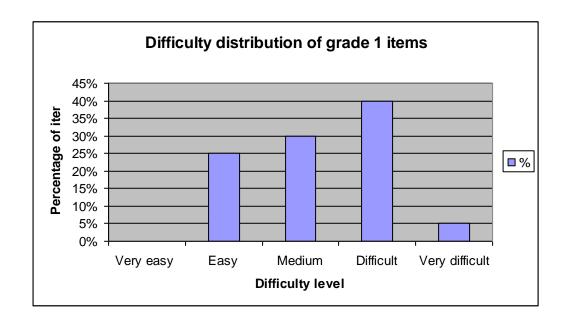
The items

A total number of twenty items were selected for each grade level test.

These items ranged from items addressing the minimum requirements of the curriculum for the specific grade to items pitched at the highest level of difficulty appropriate for the specific grade. The following is examples of grade 1 items:
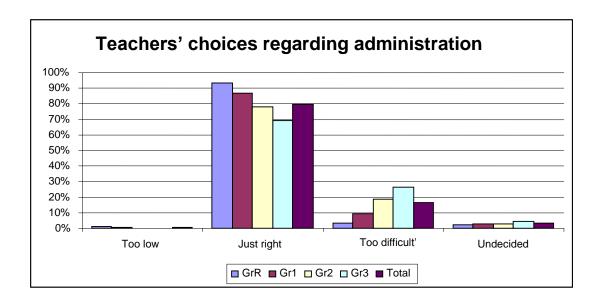
8 + 5 = ☐

October

S	M	T	W	T	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	**16**	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

Look at the calendar. Spino's birthday is a week later. Colour her birthday on the calendar.

The difficulty distribution of items was slightly skewed towards the difficult side for all the grade level tests. The grade 1 items' distribution is used to demonstrate the point in the graph below.
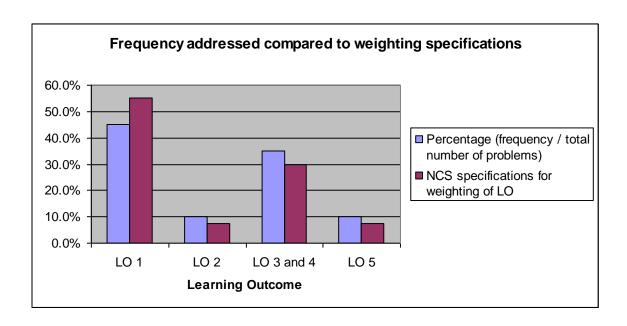
Difficulty distribution of grade 1 items

A bar chart titled "Difficulty distribution of grade 1 items". The vertical axis is labelled "Percentage of iter" ranging from 0% to 45% in 5% increments. The horizontal axis is labelled "Difficulty level" with categories: Very easy, Easy, Medium, Difficult, Very difficult. The legend shows "%". Bar values: Very easy ≈ 0%, Easy ≈ 25%, Medium ≈ 30%, Difficult ≈ 40%, Very difficult ≈ 5%.

However, on the feedback forms most teachers indicated that the items were set at the "just right" level, being not too easy or difficult.

Teachers' choices regarding administration

A bar chart titled "Teachers' choices regarding administration". The vertical axis ranges from 0% to 100% in 10% increments. The horizontal axis has categories: Too low, Just right, Too difficult', Undecided. The legend shows: GrR, Gr1, Gr2, Gr3, Total.
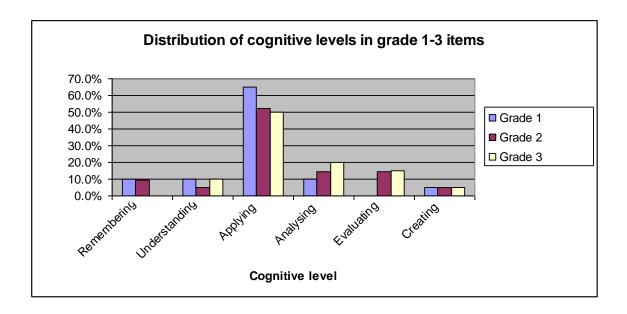
The item developers in this project avoided a potential stumble block in developing common assessment tasks, that of "testing to the current level of the learning" instead of "testing to the expected level of learning". Common assessment tasks can be utilized to set a standard for the level expected in learners' performance, and therefore can send a counterproductive message when pitched either too low or too high. The ideal is to have a

good balance of very easy, easy, moderate, difficult and very difficult items. What will then be considered a good balance? Once again, that will depend on what the project aim to achieve. If the project aims to provide teachers with examples of assessment items set at the minimum requirements, as stated in the Assessment Standards, for the grade level, more easy to moderate items will be required. If the project want to provide teachers with a grade appropriate items that goes beyond the minimum requirements more moderate to difficult items will be required. Another option is to systematically raise the difficulty level from year to year, starting with the minimum required difficulty level for the grade and progressing to the highest grade appropriate difficulty level.
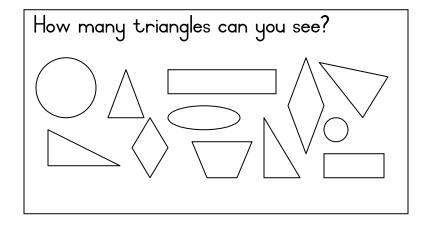
The NCS provides us with guidelines (although sometimes rather vague) regarding the level at which we should pitch the items. The item developers in this project stuck closely to the NCS to the extent of keeping the weighting of the Learning Outcomes addressed by the items to that specified in the curriculum. This ensured that the bulk of the items addressed Learning Outcome 1 (Numbers, Operations and Relationships), while simultaneously ensuring that the other Learning Outcomes are not ignored. Once again the grade 1 items distribution is used to illustrate the point.

The item developers furthermore ensured that the items lie on a variety of cognitive levels, with the focus shifting to the higher cognitive levels in grade 3. To illustrate this point the graph shows the distribution of the grade 1 to 3 items per cognitive level as defined in Bloom's Revised Taxonomy.

Distribution of cognitive levels in grade 1-3 items

The following are examples of a lower cognitive level item and a higher cognitive level item in the grade 1 assessment task.

How many triangles can you see?

There are three (3) girls. How many legs and eyes do they have together? Write a sum on the line to show how you worked it out.

District officials' and teachers' perceptions

Generally, teachers across the grades agreed that the project was of great benefit to the teaching and learning of Numeracy. Table 2 shows what teachers said about how useful the common assessment tasks were to them.

Table 2: Teachers' general perceptions of the Challenge (n = 3 152)

	Strongly agree	Agree	Disagree	Strongly disagree	No response
The tasks gave us ideas for trying new things in our classrooms.	36%	57%	3%	1%	2%
The tasks made us aware of new points to keep in mind when assessing written activities (and practical activities for Grade R).	29%	59%	4%	1%	7%
The tasks opened our eyes to aspects of Mathematics that we are neglecting.	30%	54%	13%	2%	1%
The tasks helped us identify the strengths and weaknesses of our learners.	38%	55%	4%	1%	1%
The tasks helped us identify our own strengths and weaknesses in teaching Numeracy.	33%	57%	7%	1%	2%

Conclusion

As a whole the project achieved its goal to provide teachers with a variety of grade appropriate assessment activities, and seems to indirectly and positively influence learner performance in Numeracy

As a common assessment task this project provides annual information regarding:

- The areas of weaknesses or strengths in Numeracy teaching, e.g. the need for intervention in the teaching of Learning Outcome 3 (shape and space).
- The expected standard of Numeracy teaching and learning per grade. The assessment tasks provide teachers with examples of items expected to be mastered at each level and thus, a measuring stick against which to assess the level of teaching and learning in their own classrooms.

However, to reap the full formative benefit of the project the following supportive actions should be implemented:

- All schools should be trained on the administration, scoring and reporting procedures. This training should include making teachers aware of the full value of the summary statistics in identifying the strengths and weaknesses in Numeracy teaching and learning in their classes.
- Following the administration of the tasks, teacher training should be conducted in all districts to "unpack" the assessment tasks. Such training could cover the specific Assessment Standards addressed by the items, possible ways to solve the problems, ways to teach the concept or skill addressed in the items, as well as other possible ways to assess the concept or skill (i.e. a variety of ways to ask the same question). It would be unrealistic to suggest unpacking the complete task for each grade, but at least the items found to be problematic by learners and/or teachers should be addressed.
- The data gathered regarding the districts and schools should inform interventions at school and district level during the following year, taking in

consideration how the data was gathered and using other supportive information from other sources.

However, due to the lack of standardization comparisons of learner, class, school, district and the province performance can not be made. And therefore the second aim of the project, to provide performance data, was only partially achieved.

All in all, how successful the use of teacher administrated common task assessments is viewed, largely depends on what the project wants to achieve through using these assessment tasks.