Routledge
Taylor & Francis Group

Southern African Linguistics and Applied Language Studies

ISSN: 1607-3614 (Print) 1727-9461 (Online) Journal homepage: http://www.tandfonline.com/loi/rall20

The role of monitoring and evaluation in six South African reading programmes

Denver Grigg, Jenna Joffe, Alicia Okeyo, Debbie Schkolne, Nicola van der Merwe, Mlungisi Zuma, Chao Mulenga, Adiilah Boodhoo & Joha Louw-Potgieter

To cite this article: Denver Grigg, Jenna Joffe, Alicia Okeyo, Debbie Schkolne, Nicola van der Merwe, Mlungisi Zuma, Chao Mulenga, Adiilah Boodhoo & Joha Louw-Potgieter (2016) The role of monitoring and evaluation in six South African reading programmes, Southern African Linguistics and Applied Language Studies, 34:4, 359-370, DOI: 10.2989/16073614.2016.1262270

To link to this article: http://dx.doi.org/10.2989/16073614.2016.1262270

Published online: 20 Dec 2016.

Southern African Linguistics and Applied Language Studies 2016, 34(4): 359–370

SOUTHERN AFRICAN LINGUISTICS
AND APPLIED LANGUAGE STUDIES
ISSN 1607-3614 EISSN 1727-9461
http://dx.doi.org/10.2989/16073614.2016.1262270

The role of monitoring and evaluation in six South African reading programmes

Denver Grigg, Jenna Joffe, Alicia Okeyo, Debbie Schkolne, Nicola van der Merwe, Mlungisi Zuma, Chao Mulenga, Adiilah Boodhoo and Joha Louw-Potgieter*
Organisational Psychology, University of Cape Town, Rondebosch, South Africa
**Corresponding author email: Joha.Louw-Potgieter@uct.ac.za*

Abstract: In this article we focus on six reading programmes and ask: Do these programmes work insofar as they improve the reading ability of programme participants? We apply programme evaluation methods and content to these programmes to answer this question. Specifically, we use an approach that identifies the following five different levels of evaluation, namely: programme need; programme theory; programme process and implementation; programme outcome and impact; and programme cost and efficiency. We then add appropriate evaluation questions and research designs applicable to each level. We conclude by providing suggestions to reading programme staff on how to improve monitoring (data collection) in order to strengthen the evaluability of their reading programmes.

Introduction

The poor state of the reading ability of South African learners as assessed by national and international measures has been well-documented (see the Department of Basic Education's (DBE) Annual National Assessments (ANAs) 2013; the Progress in International Reading Literacy Study (PIRLS), as reported by Howie et al. 2011, and Spaull 2014; and the Regional Southern and East Africa Consortium for Monitoring Education Quality Report (SACMEQ) of 2007, as reported by Spaull 2012). In one of the most recent assessments, Van der Berg et al. (2016) showed that 60% of South African children do not learn to read for meaning in any language by the end of Grade 3. These authors prioritised four binding constraints for this state of affairs, namely weak institutional functionality, undue union influence, weak teacher content knowledge and pedagogical skill, and wasted learning time and insufficient opportunity to learn.

Although no empirical evidence exists about the number of in-school and after-school additional reading programmes, it is common knowledge that a multitude of non-governmental organisations (NGOs) have stepped in to assist with teaching learners to read or to improve learners' reading ability. To date, no systematic review exists regarding the effect of these programmes. In this article we focus on six different reading programmes and ask the important question: Do these programmes work? We apply programme evaluation methods and content to these different reading programmes to answer this question. We conclude the article by providing suggestions for NGOs on how to improve the monitoring of their reading programmes.

First we provide a brief overview of what programme evaluation is and of a specific evaluation framework that was used in the six programme evaluations that will be discussed here.

What is programme evaluation?

A programme is a managed, organised set of activities aimed at bringing about a change in people or their circumstances (adapted from Rossi, Lipsey and Freeman 2004: 29). Programme evaluation is a systematic investigation which uses social science methods to enable us to make a judgement on whether or not a programme is bringing about change (Louw-Potgieter 2011). Programme monitoring means tracking progress during programme roll-out and the change in programme participants (Louw-Potgieter 2011).

What do these definitions mean for deciding whether or not reading programmes work? Most reading programmes contain specific activities and the designers of these programmes usually

assume that these activities will lead to better reading ability. Programme evaluators test these assumptions by formulating evaluation questions. These evaluation questions are then answered by using research designs to guide data collection and then by analysing the relevant data by means of quantitative and/or qualitative methods.

In the evaluations that follow, we have used a specific framework for programme evaluation, namely Rossi et al.'s (2004) hierarchical levels of evaluation. These levels, their main evaluation questions and an appropriate evaluation design to guide data collection are presented in Table 1 (please note that the hierarchy starts at the bottom of the table).

From Table 1 it is clear that when we assess need, design and theory, process and implementation, and cost, we tend to use a measure at a specific time. The evaluation design for this is called a descriptive design and we use this type of design to guide data collection when we want to know about the characteristics of a group or a phenomenon (Marlow 2005). For instance, when we evaluate process, we usually want to know who participated in the programme at a specific time and if this group of participants received enough of the programme.

Level 4 in Table 1 is quite different from the rest of the evaluation levels. Here we are concerned with assessing change and whether or not we can attribute this change to the programme. For the

Table 1: Evaluation levels, questions and design (Adapted from Rossi et al. 2004)

Level of evaluation	Main evaluation questions	Evaluation design
5. Cost and efficiency	Do the benefits of the programme justify its costs to society and participants?	Measure at a specific time
4. Outcome and impact*	Are the participants who got the programme better off after the programme than they were before the programme?	Measure of participants before and after the programme (also add measures during the programme)
	And better off than those who did not get programme?	Measure of participants before and after the programme, and measure of non-participants at the same times
	Was it the programme that caused this change?	Measure of participants before and after the programme, and measure of non-participants at the same times. Both groups assigned randomly to participate or not.
3. Process and implementation	Who received the programme and were these the intended participants? Did they receive enough of the programme?	Measure at a specific time
	Who delivered the programme? Were they well-trained and did they deliver the programme as intended?	Measure at a specific time
	Were there sufficient resources to deliver the programme as intended?	Measure at a specific time
2. Design and theory	Is the theory of change which underlies the programme consistent with existing social science knowledge and internally consistent	Measure at a specific time
1. Need	Does the programme correctly define the problem it addresses?	Measure at a specific time

*Impact is a word that is common in everyday vocabulary. However, in programme evaluation impact does not simply mean the influence that the programme has had on the participants. In programme evaluation impact is a technical term that refers to the causal link between the programme and the change in participants, in other words, how the programme and not anything else brought about participant change.

first evaluation question, whether the participants are better off after the programme than before, we need to collect data from before the start of the programme and at the end of the programme. The design that guides us here is called a quasi-experimental design and this means that the group/s under evaluation are assigned to the programme by the programme staff (Cook 2015). If only one group of participants is evaluated, it is called a single group quasi-experimental design. Here we collect performance data from all or some of the participants (also called the intervention group) before and after the programme. We can also refine our data collection with repeated measures (more performance measures) during the programme. These measures are used to track performance improvement as the participant moves through the programme. We use another quasi-experimental design to guide us in our data collection when we want to answer the second evaluation question (Are participants better off after the programme than before, and better off than non-participants?). Here we use the same recipe as above, but we add another group, non-participants (also called a non-equivalent group or a comparison group), and measure their performance before and after the programme. To answer the final performance question, namely whether we can attribute the change in participant performance to the programme, we need a strong design that will help us to attribute causality. An experimental design using randomised assignment to the participant group (also called the experimental group) or the non-participant group (also referred to as the control group) is appropriate here. In other words, we assign people who need the programme randomly (by chance) into either the participant group or the non-participant group and then we track their performance (see Kirk 2003 on experimental design). We assume that these two groups are similar in performance before the programme starts (their pre-test scores are the same). If their post-test scores are different and if the post-test score of the participant group is higher than the non-participant group, then we can assume with reasonable certainty that the programme caused the change. (Please note that we have not discussed the full range of experimental and quasi-experimental designs here. Readers are referred to Cook 2015 for this).

So what can monitoring and evaluation add to reading programmes? Why do we think it is important to track progress and assess whether reading programmes work or not? The answer is that it is mainly for funding, programme design and motivational reasons. Most funders require evidence that programme staff are doing what they promised to do and that their programmes are bringing about the change described in the programme goals. Programme managers, on the other hand, benefit from knowing whether they are doing the right thing and that specific programme activities are changing people's lives (or reading ability). Finally, both programme managers and volunteers who implement the programme, will be motivated by evidence that shows that their time and effort are paying off in terms of participant improvement.

Applying programme evaluation principles to six different reading programmes

In 2007, the University of Cape Town (UCT) introduced a Master's programme in Programme Evaluation, consisting of a 50% coursework component and a dissertation making up the other 50%. For the dissertation, students have to evaluate a real-life programme. The six reading programmes that we discuss here have been evaluated for dissertation purposes by six Masters' students (the first six authors of this article) and supervised by the last three authors. The reading programmes are: the Shine's Centre's Literacy Hour (Schkolne 2014) as implemented by Shine, the Shine Centre's Literacy Hour (Grigg 2012) as implemented by the Life Matters Foundation (LMF), Wordworks' Early Literacy Programme (Okeyo 2015), the Help2Read Programme (Joffe 2015), the literacy part of the Anna Foundation's 3Rs Programme (van der Merwe 2015) and Living Through Learning's (LTL) Coronation Adventure Reading Room Programme (Zuma 2016). These six reading programmes constitute all the reading programmes that have been evaluated in the Master's programme since 2007.

All six programme managers granted permission to the students to evaluate their programmes. All evaluations were given ethics clearance by the Ethics in Research Committee of the Commerce Faculty of the University of Cape Town. The evaluations were performed in the year indicated above and reflect the state of the programme at that date and do not include later changes or improvements made to the programmes. It should also be noted that some evaluations focused on a specific part or group of programme participants and not on all parts or participants. Below are brief descriptions of each programme.

The goal of the Shine Centre's Literacy Hour programme is to increase the number of children from poor socio-economic backgrounds who complete Grade 3 with grade-appropriate literacy scores. Shine identifies grade 2 and 3 learners who are struggling to read and write by means of a literacy assessment. These children are then offered individualised reading and writing practice for one hour twice a week by volunteers using a variety of reading material, games and literacy activities designed by Shine. During the hour, 15 minutes are spent on each of the following programme activities, namely paired reading, shared reading, have-a-go-writing and word play. Learners receive a motivational praise note at the end of the session. The Shine Centre also franchises other organisations to present its programme. Grigg (2012) conducted a process and outcome evaluation of the Shine programme as implemented by LMF. The outcome evaluation assessed whether or not the reading ability of Grade 2 learners in two Western Cape (WC) schools improved. He used a single group, quasi-experimental design with a before measure and two measures during the programme. Schkolne (2014) conducted an outcome evaluation of Shine's Literacy Hour programme, as implemented by Shine itself, in eight Western Cape schools, focusing on the participants in the 2011–2013 programmes. She used a quasi-experimental, regression discontinuity design with a before measure in Grade 1, and two measures during the programme when participants were in Grades 2 and 3.

The goal of Help2Read's programme is to address child literacy in underprivileged areas (and the lack of skills development and unemployment) in South Africa. Help2Read targets Grade 2, 3 and 4 learners who spend 30 minutes twice a week in school with a trained volunteer or a trained stipend tutor who uses Help2Read's book boxes as learning materials. Programme activities for each session are reading, comprehension and conversation, identifying letters, words and sounds, and playing literacy games. Joffe (2015) conducted an outcome evaluation of this programme in two Western Cape schools, focusing on Grade 2, 3 and 4 learners on the 2015 programme. She used a quasi-experimental, non-equivalent group design with a before measure and one measure during the programme.

The 3Rs programme (Reading, Running and Right-ing) aims to address the low levels of self-esteem plus the educational and social needs of rural children and to offer these children lifelong learning in order to build self-efficacy and self-respect. The 3Rs programme targets Grades 1–12 children on farms who spend two hours per week, after school, with trained, local women who live on the farm and are paid by the farmer. Age-appropriate learning materials are provided by the Anna Foundation. Programme activities consist of one hour of academic work (that can include literacy or numeracy) and one hour of either sport or drama. For the literacy part of her dissertation, van der Merwe (2015) conducted an outcome evaluation, focusing on Grade R–12 learners on five Western Cape farms. She used a quasi-experimental, single group design with one before measure and two measures during the programme.

The high-level goal of the Wordworks' Early Literacy Programme is to provide young children with the support they need to learn to read and write successfully. The programme targets learners in Grades R and 1 who spend two hours per week in school with trained volunteers who use Wordworks' learning packs. Programme activities consist of talking and doing, reading (sounding out words), writing and drawing, and games with sounds, letters and words. Okeyo (2015) performed an outcome evaluation, using a quasi-experimental, single-group design that focused on the Grade 1, 2014 participants of a 12-month-long programme and the Grade 1, 2014 and 2015 participants of a six-month-long programme. The 12-month programme had one before measure and two measures during the programme, and the six-month programme had one before measure and one measure during the programme.

The ultimate goal of LTL's programme is to build a proper educational foundation for primary school learners in South Africa by improving and developing literacy in disadvantaged schools, building confidence of learners in reading and writing and equipping teachers with effective teaching skills to administer the programme successfully in schools. The programme targets Grades 1, 2 and 3 learners who spend 75 minutes per day, for four or five days in school with a trained teacher in a special reading room equipped with LTL's learning material. An LTL teaching assistant is available twice per week to provide support to the teacher. The programme activities consist of basic recognition of familiar sounds, blending familiar sounds and letters to form three-letter words by

means of a sliding process, recognition of more unfamiliar sounds, and blending unfamiliar sounds and letters to form three-letter words. Zuma (2016) conducted an outcome evaluation using a single group, quasi-experimental design with four measures during the programme. The participants were Grade 1 learners in 18 Western Cape schools on the 2015 programme.

The good news is that, on average, all six programmes showed that participants had improved their reading ability. In the next section we would like to turn this into even better news by showing how programmes could produce stronger evidence of improved results by making slight changes to programme monitoring.

Improving basic programme monitoring

In an ideal world, programme evaluators would be involved with programmes from the very start. This early involvement would provide programme managers with useful evaluation questions and strong designs which would guide collection of suitable data to answer these questions. However early involvement rarely happens and often programme evaluators are faced with a lack of evaluation questions, no evaluation design and minimal data. On the other hand, lots of inappropriate data are sometimes available. All six programmes we evaluated had data, some of them complete and accurate data, but we are of the opinion that this state of affairs could be improved so that programme managers can demand even more useful evaluations.

Table 2 reflects the secondary data each of the six programmes had collected. These data are organised according to the evaluation levels in the hierarchical model described earlier.

From Table 2 it is clear that programmes did not have data on needs assessment (or that the students did not request these data). The great need for reading programmes has been well documented and we would advise programme managers not to repeat a needs assessment, as it is a costly exercise. However, it may help programme staff to have some indication of how many learners in how many schools the programme intends to cover in a specific year. Shine and LTL provided data on the number of learners they intended to reach in a specific year. LTL also showed how these numbers related to realistic budgeting and human resource planning.

What is interesting here is that only Shine, Help2Read and Wordworks had well-documented programme activities. From the student dissertations we know that evaluators had to construct programme activities from websites, annual reports, interviews, etc. Documented programme

Table 2: Secondary data available per programme

Level	Shine	Shine Franchise	WordWorks	Help to Read	3 Rs	LTL
Need						
Number	✓					✓
Design & theory						
Programme activities	✓		✓	✓		
Programme theory						✓
Process & implementation						
Selection	✓	✓	✓	✓		
Enrolment	✓	✓	✓	✓	✓	✓
Attendance	✓	✓	✓		✓	✓
Presenter	✓	✓	✓	✓	✓	✓
Comparison group	✓			✓		
Control group						
Outcome & impact						
First measure	✓	✓	✓	✓	✓	✓
Repeated measures	✓	✓	✓	✓	✓	✓
Sustainability measure						
Cost						
Per participant						✓
Comparison other programmes						

activities form the basis of a programme theory. Evaluators are rarely experts in reading and need to know what the participants experience on a programme and why specific activities were included in the programme. Programmes also change over time as activities are added or discontinued. Sophisticated evaluations can indicate whether these changes affected the outcomes of the programme participants.

Only one programme, LTL, had a documented programme theory. This was in the form of a complex logical framework and outcomes covered changes from Grades 1–6. Our advice to programme managers is to spend sufficient time on constructing a programme theory. Such a theory should indicate, at a glance, what activities the participants engage in and how these activities will change the participant outcomes. Below are two samples of such "at a glance" programme theories, namely Shine's programme theory as constructed by Schkolne (2014) (see Figure 1) and LTL's logical framework for Grade 1 learners as simplified by Zuma (2016) (see Figure 2).

From Table 2 we saw that Implementation/Process data were available. However, the quality of the data was often problematic. The evaluators have documented this in their dissertations and common problems were missing data (whether for single participants or a whole group of participants, or for a specific repeated measure), outliers within the data (e.g. performance expressed as a percentage and a participant obtaining a score of 210%), inaccessible data and data that were difficult to retrieve. We would like to commend LTL on providing their evaluator with clean, accurate data and recommend that programme managers visit the programme administrator to learn about excellent data management. We would also like to alert funders that when they require a monitoring and evaluation aspect to funding proposals, they actively fund data management.

We suggest the following refinements to Implementation/Process data collection:

- It matters how participants are selected into a programme aimed at improving reading ability. Often, teachers identify participants. While we do not question teachers' experience, this is not good practice, as different teachers may use different selection criteria. For instance, this could mean that learners with low self-esteem (instead of learners with poor reading performance) end up in a reading programme.
- If a programme specifies a specific target group (e.g. Grade 1 learners in under-privileged schools who perform below age-appropriate reading levels), enrolment forms should not only document names and surnames but also date of birth, school name and quintile, class and class teacher, and reading performance before the programme starts. It sounds uncharitable, but programme managers should be aware that every participant who does not fall into the target group is attending a programme that was not developed for him/her and uses resources that could be better applied to a learner who is specifically targeted by the programme.
- It is important to collect attendance data for participants, as we need to know how much of the programme they received.
- Although we steer away from evaluating individuals, we suggest that the names of volunteers/teachers/presenters be linked to specific participants/groups in order to check whether the programme is delivered as intended by all and produces the same results for comparable participants/groups.
- For presenters, we suggest that their names and surnames, whether they attended specific training sessions or not, whether they used programme-specific learning material, and their attendance of reading sessions in/after school, be documented (supervision records could be designed to check the latter two measures). It should also be documented whether presenters are teachers, teaching assistants, or volunteers.
- Where there are comparison or control groups (those learners who do not take part in the programme), the same demographic data mentioned in the first three bullet points need to be collected for these learners.

Outcome/impact data form the heart of an evaluation, as these data indicate whether or not participants changed and whether or not this change can be attributed to the programme. In Table 2 we categorised outcome data as *first measure* (these data are usually collected during the first programme session, or sometimes before the programme starts, as a selection measure), *repeated measures during the programme* (more than one performance measure while the programme is

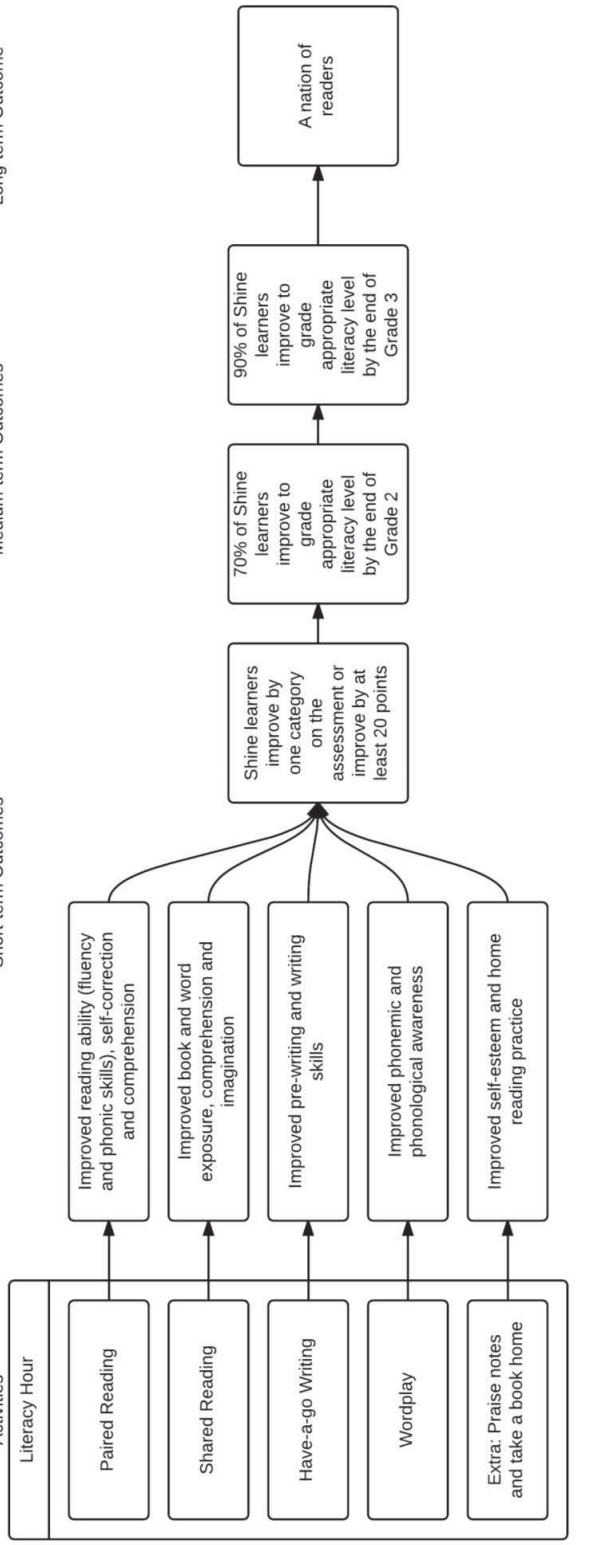

Figure 1. Shine's Programme Theory

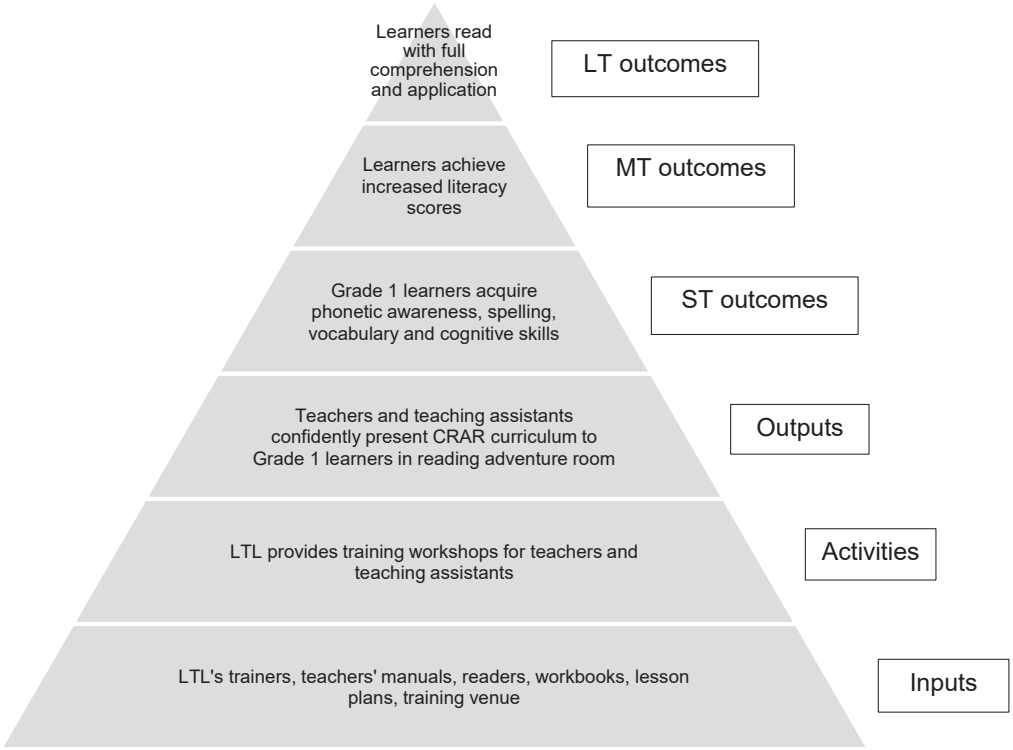

Figure 2: Programme Theory of LTL's CRAR Programme

running, inclusive of end-of-programme measures) and *sustainability measures* (measures taken 6 or 12 months after the participants exited the programme). We suggest the following refinements for collecting outcome data:

- In order to measure change, a measure before the programme starts or at the start of the programme and a measure (or several measures) during the programme are required. In the student dissertations, the terms, diagnostic, baseline or pre-test, are used for before measures and the term, post-test, for measures during the programme or at the end of the programme. It is not the intention to provide technical definitions of these terms here. Suffice it to say that a measure is needed before the programme starts, and often this measure is also used to assess the level of a learner's reading ability. All six programmes used before measures and repeated measures during the programme.
- When teaching Grade 1 learners to read, an assessment of reading ability before the programme may not provide meaningful data, as these learners often cannot read at this stage of their school career. We suggest that repeated measures during the programme be used for participants who are learning to read.
- It is crucial to collect complete data (i.e. data for each learner for each repeated measure). When evaluators conduct an outcome evaluation, learners with incomplete data are usually not included in the group whose data are analysed. A smaller participant group complicates data analysis and makes it more difficult to say with certainty that learners have improved and that this improvement is due to the programme.
- As also mentioned under Process/Implementation, outcome data for comparison groups need to be collected too. As these groups do not receive the programme, a measure of reading ability before the programme and one after the programme are sufficient.

From Table 2, it is clear that LTL was able to provide cost data per learner. It could be that cost data was available for the other programmes, but that students did not request it. Costing a programme is complex and we refer programme managers who want to add this level of data to Rossi et al. (2004: 331–366).

Adding bells and whistles to programme monitoring

When examining the monitoring data of the six programmes, it became clear to the evaluators that much more could be done with these data. We call these data bells and whistles, as simple additions in data collection can sometimes make a big difference to an evaluation. Below are such additions:

- Shine uses a standardised selection measure at the end of Grade 1 to identify learners for the reading programme that starts in Grade 2. Learners are selected into the programme based on their reading performance on this measure. If the cut-off point for selection is adhered to, it means that only those learners who are in need of the programme end up on the programme.
- Furthermore, Shine was able to adhere strictly to this selection method and enforce the cut-off point for selection. The programme manager went a step further and collected performance data not only of the participants, but also of the non-participants who were weak readers but fell just above the cut-off for programme participation. This enabled Schkolne (2014) to compare participants and weak non-participants and provide robust evidence that the programme caused the improvement in participants' reading performance.
- Although we said earlier that a before measure for participants who are learning to read is not meaningful, it may help in cases where non-English-speaking participants are learning to read in English. LTL (Zuma 2016) conducts an assessment in basic English consisting of picture matching, formation of sounds, filling in missing sounds, doing puzzles and following mazes before the participants start to learn to read.
- Participant attendance matters. We know that participants who receive more of a programme show more improvement in their outcomes (Bowie et al. 2006). Programme staff could add a layer of sophistication to their evaluations by using attendance data to split participants into two groups (high and low attenders) and assess whether there was any difference in their performance scores. The 3Rs' attendance data enabled van der Merwe (2015) to show that participants who were exposed to more of the programme showed greater literacy improvements than participants who received less of the programme.
- Apart from designing their own repeated measures for outcomes, both Shine and LTL added age-appropriate standards (what level of reading ability a learner should attain at a specific age). Shine used their own age-appropriate standards and class teachers provided LTL with data of the learners' performance on Curriculum Assessment Policy Statements (CAPS) reading assessments. Whether programmes use CAPS data, ANA data (if these are implemented again) or their own age-appropriate standards, this addition is strongly recommended. Such an age-appropriate standard enables the evaluator to assess whether the learners have improved to a level that is regarded as suitable for their age. Often we can say that learners have improved, but this improvement may still fall below the reading standard they should have achieved for their age. We acknowledge that such a standard should be applied with caution in schools where there are many under or over-age learners.

From the additions we have discussed here, it is clear that small improvements in the design of measures and data collection would enable evaluators and programme managers to claim with more certainty that reading programmes work.

Adding scientific evidence to programme monitoring

There is ample evidence in the scientific literature on what works in reading programmes. We suggest that programme managers note this evidence and incorporate it into their reading programmes.

- When learners are assessed for their reading ability at the end of the programme, we may see improvement and even age-appropriate improvement. However, we do not know whether these improvements are sustainable. We recommend an outcome measure for ex-participants half-way through the next grade in order to assess sustainability.

- All six programmes discussed here use a phonics-based approach (or an approach that includes phonics) where reading is taught by systematic letter-sound correspondence. There is sufficient scientific evidence to show that a phonics-based approach works better than whole-word or whole-language approaches (National Institute of Child Health and Human Development 2000; Dehaene 2010). (Whole-word or whole language approaches refer to approaches where learners are taught to recognise direct associations between whole words or sentences and their meaning). This distinction is important, as some older teachers may still adhere to whole-word or whole-language teaching methods and undermine a phonics-based programme.
- Simple reading material without too many illustrations is most effective in teaching a phonics-based approach for learners who are learning to read (Dehaene 2010). LTL's workbook is an excellent example of non-distracting reading material that is aligned with the CAPS curriculum.
- While it is acknowledged in developed countries that professionals (i.e. teachers) are best suited to teach learners to read or to improve their reading (Reynolds, Wheldall and Madelaine 2010) this may not always be feasible in developing countries. South African teachers in under-privileged schools, who are poorly trained in literacy, often have to cope with large classes (more than 40 learners per class) where learners have a wide range of reading abilities. In such cases, well-trained teaching assistants (LTL) or volunteers (Shine, Shine LMF, Help2Read, WordWorks, 3Rs) often do the job as well as professionals. Wasik (1998) and Elbaum et al. (2000) concluded that well-trained instructors who received quality training and adequate supervision were as effective as professionals.
- The four programmes that focus on reading improvement (Shine Centre and Shine LMF, Help2Read, and 3Rs) used one-on-one instruction. Elbaum et al. (2000) and Vaughn et al. (2003) compared the same reading programme using 1:1, 1:3 and 1:10 instruction. They found that 1:3 instruction was as effective as 1:1 instruction. Allocating more learners of the same reading level to a single volunteer could cut the cost of reading programmes significantly. The 3Rs programme, implemented in a rural context by farm women, is a case in point of innovative use of limited human resources spread over a wide range of grades.
- Learners in the Shine programme exited the programme when they reached age-appropriate reading ability, while the Wordworks programme showed that a six-month programme may be long enough for sufficient reading improvement. But what happens to learners who stay on the programme for the full duration and who do not improve? Schkolne (2014) identified ten learners who did not improve after two years on the Shine programme. We suggest that programme managers alert teachers/teaching assistants/volunteers to the possibility of lack of improvement in a small group of programme participants. Morgan (2005) described a bi-directional relationship between reading failure and social factors. In simple terms this means that poor reading ability may be associated with social and interpersonal problems in learners, and that learners who experience social and interpersonal problems outside of school may end up having reading problems. Shine LMF has a counselling and life-skills programme to deal with such issues. Programme managers may wish to implement such programmes or include referral skills (to social workers or psychologists) in their presenter training.

In this sub-section, we have addressed some research evidence for guidance for programme managers. It should be noted that we have left out the issue of early intervention versus later remediation and the focus on 'reading to learn' that usually surfaces in Grade 3 onwards.

Conclusion

From these evaluations and based on discussions with programme managers about this publication, it is clear that there is a great need for standardised assessments for young children's reading ability. Such national norms will enable NGOs to make better decisions regarding programme participation (and exit), general reading progress and the attainment of age-appropriate reading ability. It is costly to create standardised assessments and this should be done at national level. In their work for the DBE, van der Berg et al. (2016) identified national standards for reading in both English and mother tongue language for Grades 1–4 learners as one of the Department's priorities for improved literacy.

We wish to commend the programme staff of the six programmes we have discussed here. They are doing a sterling job in difficult circumstances. As more focus is placed on how to improve the poor literacy of South African children, we predict that these and other service providers outside the educational system may play an even bigger role in getting our children to read with fluency and comprehension. Any programme that can show that it, and not anything else, improves reading ability may find itself in high demand and at the receiving end of substantial funding. Programme monitoring and evaluation will enable programme managers to show what works and what does not and respond to this increased demand for literacy services.

References

Bowie L, Garrett B, Kinukawa A, McKinney K, Moore KA, Redd Z, Theokas C, Wilson B. 2006. *Program implementation: What do we know?* Washington, D.C.: Child Trends.

Cook TD. 2015. Quasi-experimental design. Available at: http://onlinelibrary.wiley.com/doi/10.1002/9781118785317.weom110227/abstract;jsessionid=193BB438139E03764794B9869D880AF4.f04t01?userIsAuthenticated=false&deniedAccessCustomisedMessage [Accessed 29 August 2016].

Department of Basic Education. 2013. *Report on the annual national assessment of 2013*. Available at: http://www.education.gov.za/Portals/0/Documents/Reports/2013%20Ana%20Diagnostic%20email%20Version.pdf?ver=2014-02-27-142226-293 [Accessed 29 August 2016].

Dehaene S. 2010. *Reading in the brain. The new science of how we read*. New York: Penguin Books.

Elbaum B, Vaughn S, Tejero Hughes M, Moody SW. 2000. How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology* 92: 605–619.

Grigg D. 2012. A theory, implementation and short-term outcome evaluation of Lifematters Foundation's literacy intervention. Master's dissertation, University of Cape Town, South Africa.

Howie S, van Staden S, Tshele M, Dowse C, Zimmerman L. 2011. Progress in international reading literacy study 2011: South African children's reading literacy achievement summary report. Centre for Evaluation and Assessment. University of Pretoria. Available at: http://w.ac.za/sitefiles/File/publications/2013/PIRLS_2011_Reeb.upport_12_Dec.PDF.

Joffe J. 2015. An outcome evaluation of the Help2Read Programme. Master's dissertation, University of Cape Town, South Africa.

Kirk RE. 2003. Experimental design. Available at: http://onlinelibrary.wiley.com/doi/10.1002/0471264385.wei0201/abstract?userIsAuthenticated=false&deniedAccessCustomisedMessage [accessed 29 August 2016].

Louw-Potgieter J. 2011. Monitoring lecture slides. University of Cape Town, South Africa.

Marlow CR. 2005. *Research methods for generalists in social work*. New York: Thomas Brooks/Cole.

Morgan M. 2005. Reading to learn. A major social and educational issue. In: Kennedy E, Hickey TM (eds), *Learning to read and reading to learn.* Dublin: Reading Association of Ireland. pp. 1–7.

National Institute of Child Health and Human Development. 2000. *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instructions.* NIH Pulbication No. 00-4769. Washington, D.C.: Government Printing Office.

Okeyo A. 2015. Strengthening foundational literacy: A process and outcome evaluation of the Wordworks Early Literacy Programme. Master's dissertation, University of Cape Town, South Africa.

Reynolds M, Wheldall K, Madelaine A. 2010. Components of effective early reading interventions for young struggling readers. *Australian Journal of Learning Difficulties* 15: 171–192.

Rossi PH, Lipsey MW, Freeman HE. 2004. *Evaluation: A systematic approach* (7th edn). Thousand Oaks: Sage.

Schkolne D. 2014. An outcome evaluation of the Shine Centre's Literacy Hour Programme. Master's dissertation, University of Cape Town, South Africa.

Spaull N. 2012. South Africa at a glance. SACMEQ at a glance series. Research on Socioeconomic Policy (RESEP). Available at: http://resep.sun.ac.za/wp-content/uploads/2012/07/Spaull-2012-SACMEQ-at-a-Glance-10-countries.pdf [accessed 29 August 2016].

Spaull N. 2014. Accountability in South African Education. Ch. 4 in 'Transformation Audit 2013: Confronting Exclusion'. pp. 51–65. Institute for Justice and Reconciliation. Cape Town. Available at: http://transformationaudit.org/blog/wp-content/uploads/2014/03/Accountability%20in%20South%20African%20education.pdf [accessed 29 August 2016].

Van der Berg S, Spaull N, Wills G, Gustafsson M, Kotze J. 2016. Identifying binding constraints in education. Available at: http://resep.sun.ac.za/wp-content/uploads/2016/06/PSPPD_BICiE-email-01062016.pdf [accessed 29 August 2016].

Van der Merwe N. 2015. An evaluation of the Anna Foundation's 3Rs Programme: Does the Reading, Running and Right-ing Programme ultimately increase learners' self-esteem? Master's dissertation, University of Cape Town, South Africa.

Vaughn S, Linan-Thompson S, Kouzekanani K, Bryant DP, Dickson S, Blozis SA. 2003. Reading instruction grouping for students with reading difficulties. *Remedial and Special Education* 24: 301–315.

Wasik BA. 1998. Volunteer tutoring programs in reading: A review. *Reading Research Quarterly* 33: 266–291.

Zuma M. 2016. An outcome evaluation of Living Through Learning's Coronation Reading Adventure Room Programme. Master's dissertation, University of Cape Town, South Africa.