Prinsloo, CH (Dr)
Email: chprinsloo@hsrc.ac.za.
Human Sciences Research Council (HSRC), Education and Skills Development (ESD) Research Programme

Harvey, JC (Ms)
Human Sciences Research Council (HSRC), Education and Skills Development (ESD) Research Programme

DOI: http://dx.doi.org/10.18820/2519593X/pie.v34i4.1
ISSN 0258-2236
e-ISSN 2519-593X
Perspectives in Education
2016 34(4): 1-14
© UV/UFS

The viability of individual oral assessments for learners: Insights gained from two intervention evaluations

Abstract

It is essential for learners to develop foundational literacy skills, ideally, in the first grade of formal education. These skills are then firmly entrenched and can be expanded in the following grades to form a basis for all future academic studies. Appropriate assessment practices and tools to aid this process can inform the achievement of quality education. Assessment and the curriculum are intertwined concepts in relation to teaching and learning. Through assessment, it can be established if all learners have attained curriculum content, knowledge and proficiencies in a given year. Furthermore, assessment can assist in advising teachers on which specific areas learners are struggling with as well as provide insight for remedial measures. Together, this can offer ways to improve education. In this article, individual oral assessment using the Early Grade Reading Assessment (EGRA) tool is discussed based on two recent impact evaluations of teacher interventions. Each intervention conceptualised its own theory of change to improve learner language and literacy development. The interventions also differed in relation to the target language; English as First Additional Language and Setswana as Home Language. Despite these differences, using the EGRA tool in both intervention evaluations allowed for a discussion on its usefulness in South Africa. This was done with regard to suitability, reliability and validity, assistance to educators, amendments and suggestions to overcoming challenges related to practicalities. In conclusion, recommendations for improving education and the development of literacy in South African schools are made.

Keywords: *Early Grade Reading Assessment (EGRA), individual assessment, intervention evaluation, oral assessment, teaching and learning*

1. Introduction

Studies on early literacy development in developing countries have shifted the focus from access to education to improving the quality of education (Wagner, 2010; Davidson & Hobbs, 2013). Quality education, rather than enrolment rates or the number of years of education attendance, is necessary for effective learning in schools and learners' subsequent academic and employment trajectory. Receiving quality education is also linked to lessening learner attrition (Gove & Wetterberg, 2011; Davidson & Hobbs, 2013). South African learner achievement and dropout seems to

be more of a concern in the higher grades, i.e. grade 9 upwards, rather than in the earlier grades (Department of Basic Education, 2011, 2013). However, poor achievement and this increase in dropout is theoretically related to poor literacy acquisition during grades 1 to 3 because reading ability lays the foundation for all future learning (Prinsloo *et al.*, no date; Gove & Wetterberg, 2011; Green *et al.*, 2011). Therefore, quality literacy instruction in the first years of formal education is crucial for each individual learner, as it influences all aspects of their future and improves the education system as a whole (Heckman, 2000; Cunha *et al.*, 2006). This article examines the usefulness of the Early Grade Reading Assessment (EGRA) instrument to provide necessary information on the effectiveness of learning in the early phase of literacy development in South African schools.

The quality of education can be monitored through appropriate assessment methods (Wagner, 2010; Davidson & Hobbs, 2013). The EGRA tool supposedly offers a hybrid assessment approach. This approach is based on the idea that methodologies used in large-scale education assessments can be reshaped into large enough instruments that can be used with greater speed at a reduced cost and can be adapted to local and linguistically diverse contexts. Therefore, the EGRA has the fewest sampling exclusions as it allows for development in local languages and tends to sample amongst the most disadvantaged young learners (Wagner, 2010). In addition, it was particularly designed for low-income country grade 1 to 3 learners who may not be able to read or write (Gove & Wetterberg, 2011). The foregoing explains why this tool was used in the impact evaluations designed by the Human Sciences Research Council (HSRC). Firstly, it was necessary to establish a baseline assessment to identify the effect of the interventions. This took place in the beginning of the year when most learners are still illiterate. In addition, issues such as assessment sophistication, test anxiety, concentration and attention span require a one-on-one oral testing approach (Hobbs & Davidson, 2015; Piper & Zuilkowski, 2015). Therefore, this assessment tool aligns well with the purpose of the impact evaluation studies referred to in this article: To assess the literacy skills developed by South African grade 1 learners. A brief background of literacy education in South Africa is given as well as of the different types of assessment available before focusing on the interventions.

2. Background

Policies regarding literacy instruction in the South African curriculum are informed by the theory of additive bilingualism. Within this theory, it is argued that learners should be taught in their mother tongue/home language (HL) for as long as possible while an additional language is taught as a subject to complement rather than replace the HL. The principle is that a strong literacy base in their HL better enables learners to acquire literacy in an additional language (Reeves *et al.*, 2008; Hoadley *et al.*, 2010). It has been indicated that HL instruction is beneficial for learners and provides an advantage over learners who are solely educated in an alternate language (Gupta, 1997; Pretorius & Mampuru, 2007).

Scholars recommend that grades 1 to 6 should be taught in the HL, with an additional language as a subject, for successful learning of and later through, a second language (Thomas & Collier, 1997, 2001; Reeves *et al.*, 2008). However, South African primary schools more often than not switch from instruction in the African HL to instruction in English from grade 4 onwards (Pretorius & Mampuru, 2007; Reeves *et al.*, 2008; Pretorius & Currin, 2010). This early switch means that additive bilingualism instead gives way to transitional bilingualism. Transitional bilingualism refers to a short-lived form of bilingual education during which the

African HL is removed as a medium of instruction (Reeves *et al.*, 2008). It is thus important that South African learners develop an adequate level of reading ability in their HL during grade 1, which is then strengthened during grades 2 and 3. As stated above, assessment can be used to inform ways to improve the quality of instruction.

2.1 Different kinds of assessment techniques

Assessments have various purposes. With regard to classroom teaching and learning, assessments can be performed "of", "for", or "as" learning (Black & Wiliam, 1998; Lorna & Katz, 2006; Wiliam, 2011). Assessment of learning confirms the knowledge and skills learners have gained, determines if the curriculum outcomes have been achieved and may be used to relate learners' achievement to mean score levels. This is regarded as summative assessment and can be used in large-scale studies for comparison. Assessment for learning aims to inform teachers about the learning process so that their teaching and learning activities can be adjusted more immediately. This assessment is continuous in nature and involves the whole population of learners so that teachers can understand how, when and if learners apply their gained knowledge. Assessment as learning is used to develop and support the metacognition of learners. It does so by recognising the learner as the connector between assessment and learning: Learners themselves monitor their learning and use this feedback to alter their own understanding. Here, teachers assist students to develop monitoring skills (Lorna & Katz, 2006; Wiliam, 2011). Assessment for and as learning form part of formative assessment which refers to assessments taking place during teaching to inform the teacher regarding the teaching process to allow them to make appropriate changes (Lorna & Katz, 2006; Gove & Wetterberg, 2011; Wiliam, 2011).

In essence, all three of the different types of assessment (of, for, and as) occurring in the classroom setting are valuable and necessary as they serve different purposes. The purpose of the assessment can therefore inform the most appropriate choice of assessment method (Brown, 2004). When conducting an impact evaluation of a literacy intervention, the purpose of the assessment is to quantify learner progress, i.e. an assessment of learning, which requires a summative form of assessment. For this purpose, there are various assessment methods that can be used when using an early-grade learner sample including written, criterion-referenced, standardised (normed) and individual oral. In addition to aligning the purpose of the assessment and its method, the reliability, validity and fairness of the assessment method must be ensured.

A reliable assessment is one which is consistent and predictable (Hubley & Zumbo, 1996; Wagner, 2010). Various checks can assess this, including assessment using the same test (test-retest reliability), ensuring the assessor marks reliably to a defined standard (intra-rater reliability) and/or assessment by different assessors (inter-rater reliability). In addition, separate items on the test can be checked for their association to each other (inter-item reliability) (Hubley & Zumbo, 1996; Brown, 2004; Wagner, 2010). The assessment must also be valid. This refers to the degree that the test items actually measure what the assessment instrument aimed to assess (Hubley & Zumbo, 1996; Wagner, 2010). In this regard, the assessment method can be checked for validity by ensuring its items or components indicate all possible items (content validity), it estimates and predicts the measured criterion (concurrent and predictive criterion referenced validity, respectively) and/or reflects the underlying construct to link to a model or theory (construct validity) (Hubley & Zumbo, 1996). All assessment measures, particularly in educational assessment, should be reliable and valid.

2.1.1 Reliability and validity of the EGRA

The EGRA subtests have demonstrated sufficient reliability and validity in nearly 30 years of research (Davidson & Hobbs, 2013). With regard to South Africa, RTI International together with the South African Department of Education and the Molteno Institute of Language and Literacy (MILL) conducted a study in 2009 to assess the efficacy of the systematic method for reading success (SMRS). The study used the EGRA tool to evaluate the reliability of the subtests letter sound, word recognition, passage reading and passage comprehension questions. Their statistical analyses indicated that the tool is highly reliable ($\alpha = .95$), with the Cronbach alpha scores for each subtest being above 0.90. In addition, high levels of correlation between scores on each subtest were indicated by simple bivariate Pearson correlations ($p < .001$) (Piper, 2009; Gove & Wetterberg, 2011).

In the two impact evaluations undertaken by the HSRC, the Early Grade Reading Assessment (EGRA) tool was used as an individual, oral, summative assessment based on the documentation of its psychometric properties in the foregoing paragraph. The widespread successful translation/adjustment into various additional languages appropriate to contexts similar to that of South Africa provided the necessary support for this decision. It was necessary to rely on this as item-level data could not be captured due to cost constraints. Marking could also not be repeated by the same or different test administrators (TAs) to document retest and inter-rater reliability. The availability of norms appropriate to the learners assessed was also not required as learner proficiency was only compared with their own subsequent achievement over time.

2.2 Interventions

The first intervention was developed by siyaJabula siyaKhula (sJsK), a non-profit organisation which aims to assist learners through working with school communities (Harvey *et al.*, 2015; Prinsloo *et al.*, 2015). This intervention was administered between 2013 and 2015. The second intervention was the 3ie-funded Early-Grade Reading Study (EGRS) performed by the Department of Basic Education (DBE). This is a two-year intervention of which the first-year implementation has been completed (2015) while the second-year is currently being concluded (2016) (Taylor *et al.*, 2016). Details of each intervention as well as the methods used during the impact evaluation are presented below.

3. Impact evaluation methodology

The HSRC conducted impact evaluations of the two literacy interventions mentioned above, which are aimed at improving the reading achievement of early grade learners. In both impact evaluations, contextual survey questionnaires were completed and learner achievement baseline scores for the treatment and control schools were established. Fieldworkers who performed the data collection underwent extensive training that not only involved administration and context information but also involved practise sessions, simulations and a final selection process. This ensured standardisation of administration. The contextual instruments consisted of self-report background questionnaires that the school principals, schoolteachers and parents/caregivers completed. Both impact evaluations used the EGRA tool to assess grade 1 and 2 learner reading achievement, among other tests and grades dependent on the interventions' theory of change.

3.1 Intervention 1: sJsK (Harvey *et al.*, 2015; Prinsloo *et al.*, 2015)

3.1.1 Context

The intervention was administered in schools within Limpopo, South Africa, between 2013 and 2015. The intervention targeted English literacy as the First Additional Language (FAL). The theory of change was the Learner Regeneration Methodology's Corkscrew Model of Literacy Development to develop phonics and sentence structures, comprehension and literacy.

The intervention trained teachers (9 × 1-2 hour sessions) as well as community members (3 days × 5 weeks and 4 weeks in-classroom training) as reading facilitators in regenerating learners' gaps in and automating English literacy foundations. Learners were grouped according to ability with one facilitator assigned to ten learners. The facilitators (6 × 1 hour sessions per semester) used designed lesson plans, individualised for each grade, based on the developed, aligned and graded materials (dual-language readers, workbooks, wall charts and handouts).

3.1.2 Participants

Two cohorts of grade 1, 4 and 7 learners, whose home languages were mostly Xitsonga or Tshivenda, took part in the intervention with the first cohort beginning the intervention in 2013 and the second cohort in 2014. The EGRA subtests as part of the impact evaluation were administered orally to all grade 1 learners from each school (N = 1 085).

3.1.3 Instruments

EGRA letter sound (110 letters in one minute), word recognition (50 words in one minute) and non-word decoding (50 non-words in one minute) were selected to assess phonics as well as decoding fluency at letter and word level. These are accepted as foundational skills necessary for the later development of literacy, comprehension and academic language proficiency (Gove & Wetterberg, 2011; Davidson & Hobbs, 2013; Pretorius, 2013). Additional literacy instruments to assess vocabulary, reading age, sentence and paragraph reading, comprehension skills and writing skills were administered as the learners progressed from grade 1 to 7.

3.1.4 Procedures

The sJsK intervention matched control schools to the treatment schools and the sJsK facilitators and their coordinators scored the assessments. As the assessments were not performed by the HSRC, the HSRC monitored between 30 and 40% of assessments to guard against possible manipulation and therefore for quality assurance purposes. Equivalence analysis demonstrated the absence of assessment bias.

3.1.5 Brief findings of the impact evaluation

At all grade levels, learners from treatment schools showed stronger gains on test instrument assessment. However, these gains were less evident for grade 7 learners. This is unsurprising, as they will have the larger conceptual gaps requiring regeneration. In relation to the EGRA subtests, a lag was evident at grade 1 level for word recognition and non-word decoding. A full discussion is presented in Prinsloo *et al*. (2015) and Harvey *et al*. (2015).

3.2 Intervention 2: 3ie-funded EGRS (Taylor *et al.*, 2016)

3.2.1 Context

This intervention was implemented in the North-West, South Africa during 2015 and 2016. The intervention targeted literacy in Setswana as the First Language. The theory of change posited that eventual academic language proficiency is premised on adequate vocabulary, phonics-based decoding fluency, phonemic awareness and comprehension.

The study provided for three different intervention types, as reflected in table 1, aimed at improving different aspects of typical classroom teaching (see Taylor *et al*., 2016 for a full description). Teacher training focused on how to implement the programme, teach reading acquisition and effectively use available materials (government provided workbooks and curriculum pacing). Intervention 1 therefore involved teacher training, scripted lessons plans and graded readers whereas intervention 2 had the same format but with the addition of specialised reading coaches on a monthly basis. The coaches modelled, observed, critiqued/advised and gave feedback regarding the interventions' "ideal" lessons. In intervention 3, the community literacy facilitators were parents or other caregivers of the learners who received information on the school's current achievement level along with a tool to identify and monitor their child's reading skill stage and development according to specified milestones. These facilitators were also able to partake in facilitated group discussions with teachers to align targets, actions and responsibilities in assisting learners. Table 2 overviews the sub-tests employed at each evaluation stage.

Table 1: Overview of EGRS intervention types (Number of participating schools in brackets)

Intervention contents	1 (50)	2 (50)	3 (50)	Control (80)
Teacher training (two days per semester)	✓	✓		
Teacher training (monthly coaching)		✓		
Scripted lesson plans	✓	✓		
Graded readers	✓	✓		
Community literacy facilitators			✓	

3.2.2 Participants

One cohort of grade 1 learners from the above schools, along with their teachers, began the intervention in 2015 and continued it as grade 2 learners in 2016. Twenty learners from each school were randomly selected to take part in the impact evaluation (N = 4 600).

3.2.3 Instruments

Table 2: Subtests used as part of impact evaluation as progressed over the two years

	Baseline (Wave 1) Feb 2015	Midline (Wave 2) Oct/Nov 2015	End-line (Wave 3) Oct/Nov 2016
Oral pictorial/vocabulary	✓		
Digit span or auditory sequencing	✓		
Phonemic awareness	✓	✓	✓
EGRA Letter Sound	✓	✓	✓
EGRA Word Recognition	✓	✓	✓
EGRA Non-word Decoding		✓	✓
Sentence reading (Setswana)	✓	✓	
Paragraph reading (Setswana)		✓	✓
Writing (Setswana)		✓	✓
Reading (English)			✓

3.2.4 Procedures

The 3ie-funded EGRS study used a randomised control trial (RCT) design and an independent service provider performed and scored the assessments. This limited potential bias or manipulation. However, the HSRC still monitored 10% of data collection for quality assurance purposes. In order to conserve the benefits of an on-going single-blind study, the HSRC and the service provider remain unaware of which schools are receiving one of the intervention types and which schools are control schools.

3.2.5 Brief findings of the impact evaluation

At the point of this article, only the midline results were available and are considered preliminary. These indicated a significant, albeit small, impact for both intervention 1 and intervention 2. For further insights and a discussion of the survey instrument findings, see Taylor *et al*. (2016).

4. Findings and discussion related to the use of the EGRA tool

There are several challenges concomitant with assessing literacy in early grade learners. The design of the EGRA tool itself as well as careful administration allow for mitigation of these challenges. Firstly, test anxiety has long been noted as a negative influence in language learning (Hewitt & Stephenson, 2012). This can be further exacerbated in the case of learners from disadvantaged backgrounds, as in the case of the learners who participated in the two interventions. In this context, learners' exposure to language, reading materials and other educational resources within the home is limited which negatively influences their language and self-regulation development (Howard & Melhuish, 2016). Given the lack of stimulation within the home, there is added pressure to develop these skills at school, which augments learner anxiety in test situations. Although only applicable to the sJsK intervention, it is also worth noting that learners may be assessed in a language other than their HL depending on their school and other factors. Not only can this give an inaccurate measure of their literacy ability but it can also accentuate their test anxiety (Phillips, 1992; Hewitt & Stephenson, 2012). The EGRA tool is

designed to assess learners who may be illiterate which ensures that the pitch of the subtests is at the right level (Gove & Wetterberg, 2011). The one-on-one aspect of individual testing removes any competitive pressure learners may experience. It also provides a space for the test administrator (TA) to put the learner at ease, discussed at a later stage.

To relate test anxiety more specifically to the processes of early-grade (grade 1) impact evaluations, it is noted that the baseline assessment usually takes place during the first weeks of schooling. This is in order to provide a comparative measure for the results of the post-intervention assessment. It is expected that learners are for the most part pre-literate. This understandably increases their anxiety when asked to individually perform a not-yet-learned face-to-face task with a TA who is not their teacher. In addition, formal education and assessment are still new experiences for these learners. Learner anxiety may have a detrimental influence on their test performance, which may have other implications including incorrect assumptions that the test instruments are too difficult (a false floor effect which is discussed at a later stage) and, if the intervention has an effect, making the intervention appear more effective than in reality. However, careful consideration and training of the TA can assist in this challenge. In addition, quality assurers from the HSRC conducted random visits during the data collection period and made note of any test anxiety shown by the learners and what the TA did to assist.

As indicated, the recruitment, selection and training of the TAs are important factors. In order to offset test anxiety and put the learner at ease, the TA must be skilful in working with children and establishing rapport. Here a fine balance must be maintained; the TA must also refrain from inadvertently coaching a particular learner(s) as this violates the standard manual. With regard to language, the TA must be well versed in the HL of the learner, the language of teaching and learning of the school and possibly English for training purposes. This is particularly important during an oral assessment, as the TA must be able to communicate effectively so that the learner is at ease and fully understands what is expected. Furthermore, the TAs must be willing to be trained in the method used during the impact evaluation and administer the assessments without deviation or personal inflection. In addition, the EGRA subtests are timed and these measures must be strictly retained. These are important considerations for quality assurance, as the tests must be administered in the same manner to each learner without change or manipulation. Therefore, the HSRC hosted training workshops for TAs, which included role-plays as well as simulations at nearby schools. The latter were included to ensure that the TAs were fully proficient but also to develop their familiarity with the tests. It is not beneficial to the learners if the TA is unsure of themselves or if they robotically intone each instruction.

Concerning reliability and validity, the EGRA tool requires intra-rater reliability if it is a single TA assessing all learners at that point in time. In addition, the oral nature can mean that there is no paper-and-pen record of learner performance; there are only the recorded assessment scores (Gove & Wetterberg, 2011). In the impact evaluations performed by the HSRC, the fieldworkers were extensively trained and visited the schools in pairs. Learner assessments were also well recorded, by means of crossing out incorrect responses and scoring on individual learner assessment packs each assigned a unique learner code. In order to ensure validity, pilot assessments were performed before each data collection period. Therefore, the quality of the impact evaluations was assured by the HSRC by pilot assessments, intensive training, fieldwork monitoring and statistical analysis to assess for bias or manipulation. This served to support that in the South African context, the EGRA

subtests are authentic, in that they measure what they intend to measure rather than what would be easy to assess (Brown, 2004). As stated previously, it was not possible to calculate the reliability coefficients from the intervention studies. The use of this assessment method enabled the TA to separate the inability to read from a lack of content knowledge. Therefore, rather than comprehension skills, it was the accuracy of reading that was assessed which indicates their knowledge of phonics and decoding skills and at most, some vocabulary.

Accuracy is an additional concern when assessing such young learners, as they are quick to develop literacy skills. The exponential increase in skills can lead to a large percentage increase between baseline and end line measurements. This then makes it appear that the intervention was more effective than in reality. This is also linked to floor and ceiling effects. However, much of the foregoing is obviated by implementing a (randomised) control trial, in which the relative proficiency growth of learners exposed to interventions is compared against the lower growth for learners not exposed to the interventions. The test instruments were designed in such a way that the range of responses reported by the instrument is rather wide to avoid very low (floor) or very high (ceiling) achievement scores. If the scores are all very low and demonstrate none of the skills being assessed, then little further information from the test can be derived. The same is true of the ceiling effect. Such careful development of the testing instruments ensures that the pitch of the instrument is at the correct level. The use of the EGRA tool offsets the mentioned challenge as it is specifically aimed at low-income countries whose learners perform below the “floor” of previous international assessments. This assessment was designed specifically for grade 1 to 3 learners to assess the foundational skills necessary for the development of reading and therefore does not assume any reading or writing skills on the part of the learner (Gove & Wetterberg, 2011). When planning the current impact evaluations, it was envisaged that as cohorts entered higher grades, additional assessment subtests would be included. This is to measure developed proficiencies not assessed well any longer by the standardised EGRA tool, as used in the current evaluations, as the EGRA sub-tests cannot be changed.

Another challenge of early grade reading assessment in South Africa is related to the standardisation of the curriculum language of learning and teaching (LoLT). The LoLT for the early grades can be any one of the 11 official South African languages and is usually the HL of the majority of learners in the area. However, despite their official status, the standardisation of African languages is subject to a complex debate between linguists, educators and policy makers. Although there is insufficient space here to give this debate due consideration, it is worth noting that the resultant controversy has further diminished the use and perceived value of African languages in education (see Webb, 1999, 2004, 2010; Barkhuizen, 2002; Wa Kabwe-Segatti, Lafon & Webb, 2008; Webb, Lafon & Pare, 2010). Despite this complex situation, the EGRA tool is able to overcome this challenge by virtue of its design and the fact that it can be adapted to accommodate local varieties/dialects of the home languages, besides already having been translated into the official African languages in South Africa. In addition, and for the moment, the word recognition and non-word decoding subtests, in the authors’ experience, albeit quite basic, remain quite difficult for most rural learners in the grade 1-3 range, where the evaluations being discussed were pitched.

With regard to practicalities, individual oral assessment is time consuming and is dependent on the physical facilities and human resources available, as either the assessment must take place in a separate quiet room or the rest of the class must be disciplined to sit and work quietly without interrupting the on-going assessment. Additionally, conducting many

consecutive individual oral assessments is a tiring experience. To illustrate the time needed, the entire learner assessment battery including the EGRA required approximately four and a half hours for assessing 20 learners in the 3ie-funded impact evaluation. Of that time, the EGRA subtests needed approximately 4-6 minutes per learner. What was noted during the implementation of data collection was that it is important to keep up the motivation of the TA so that each child is treated equally without deviating from the administration manual or reducing enthusiasm. This does not only ensure quality assessment but also equal treatment of all learners independently of when they are assessed. Together with the aforementioned, a high sophistication on the part of the TA is thus required, which also requires high payment and high training.

5. Recommendations

Although the standard EGRA can be and was used in the South African context, it is suggested that a sub-committee for each African language be formed. Each committee can then look in-depth into the current relevance and correctness with particular attention to alignment with high-frequency patterns and occurrences in real language utterances, schoolbooks, etc. Such a process will enable equitable assessment of all African languages regardless of complexity. This will also assist in preparations to measure very proficient learners by, for example, using longer stimulus words or, for that matter, more words. However, it is important to note that one could also record the seconds (below the allowed 60 seconds for each of the EGRA subtests) that it takes such proficient learners to complete all letters or words. It is also recommended that the standard EGRA subtests be used rather than additional or amended scales as learner achievement levels in rural underperforming schools is low and will remain so for a while still.

The observations made regarding the use of the EGRA tool can be further extended to include classroom practice. Concerning assessment, reading curriculum-based measures (reading CBMs) are a type of continuous assessment with several benefits: 1) they are curriculum neutral, 2) they can be administered and scored quickly to provide feedback on learner performance to teachers who then implement remedial measures and 3) content is designed to assess learning goals rather than instructional levels. One such CBM is the EGRA toolkit (Gove & Wetterberg, 2011). In using the EGRA tool to inform teachers, there are several scores that can be used, namely decoding fluency, power of achievement and speed of effort. Decoding fluency is given by calculating the correct ratio of the letter sound, word recognition and non-word decoding subtests as follows:

- EGRA LS Fluency = EGRA LS Correct Letters/EGRA LS Reached Letters * 100
- EGRA WR Fluency = EGRA WR Correct Words/EGRA WR Reached Words * 100
- EGRA NWD Fluency = EGRA NWD Correct Words/EGRA NWD Reached Words * 100

The correct letters or words also indicate the power of learners' achievement whilst the reached word count gives the speed of learners' reading effort. A next level would be to look diagnostically at each sub-item (letter, word, letter combination) and identify which ones each learner flies through or stumbles over consistently. The teacher can then concentrate in subsequent lessons on the latter.

However, this is not without challenges as South African classes can be quite large which exacerbates the need for discipline, particularly when an individual learner is being assessed, as well as the time needed to assess all learners. The improvement of teacher training is thus

recommended as well as the appointment of reading assistants. This training must not only focus on how to teach areas such as phonics or spelling but also the different structures of each language, their semantic organisations and discourse structure. Training programmes must thus foster content knowledge, language knowledge, pedagogy as well as flexibility and responsiveness within teachers (Moats, 2009). This is further motivation for sub-committees for each African language who can assist in creating such a programme.

The above recommendations for sound teacher training must be accompanied by curriculum and policy changes. Within the Curriculum and Assessment Policy Statements (CAPS) for grades R to 3, the grade 1 term 1 curriculum allows 4.5 hours per week for reading which includes phonics, shared reading and group reading. In the following terms, this curriculum progresses to include paired reading where learners read books to one another and independent reading. However, the curriculum does not allot time to one-on-one reading teaching. Despite this, by the end of the first grade learners are assessed in terms of phonics, decoding and reading ability. It is recommended that the curriculum undergo revision to accommodate one-on-one reading instruction as well as individual oral assessments. Furthermore, adequate language development materials need to be provided by government departments that are designed and graded at the correct pitch for South African learners.

6. Conclusion

Individual oral assessment as used in the EGRA tool is suitable for application in low socio-economic communities, is linguistically flexible for the diverse language context of South Africa and is specifically developed for the assessment of early grade learners' reading ability. These aspects of the EGRA tool meet several challenges of assessing early grade learners. However, the test also has several practical requirements in order to be an effective instrument. The most important of which is training in the administration method ensured by, as used during the impact evaluations, TA simulation training, modelling, role-play, followed by moderation and quality assurance. The EGRA tool can also be used within the classroom as a CBM as a continuous assessment. In addition, recommendations are made for teacher training, curriculum changes and government responsibility in ensuring the adequate provision of reading development materials. By doing so, the quality of literacy education in South Africa will be improved.

References

Barkhuizen, G.P. 2002. Language-in-education policy: Student's perceptions of the status and role of Xhosa and English. *System*, 30(4), 499-515. https://doi.org/10.1016/S0346-251X(02)00051-9

Black, P. & Wiliam, D. 1998. Assessment and classroom learning. *Assessment*, 5(1), 7-74. https://doi.org/10.1080/0969595980050102

Brown, S. 2004. Assessment for learning. *Learning and Teaching in Higher Education*, 1, 81-89.

Cunha, F., Heckman, J.J., Lochner, L. & Masterov, D.V. 2006. Interpreting the evidence on life cycle skill formation. In E.A. Hanushek & F. Welch (Eds.). *Handbook of the economics of education*. pp. 697-812. Amsterdam: Elsevier.

Davidson, M. & Hobbs, J. 2013. Delivering reading intervention to the poorest children: The case of Liberia and EGRA-Plus, a primary grade reading assessment and intervention. *International Journal of Educational Development,* 33(3), 283-293. https://doi.org/10.1016/j.ijedudev.2012.09.005

Department of Basic Education (DBE). 2011. *Report on dropout and learner retention strategy to portfolio committee on education*. Available at http://www.education.gov.za/Portals/0/Documents/Reports/REPORT ON DROPOUT AND ETENTION TO PORTFOLIO COMMITTEE JUNE 2011.pdf?ver=2015-03-20-120521-617 [Accessed 1 June 2016].

Department of Basic Education (DBE). 2013. *Macro indicator report*. Available at http://www.wcedcurriculum.westerncape.gov.za/ [Accessed 1 June 2016].

Gove, A. & Wetterberg, A. (Eds.). 2011. *The early grade reading assessment: Applications and interventions to improve basic literacy*. Chapel Hill: Research Triangle Institute (RTI) Press.

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), 109-122. https://doi.org/10.4102/sajce.v1i1.80

Gupta, A.F. 1997. When mother-tongue education is not preferred. *Journal of Multilingual and Multicultural Development*, 18(6), 496-506. https://doi.org/10.1080/01434639708666337

Harvey, J., Prinsloo, C.H., Thaba, W. & Moodley, M. 2015. *Hope for rebuilding language foundations: Part 2 - Replication in a second cohort. Human Sciences Research Council Evaluation Report: siyaJabula siyaKhula's Learner Regeneraion Project, Vhumbedzi and Malamulele North-East, Vhembe, Limpopo. Evidence after one year*. Pretoria, South Africa: HSRC.

Heckman, J.J. 2000. Policies to foster human capital. *Research in Economics*, 54, 3-56. https://doi.org/10.1006/reec.1999.0225

Hewitt, E. & Stephenson, J. 2012. Foreign language anxiety and oral exam performance: A replication of Phillips's MLJ study. *The Modern Language Journal*, 96(2), 170-189. https://doi.org/10.1111/j.1540-4781.2011.01174.x

Hoadley, U., Murray, S., Drew, S. & Setati, M. 2010. *Comparing the learning bases: An evaluation of foundation phase curricula in South Africa, Canada (British Columbia), Singapore, and Kenya.* Pretoria, South Africa: Umalusi: Council for Quality Assurance in General and Further Education and Training.

Hobbs, J. & Davidson, M. 2015. Expanding EGRA: The early grade literacy assessment and its contribution to language instruction in Liberia, in *UKFIET Conference proceedings*.

Howard, S.J. & Melhuish, E. 2016. An early years toolbox for assessing early executive function, language, self-regulation, and social development: Validity, reliability, and preliminary norms. *Journal of Psychoeducational Assessment*, 1-21. https://doi.org/10.1177/0734282916633009

Hubley, A.M. & Zumbo, B.D. 1996. A dialectic on validity: Where we have been and where we are going. *The Journal of General Psychology*, 123(3), 207-215. https://doi.org/10.1080/00221309.1996.9921273

Lorna, E. & Katz, S. 2006. *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning.* Available at http://www.wncp.ca/media/40539/rethink.pdf. [Accessed 15 July 2016].

Moats, L. 2009. Knowledge foundations for teaching reading and spelling. *Reading and Writing*, 22(4), 379-399. https://doi.org/10.1007/s11145-009-9162-1

Phillips, E. M. 1992. The effects of language anxiety on students' oral test performance and attitudes. *The Modern Language Journal*, 76(i), 14-26. https://doi.org/10.1111/j.1540-4781.1992.tb02573.x

Piper, B. 2009. *Integrated education program: Impact study of SMRS using early grade reading assessment in three provinces in South Africa*. Chapel Hill: Research Triangle Institute (RTI) Press.

Piper, B. & Zuilkowski, S.S. 2015. Assessing reading fluency in Kenya: Oral or silent assessment? *International Review of Education,* 61(2), 153-171. https://doi.org/10.1007/s11159-015-9470-4

Pretorius, E.J. 2013. Butterfly effects in reading? The relationship between decoding and comprehension in Grade 6 high poverty schools. *Journal for Language Teaching*, 46(2), 74-95. https://doi.org/10.4314/jlt.v46i2.5

Pretorius, E.J. & Currin, S. 2010. Do the rich get richer and the poor poorer? The effects of an intervention programme on reading in the home and school language in a high poverty multilingual context. *International Journal of Educational Development*, 30, 67-76. https://doi.org/10.1016/j.ijedudev.2009.06.001

Pretorius, E.J. & Mampuru, D.M. 2007. Playing football without a ball: Language, reading and academic performance in a high-poverty school. *Journal of Research in Reading*, 30(1), 38-58. https://doi.org/10.1111/j.1467-9817.2006.00333.x

Prinsloo, C.H., Harvey, J., Thaba, W. & Moodley, M. 2015. *Hope for rebuilding language foundations: Part 1 - Following a cohort into a second year. Human Sciences Research Council Evaluation Report: siyaJabula siyaKhula's Learner Regeneration Project, Vhumbedzi and Malamulele North-East, Vhembe, Limpopo. Evidence after two years*. Pretoria, South Africa: HSRC.

Prinsloo, C.H., Ramani, E., Joseph, M., Rogers, S., Mashatole, A., Lafon, M., Webb, V., Mathekga, P., Mphahlele, T., Mokolo, F., Somo, L., Tlowane, M., Bopape, S. & Morapedi, M. n.d. *An inter-province study of language and literacy paradigms and practices in foundation phase classrooms in Limpopo and Gauteng*. Pretoria, South Africa: HSRC.

Reeves, C., Heugh, K., Prinsloo, C.H., Macdonald, C., Netshitangani, T., Alidou, H., Diedericks, G. & Herbst, D. 2008. *Evaluation of literacy teaching in primary schools of Limpopo province*. Pretoria, South Africa: HSRC.

Taylor, S., Cilliers, J., Fleisch, B., Prinsloo, C.H., van der Berg, S. & Reddy, V. 2016. *The Early Grade Reading Study: A midline report*. Pretoria, South Africa: Department of Basic Education (DBE).

Thomas, W.P. & Collier, V. 1997. *School effectiveness for language minority students.* Washington, DC: National Clearinghouse for Bilingual Education.

Thomas, W.P. & Collier, V.P. 2001. *A National study of school effectiveness for language minority students' long-term academic achievement*. Available at http://www.usc.edu/dept/education/CMMR/CollierThomasExReport.pdf. [Accessed 29 June 2016].

Wa Kabwe-Segatti, A., Lafon, M. & Webb, V. (Eds.). 2008. The standardisation of African languages: Language political realities. *Proceedings of a CentRePoL workshop* held at University of Pretoria on March 29, 2007, supported by the French Institute for Southern Africa. 1-141.

Wagner, D.A. 2010. Quality of education, comparability, and assessment choice in developing countries. *Compare: A Journal of Comparative and International Education*, 40(6), 741-760. https://doi.org/10.1080/03057925.2010.523231

Webb, V. 1999. Multilingualism in democratic South Africa: The over-estimation of language policy. *International Journal of Educational Development*, 19, 351-366. https://doi.org/10.1016/S0738-0593(99)00033-4

Webb, V. 2004. African languages as media of instruction in South Africa: Stating the case. *Language Problems and Language Planning*, 28(2), 147-173. https://doi.org/10.1075/lplp.28.2.04web

Webb, V. 2010. The politics of standardising Bantu languages in South Africa. *Language Matters*, 41(2), 157-174. https://doi.org/10.1080/10228195.2010.500674

Webb, V., Lafon, M. & Pare, P. 2010. Bantu languages in education in South Africa: An overview. Ongekho akekho! - the absentee owner. *The Language Learning Journal*, 38(3), 273-292. https://doi.org/10.1080/09571730903208389

Wiliam, D. 2011. What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14. https://doi.org/10.1016/j.stueduc.2011.03.001