

This report details the impact of the intervention and the extent to which child characteristics, teacher/classroom variables and the child's home-learning environment predicted improvements in the main outcome variables.

BACKGROUND TO THE PROGRAMME

Early Childhood Development (ECD) teachers have a critical role to play in providing language- and print-rich learning experiences for young children. There is a need for evidence-based, contextually relevant teacher development programmes and resources that support quality language teaching.

The Little Stars programme was carefully designed to meet the specific needs, characteristics and circumstances of early childhood education in low-resource contexts in South Africa. To this end, it has the following features:

- is designed for contexts where children may have had limited language learning opportunities and experiences with books in their early years.
- uses a wide variety of stories by different authors and illustrators to reflect a range of cultures and languages. The stories, published by Book Dash, African Storybook and Nal'ibali, are freely available through Creative Commons licences.
- supports teaching and learning in the mother tongue.
- encourages teachers to embrace the diversity of languages children bring to the class and supports learning in an additional language.
- recognises the importance of connecting home and pre-school environments.
- enables children and teachers to explore the purpose and meaning of print in their lives.

The oral storytelling, which is a focal aspect of this programme, provides a bridge to written language for children. Hearing a story and then seeing the same story in written form opens the door to the relation between printed and spoken words.

THE INTERVENTION

The programme aligns with the National Curriculum Framework, with stories linked to each of the six Early Learning and Development Areas. It has a play-based approach to teaching and learning and follows a similar structure to the widely used Wordworks Grade R story-based home language teaching programme (Wordworks, 2023b). It was designed around key principles of early learning that include the following: nurturing and responsive relationships are crucial for learning and language development; children learn best when new learning has meaning and is connected to something familiar; children learn by being active and using all their senses; children make meaning through stories and play; children learn best when they are encouraged to interact, share ideas and ask and answer questions (O'Carroll et al., 2023).

Materials

The Little Stars programme includes 18 story packs providing a total of 36 weeks of teaching. The programme comprised five activities per week with one daily teaching

activity (in a two-week cycle/routine), that required only 15–30 minutes per day and could be integrated into an existing daily teaching programme. Each two-week cycle begins with telling a new story, followed by retelling the story, focusing on target vocabulary. Children were then engaged in activities related to each story: a song or rhyme to reinforce vocabulary introduced in the story; listening for beginning sounds in words used in the story; role playing the story to provide opportunities to use new vocabulary and phrases; retelling the story with sequence picture cards to build understanding of narrative structure; shared writing and reading a big book based on the oral story to teach print concepts; children drawing their favourite part of the story and creating 3D objects related to the story; little books for children to take home to retell the story. Each Little Stars story pack includes: a Teacher’s Guide with guidelines for daily activities in a two-week cycle; a Big Book, 2D puppets, sequence pictures and a photocopyable little book. These materials are Creative Commons licensed (Wordworks, 2023a).

Training of teachers

The teachers in the intervention group attended an orientation workshop to find out about the Little Stars programme and the research project and to clarify expectations. Following this, they participated in two full-day training workshops and then four monthly workshops (2.5 hours each, held quarterly between February and August 2022). The training ended with a graduation event that included opportunities for feedback and reflection. Two teachers in the sample attended less than half of the workshops and were excluded from the study. The attendance rate for the remaining teachers was excellent, with an average of 6.2 out of 7 sessions. The control group received the same training after the conclusion of the trial (from August to December 2022).

Fidelity to programme

The trainers visited the 26 teachers in the intervention group in July 2022. As trainers could not observe teachers doing all 10 activities in the two-week cycle, we asked teachers to do one of the main teacher-led activities: *Storytelling*, *Sequence pictures* or *Reading a Big Book* and a child-led activity: *Drawing and emergent writing*. Using a rating scale of 1 (poor) - 5 (good), the trainers rated the teachers on these items. The results of the fidelity assessment are reported below.

EVALUATION OF THE PROGRAMME

To explore the effectiveness of the programme, we addressed the following main research questions:

1. Is the Little Stars classroom programme perceived as being useful and feasible for teachers with limited formal training to implement effectively in under-resourced contexts?
2. What is the effect of a resource-based training programme on teachers’ teaching practices and on interaction in the classroom?
3. What is the effect of the story-based intervention programme on children’s early literacy and language skills?

Methods

Design: The design is a randomised controlled trial with two separate language groups: one comprising isiXhosa-speaking children and educators, the other comprising Afrikaans-speaking children and educators. Early Childhood Development (ECD) centres were randomly assigned either to the treatment (Little Stars) condition or to a business-as-usual wait-list control condition. One ECD centre in each language group had two classrooms and these classrooms were assigned to the same group.

Recruitment and consent took place before randomisation. A larger group of teachers and principals were invited to an information session about the project and trainers explained the consent form in the participants' home language. They were given an option of completing the form in their home language or English. Teachers who consented to participate explained the project to parents and only children whose parents had completed consent forms were included in the study. The study was approved by the Health Research Ethics Committee, Stellenbosch University, in May 2021 [N21/05/047].

Study setting: The ECD centres were located in low-income suburbs of Paarl and Wellington, rural towns within 80 kms of Cape Town, as well as Khayelitsha, a township established during the apartheid era in Cape Town. All sites are in the Western Cape province of South Africa. Centres were recruited through local ECD fora and NGOs (Inceba Trust, Sikhula Sonke, Ikamva Labantu).

Child participants: Participants were 156 isiXhosa-speaking and 154 Afrikaans-speaking children (~25% of the children in each class) recruited from the 51 pre-Grade R classrooms in the participating ECD centres. Children who failed the hearing test and/or WHO disability screening (see below) were excluded from the study. In addition, all of the child data from one ECD centre were excluded from analysis where the practitioner attended 50% or less of the training, was not the main teacher in the 4- to 5-year-old class, left the centre within 6 months of the start of the intervention, or where the centre closed part-way through the study. Child data were excluded from analysis where children moved or left the centre. The final sample for data analysis comprised 112 isiXhosa-speaking and 135 Afrikaans-speaking participants across 25 isiXhosa (13 intervention, 12 control) and 27 Afrikaans ECD centres (14 intervention, 13 control). Reasons for exclusion of data are detailed in Table 1. All participants with data included in the final analysis had ELOM assessment scores at Times 1 and 2, but the number of children completing other measures varied.

Table 1

Sample exclusion and attrition numbers by language and study group

	isiXhosa		Afrikaans	
	control	intervention	control	intervention
Children assessed at baseline	68	88	75	79
Children excluded from study (failed hearing test/WHO disability screener)	3	3	1	1

ECD centre removed from study	2	10	0	5
Children lost/removed by endline	19	25	7	12

Teacher participants and ECD characteristics: The baseline sample consisted of 56 teachers (29 isiXhosa, 27 Afrikaans) from ECD centres. Participants were recruited through local ECD forums and NGOs in Paarl, Wellington and Khayelitsha (Inceba Trust, Sikhula Sonke and Ikamva Labantu). These NGOs support and train principals and ECD teachers and conduct regular visits to ECD centres to provide them with support and monitor their progress. The final sample included 51 teachers (25 isiXhosa, 26 Afrikaans). The teachers' qualifications varied: 21% had no qualifications, matric, or Level 1; 73% had an NQF Level 4 or 5 qualification and 6% had an ECD diploma or NQF Level 6. Their average length of teaching experience was: 7 years, with 20% of the sample in their first or second year of teaching.

The ECD centre fees varied between R120–R650/month (an average of R363 per month).

Child measures, data collection and coding protocols: In February 2022 (Time 1) trained and accredited assessors conducted classroom observations and assessments of the children. The programme was designed to last for the full school year (36 weeks). Covid delayed the programme start and ethics approval required the control group to access the intervention in the same calendar year as the intervention group, thus the evaluation was conducted at 26 weeks in August 2022 (Time 2). Validated processes to measure age, height, gender, and language spoken were followed.

Few standardised instruments for the assessment of general ability, language and literacy are available in South Africa, particularly in African languages. The assessments used are detailed below and were selected or developed specifically for this study, because they were available (or deemed suitable for translation and use in this context) in isiXhosa and Afrikaans. All assessments were administered, and responses transcribed and scored by native speakers of isiXhosa and Afrikaans, as appropriate. Sessions including assessments that required post-scoring were recorded. Children were assessed individually in a space in the ECD Centre and all responses were written down on the test forms and voice-recorded on the tablets. Different vocabulary assessments were administered at pre- and post-test, as detailed below.

Screening measures:

Hearing. Before the study commenced, a hearing screening was conducted by a qualified audiologist on all child participants. This screening revealed that two children had hearing loss; they were referred for a diagnostic audiology and excluded from the final sample. Twenty-six percent of children in the sample had middle ear problems and were referred for medical treatment and re-screening. Their data were not excluded from the final sample.

Disability screen. All sampled children were screened at both assessment points for disabilities likely to affect performance on the ELOM using four modified

questions from the World Health Organisation Ten Point Disability Screen (Durkin et al, 1995).

1. WHO Screen: “Compared with other children, does the child have difficulty seeing, either in the daytime or at night?” ELOM modification: “Did this child seem to have difficulty seeing?”
2. WHO Screen: “Does the child appear to have difficulty with hearing?” ELOM modification: “Did this child appear to have difficulty with hearing?”
3. WHO Screen: “When you tell the child to do something, does he/she seem to understand what you are saying?” ELOM modification: “When you told this child to do something, did he/she seem to have difficulty understanding what you were saying?”
4. WHO Screen: “Does the child have difficulty in walking or moving his/her arms, or does he/she have weakness and/or stiffness in the arms or legs?” ELOM modification: “Did this child have difficulty in walking or moving his/her arms, or did he/she have weakness and/or stiffness in the arms or legs?”

Children who failed 2 or more items on the WHO screening on were excluded.

Background measures: child:

General ability and readiness to learn. The Early Learning Outcomes Measure (ELOM) is a standardised tool suitable for measuring the effects of early learning programmes and children’s readiness to learn in children aged 50-69 months (Snelling et al., 2019). It assesses five domains: gross motor development, fine motor coordination and visual motor integration, emergent numeracy and mathematics, cognition and executive functioning, and emergent literacy and language (ELL). Content and construct validity, reliability and cross-cultural fairness have been established (Anderson et al., 2021). We used the total scores, which correlate well with the WPPSI (Anderson et al., 2021; Wechsler, 2012) as an index of general ability and readiness to learn.

Social-emotional functioning. Children’s social-emotional functioning was measured through interviews with the participating children’s teachers using the Social-Emotional Rating Scale of the ELOM (Snelling et al., 2019). This scale assesses two aspects: (1) Social relations with peers and adults: including the ability to cooperate without prompting; to work with peers in group activities; to resolve problems without aggression; to seek support, assistance and information from familiar adults; (2) Emotional readiness for school: including the ability to communicate with adults; appropriate expression of needs and feelings; willingness to do things without help; ability to adjust to changes in class or home routine; confidence in new experiences; and initiating activities.

Home learning environment. The ELOM Home Learning Environment Tool (HLE) (Dawes et.al, 2023) was used to measure the time a caregiver spends with the child during the week and at weekends, early learning resources and activities in the home. The HLE interview was completed by telephone interviews with caregivers for just under half of the sample (n = 66).

Initial vocabulary. The Cross-linguistic Lexical Tasks (CLT) was administered to collect data on noun and verb comprehension and production. It was developed as a cross-linguistically and cross-culturally comparable tool for the lexical assessment of children (Haman et al., 2017). It is available in isiXhosa and Afrikaans and has been

used in studies in South Africa in mono- and multilingual populations (Potgieter & Southwood, 2016). Each language version comprises a unique combination of target words and detractors customised for the linguistic and cultural context of the country where the language is spoken (Haman et al., 2017)

Administration and scoring. The CLT consists of four sections: noun comprehension, verb comprehension, noun production, and verb production. Each section has 30 test items and 2 practice items, with a total maximum score of 120 points, and a maximum score of 60 points for comprehension and production. The comprehension sections are picture recognition tasks where the child hears the target word embedded in a prompt question and has to choose the matching picture on a page with four pictures (e.g., *Who is digging?* for verb comprehension tasks and *What is this?* for noun comprehension tasks). The production sections are picture naming tasks where the child sees one picture at a time and answers the prompt question from the experimenter with a word (e.g., *What is that?* for noun production tasks and *What is he/she doing?* for verb production tasks). The target words are randomly distributed within each section and not according to complexity or AoA. The assessment takes approximately 15 minutes per participant. The scoring guidelines of Bonacker et al., (2016) were followed to score the production sections. An answer was scored correct if the child provided the correct target word, a regional variant, a word that corresponds with the picture and is more specific than the target word (e.g.), or mispronunciations if the word was still recognisable as the target word. Because of the prevalence of code switching in the study population, loan words or English versions of the target word were scored correct. Words were scored incorrect if the child provided a word that is too general, a paraphrase, or a word from another word class (e.g., a noun in the verb section and vice versa). Only target responses were accepted as correct in the comprehension sections.

Reliability. Inter-scoring reliability was good. For the isiXhosa group, all samples (100%) were scored independently by two researchers and all disagreements were resolved through consensus. Agreement was 94% for noun and verb production and 100% for noun and verb comprehension. For the Afrikaans group, 16 (10%) randomly selected samples were independently scored by a native speaker familiar with the CLT. Agreement was 98% for noun and verb production and 100% for noun and verb comprehension. Internal consistency for both samples was good: Cronbach's alpha, isiXhosa, $\alpha=0.77$; Afrikaans, $\alpha=0.88$.

Background measures: teacher and classroom:

Teacher variables. We used the following metrics as indicators of teacher experience and training: practitioner age, practitioner qualification, years spent teaching, class fees, and estimated size of class.

Classroom variables. Observers rated the quality of language, literacy, and learning activities in each classroom using subscales from versions of the Early Childhood Environment Rating Scale, which was developed to assess curriculum and pedagogy in preschool classrooms using the English curriculum (Sylva et al., 2006). Observations were conducted at Time 1 and Time 2. We used the Time 2 scores as predictors in our analyses as a proxy of the highest quality of support provided to children.

Early Childhood Environment Rating Scale – 3 (ECERS-3). We selected two of the six subscales of the ECERS-3 (Harms et al., 2014): Language and Literacy, and Learning Activities. The Language and Literacy subscale includes ratings of the extent to which practitioners help children expand vocabulary, encourage them to use language, use of books and familiarity with print. The Learning activities subscale includes items to assess support for development of fine motor skills, dramatic play, understanding maths and written numbers.

Early Childhood Environment Rating Scale – Extension (ECERS-E). We assessed classrooms on the Literacy subscale from the ECERS-E (Siraj-Blatchford et al., 2010), which covers: Print in the environment; book and literacy areas; adult reading with the children; sounds in words; emergent writing/mark-making; and talking and listening. To obtain a more culturally-sensitive measure of quality classroom practice, evidence for adults engaging in storytelling was included in the book reading code, because the former is not resource-dependent.

Primary outcome measures: child language

Multilingual Assessment Instrument for Narratives (MAIN). The MAIN (Garagina et al, 2019) was developed as a tool for the assessment of narrative abilities of children aged 4 to 9 years in multilingual populations and from diverse cultural backgrounds and has been validated for use in South Africa (Klop & Visser, 2020). The MAIN comprises four stories with parallel structure, each contains three distinct episodes portrayed across a sequence of six colourful pictures. Each episode includes five components: goal, attempt, outcome, and one internal state terms relating to the initiating event and one internal state term relating to the reaction to the outcome. It was administered at Time 1 and Time 2. Children completed one story at pre-test (Time 1) and one at post-test (Time 2) (Cat, Dog: order counterbalanced within group).

Administration and scoring. Narrative retells were elicited as per the MAIN protocol (Gagarina et al., 2019). Children were first presented with the six picture sequence depicting the whole story, and then asked to tell the story, seeing two pictures (representing one episode of the story) at a time. A story structure complexity score was calculated by categorising each episode for structural complexity: 1 point was awarded for sequences that included no goal, 2 points for incomplete episodes that included a goal but lacked an attempt, and three points for complete episodes (goal, attempt, outcome), resulting in a maximum score of 9 (Maviş et al., 2016). After the retell, comprehension was assessed with 10 open-ended questions that assessed the goals in each episode, understanding of the internal state terms related to the initiating event and reactions, and a question tapping theory of mind. The maximum possible score was 10.

Reliability. Inter-rater reliability for transcribing and coding the MAIN production indicated good reliability. A native speaker for each language retranscribed 26 (10%) of samples. Word level agreement was 96% for isiXhosa and 99% for Afrikaans. All samples were coded by two independent researchers who reached good agreement on structural complexity: isiXhosa =90%; Afrikaans = 91%.

Proximal Vocabulary Assessment (PVA). The PVA was developed to assess learning of vocabulary taught in the programme. It was administered at Time 2 only. Seventeen target words (7 nouns, 7 verbs and 3 adjectives) were identified from the wordlists for each picture book used in the programme. Two words, a noun and a verb,

were used as training items; the other 15 words were used to compile expressive and receptive tasks for each item.

Administration and scoring. The expressive task was administered first. The child was shown one picture at a time and answered the prompt question from the assessor with a word (e.g., *What is that?* for a noun or *What is he/she doing?* for a verb). Pictures from the programme story books were not used. Responses were coded as follows: 3 points for a correct answer, 1 point for a synonym or code switching, 0 points for an incorrect answer (max = 45). After completion of the expressive task, the assessor presented the receptive task. The same target words were presented with three foil images with a similar theme as the target word on one page. This task was only scored if the child provided an incorrect answer in the expressive task (1 or 0 points) and could obtain 1 point for correct and 0 points for an incorrect answer for this task (max = 15). The assessment took 10 minutes per child. The maximum score for the test is 45 points.

Reliability. Inter-rater reliability for transcribing and coding the responses in the expressive task indicated good reliability. 100% of samples were coded by two independent researchers who reached good agreement. Any disagreements were resolved by discussion.

Primary outcome measures: emergent language and literacy

Early Literacy Protocol (ELP). The ELP was designed in the Stellenbosch University Division of Speech, Language and Hearing Therapy to assess print awareness and phonological awareness. It is available in isiXhosa and Afrikaans (as well as English) and has been used in clinics for more than 15 years. We selected tasks that were relevant for this age group. The print awareness items assess print concepts, environmental print, book concept and orientation, and reading orientation. The phonological awareness items assess syllable segmentation, word synthesis, syllable synthesis, and identification of phonemes at the start of words.

Administration and scoring. The phoneme oddity task in the original ELP was similar to item 26 in the ELOM, so we created a 4-item phoneme identification task specifically for this study. All items were verbs because all nouns in isiXhosa have vowel articles (e.g., *ipensile*; *umama*). Recent work has included similar items for Grade 1 children (Wills et al., 2022). Different to our pre-registration, we did not include the 4 items to tap word synthesis in the phonological awareness score, because isiXhosa is an agglutinating language and this may not be a sensitive index of language ability; notably, this group obtain lower scores at Time 2 than Time 1 on word synthesis. Items are scored 1 point for correct (0 for incorrect) and summed to produce a print awareness score (max=10) and a phonological awareness score (max=12). The ELOM was administered and scored according to the clinical protocol.

Emergent Language and Literacy from ELOM. As noted above, the ELOM is a standardised tool suitable for measuring the effects of early learning programmes and children's readiness to learn in children aged 50-69 months (Dawes et al., 2020; Snelling et al., 2019). We used scores from the domain of emergent literacy and language (ELL), which assesses how well children are able to communicate effectively and use language. Items in this domain cover assessment of: ability to speak in full sentences and relate a logical account of events with correct language usage; naming of common objects; understanding of a story that is told to them; and recognition of

sounds in words. This assessment was administered and scored according to the manual.

Analysis plan

When tests are translated and adapted for use in ethnolinguistic samples that are different from those on which the source test was developed, it is necessary to undertake psychometric analyses to establish their conceptual, construct, and metric equivalence to the source (Hambleton & Zenisky, 2010; van de Vijver & Tanzer, 2004). This is lengthy and costly procedure, particularly in a country like South Africa with its many languages and within language variations. At the time this study was conducted, the equivalence of the CLT, the MAIN, and the ELP across languages had not been established. In addition, IRT corrected (standardised) scores were not yet available (Bortolotti et al., 2012) and, in any event, IRT scores are only comparable within (e.g., longitudinally) and not across languages (Leon & Singh, 2017). With these limitations in mind, we decided to use participants' total scores on each instrument for our analyses. The same procedure was used in analyses of receptive vocabulary data from the Young Lives longitudinal study of the development of cognitive skills from five to sixteen years of age in four countries and multiple languages (Tredoux & Dawes, 2018).

Attrition and imputation of missing data

- Comparisons of baseline measures for final sample and attritions
- Explanation of treatment of missing data. Overall, only 3.36% of cells contain missing values. The table below shows where the majority of these missing values are found. The HLE variables contain the most missing values.

We conducted all data analyses using *SPSS*. Our a priori analysis approach began with investigation of distributional characteristics of our variables followed by preliminary analyses to examine classroom-level attrition and child-level non-response (missingness), initial equivalence, and potential cohort differences. We used a research-informed approach to multiple imputation for multilevel data.

RESEARCH QUESTION ONE: Is the Little Stars classroom programme perceived as being useful and feasible for teachers with limited formal training to implement effectively in under-resourced contexts?

To address this Research Question, we investigated whether the teachers perceived the training and classroom programme as useful, valuable, and feasible to effectively implement in under-resourced contexts. Furthermore, we explore the teachers' reports about what they have learned and their perceptions of how the programme has influenced their attitudes and classroom practice. The results are reported under three separate subheadings, below. The feedback reported here is from teachers in the intervention group, comprising 26 teachers, who received training between February and August. Training attendance was excellent, with an average attendance rate of 6.2 out of 7 sessions.

Results

1. Immediate reactions to the training

Teachers rated the facilitation very positively, with 86% considering it excellent.

Over 80% of teachers rated the training as worth their time and highly valuable, indicating their positive perception of its worth.

“There were things that I didn't know. I learnt so many things that I did not know before, now I know better.”

Teachers showed high motivation and preparedness throughout the training, with 90% rating their motivation and 92% rating their preparedness as 6 or 7 on a scale of 1 to 7.

“I've gained a lot, because it was so difficult for me to teach language and now I can't wait for Monday”

“I feel Inspired and motivated about what I have learnt over the past two days.... I feel confident that I will be able to apply it.”

2. Teacher insights and learning

Teacher feedback indicated that the programme offered them more than just training on how to follow the programme. It also provided them with valuable knowledge about how children learn, and how language and early literacy develops. Their feedback revealed a variety of things they learned. Here are some examples:

“... children start learning by pretending to read the big book.”

“I didn't know that young children are able to answer questions the way they do. This encouraged me to engage children more.”

“I learned to have physical resources like the milk in the bottle helps to make the activities more effective instead of only making use of pictures.”

“If the child don't want to talk or answer, wait until feel free to talk.”

“...they learn through seeing and doing, finally touching.”

The importance of **engaging the children for learning** was a recurring theme. Teachers' feedback showed that the program made them realise the importance of making learning fun for children and the value if children are actively involved during learning activities.

“That when a child enjoy something they learn much better”

“I loved the way we were taught how children learn from play and how to build vocabulary.”

Some teachers' feedback reflected a shift in their attitude and the way they intended to **interact** with the children in their classes, as can be seen from the following responses:

“I don't have patience for a child who speaks slow. Now I am able to.”

“To motivate them give them a chance to explore [for] themselves.”
“I always praised the children previously but now I'm doing it much more.”

Finally, while reflecting on their learning many teachers also referred to the importance of preparing for language and literacy teaching:

“I must prepare and be ready before reading my story.”
“You have to prepare yourself by learning the story by practising before presenting to the class.”

3. Putting the programme into practice

In addition to gaining new insights and shifting attitudes, teachers' feedback indicated that they were using what they learnt in training and putting the programme into practice:

- Teachers reported using 70% of the stories provided, and the majority implemented the recommended activities at different time points and the extent to which they implemented activities improved over time. This was in spite of short time frames between training due to time pressure to allow for training for the control group teachers.
- They reported challenges around the little books and roleplay activities in the programme:
 - o Some teachers had a challenge with photocopying the little books and others reported that children found it difficult to fold the little books.
 - o Teachers also mentioned challenges related to roleplay, such as ensuring equal opportunities for all children and the activity taking a long time to complete.

Despite these specific challenges, teachers reported that the children genuinely enjoyed roleplay and acknowledged the benefits of using the little books, noting their positive impact on language development and vocabulary.

- Overall, most teachers perceived the activities to have worked well and reported a growing positive perception of their effectiveness.

“[It was interesting...] the way the children are able to tell the story with the help of the little books. It showed they were listening even though it sometimes doesn't feel like it.”

At three different time points during the study, the teachers were asked to rate how well different activities worked in their classroom on a scale from 1 = **‘very challenging’** to 7 = **‘worked very well’**. Ratings improved as they progressed through the training: at Time 1, an average 49% of teachers gave a rating of 5-7, at Time 2, this increased to 66% and by Time 3, 83% of teacher ratings were between 5 and 7.

- Teachers were impressed with the programme's resources and activities, finding them fun, interesting, and easy to implement. They also saw the value of making their own inexpensive resources:

“What is interesting is to use things not expensive.”

- They highlighted the programme's support for children's learning and language development.

“Role play helps the children with language and words”

Many of the teachers' responses regarding change in their classroom practice, related to their **storytelling** and the way they use stories in the classroom:

- Teachers now planned to use a single story across different activities over an extended period.
- They planned to enhance children's experiences of stories by utilizing materials such as pictures and puppets.
- Teachers intended to actively involve children in storytelling and make learning interesting for them by asking questions before sharing the story, providing opportunities for children to tell their own stories and play characters, and incorporating questions about the story to engage them.

“I could not tell a story before but now I can even tell it using just pictures.”

Significance and Implications

These research findings indicate that the Little Stars training and classroom programme is perceived as highly valuable and feasible for teachers in under-resourced contexts. Teachers demonstrated high motivation, preparedness, and self-reported implementation of the programme's resources and activities. The programme positively influenced teachers' perceptions of language and early literacy teaching, as well as their understanding of how children learn.

The positive outcomes of the Little Stars programme suggest its potential for strengthening teaching practices, particularly in the areas of storytelling, role play, and creating engaging learning experiences. Teachers expressed intentions to be more prepared for teaching activities and to incorporate the programme's strategies and resources into their future teaching, leading to a shift in their approach to language and early literacy instruction as well as to the children in their class.

RESEARCH QUESTION TWO: What is the effect of a resource-based training programme on teachers' teaching practices and on interaction in the classroom?

The trainers visited the 26 teachers in the intervention group in July 2022. As trainers could not observe teachers doing all 10 activities in the two-week cycle, we asked

teachers to do one of the main teacher-led activities: *Storytelling*, *Sequence pictures* or *Reading a Big Book* and a child-led activity: *Drawing and Emergent Writing*.

Using a rating scale of 1–5, the trainers rated the teachers on the questions below. It was encouraging to see that 75% of teachers achieved an overall score of 4–5. Of the remaining 25% teachers, 17% achieved a score of 3, and 8% achieved a score of 1–2.

Observing teaching in the classroom

In addition to the class visits by the Little Stars trainers, independent trained observers visited to observe: the teacher and children, the classroom environment and teaching practices.

The observers used the Early Childhood Environment Rating Scales (ECERS) to guide their observations. We had selected two of the six subscales of the ECERS-3 (*Language and Literacy*, and *Learning Activities*), and one of the four subscales of the ECERS-E (*Literacy*).

For each item, the ECERS guidelines provide indicators that the observer should look for. The observer ticks each indicator they observe and calculates a score for each item on a seven-point scale: inadequate (1–2); minimal (3–4); good (5–6); excellent (7). Some indicators are easy to achieve, while others require more resources and a more skilled teacher.

For example, here are some of the indicators for the ECERS-E item *Adult reading with the children*:

- Adults rarely read to the children.
- Adults reading with children daily.
- There is some involvement of the children during reading times (for example, children are encouraged to join in with repetitive words and phrases in the text, adult shares pictures with the child/ren or asks simple questions).
- Children take an active role during reading times, and the words and/or story are usually discussed.
- Children are encouraged to think about and consider ‘what if’ questions, and/or

As can be seen from the example above, the focus is on storybook reading and oral storytelling, and so we created additional items for ECERS-3 and ECERS-E by replacing ‘reading books’ with ‘telling stories’. Teachers could get credit for either story book reading or oral storytelling.

Results and conclusions

The results from trainer visits and classroom observations indicated the following:

- Teachers from both the Afrikaans and isiXhosa groups scored similarly (there were no significant differences). This suggests that the programme take-up was not specific to one context.

- There was a range of scores on all measures. Only two of the teachers in the intervention group achieved low scores on their use of the programme, which suggests good programme take-up overall.
- On average, the intervention group teachers' quality of teaching improved more than that of teachers in the control group, and on two of three subscales these improvements were significant.
- The intervention group showed a greater increase in the percentage of items that improved by one point on the scale.
- Analysis of items suggested that:
 - Changes in average scores were driven by changes in items related to literacy and language, with some positive shifts in indicators that related to 'maths talk'.
 - Teaching practices did not shift as much as we had expected for *Drawing and emergent writing*.

These findings indicate good take-up of the Little Stars programme in two contexts (Paarl and Khayelitsha). They show that resource-based training is effective in improving teaching practices and interaction in the classroom.

RESEARCH QUESTION 3: What is the effect of the story-based intervention programme on children's early literacy and language skills?

This research question had a number of sub-questions, which are detailed and addressed below.

Research Question 3A: Relative to a control group, does the Little Stars story-based intervention result in gains in children's early development, narrative production and comprehension, and early literacy skills (print concepts and phonological awareness)?

Implementation quality is a measure of the extent to which practitioners in the intervention group implemented the Little Stars programme as intended and as rated by the Wordworks trainer (did the practitioner use the resources and follow the two week cycle; was she prepared and did she follow the steps in the teacher guide; were the children involved in the activity).

Implementation quality significantly predicted improvement in ELOM Total and cognitive and executive functioning (CEF), as well as tests of narrative skill (MAIN) and print awareness. Implementation quality showed a positive but non-significant trend towards predicting improvements in fine motor control and visual-motor co-ordination. This means that in the classrooms where teachers implemented the Little Stars programme well/as intended/effectively, children made the greatest gains in their overall development, and in particular in their cognitive and executive functioning and their fine motor control and visual-motor co-ordination.

In addition to targeting specific language and early literacy skills, the programme targets teacher-child interaction, which is linked to better child outcomes in general. The programme includes weekly drawing and fine motor activities to develop

foundational skills for writing. Within each two-week cycle there is a ‘Learning to listen’ activity which targets auditory processing, working memory, inhibitory control – all skills assessed in tasks in the CEF Domain.

Children’s narrative skills and concepts about print also improved more for children in high quality implementation classrooms than for those in classrooms where the programme was not implemented, or not implemented effectively. The programme did not have a significant effect on phonological awareness and narrative comprehension or the Emergent Language and Literacy domain.

All significant effects were measured for effect size using Cohen’s *d*, and all significant findings had a small to medium effect, suggesting a moderate contribution to performance improvement over time. To identify the exact contribution of the programme to ELOM total score, the beta represented by Implementation Quality was multiplied by the average implementation score performance and by the maximum average imputed value. This produces an estimate of the average effect of the programme and the peak effect of the programme. In this case the average implementation score was 32.63 and the maximum was imputed to be 45.80. This means that the average contribution of the programme was 3.59 ELOM points and the peak contribution was 5.04 ELOM points. This is equivalent to approximately 3 to 5 months of additional age equivalent development to those who benefitted from the intervention.

RESEARCH QUESTION 3B: What programme, teacher and home environment factors predict changes in scores?

The characteristics of the sample (including fee levels and class size), teacher variables (qualification, years of experience), quality of teaching (as measured by an observational rating scale) and the extent to which the teaching programme was implemented as per the guidelines (implementation quality) are outlined above. Also included is a description of the children in the sample and their home learning environment. We were interested in understanding whether any of these variables predicted children’s progress on the ELOM and other language measures. We used a statistical technique known as multi-level modelling and included the following variables as predictors.

Teacher/classroom variables: language of instruction, implementation quality, quality of teaching (as measured by the ECERS scales), fee levels and class size.

Home Learning Environment variables (HLE): the total amount of time that caregivers reported actually spending on 8 types of activities (activity); the total amount of time that caregivers reported having to spend time with their children during the week and during the weekend (time); the total number of books and toys that caregivers reported having in the home (resources).

We were interested in the potential for each of these variables to contribute to children’s progress on measures of narrative, vocabulary, print awareness and

phonological awareness, ELOM Total scores, as well as all domains except Gross Motor Development (the programme did not target this domain). The main findings were as follows.

Children in Afrikaans and isiXhosa classes made similar gains across all assessment tasks, except for ELOM Emergent Numeracy and Maths, and phonological awareness, where the Afrikaans children made greater gains than the isiXhosa children. Afrikaans children also performed significantly better on a test of the vocabulary taught during the programme. The programme did not target Emergent Numeracy & Mathematics (ENM), but ENM includes items that focus on vocabulary for early maths concepts that are language-based. The Afrikaans centres generally charged higher fees and had more experienced and qualified practitioners than isiXhosa centres. The relative gains in maths scores may reflect better maths teaching and more maths related resources and games. Afrikaans children had lower initial scores than isiXhosa children on syllable blending tasks so there was more room for improvement. In the phoneme identification task, the isiXhosa children's scores were lower at pretest and they made limited progress. These differences may partly reflect different language structures. Other studies with children who speak African languages have found similarly low scores on phoneme level tasks, even in children a year or two older than this sample.

For other skills, there was evidence that ECD Centre fees and class size were associated with greater gains. Children in smaller classes made greater gains on ELOM Emergent Language and Literacy and oral narrative skill. The smaller classes may enable more interaction between practitioners and children. Children in centres with lower monthly fees made more progress in emergent numeracy and maths, as well as fine motor control and visual motor co-ordination. This might reflect greater relative benefits in more deprived contexts.

In the classrooms with the best overall quality of teaching (as measured the Literacy subscale of the ECERS-E), children made the greatest gains in overall development (ELOM total), fine motor control & visual motor integration (FMC&VMI) and narrative skill (MAIN). This indicates that a basic level of quality is required for programme to be effective.

The children who had fewer home learning resources made greater gains in overall development (ELOM Total) and fine motor control and visual motor integration (FMC&VMI), possibly reflecting greater relative benefits for children from more deprived contexts. Children whose parents/caregivers reported spending more time with them during the week and at weekends made greater gains in ELOM Emergent Language and Literacy. s

RESEARCH QUESTION 3C: What child characteristics predict change in scores?

The child variables of interest were: age, gender, height for age (as an indicator of child's growth and nutritional status), socio-emotional functioning (SEF), years in the programme, baseline vocabulary scores.

Child gender was not a significant predictor of progress when other variables were taken into account. Children’s age was not a significant predictor of progress on any tasks except comprehension of oral narratives where the younger children improved more.

In order to explore whether children’s baseline vocabulary scores influenced their progress on any of the measures, we analysed the language groups separately, controlling for age. CLT scores were found to be moderately or strongly correlated with the PVT endline score, suggesting that vocabulary at baseline is related to vocabulary at endpoint, regardless of whether children were in the intervention or control group. For the most part, these correlations were strongest for the Afrikaans children and weaker for the isiXhosa children.

		Afrikaans				isiXhosa			
Partial correlations by Language and Group		CLT Noun and Verb Production		CLT Noun and Verb Comprehension		CLT Noun and Verb Production		CLT Noun and Verb Comprehension	
		Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control
PVT Total	Correlation	0.79	0.51	0.74	0.67	0.37	0.43	0.35	0.44
	Significance (2-tailed)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.01	<0.01

Children who had been in an early learning programme for longer scored higher on the vocabulary measure at endline but improved less on ELOM Emergent Language and Literacy. This may be because children who had been in an early learning programme for a shorter period had more to gain.

CONCLUSIONS

The Little Stars programme was effective in improving many aspects of the children’s emergent language and literacy. In addition, children’s progress was influenced by the quality of the teaching offered, and the quality of the children’s Home Learning Environment. These findings have important practical implications for the provision of sound and effective language interventions in ECDs. In particular, the results suggest that the programme is likely to be most effective in contexts where there are also efforts in place to improve teaching quality and, if possible, the children’s learning environment more generally.