



South Africa 2024 Thrive by Five Index Non-Enrolled Enumeration Area Selection

Gunilla Pettersson Gelande, Sonja Giese, Servaas van der Berg, Caylee Cook

Selection of enumeration areas (EAs) for the non-enrolled sample

Background

For the 2024 round of the Thrive by Five Index, the non-enrolled child sample will be a ‘test’ sample to explore the feasibility of reaching non-enrolled children, and it will be notably smaller than the enrolled sample.

The target population for the non-enrolled sample are children aged 50-59 months not currently enrolled in any early learning programme (ELP) who come from the lower end of the socio-economic spectrum.

What we want to be able to do

Compare the key child development outcomes measured by the Index: early learning (proportion on track/average ELOM scores), social and emotional functioning (proportion who meet the standard/average scores) and physical growth (proportion stunted) for non-enrolled and enrolled children aged 50-59 months.

Ideally, the non-enrolled sample would consist only of children with no exposure to an ELP but based on experience and the findings from the 2024 Index Pilot, it would not be feasible to identify enough children with zero ELP exposure in practice within a feasible resource envelope.

There is also a disability eligibility criterion according to which children who have difficulties (a) seeing (with or without glasses), (b) hearing (with or without a hearing aid) or (c) walking or moving their arms or have weakness or stiffness in the arms or legs; and/or (d) significant difficulties understanding what people say, are ineligible for the sample.

Constraints/challenges

1. *Need to create sub-sample of enrolled children from the lower end of the socio-economic spectrum to compare to the non-enrolled sample.* This would be the group of main interest. Further, it would not be realistic to expect to access households in the top end of the socioeconomic spectrum (e.g. experience from 2022 WC stunting baseline survey and 2022 Population Census).
2. *By design, children in the non-enrolled sample will currently not be enrolled in an ELP but many have been exposed to an ELP at some prior point in time.*

Discarded approaches

Option 1: *Simple comparison* of the key child development outcomes across *statistically representative samples of non-enrolled and enrolled children* from the lower end of the socioeconomic spectrum. [Not possible because of constraint 2 and budget].

Option 2: Comparison of the key child development outcomes across *two constructed groups of children with limited/less and extensive/more ELP exposure* from the lower end of the socioeconomic spectrum.

- would need to create samples of currently non-enrolled and enrolled children from the lower end of socioeconomic spectrum.
 - ✓ after the enrolled sample enumeration areas (EAs) have been determined, could identify *X* of the poorest EAs and sample non-enrolled children in the same areas as the sample ELPs/enrolled children.
- would need to account for the fact that among children in the non-enrolled sample, a relatively small proportion would have zero ELP exposure, and that among children in the enrolled sample, although all currently attend an ELP, their degree of ELP exposure will vary. After obtaining the samples of *currently* non-enrolled and enrolled children:
 - ✓ construct a measure of ELP exposure (time in ELP, distribution over time, type of ELP) and define a cut-off to allocate children from both samples into two groups: limited/less and extensive/more ELP exposure and compare the key outcomes across the two groups [highly demanding in terms of data collection.]

Option 3: Construct a measure of the degree of ELP exposure and use as an explanatory variable in regression analysis of factors that influence the key child development outcomes.

- Weaker from result robustness and messaging perspectives.
- In a sense, this would make the non-enrolled sample less important for this research purpose since a measure of ELP exposure could be constructed for children in the enrolled sample only, although there would be no “never attended an ELP” cases.

Approach tested during the pilot

Option 4: Comparison of the key child development outcomes across statistically non-representative but indicative samples of non-enrolled and enrolled children from the lower end of the socioeconomic spectrum.

- after the enrolled sample enumeration areas (EAs) have been determined, *identify 45 EAs from the bottom three socio-economic quintiles as measured by Grade 3 enrolment weighted school quintiles, and sample non-enrolled children in the same areas as the enrolled children.*
- the *selected EAs should capture three types of areas: informal, rural and urban formal.*
- for a *household to be eligible for inclusion in the non-enrolled sample*, it must: (a) have a *child aged 50-59 months* at the time of the survey who (b) *who did not attend any early learning programme (ELP) in 2024 or in 2023* and (c) *who does not have any of the four specified disabilities*. An ELP is defined as: pre-school, nursery school, crèche, ECD centre, mobile ECD service, Grade R or playgroup. (Key assumption: for children who attended an ELP before 2024 any effects on the key child development outcomes have faded out.)
- sample the same number of non-enrolled as enrolled children in each ELP cluster.
- during the 2024 Index pilot random walk, it was documented what proportion of the visited households have children aged 50-59 months; what proportion of these attended an ELP in 2024; and what proportion of these attended an ELP in 2023. This was in anticipation of there likely not being enough eligible children to reach the intended sample size if those who attended an ELP in 2023 were excluded.

Pilot findings

Option 4 was tested during the 2024 Index Pilot in April-May 2024. During the pilot households generally opened the door and engaged with the fieldworkers.

Challenge 1: Initially, an interval of two homes was used during the random walk but this was changed to an interval of one as it was difficult to locate enough households with children aged 4 to 5 years (the Household visit screening form (first visit) asked for this age group rather than only 4-year-olds to avoid incorrectly excluding households with potentially age-eligible children).

Challenge 2: In households with 4-year-olds, the main challenge was to find children who met the ELP attendance eligibility criterion defined as “not having attended an ELP in 2024 or in 2023”. Based

on the initial ELP attendance criterion, 1.3% of households visited had eligible children and the ELOM assessment and primary caregiver interview were administered, see Table 1. This proportion was considered too low, and it was deemed necessary to relax this eligibility criterion to “did not attend an ELP in 2024” (i.e. children who attended an ELP in 2023 are considered eligible), this significantly improved the proportion of eligible children, see Table 1.

For the upcoming data collection for the non-enrolled sample, several activities are expected to boost the identification of eligible children.

(a) The poorest EAs will have been selected where a larger proportion of children are likely to not have attended an ELP or to only have limited attendance during 2024;

(b) an Index awareness campaign will have been conducted in the relevant communities;

(c) Provinces and EAs with relatively larger proportions of (non-enrolled) children aged 3 to 5 will be covered; and

((d) the full ELP listing will have been completed making it possible to target areas where there are relatively fewer ELPs. Whether this is an option will depend on the relative timing of the ELP listing and HH screening.)

Snowball sampling may also be considered as households in the pilot sites were able to refer assessors to other eligible households. It may also be possible to identify eligible children through ELPs / NGOs, although this might introduce some sample bias.

Table 1 Non-enrolled sample household visit form: total visits and eligibility counts and percentages

<i>NOTES:</i> The counts and percentages in this table are preliminary as in some cases, enumerators did not administer the HH visit form.			
	Counts by ELP attendance eligibility criterion		
	Total (1. + 2.)	1. Initial approach --Did not attend an ELP in 2024 or in 2023	2. Relaxed approach --Did not attend an ELP in 2024
SCREENING:			
Doors knocked on	1,221 (TBC)	1,111 (TBC)	110 (TBC)
Households available (on the day & 28 revisits)	795 (TBC) 65% of doors knocked on	721 (TBC) 65% of doors knocked on	74 (TBC) 67% of doors knocked on
Households available + have children aged 4-5 years (<i>first layer screening question</i>)	201 (TBC) 16.5% of doors knocked on	193 (TBC) 17.4% of doors knocked on	8 (TBC) 7.3% of doors knocked on
Households available + children aged 4-5 years + did not attend ELP in 2024 (<i>first layer screening question</i>)	39 (TBC) 3.2% of doors knocked on	35 (TBC) 3.1% of doors knocked on	4 (TBC) 3.6% of doors knocked on
Households available + children aged 4-5 years + did not attend ELP in 2023 (<i>first layer screening question</i>)	27 (TBC) 2.2% of doors knocked on	27 (TBC) 2.4% of doors knocked on	N.A.

NOTES: The counts and percentages in this table are preliminary as in some cases, enumerators did not administer the HH visit form.			
Counts by ELP attendance eligibility criterion			
	Total (1. + 2.)	1. Initial approach --Did not attend an ELP in 2024 or in 2023	2. Relaxed approach --Did not attend an ELP in 2024
Child's primary caregiver available	22 (TBC) 1.8% of doors knocked on	18 (TBC) 1.6% of doors knocked on	4 (TBC) 3.6% of doors knocked on
PRIMARY CAREGIVER INTERVIEW (detailed questions about child age, ELP attendance, any disability):			
Households available + child 50-59 months (<i>exact age criterion</i>)	17 (TBC) 1.4% of doors knocked on	14 (TBC) 1.3% of doors knocked on	3 (TBC) 2.7% of doors knocked on
Households available + child 50-59 months (<i>exact age criterion</i>) + meet shown ELP attendance eligibility criterion	16 (TBC) 1.3% of doors knocked on	13 (TBC) 1.2% of doors knocked on	3 (TBC) 2.7% of doors knocked on
Households available + child 50-59 months (<i>exact age criterion</i>) + meet shown ELP attendance eligibility criterion + meet disability criterion (%)	16 (TBC) 1.3% of doors knocked on	13 (TBC) 1.2% of doors knocked on	3 (TBC) 2.7% of doors knocked on
Households available + child 50-59 months + meet shown ELP attendance eligibility criterion + meet disability criterion + were assessed (%)	16 (TBC) 1.3% of doors knocked on [7 additional children assessed but enumerators did not administer HH visit form to all hhs]	Unknown, see Notes in first row of this table.	Unknown, see Notes in first row of this table.

EA selection rules

To select the **45 EAs** in three provinces – **Gauteng, KwaZulu-Natal** and **Western Cape** – for the non-enrolled sample the following steps were used.

1. Across the 3 provinces, exclude EAs that have fewer than 15 enrolled Grade 3 students per square kilometre (reliable, recurrent administrative data).
3. Define 3 area types – informal, rural and urban formal – for the Index using StatsSA 2011 Census sub-area type: (a) urban formal = formal residential areas; (b) informal = informal residential areas; and (c) rural = farms + small holdings + traditional. (StatsSA 2021 Census classifications not available in time).
4. Cutoffs for an EA to count as:
 - a) Informal, needs to consist of 10% or more informal residential areas. (Before the fieldwork, the survey firm will examine up-to-date maps of each relevant EA to locate the informal areas

and only cover those for the non-enrolled sample (and cover the informal areas located furthest away from the ELPs listed for the Enrolled sample, if the listing timing allows for it).

b. Rural, needs to consist of 20% or more farms, small holdings and traditional areas and not already be selected as informal (and cover the rural areas located furthest away from the ELPs listed for the Enrolled sample, if the listing timing allows for it).

c. Urban formal. needs to consist of 60% or more formal residential areas and not already be selected as informal or rural (and cover the urban formal areas located furthest away from the ELPs listed for the Enrolled sample, if the listing timing allows for it).

5. If more than 15 EAs exist for a given area type (informal, rural, urban informal), the EAs will be ranked by Grade 3 enrolment per square kilometre and the EA with the higher cut off values selected.

Using these rules, the 45 EAs in Table 2 were selected for the non-enrolled test sample.

After up-to-date maps showing structures have been prepared for each selected EA, these will be visually inspected to ensure there have not been any major changes since 2011. E.g. large outmigration from an EA resulting in very low population (and child) density. If such cases are identified, they will be replaced by the next EA on the list created using the EA selection rules as relevant.

Table 2 Preliminary EAs selected for the 2024 Index test non-enrolled sample

prov	dc	select	grade 3 enrolment weighted school quintile	area_sq_km	grade 3 enrolment per square km
Gauteng	City of Johannesburg	INFORMAL	1	5	124
Gauteng	City of Johannesburg	INFORMAL	2	7	27
Gauteng	City of Johannesburg	INFORMAL	2	3	319
Gauteng	Sedibeng	INFORMAL	2	4	71
Gauteng	West Rand	INFORMAL	3	1	74
Gauteng	West Rand	INFORMAL	3	1	287
Gauteng		INFORMAL	3	11	44
KwaZulu-Natal	eThekweni	INFORMAL	3	9	47
KwaZulu-Natal	eThekweni	INFORMAL	3	9	67
KwaZulu-Natal	eThekweni	INFORMAL	3	3	127
KwaZulu-Natal	iLembe	INFORMAL	3	4	34
KwaZulu-Natal	Umgungundlovu	INFORMAL	3	6	66
Western Cape	City of Cape Town	INFORMAL	2	1	280
Western Cape	City of Cape Town	INFORMAL	3	5	134
Western Cape	City of Cape Town	INFORMAL	3	1	157
Gauteng	City of Tshwane	RURAL	1	22	34

South Africa Thrive by Five Index 2024.

prov	dc	select	grade 3 enrolment weighted school quintile	area_sq_km	grade 3 enrolment per square km
Gauteng	City of Tshwane	RURAL	1	9	72
Gauteng	City of Tshwane	RURAL	1	8	78
Gauteng	City of Johannesburg	RURAL	3	57	15
Gauteng	City of Johannesburg	RURAL	3	45	31
Gauteng	City of Tshwane	RURAL	3	9	100
Gauteng	Ekurhuleni	RURAL	3	25	36
Gauteng	Sedibeng	RURAL	3	27	18
Gauteng	Sedibeng	RURAL	3	8	52
KwaZulu-Natal	Umkhanyakude	RURAL	2	12	24
KwaZulu-Natal	eThekwini	RURAL	3	42	17
KwaZulu-Natal	eThekwini	RURAL	3	31	18
Western Cape	Cape Winelands	RURAL	1	2	125
Western Cape	City of Cape Town	RURAL	2	5	200
Western Cape	City of Cape Town	RURAL	3	8	44
Gauteng	West Rand	URBAN FORMAL	1	1	273
Gauteng	City of Johannesburg	URBAN FORMAL	2	6	124
Gauteng	City of Johannesburg	URBAN FORMAL	2	6	150
Gauteng	City of Johannesburg	URBAN FORMAL	2	4	209
Gauteng	City of Tshwane	URBAN FORMAL	2	6	124
Gauteng	City of Johannesburg	URBAN FORMAL	3	2	157
KwaZulu-Natal	eThekwini	URBAN FORMAL	2	4	129
KwaZulu-Natal	eThekwini	URBAN FORMAL	3	14	55
KwaZulu-Natal	Umgungundlovu	URBAN FORMAL	3	1	295
Western Cape	Garden Route	URBAN FORMAL	1	1	246
Western Cape	West Coast	URBAN FORMAL	1	1	181
Western Cape	City of Cape Town	URBAN FORMAL	2	3	186
Western Cape	Garden Route	URBAN FORMAL	2	2	224
Western Cape	Central Karoo	URBAN FORMAL	3	4	34
Western Cape	City of Cape Town	URBAN FORMAL	3	6	234