ikapadata

helping you understand, fast

INNOVATION EDGE

Thrive by Five Index 2021
Fieldwork Report

(formerly South African Early Years Index)

30 November 2021

Table of Contents

1. Introduction

The cumulative effect of South Africa's investments in early childhood development services – good health care and nutrition, parenting support, high quality early learning programmes, and child safety and protection – can be determined by measuring the proportion of children who are developmentally 'on track' by the age of 5 years.

No national data currently exists to track this important indicator. The South African Thrive by Five Index (formerly the SA Early Years Index) intends to address this gap. The Index will monitor trends over time in the proportion of preschool children who are on track for age in key areas of development. The aim is to use the data to support efforts to ensure that more children receive the nurturing care and services they need to help keep them on track. The Index was initiated as a partnership between Innovation Edge and First National Bank (FNB) and will be launched as a collaborative effort between private, public and non-profit sectors. The anchor government department for the Index is the Department of Basic Education.

The sample consists of a nationally representative selection of children aged 50-59 months who are enrolled in various types of ELPs. For each child, the following developmental domains were assessed:

- Learning (measured using the ELOM 4&5 direct assessment tool)
    - Gross Motor Development: This domain assesses the abilities required to control the large muscles of the body. Ideally, we would want to see children showing good control and coordination in their large movements.
    - Fine Motor Development and Visual Motor Integration: This domain assesses the abilities required to control the small muscles of the body and to coordinate fine motor movements with visual information that is perceived by the eyes. This allows coordinated movements such as copying shapes.
    - Emergent Numeracy and Mathematics: This domain assesses the ability to understand number concepts, symbols, shapes, and size.
    - Cognition and Executive Functioning: This domain assesses the ability to think critically, solve problems, form concepts, attend to instructions, and control impulses.
    - Emergent Literacy and Language: This domain assesses the ability to communicate effectively, use language and demonstrate listening comprehension.

- Social and Emotional Functioning (measured using the ELOM Social and Emotional Functioning Rating Scales): This domain covers self-care, social relations with peers and adults, and emotional readiness for school.
- Height for Age (measured using stadiometers): This domain measures whether a child is at risk of stunting. Stunting is a condition that arises from prolonged under-nutrition and it affects physical and brain development.

The child outcomes data is complemented by baseline assessments (audit) of selected Early Learning Programmes with respect to the type of early learning services provided, staff qualifications and training, resources and funding, quality of the learning environment, as well as access to government support services. Programme quality was measured using the ELOM Programme Quality Assessment Tool and interviews were conducted with principals and practitioners.

Data collection for the Index took place between 20 September and 19 November 2021. This fieldwork report describes field work preparation and management and provides an overview of key challenges.

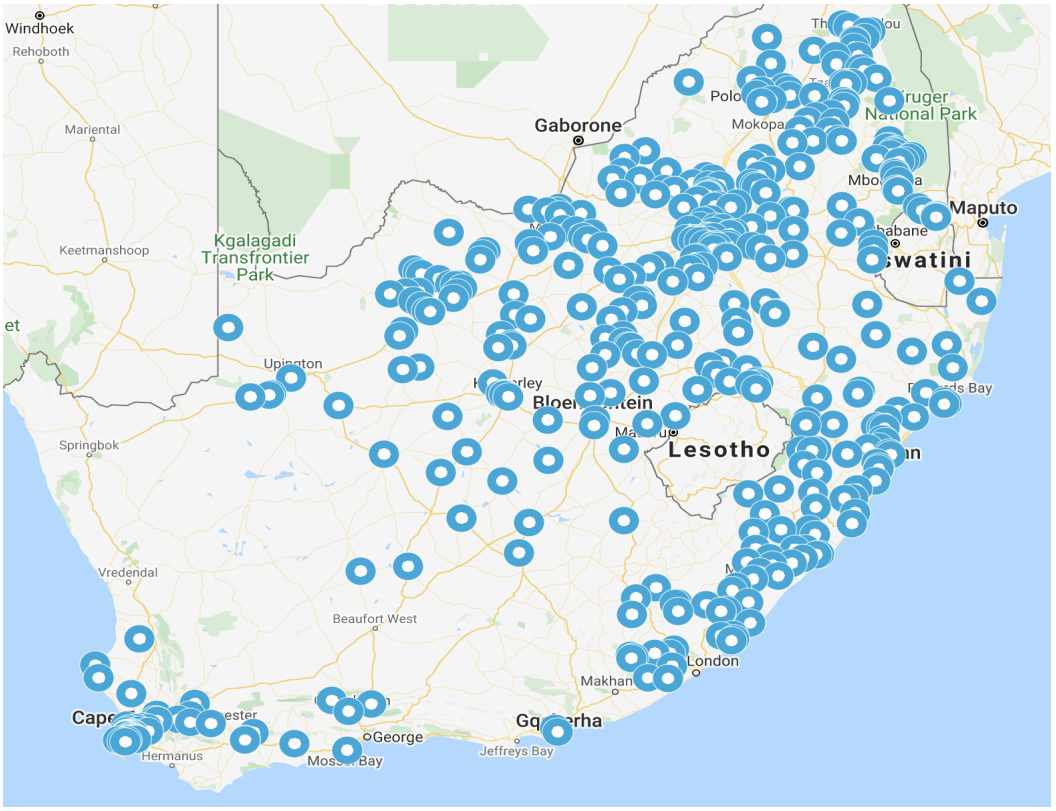
2. Sample

Sampling Strategy

The sample consists of a nationally representative selection of children aged 50-59 months who are enrolled in various types of ELPs (i.e., playgroup, preschool or crèche, day mother programme or in a toy library group). The aim is a representative sample both at a national and provincial level with a confidence interval (CI, "margin of error") of five percent points at a confidence level of 95%.

In the absence of a complete sampling frame for either children or ELPs, it was decided to cluster ELPs via primary schools for random sampling. In each province, 48 schools (432 nationwide) were randomly selected to be used as clusters. Clusters are defined as a 10km radius around a school, or, where this impractical because there were not enough ELPs in the vicinity of the schools, as the ward in which the school is located. The school sample in each province is stratified by socio-economic quintiles (1 to 5). Child assessments (ELOM) were to be conducted at 3 randomly selected ELPs per cluster, resulting in a sample of 144 ELPs per province (1296 nationwide), of which 522 randomly selected ELPs were audited (an average of 1.2 audits per cluster).

Figure 1. Map of Sample School Clusters



Participating children (four per ELP) would be randomly selected from the pool of children who were present on the day when consent forms were distributed and in attendance on the day assessments were undertaken. This means that ideally twelve children per school cluster, 576 children per province, and 5184 children nationwide would participate in the study.

The result of this approach is a multistage cluster sampling design that takes into account what we know about variances and intra-cluster correlations in ELOM scoring from previous assessments while remaining practical and feasible in the absence of a complete sampling frame .

Table 1. Sampling Strategy for SAEYI 2021

Sampling considerations	Per Province	National
Primary schools randomly selected from DBE database (these schools will be used to help identify ELPs)	48	432
Early Learning Programmes to be included in child outcomes sample (3 ELPs randomly selected per school cluster)	144	1296
Randomly selected children to be assessed (4 per ELP - 2 boys and 2 girls who fall within the targeted age band)	576	5184
Early Learning Programmes to be included in the Audit sample (average of 1.25 out of every 3 ELPs selected)	58	522

The sampling strategy allows the Index to disaggregate findings on the proportion of children who are on track for each developmental domain / sub-domain by sex, province and income quintile.

Sampling Frame

To build a sampling frame of suitable ELPs, we made use of the Vangasali dataset of ELPs by identifying all ELPs that are in the same ward as the sampled schools. If there were less than three ELPs in a ward we contacted the sampled school and the ELPs that do exist in the ward to identify additional ELPs in the ward. We combined this list of ELPs with a list prepared in 2020[1] that used the same approach of calling sampled schools and ELPs for referrals. This second list included ELPs that were previously audited in 2020. As a result, we were able to create a list of 4,301 ELPs – 3,185 ELPs from the Vangasali dataset and 1,116 ELPs from the 2020 ELP list. A few duplicates between the two lists did exist but these were minimal.

The next step was to call each ELP on this list in order to introduce the research as well as to confirm both their willingness and eligibility to participate. The aim was to create a sampling frame of a minimum of 4 confirmed ELPs per school cluster, from which 3 ELPs per cluster could be randomly selected for inclusion in the sample, with 1 additional ELP as a backup. ELPs were considered eligible if they (1) operated for more than 8 hours per week, and (2) had at least 6 children aged 50 to 59 months (4 years old) in regular attendance who spoke at least one of the official South African languages as their home language.

[1] Preparations for fieldwork and some audits were undertaken already in 2020 when the study was originally to be rolled out, but Covid-19 interrupted these original plans.

Each ELP was called up to 5 times before they were dropped from the list as being permanently out of reach. ELPs in the call list were ordered by their source such that ELPs sourced via schools or other ELPs appeared at the top of the cluster in randomized order, followed by ELPs sourced from the Vangasali database, also in randomized order. Because the Vangasali database is crowd sourced, the assumption is that many of the ELPs listed may have inaccurate details. The rationale behind this ordering therefore is to give preference to ELPs that were either visited or confirmed in 2020 that are known to exist. In this way, we could also avoid calling duplicate ELPs that were also appearing in the Vangasali database.

In the end we were able to reach 1,655 ELPs, which is just over a third (38%) of the total ELPs on the list. The remaining 2,646 ELPs either had invalid numbers (i.e. number does not exist), or the wrong number (i.e. it did not belong to anyone affiliated with the ELP), or went straight to voicemail each time. From the ELPs that were contacted, 1,013 ELPs (61%) confirmed their eligibility and willingness to participate. The rest of the ELPs either refused (95), or were ineligible (212), or asked to be called back but did not answer (335).

The result was an average of 2.3 ELPs per school cluster. However, these were not evenly distributed across provinces (see Table 2). For example, the Northern Cape started out with a distinct disadvantage of having the fewest ELPs in the calling list (196). Less than 50% of those ELPs were reached and only 59 ELPs were ultimately confirmed for this province. In contrast, the more densely populated provinces of Gauteng and Western Cape had 155 and 167 ELPs confirmed, respectively.

Table 2. ELP Calls Response Rates

Province	ELPs in Call List	ELPs Reached	ELPs Confirmed	Confirmed as % ELPs Reached
Eastern Cape	299	123	79	64%
Free State	216	127	79	62%
Gauteng	1,117	344	155	45%
KwaZulu-Natal	307	160	95	59%
Limpopo	516	216	161	75%
Mpumalanga	438	160	108	68%
Northern Cape	196	83	59	71%
North West	514	173	110	64%
Western Cape	698	269	167	62%
Total	3,641	1,655	1,013	61%

The remainder of the sampling frame was created during the fieldwork preparation phase. Fieldworkers visited each confirmed ELP in a cluster to verify the ELP details, including recording the GPS location, and to distribute ELOM consent forms for children to take home to their parents. If the cluster had fewer than three confirmed ELPs, the fieldworker would request a referral from the ELPs they visited in order to complete the cluster. As a result, an additional 463 confirmed and eligible ELPs were added to the sample.

Even with this strategy, we still faced challenges in the Northern Cape. Some clusters were so remote that there was either no ELP in the area, or only a single ELP serving the entire area. These were mostly farming communities with very large distances between towns or villages. For this reason, eight clusters were replaced with different randomly selected clusters in the same quintile category and eight extra clusters were added in case we faced the same situation in the newly added clusters.

3. Instruments

ELOM 4&5 Direct Assessment

The standardised ELOM 4&5 tool was used to conduct assessments. The tool consists of 23 assessment items in five development domains. The Gross Motor Development domain consists of four items, Fine Motor Coordination and Visual Motor Integration consists of four items, Emergent Numeracy and Mathematics has five items, Cognition and Executive Functioning consists of four items, and Emergent Literacy and Language has six items.

Previously, assessors were required to consult the ELOM Direct Assessment Manual in order to follow the instructions specific to each task. However, a fully digitised version of the ELOM tool was developed that incorporated all instructions as they appear in the manuals. The instructions were also colour coded such that materials the assessor needed for each item was coded in blue, the words that the assessor must say out loud to the child were coded in bold black, the actions the assessor needs to take for each item were coded in italics, and text related to hints or reminders for the assessor were coded in green. Finally, stop rules were hard coded into the tool such that if the child did not succeed on a specific item, the system automatically progressed to the next item. Stop rules were colour coded in red.

The tool also includes questions about the basic details of the child, including the number of years the child has been attending the ELP. However, this question was deemed insufficient to determine the child's exposure to the programme during Covid-19 and two additional questions were added: *(1) When was this child*

enrolled? (2) In the past term, how many times a week on average has this child attended class?

Finally, the translations of the English ELOM into all South African languages were uploaded onto the digital form.

ELOM Socio-Emotional Rating Scale

This tool contains two short rating scales (consisting of 13 items) in which a child's teacher rates a child's social relations and emotional functioning. It is intended to be used alongside the ELOM Direct Assessments to measure aspects of a child's behaviour that cannot be reliably assessed by a stranger in a testing situation.

Children are assessed in terms of their social relations with peers and adults, their emotional readiness for school, and their self-care. Each domain assesses the following skills:

- Social relations with peers and adults: The ability to work well with peers in group activities; the ability to resolve problems with peers without aggression; the ability to cooperate with peers without prompting; the ability to seek support and assistance from familiar adults; the ability to seek information or explanation from familiar adults; and the ability to take initiative in creating cooperative activities with a familiar adult through cooperative activities.
- Emotional readiness for school: Ability to communicate with adults; appropriate expression of needs and feelings; independence and willingness to do things without help; ability to adjust to changes in class or home routine; confidence in new experiences; and initiating activities.
- Self-care: Extent of independent toilet use.

ELP Baseline Assessments

The set of tools used to conduct baseline assessments of ELPs consisted of four instruments: Principal Interview, Practitioner Interview, Quality Assessment, Environmental Observation. Several changes were made to the instruments since they were first used in 2020 when we commenced fieldwork for the baseline before fieldwork was stopped by the pandemic.

Principal Interview. Apart from the basic details of the ELP, the Principal Interview covers questions about the operational aspects of the programme. Several sections of the tool had substantial changes, particularly questions on programme registration with DSD, as an NPO, or as a private business. Based on the previous assessments conducted in 2020, we found that most ELP staff did not necessarily understand or perceive a difference between programme registration and registration as a partial care facility. They could, however, confidently say whether

they were registered with DSD or not. Changes to the tool made it easier for principals to respond to questions about registration without requiring the assessor to go into lengthy explanations of the DSD registration process. This also enhanced data quality as the questions were presented more intuitively.

Data quality was further bolstered by incorporating check questions for every section that required counting either numbers of staff or children. For example, when asked how many of the ELP's registered children are male vs female, these two values would then be summed in the backend, and then compared with the figure the principal provided for the total number of children registered with the ELP. If these two values did not match up, the assessor would receive an alert instructing them to go back and check the principal's responses and either correct the total number of registered children, the total number of girls, or the total number of boys registered with the ELP. The assessor would not be allowed to proceed to the next question until these responses were corrected and these two values matched up.

Additional sets of questions were also added about the programme's operating hours as well as the type of services being offered. Because the majority of questions in the Observation and Quality tools were geared towards conventional ELPs, the questions around type of services were used to screen out mobile ELPs, pure toy libraries, and day mothers (programmes located in a home for fewer than 6 children).

Finally, a new set of questions related to the effects of Covid-19 on staffing, salaries, and children attendance were added as well.

Practitioner Interview. This instrument was used for an interview with one of the ECD practitioners working at the ELP (preferably the practitioner whose class was observed, see below). It consists of questions about the practitioner's demographics, qualifications, educational background, teaching methods and attitudes toward "Learning through Play".

The main changes to the 2020 tool involved removing questions that were considered a better fit for either the Principal Interview or the Observation tool. For example, questions on the availability of learning and teaching support materials were moved to the Observation instrument. A set of questions assessing practitioners' pedagogic beliefs about child agency and free play were also incorporated. These "Learning Through Play" questions were adapted from the National ECD Census instruments.

Environment Observation. This instrument was used to capture observations about the infrastructure (space, electricity, heating, sanitation etc) of the ELP, the availability of toys and teaching materials, and the outdoor space.

While the overall sections of the tool remained the same as the one used in 2020, a few questions were added to several sections in order to add more detail. For example, a question on the ELP's main source of power and energy were split into three questions differentiating between sources for heating, cooking, and lighting, respectively. Similarly, questions about access to the ELP were expanded from a single question about the presence of a fence around the ELP's premises to adding two more questions about the presence of a lockable gate and the presence of security or someone checking who enters the facility.

Programme Quality Assessment. The Quality Assessment tool consists of 22 items and 5 sub-scales: Learning Environment (6 items), Learning Assessments (2 items), Relationships and Interactions (4 items), Curriculum (5 items), and Teaching Strategies (5 items). The tool originally consisted of 23 items. Data from the 2020 assessments was used to conduct Exploratory Factor Analysis to determine whether the questions in the tool show internal validity and reliably measured the same construct of "quality". As a result, only one question from the Curriculum sub-scale – Children are familiar with the daily routine – was dropped from the tool as it displayed low levels of internal consistency with the rest of the items in the tool.

4. Recruitment and Training

Training for the Index 2021 took place in three different locations – Cape Town, Bloemfontein, and Johannesburg. Separate training workshops were held for the ELOM and ELP Baseline assessments over a period of three weeks in August and September.

Training for the ELP baseline assessments was held in Cape Town on 30 August – 2 September. A total of 21 candidates were invited to training and 19 candidates attended. The workshop was led by Elizabeth Girdwood and Linda Biersteker from Innovation Edge. Sonja Giese from Innovation Edge, Janeli Kotze from Department of Basic Education, and Liezl Worship (independent trainer) provided training support, alongside six ikapadata staff members including a Senior Research Consultant (Lameez Alexander), a Research Consultant (Thabisile Seme), and four Research Assistants (Bulelani Gwampi, Ornella Kasonga, Mahlatse Maeko, Thandeka Mcemani).

The fieldworker training for the ELP Baseline Assessments consisted of a four-day training workshop. The aim of the training was to equip the fieldworkers with the necessary skills and know-how to successfully assess an Early Learning Programme. The training also served a dual purpose of final piloting of the instruments and research protocols. On the first day, emphasis was placed on the Quality Assessment, Principal Interview and Practitioner Interview instruments. These instruments are complex and require considerably more time to administer than the

Environmental Observation instrument. The next two days were spent visiting Early Learning Programmes in the Cape Town Central Metro area. The primary objectives of the ELP visits were threefold:

1) Ensure that all fieldworkers left the training with a similar understanding and reference point of what a high performing ELP should look like, regardless of individual differences in previous ELP knowledge and experience.
2) Strengthen fieldworkers' understanding of the nuanced differences in criteria when rating the quality of the ELPs activities as inadequate, basic, or good.
3) Prepare fieldworkers for what they could expect in the field and practical tips for assessing a class of young children without being overly intrusive (e.g. sitting down on the floor rather than standing up or hovering).

For the ELOM assessment training, candidates were split across three locations. Fourteen accredited assessors from Western and Eastern Cape attended training in Cape Town on 25-26 September. In Bloemfontein, 13 new candidate assessors were invited to training on 5-11 September – six from Free State, five from Northern Cape, and two from Eastern Cape. An additional three accredited assessors attended training on 10-12 September. The largest training was held in Johannesburg, where 15 new candidate assessors attended on 5-11 September and 17 accredited assessors attended on 10-13 September.

The ELOM trainings were led by several ELOM facilitators from Innovation Edge: Elizabeth Girdwood and Liezl Worship in Cape Town; Pumla Mkhiva and Ernestina Letsela and Marieta Visser in Bloemfontein; Pearl Tom, Miranda Rikhotso, Mphokazi Malangeni, and Paphama Mateza in Johannesburg. Elizabeth Girdwood was also present at the new candidate training sessions in Bloemfontein and Johannesburg.

On the first day of the new ELOM assessors training, the recruits were introduced to the ELOM Tool Kits as well as the ELOM Direct Assessment Manuals. Recruits spent the majority of the afternoon working in pairs to practise the administration of each item in the ELOM assessment. The day ended with a demonstration of Covid-19 protocols to be followed when conducting ELOM assessments. For the remaining four days of training, recruits conducted practical assessments at selected training ELPs in close proximity to the training venues. On the final day of training, recruits completed the ELOM interrater reliability test. The digital test consisted of 88 questions, each one accompanied by a short video clip demonstrating a particular assessment item. Recruits were required to score the child's performance in the video clip. These scores were then compared to the correct score for each item. Recruits were expected to achieve 90% or higher in order to pass the test. The test results, combined with the ELOM facilitators' observations of assessors during the practical assessments, were used to evaluate whether the recruit was ready to conduct assessments on their own. The final decisions were made by Linda Biersteker and Elizabeth Girdwood from Innovation Edge. Only one out of a total of

28 new assessors did not meet the minimum requirements to pass the test and was therefore not accredited.

For the one-day accredited ELOM assessors training, the Digital ELOM Refresher – an online training platform newly developed by Innovation Edge – was used to train accredited assessors to help them become reacquainted with the tool and refresh their memories of what the ELOM assessment entailed. The online platform consists of seven modules – an introductory module, five content modules (one per ELOM domain), and the final interrater reliability test (the same 88 questions that were used for new recruits). Each content module was accompanied by a video demonstrating the administration of assessment items associated with the particular ELOM domain as well as a short quiz. Before taking the final interrater reliability test, accredited assessors were given the opportunity to work in pairs and practice the administration of each item in the ELOM assessment.

Due to the Covid-19 pandemic, a few accredited assessors lacked recent experience of administering ELOM assessments. For these twelve candidates (nine in Johannesburg and three in Bloemfontein), an additional morning was spent at a training ELP in order to conduct practice assessments with children.

Each training was combined with a session on fieldwork logistics led by the ikapadata team. During this session, the candidates were provided with a short overview of the project's purpose and sampling strategy, procedures for confirming visits with ELP principals or administrators, the use of SurveyCTO on the tablets, and Covid-19 protocols to be followed at all times when in the field.

A final team of 18 ELP baseline assessors and the strongest 54 accredited ELOM assessors were selected for fieldwork based on their performance during the training. These 72 individuals were assigned to 18 teams (two per province), each consisting of one ELP baseline assessor and three accredited ELOM assessors.

5. Fieldwork

Preparation & Scheduling

Pre-visit ELP Calls. Eight call agents working remotely from their homes contacted ELPs in the sampling list between 7 July and 26 August. Call agents were provided with a detailed calling script outlining exactly what to say to introduce the research to the principals. They also participated in a one-day training, which included scripted mock interviews with ikapadata research assistants. All candidates were required to submit their completed mock interviews and their responses were compared to the correct interview responses. Candidates were then given additional training on any questions where they entered the incorrect response.

Finally, a list of Frequently Asked Questions were created to help them address any issues or concerns raised by principals that are not covered in the calling script.

All principals were interviewed in their preferred spoken language. The interview instrument was coded such that the call agents' language was matched with that of the person being interviewed. The call agent could only proceed with the call if there was a match. Calls were recorded and a random selection of calls were checked each week for data quality purposes. Principals of all confirmed and eligible ELPs were then sent an automated email thanking them for their participation along with a copy of the official letter from the National Department of Basic Education explaining the research (translated in 11 official languages). Principals could also request that the letter be emailed to them before proceeding with the interview. The call centre confirmed a total of 1,301 eligible ELPs.

Consent Forms Distribution. 18 fieldworkers (2 per province) were assigned to visit 24 clusters each between 2-27 August. Initially, fieldworkers were required to complete 1.5 to 2 clusters per day. However, because of lengthy distances between some clusters, another 2 weeks of fieldwork were added from 6-17 September. In addition, some ELPs initially confirmed over the phone had to be replaced either because of refusals or because one ELP was quite far from the other two ELPs in the cluster. There were also some duplicate ELPs with different names that were initially confirmed over phone and then confirmed in field (in this case the in-field record was retained).

Each principal was informed that assessors would be visiting them between September and October to assess the children as well as to conduct interviews with the principal and practitioner. The fieldworkers also verified the ELP details (principal details, address, phone numbers) as well as the number of four and five year old children registered with the ELP. Sufficient consent forms were then provided for all the four year old children at the ELP.

At the end of this period, 1341 ELPs were visited and confirmed eligible with an average of 3 ELPs per cluster. Because some clusters in the Northern Cape needed to be replaced, consent form distribution in this province continued into the fieldwork period. The final sample consisted of 1367 confirmed ELPs.

Table 3. Final Sample of Confirmed ELPs

Province	Total Confirmed ELPS
Eastern Cape	157
Free State	147
Gauteng	143
KwaZulu-Natal	145
Limpopo	165
Mpumalanga	151
Northern Cape	159
North West	152
Western Cape	148
Total	1,367

In-field Sampling

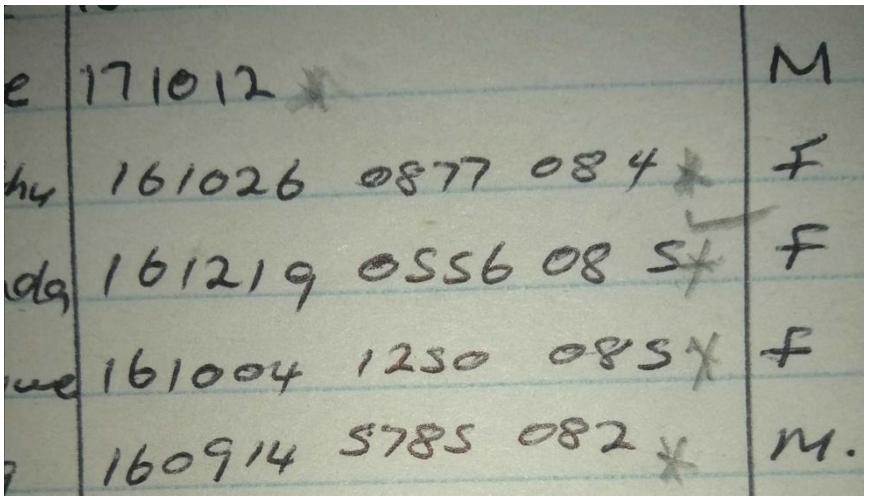
ELOM assessors were responsible for selecting the classroom from which children would be randomly selected for the assessments. ELP Baseline assessors were then required to conduct Quality assessments and interview the practitioner of the same classroom. For classroom selection, assessors were told to select the classroom with the largest number of 4 year old children (i.e. the most 4-year old boys and girls). This was to ensure that there were sufficient eligible children for the assessments, which is something we could not have controlled had the selection of classrooms been completely randomised.

Once the classroom was selected, assessors completed a sampling form for the random selection of children in that classroom. In order to complete the form, assessors were required to follow a series of steps. First, they had to request the classroom register with the children's names and birth dates listed. Next, they had to go down the list and tick children with any severe impairments[2], children whose parents had refused consent, and children whose home language did not match the language in which they were trained to conduct assessments. These children would be excluded from the assessments. Now they were ready to complete the sampling form. In the form itself, they were provided with two cut off birth dates. All children born before and after these cut off dates were to be excluded as well. Finally,

[2] The following impairments prevent a child from being eligible for an ELOM assessment: eye sight, hearing, difficulty understanding instructions, difficulty in walking or moving their arms or legs.

children who were not present on the day were also excluded. The remaining children on the list were considered eligible and used for random selection.

Figure 2. Example of Classroom Register



The sampling form was configured to randomly select two boys and two girls (or three girls and one boy if there was only one boy present, etc.). Assessors were required to enter the total number of eligible boys and girls present on the day of assessment. Thereafter, they would assign each eligible girl a number, starting with the number one, up to and including the total number of eligible girls present. The form would then display two randomly selected numbers between one and the total number of eligible girls present. The assessor would then select the two girls with the corresponding numbers. The process was then repeated for boys.

Fieldwork Protocols

Team leaders were instructed to plan their visits with ELP principals up to one week ahead of time, typically on a Thursday or Friday afternoon after completion of fieldwork. They would inform the principal of the name of the assessor that would be visiting their ELP as well as request that the principal prepare any documentation related to the ELP's registration, staff records, and children's registers. The principals were called again the day before the scheduled visit to confirm if the ELP was still available, as well as to check the ELP's operating hours, and the daily schedule of the 4-year old children (in preparation for the assessments). If the ELP was no longer available, the team leader was required to reschedule or replace the ELP with another ELP close by (where possible).

The teams were instructed to arrive at each ELP between 7.30am and 7.45am, and to begin working from 8am. Teams mostly managed to always arrive on time, however the start time varied and was often delayed due to children arriving late at school (sometimes as late as 10am). In such cases, assessors were instructed to assess any children that were eligible and were already present to ensure that they

could complete four assessments before children left the ELP for the day (typically between 12-1pm). Wherever possible, assessments took place in a quiet space away from the other children. The setup included a seat for the child and the assessor, some form of table, and the ELOM kit. Children were introduced to the assessor who explained the purpose of the ELOM assessment to the child and provided assurance that this was not a school test and that the child could stop the assessment at any time, before the child was asked for its assent to participate in the study.

For the ELP Baseline Assessments, assessors were instructed to start with the Quality assessment as soon as possible because it required a minimum of 120 minutes of observation of children's activities in and outside of the classroom. The Principal and Practitioner interviews were to be scheduled around their respective availability for the day. The Observation form would then be completed at any point in the day when neither interviews or quality observations were possible (e.g. during children's breakfast). All forms were to be completed to indicate cases where the principal or practitioner may have been unavailable so we could track which ELPs we would need to return to during mop ups.

Figure 3. Makeshift ELOM Assessment Setup

Covid-19 Protocols

Assessors were expected to adhere to the following general precautions when in field:

1. Cloth masks are to be worn at all times covering both the nose and the mouth
2. Hands must be sanitised before entering the ELP premises
3. Social distancing to be practised by maintaining a 1.5m distance from any other person (both staff and children) at all times

Covid-19 protocols specific to conducting ELOM assessments included:

1. Assessments were to be conducted in a well ventilated space where the door or windows could be left ajar
2. Assessors must sanitise their hands before and after each assessment, as well as upon departure from the ELP
3. Children were provided with plastic face shields to be used throughout the assessment over the cloth masks (it was permitted for them to use only cloth masks when doing gross motor activities such as tossing bean bag task)
4. All kit items, the child's plastic face shield, as well as any tables and chairs used during the assessment, were to be sanitised between assessments
5. Should the child present with any symptoms (e.g. coughing) during the assessment, the assessor was to stop the assessment and inform the child's practitioner

In addition, all assessors were provided with their own Personal Protective Equipment that they were expected to use during fieldwork. This included:

1. Adult Face Shield
2. Child Face Shield (for use during ELOM assessments)
3. Sanitiser (for hands and ELOM kit items) and microfiber cloth (for wiping down desk / tables used during ELOM assessments)

Because assessors were going to be working in teams of four and sharing the same vehicle on a daily basis, the following additional Covid-19 protocols were emphasised:

1. Before heading to the ELP, a temperature check must be conducted (team drivers were provided with a handheld thermometer); if the person showed signs of a fever, they were to visit their nearest Covid testing facility for a rapid test; if the test result was positive, they were to return home immediately and self-isolate.
2. Each fieldworker completed a daily "Health Check" form that they were required to submit for screening purposes and would alert office staff to individuals with any positive symptoms.

3. A face mask was to be worn in the vehicle at all times, and hands were to be sanitised before entering the vehicle in the morning and in the afternoons.
4. If a person felt ill or showed any symptoms prior to heading to field in the morning, these were to be reported immediately and the person would stay out of field until they'de recovered; a negative rapid Covid test result would be required before they would be allowed to re-join their team.

No fieldworker contracted the COVID-19 virus during the course of fieldwork and no known cases were reported subsequent to their visits at any of the participating ELPs.

Data Collection & Mop-Ups

The teams were required to visit one school cluster a day and complete the baseline assessment at one ELP as well as four child assessments at each of the three ELPs in the cluster. With two teams and 48 school clusters per province, the minimum fieldwork period is 24 days or five weeks, plus one extra week to account for the October school holidays when ELPs are either closed or children are less likely to attend. An additional two weeks was then added for teams to return to any incomplete clusters with missing child assessments or baseline assessment forms.

After 6 weeks in the field, we had completed 396 out of 522 Baseline Assessments (76%) and 4,307 out of 5,184 valid ELOM assessments (83%). The majority of provinces had completed 40 or more Baseline Assessments, with NW and WC having completed just over 50 each. Eastern Cape, Northern Cape, and Kwa-Zulu Natal were trailing behind with 38, 39, and 35, completed assessments respectively. At this point, each province had also collected more than 450 ELOM assessments, except for the Eastern Cape with just over 400 assessments.

Limpopo was the only province who had completed the target 576 assessments ahead of time. This was mainly because the teams were not hampered by the October school holidays and could continue collecting data as usual. In the rest of the provinces, only those team members who were conducting baseline assessments remained in the field. The Nelspruit team in Mpumalanga did not go to field at all due to ELP closures during the school holidays.

The Mpumalanga Emalahleni team was the next team to complete their target of 576 assessments. Unfortunately, the isiZulu ELOM form had an error in one of the tasks. Item 23 instructs the child to say out loud the word, from a list of three words provided, starting with a particular letter of the alphabet. The assessor then selects the word in the form that the child says. In the isiZulu form, the list of words displayed for the letter "D" was incorrect in the question itself, but correct in the list of words the assessor needed to select in order to score the child's response. This led the Emalahleni team to provide the incorrect list of words to the child. They did, however, select the correct word from the choice options when the child

responded with the word that started with a "D", despite that word being different from the word in the choice option.

This error was only detected on 20 October and affected 189 ELOM assessments. The form was immediately corrected for fieldwork the next day. The incorrect words in the isiZulu form were in fact the same words used for the two practice rounds before the child is actually scored, so it is possible that the incorrect words provided in the isiZulu form are of equal quality or difficulty as the words in the correct translation. This error would have to be taken into account when conducting analysis and a test to check for statistically significant differences between the scores of this team before and after the translation was corrected. As a result, the Mpumalanga teams remained in the field in order to conduct extra assessments that could possibly be used as replacements if needed.

For the final two weeks of mop ups, some fieldwork teams were reassigned from provinces that were further ahead to provinces that were falling behind. One team of four ELOM assessors were relocated to the Northern Cape, and one team of three ELOM assessors and one ELP baseline assessor was relocated to the Eastern Cape. Unfortunately, there were no surplus isiZulu speaking assessors that we could relocate to Kwa-Zulu Natal.

Table 4. Final SAEYI Sample

Province	Valid ELOM Assessments	Completed Baseline Assessments[3]
Eastern Cape	599	62 (0)
Free State	576	58 (2)
Gauteng	570	61 (7)
KwaZulu-Natal	568	56 (7)
Limpopo	574	56 (3)
Mpumalanga	692	58 (2)
Northern Cape	601	61 (9)
North West	596	62 (9)
Western Cape	559	60 (1)
Total	5,335	534 (40)

[3] Figures in brackets for additional "incomplete" baseline assessments, for which at least one of the forms were not completed.

6. Data Capture and Cleaning

Data was submitted in real-time to SurveyCTO servers. These submissions were downloaded every ten minutes and then exported to the SAEYI Project Base in Airtable where data entry corrections were made. Two mechanisms of data quality control were utilised during data entry and after form submission.

First, entry errors were minimized by coding confirmation questions directly into the forms, particularly for questions that required entering numbers of staff or children. For example, numbers of girls and boys registered with the ELP would be cross-checked with the value previously entered for the total number of children registered with the ELP. If there was a mismatch, the assessor would be alerted to the error (with all entered responses clearly displayed) and then required to go back and correct their responses for either the total number of children, the number of girls, or the number of boys registered (or all three, depending on which responses were incorrect). The assessor would only be allowed to proceed with the form once these values matched. Similarly, constraints were placed on specific numeric questions (e.g. total number of household members), as well as multiple choice questions (e.g. "None" could not be selected if any other option was also selected).

Second, after forms were submitted and exported to Airtable, several check variables were used to flag quality control issues in both ELOM and Baseline Assessment forms. As soon as a submission was received, the relevant assessor would then be sent an automated message in Slack detailing the name of the ELP, the relevant form that was submitted, the specific issue with the submission, and in the case of ELOM assessors, the name of the child as well. These queries were then dealt with directly and resolved while the assessors were in field so that data could be corrected as quickly as possible (see Figure 2). Below is a more detailed outline of the quality checks used for the ELOM and Baseline Assessments.

ELOM Quality Checks

Data quality control messages were sent to Slack for the following ELOM check variables.

Child's Age. When the age of the child was outside of the expected ELOM age range, the assessor was asked to confirm, in the form itself, whether the date of birth entered was correct or whether they had made a mistake. If they selected that it was not correct, they were then prompted to go back and correct the child's birth date. Because the form allows you to proceed even if the child is outside of the age range (there are various reasons why this would be permitted but, to give one example, the child could be 49 months on the day of the assessment and be turning 50 months the very next day), a message was then sent via Slack to the assessor

requiring them to explain why the child was assessed. While the assessment would automatically be coded as "invalid" (and therefore excluded from counting towards the sample), the client could go back and review these submissions for final inclusion or exclusion. In total, 107 submissions were flagged as invalid due to age.

Disability screening. There are five questions in the ELOM form measuring whether a child has any severe physical impairments. These questions appear at the end of the actual assessment. However, assessors are required to check with the practitioner before selecting a child if the child has any known impairments. Should the assessor observe any impairments while conducting the assessment, they were required to end the assessment early. All submissions where any screening question was scored "yes" would be flagged and checked with the assessor if this was indeed the case or if it was an error. Because submissions for children with physical impairments are automatically coded as "invalid", we were able to catch any completed assessments that would otherwise be excluded. In total, twelve submissions were flagged as invalid due to screening.

Duration. Assessments that took longer than 75 minutes or shorter than 25 minutes were also queried. In the majority of queries the assessments were shorter than 25 minutes and the main reason provided by the assessors was that the child either struggled and many items were skipped, or in other cases, the child did really well and the assessment went faster than normal. There were only nine cases where the assessments took longer than 75 minutes, eight of which occurred during the first week of fieldwork. In the case of one assessor in KZN, the field manager had noticed that the assessor was receiving more of these alerts than his peers in the first two weeks and immediately contacted him. The assessor then explained that, due to technical challenges with his device and because he did not want to lose time, he had been using the ELOM Direct Assessment Manual to conduct the assessments, recorded the scores on his phone, and then entered the data into the digital form later on when his device was fully functional. He was reminded that manual capturing of data was not permitted under any circumstances and that his contract would be terminated should he continue to do so. In total, 32 submissions were flagged as invalid because they were shorter than 25 minutes.

ELOM Socio-Emotional Rating Scale. Missing socio-emotional ratings by ECD practitioners were also flagged. The main reason was simply that the practitioner was not available. In some cases the alert was triggered because an assessor submitted the ELOM assessment and teacher interview in two separate forms, for example if the ECD practitioner was only available at a late stage, and selected the wrong child in the form when completing the socio-emotional rating questions. This error could then be corrected. In total, 442 submissions are without a socio-emotional rating.

Baseline Assessment ELP Quality Checks

The majority of the quality control checks for the Baseline Assessment instruments were coded directly into the form. However, a few standard checks and checks that were not possible to incorporate directly were conducted after form submission.

Principal Interview. Interviews with durations shorter than 45 minutes and ELPs with more than 200 children were queried. The latter was rare, but the former did occur more frequently with the main reason being that the principals were fully prepared and had all the relevant information readily available. We also checked any submissions where the ELP's registered capacity (as per their DSD registration) and their actual number of registered children did not match. While this is more often the reality than not, it does allow us to check for any potential errors. For example, one assessor would enter (1) for the ELP's registered capacity instead of the required (-1) in cases where the capacity was unknown or the ELP was not registered. It took several alerts and a gentle reminder that she would continue to receive these alerts until she started entering the correct response of (-1) for "Don't know". The alerts stopped shortly thereafter.

Practitioner Interview. Interviews shorter than 15 minutes were queried (only 33 submissions). Specific variables that were checked included household sizes greater than 10 people and measured class size greater than $80m^2$ or smaller than $10m^2$. The larger class size occurred more frequently as many ELPs are located in community centres or halls with no clear demarcations for classroom boundaries.

Quality Assessment. Assessors were required to observe children's activities for a minimum of two hours. Submissions with durations of less than 120 minutes were flagged. This quality check occurred more at the start of fieldwork because assessors would take notes and then transfer the scores only once they were done. While note taking was encouraged (it's difficult to keep track of each observation pertaining to all the different sub-sections without taking notes), they were also required to leave the form open on the device so they could enter and update scores throughout the observation period. Very low scores for each sub-section of the tool were also queried. However, these did not raise any red flags since it was highly possible for individual ELPs to have low scores.

Observation Tool. Assessors could complete the Observation form more independently of ELP staff than the rest of the forms, which also meant they could complete it fairly quickly. Here, only two variables were queried: when an assessor indicated that the ELP had no toilet or no learning and teaching materials, both of which would be highly unusual. There was one case in Eastern Cape of an ELP having no toilet. The ELP was located in a very small shack.

Figure 4. ELP in Eastern Cape

Data Dashboard

A dashboard detailing the ongoing progress of data collection as well as an overview of all data quality checks was available to all relevant stakeholders. The dashboard visualised real time data summarised by province and team. In this way, suspicious patterns of scoring between teams could be easily detected. It was noted that the teams in Mpumalanga were scoring children relatively high for the ELOM assessments. The decision was then made to re-train the teams to ensure that they had a uniform understanding of how to score different tasks. An ikapadata field manager supervised both training sessions. Training occurred at a selected ELP and assessors were required to assess one child each. All assessors would then score each child and submit their scores. In this way, we could compare scores for the same child between assessors. While the teams benefited greatly from this re-training, there was nothing particularly concerning detected in their scores. The overall scores for Mpumalanga remained relatively high compared to other provinces. Data collected by ikapadata for another ELOM project in Mpumalanga, Limpopo, and Northern Cape showed a similar trend.

7. Challenges & Future Recommendations

Availability of Eligible Children

The irregular attendance of 4-year old children posed a significant challenge that was encountered right from the start to the very end of fieldwork. In the first week, only a few assessors reported that there were fewer than four eligible 4-year old

children to assess. For example, one assessor in the Northern Cape reported that the ELP had 35 registered children, 13 were present on the day of the visit, only three of those were eligible 4-year old children, and the rest were 5-year olds. Another assessor in KZN reported that out of the five eligible 4-year olds registered, three children were absent on the day. Unfortunately, these reports became increasingly more frequent over time. There were several times when staff would call parents to bring their children to the centre, or would even collect the children from their homes themselves, simply for the sake of conducting the assessments. The rainy season in KZN and Eastern Cape exacerbated the problem. The only provinces where irregular attendance rates were not an issue were Limpopo and Mpumalanga.

Recommendation: Consider extending the age range for the Index to include 5-year old children.

Covid-19

The irregular attendance rates were at least partly due to the impact Covid-19 had on the ECD sector in South Africa, but this was not the only challenge posed by the pandemic. Closures of ELPs during lockdown and hesitancy of ECD practitioners to let assessors into their ELPs during a pandemic made sampling probably more difficult as it might have been before Covid-19 as we struggled to reach sufficient numbers of ELPs and learners. Covid-19 protocols and the occasional absence of assessors in the field due to positive Covid screenings also slowed down fieldwork. Additionally, it remains unclear what the impact of masks and sanitising might have been on children's comfort levels and performance during the assessment.

School Calendar & Public Holidays

While the availability of eligible children was a challenge in its own right, attendance was significantly affected by the holidays during September and October. Many children do not attend ELPs during school holidays (4-8 October) since they could stay home with older siblings. Only three out of 18 fieldwork teams continued to conduct assessments during school holidays. Two teams in Limpopo were able to carry on as normal as most ELPs were fully operational with regular attendance. The third team in the Free State was less successful. There were also two public holidays during this period: Heritage Day in the first week and National Election Day in the second last week. For both these holidays, ELPs would either close or children would simply stay at home on the day and for a day before or after the holiday itself. Specifically, lower attendance was observed on the Friday before National Election Day and on the Friday after Heritage Day. Some ELPs had special Heritage Day celebrations the day before the public holiday, which also made it difficult to conduct fieldwork. Finally, towards the end of fieldwork many ELPs were busy preparing for graduation and end of year ceremonies, which meant that children and staff were not involved in regular learning activities. This made it particularly

difficult to conduct the ELP Quality assessment as this requires careful observation of a regular day of activities.

Recommendation: While it was not possible to do so this year due to Covid-19, starting fieldwork in the first week of September would help to avoid the busy end of year period (Nov-Dec) and allow for extra time lost during the week of school holidays.

Government Communication with Sector

A few assessors reported encountering difficulties with gaining access to ELPs because principals were instructed by local government officials to prevent ikapadata from collecting data. In all of these cases it was because the local official was not informed of the research and therefore did not deem it safe for principals to permit access. In two other cases, it was the ECD forum that the ELP belonged to that had advised the principal to refuse access because they had not been informed by their local government official, who in turn was also not aware. One ECD forum went as far as threatening to withdraw support for the ELP should they go against their advice and decide to participate. While the problem itself was not highly prevalent – it was limited to a few instances in Gauteng, Free State, and Northern Cape and most times it was resolved with a phone call or an email to the local official – it does suggest that important stakeholders felt excluded from the process and were willing to take drastic action as a result.

Recommendation: Create and implement a systematic framework for disseminating information about the Index, and specifically relevant data collection efforts, to local government officials, ECD Fora, and other "on-the-ground" stakeholders in the sector to increase awareness, foster inclusion, and promote goodwill.

Overlap of National ECD Census

This year, ikapadata has had the privilege and the honour of being responsible for managing data collection for both the Index and the National ECD Census. Both these projects are significant endeavours involving hundreds of fieldworkers and thousands of ELPs across South Africa. It was therefore inevitable that fieldworkers from both projects would end up visiting the same ELP. This caused confusion, and sometimes frustration, on the part of principals as they were not sure why they needed to answer similar if not the same questions twice (there were significant similarities between questions in the Principal Interviews for SAEYI and the Census). As a result, a few principals refused to repeat the interview. However, the problem was not significant or widespread since most principals were extremely gracious and generous with their time.

Recommendation: Consider a combined communication strategy for all related data collection activities in the sector.