



South Africa 2024 Thrive by Five Index

Datasets Summary: Non-Enrolled Children

Pub Version 1

Prepared by DataDrive2030 Tb52024 Technical Team, 17 June 2025

POPI Act

The 2024 Thrive by Five Index Pub Version datasets must be stored in compliance with the POPI Act, and the data must not be used for purposes other than those agreed to in the NDA.

Study objectives

The objectives of the South Africa Thrive by Five Index 2024 (2024 Index) are to:

- determine the proportion of young children in South Africa who attend early learning programmes (ELPs) that are on track for their age in three key areas of development, early learning, social and emotional functioning (SEF) and physical growth;
- monitor trends in performance gaps and gains over time and across socio-economic groups;
- explore a range of potential predictors of key child outcomes;
- track progress in the attainment of local and global development goals, including the Sustainable Development Goals (SDGs); and
- inform ECD decision-making and actions at multiple levels.

For the 2024 round of the Thrive by Five Index, the pilot non-enrolled child sample was a ‘test’ sample to explore the feasibility of reaching non-enrolled children. The target population for this pilot sample were 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.

Survey forms administered in the field

The development of the 2024 Index survey forms, including for the pilot study of non-enrolled children, was guided by the [Nurturing Care for Early Childhood Development Framework](#). The 2024 Index form development included consultations with more than 50 national and international researchers from across the required fields of expertise.

To achieve the 2024 Index study objectives for non-enrolled children, three survey forms were used to collect data for a purposively selected sample of children aged 50-59 months in low-income areas currently not attending ELPs in three provinces: Gauteng, KwaZulu-Natal and Western Cape:

1. early learning outcomes measure for ages 4 and 5 (ELOM4&5);
2. primary caregiver rating of child social and emotional functioning (SEF);
3. primary caregiver interview.

The **ELOM4&5 form** for the non-enrolled pilot sample is the same as the one used for the 2021 Index except for the addition of a small number of complementary questions and slight revisions to the disability screening questions. Unlike the form used for the sample of enrolled children, it captures child weight in addition to height.

The **rating of children’s social and emotional functioning (SEF) form** is the same as in 2021 except for two additional subscales namely, task focus and following instructions, each consisting of four questions, and that it was administered to primary caregivers instead of to practitioners.

The **primary caregiver interview form** is new for the 2024 Index. The interviews with assessed children’s primary caregivers (PCGs) were conducted face-to-face in the children’s home. It captures features of the child’s socio-economic background, home learning environment and health and key characteristics of primary caregivers.

Sample sizes

The pilot sample of non-enrolled children was purposively selected and is *not statistically representative of the underlying population*. It captures key characteristics of the non-enrolled assessed children and their primary caregivers and households, as well as key child outcomes (early learning, SEF and physical growth).

For details on the identification of households with non-enrolled children, see the [Tb52024 Non-Enrolled Enumeration Area Selection](#) note available separately.

The planned sample size for non-enrolled children was 540 but it was extremely challenging to identify non-enrolled children in the field for a variety of reasons (see *2024 Index Technical Report*). 334 non-enrolled children were identified but after removing (a) 1 child who was not within the eligible age range; (b) 1 child who did not have a PCG interview; and (c) 60 ELOMs that were invalid according to the ELOM rules (no consent, no assent, child scored zero on two or more ELOM domains, child exited assessment early, child has a disability, assessment not conducted in child's home language) the realised sample size was 272. These children reside in 262 households as the study design allowed for multi-family arrangements, e.g. cousins living in the same household.

Variable naming conventions, IDs, weights, flags & missing values

Variable labels are intended to be self-explanatory. But the appropriate survey form(s) should always be consulted when using the dataset(s) to ensure correct interpretation and usage.

Each survey form and by extension dataset, has its own variable prefix to indicate (a) that it is a collected (as opposed to derived) variable and (b) to indicate which survey form collected it to facilitate when multiple datasets are merged. The variable prefix n_ indicates that a variable was derived by the DataDrive2030 Tb52024 technical team using collected data. For example, in the ELOM-SEF dataset, the collected variables are prefixed elom_ while derived variables are prefixed n_elom_. *The derived variables include calculated ELOM scores and cutoffs.*

There are three identifiers for the 2024 Index: id_prov=province ID, id_hh=HH ID and id_child = child ID.

To merge the ELOM-SEF and PCG datasets the variable id_child should be used.

Missing values are coded as follows: . =system missing (mostly due to skip logic), .a or -1=Don't know, .b or -2=Refused, .c or -3=Not applicable and .z=Invalid response (due to tablet error in the field or inconsistent reporting due to enumerator/respondent error).

Potential remaining data errors in the datasets

The Pub Version 1 datasets have been through a two-step data checking and cleaning process, nonetheless, some data errors are likely to remain. If a data user discovers a potential data error, please fill in the provided error reporting form and submit to Simone Beukes, 2024 Index Operations Manager at simone@datadrive2030.co.za.

Before submitting an error reporting form please ensure:

- *you have checked the relevant survey forms (including skip logics);*
- *you have read the accompanying documentation (this Summary, the sampling strategy and codebooks); and*
- *unless it is a critical potential data error, to submit potential errors in batches rather than individually to facilitate the extended data cleaning.*

Table 1. Overview of the 2024 Thrive by Five Index survey forms (enrolled children)

	Variable prefix	Survey form	Respondent	Main purpose	Planned sample size	Realised sample size	Response rate
1	elom_	ELOM 4&5 (face-to-face) <ul style="list-style-type: none"> • gross motor development • emergent literacy and language • emergent numeracy and mathematics • fine motor coordination & visual motor integration • cognition and executive functioning • height measurement using stadiometer 	Child aged 50-59 months at the time of the survey visit	Capture early learning levels, overall and by domain Measurement of child height to capture physical growth status	540	272	50%
2	sef_	Child social and emotional functioning, SEF (face-to-face)	Primary caregiver of the assessed child	Capture child social and emotional functioning levels	540	272	50%
3	cg_	Primary caregiver interview, PCG (face-to-face)	Primary caregiver of assessed child	Capture child socioeconomic status, home learning environment, home stressors and primary caregiver and household characteristics that are expected to influence the key child development outcomes measured by the Index	540	272	50%