PRE-ANALYSIS PLAN

# Amendment 1: Heterogeneous Treatment Effects by Household Socio-Economic Status

**Date:** 27 April 2026

**Authors:** Jacobus Cilliers, Nompumelelo Mohohlwane, Kolosa Nonkenge, Stephen Taylor, Sharon Wolf, Zamangwe Zwane

**Status:** Pre-specified prior to fitting any of the procedures described below on the Wave-6 dataset

**Parent document:** Pre-analysis Plan: Long-term Impacts of a Successful Foundational Literacy Program in South Africa, October 2025

**Section extended:** §5.2 Heterogeneous Treatment Effects

---

## A1.1 Motivation

The original PAP (*§5.2*) specifies two moderators for the heterogeneity analysis: secondary-school quality and student gender. This amendment adds a pre-specified analysis of heterogeneity by **household socio-economic status (SES)**.

The motivation is empirical and theoretical:

- *Empirically,* results on the Wave-6 primary data show a sustained 0.13 SD treatment effect on home-language (Setswana) literacy, but no statistically detectable effect on the "other subjects" index (English, Math, Science). The pattern of point estimates (positive for Math and English, near-zero for Science) is consistent with a story in which spillovers to non-targeted subjects exist but are concentrated in a subset of the sample. Identifying *which* subset of households realizes spillovers from foundational literacy gains is policy-relevant for targeting.
- *Theoretically,* the skill-formation framework (Cunha and Heckman, 2007; Heckman, 2006) predicts that early skills become more productive when complemented by subsequent investments. In a setting where parents and households differ substantially in their capacity to provide complementary investment (books, study space, support for homework, expectations of post-secondary attainment), the long-run returns to a foundational-literacy intervention should be larger among higher-SES households. This contrasts with the standard short-run "compensatory" prediction, in which low-SES students gain more because their counterfactual trajectory is worst — but is the natural extension of skill-formation theory to longer time horizons and to outcomes other than the targeted skill itself.

We acknowledge that this amendment is written after the main Wave-6 ATEs are known. The pre-specification therefore concerns the *moderator analysis*, not the main effects. The hypothesis below is pre-stated for the heterogeneity tests in particular.

## A1.2 Hypothesis

**H1 (skill-begets-skill / complementary investment):** Among students from higher-SES households, the program has *larger* long-run impacts on outcomes other than home-language literacy (i.e., on the other-subjects index, math literacy, and English comprehension).

H1 implies, for the moderators specified in §A1.5:

- Caregiver completed matric (Grade 12) or higher → *larger* treatment effect.
- Number of books at home ≥ 11 → *larger* treatment effect.
- Household receives government grant (poverty proxy) → *smaller* treatment effect.

The corresponding null is that treatment effects do not vary along these dimensions; the alternative tested is one-sided in the direction implied by H1.

## A1.3 Sample

The estimation sample is the **Wave-6 non-attriter subsample** (`W6_status_2025 == 1`, N ≈ 1,690), further restricted to learners with complete data on the moderators specified below. Joint orthogonality of these moderators across the treatment and control arms is established in a balance test conducted before this amendment was filed (RI joint $p$ = 0.97 across 17 candidate Wave-6 covariates; see Appendix A1.A).

## A1.4 Outcomes

**Primary outcome:** *Other-subjects index*, defined as in PAP §4.1: an Anderson (2008) inverse-covariance-weighted index of English written comprehension, mathematical literacy, and science literacy, standardized to control mean 0 and standard deviation 1.

**Secondary outcomes:** Mathematical literacy and English written comprehension separately. (Science is excluded from the secondary set because the main-effect point estimate is essentially zero, leaving little for moderators to amplify.)

No new outcomes are added by this amendment; we condition the existing Family-B outcomes on a new moderator dimension.

## A1.5 Moderators

Three moderators of household SES are pre-specified:

1. **Caregiver completed matric (Grade 12) or higher** — binary indicator from the Wave-6 Learner Interview item “What is the highest level of education of your caregiver?”. Levels coded as in the survey instrument: ≥ matric vs. below matric. *Don’t know* (39.6% of respondents) is preserved as a separate indicator in the regression specification (i.e., two dummies — *matric+* and *DK* — with reference category *below matric*).
2. **Number of books at home ≥ 11** — binary indicator from the Wave-6 Reading-environment item “How many books are there in your home?”. Five-level ordinal collapsed at the median: 0 / 1–10 (54%) vs. 11–25 / 26–100 / >100 (46%).
3. **Household receives any government grant** — binary indicator from the Wave-6 Learner Interview. Functions as a poverty proxy (75% of households).

Although all three moderators are measured at Wave 6, they capture aspects of household SES that are plausibly time-invariant or only weakly responsive to a Grades 1–3 literacy intervention measured ten years earlier (caregiver education was determined long before the program; books-at-home and grant receipt are slow-moving). The balance test in §A1.3 supports the assumption that these covariates were not differentially shifted by treatment.

## A1.6 Procedure 1 — Pre-specified Linear Interactions (Confirmatory)

For each combination of moderator *W* ∈ {caregiver matric+, books ≥ 11, government grant} and outcome *y* ∈ {other-subjects index, math, English}, we estimate:

$$y\_isb = \beta_0 + \beta_1\ T\_is + \beta_2\ W\_isb + \beta_3\ (W \times T)\_isb + X'\_isb\ \Gamma + \rho\_b + \delta\_d + \varepsilon\_isb \qquad (A1.1)$$

where *T* is the treatment indicator (school-level assignment), ρ_b are strata fixed effects, δ_d are district fixed effects, and *X* contains the parsimonious control set from PAP §5.1 (gender and baseline composite literacy score). When *W* is *caregiver matric+*, the regression includes both the *matric+* and *DK* dummies and their interactions with *T*. Standard errors are clustered at the primary-school level (the unit of randomization.

The coefficient of interest is $\beta_3$. Under H1:

- *W = caregiver matric+*: $\beta_3 > 0$
- *W = books ≥ 11*: $\beta_3 > 0$
- *W = government grant*: $\beta_3 < 0$

Inference uses one-sided tests in the H1-implied direction.

**Multiple-testing correction.** Within each outcome, three moderator tests are jointly subject to Benjamini–Hochberg FDR control at α = 0.10 (the FDR level used elsewhere in PAP §4.3 for secondary outcomes). The other-subjects index is the *primary* outcome of this amendment; math and English are secondary and are reported with the same FDR procedure applied within each.

## A1.7 Procedure 2 — Generic-ML CLAN and BLP (Exploratory)

In addition to the pre-specified linear interactions, we will conduct an exploratory analysis of heterogeneity following Chernozhukov, Demirer, Duflo & Fernández-Val (2018), *"Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments"* (henceforth CDDF). The procedure is conducted on the **other-subjects index** as primary outcome.

### Step 1: Estimate CATE

Fit a causal forest (Athey, Tibshirani and Wager, 2019; `grf::causal_forest` in R) on the Wave-6 non-attriter sample, using the X-set defined below and clustering at the primary-school level (`clusters = emis_id`). Compute out-of-bag estimates of the conditional average treatment effect, $\hat{\tau}(X_i)$, for each learner.

### Step 2: Best Linear Predictor (BLP) Test of Heterogeneity

Regress $y_i$ on $T_i$, and an interaction of $(T_i - \bar{T})$ with $(\hat{\tau}(X_i) - \bar{\hat{\tau}})$. The coefficient $\alpha_2$ is the global heterogeneity test. We reject the null of no heterogeneity at $\alpha = 0.10$ (one-sided $\alpha_2 > 0$ under H1). This is the BLP test of CDDF §3.

### Step 3: Classification Analysis (CLAN)

Define the top and bottom quintiles of $\hat{\tau}(X_i)$. For each pre-specified moderator *W* in the X-set, test the difference $\bar{W}\_Q5 - \bar{W}\_Q1$. Apply BH-FDR within the X-set at $\alpha = 0.10$. This characterizes which observed covariates predict above- vs. below-median treatment-effect status, providing a description of heterogeneity that is interpretable in terms of household and student characteristics.

### Step 4: Sample-Split Aggregation

Steps 1–3 are repeated with $B = 100$ random splits of the data into estimation and testing halves. Reported p-values are aggregated across splits via the Storey-corrected median p-value, as recommended in CDDF §4.3.

### X-set for the Forest (≈14 Covariates)

**Household SES (Wave 6, plausibly time-invariant):**

- Household-asset PCA1 (first principal component of 11 binary asset items from the Wave-6 home-learning module, dropping items plausibly acquired by the child: computer/tablet, study desk, own room, mobile phone, internet)
- Caregiver matric+ (binary), Caregiver education = DK (binary)
- Books at home (continuous, 0–4 ordinal)
- Government grant, child grant (binaries)
- Caregiver employed (binary)
- Has father (binary)
- Ever school-hungry (binary)
- Commute time, minutes (winsorized at 99th percentile)

**Pre-treatment covariates (PAP §5.1 control set):**

- Baseline composite literacy score
- Student gender (female)
- Annual National Assessment pass rate (school-level, baseline)
- Community wealth index (2011 census)
- Community secondary-school attendance rate, ages 13–18 (2011 census)

**Notes on X-set construction.** *Has mother* and *Mother is not the caregiver* are excluded despite the balance test, due to (a) marginal individual-level imbalance ($p < 0.05$, single failure out of 17 tests in the balance check) in the case of *Has mother*, and (b) high missingness from the conditional skip pattern in the case of *Mother is not the caregiver*. *Setswana spoken at home* is excluded due to limited variation (96% Setswana on the non-attriter sample). District is excluded as a moderator on the grounds that ten years post-baseline it is no longer a meaningful proxy for any time-invariant characteristic (learner mobility).

## A1.8 Robustness Checks and Placebo

4. **Placebo on baseline.** Repeat Procedures 1 and 2 with the dependent variable replaced by *baseline composite literacy* (`BL_index`). As baseline literacy is measured prior to treatment, we expect to find no significant heterogeneity by any moderator. A failure of this placebo would suggest that the procedure is detecting noise rather than treatment-effect heterogeneity.
5. **Asset-index sensitivity.** Re-run Procedure 2 with (a) the asset PCA1 constructed from all 16 binary items (including child-acquired), and (b) an equal-weighted asset count instead of PCA. Results that differ qualitatively across these specifications would indicate sensitivity to index construction.
6. **Pre-treatment-only X.** Re-run Procedure 2 with only the five pre-treatment covariates from PAP §5.1, removing the eight Wave-6-measured SES variables. This reports the heterogeneity estimable purely from baseline characteristics, against which the augmented specification can be compared.
7. **Cross-validation of methods.** For each pre-specified moderator in Procedure 1, compare the linear-interaction $\hat{\beta}_3$ to the analogous CLAN difference $\bar{W}$_Q5 − $\bar{W}$_Q1 (signed appropriately). Convergent evidence across the two methods strengthens any heterogeneity claim.

## A1.9 Anticipated Limitations

8. **Power.** With N ≈ 1,500 after complete-case selection on the X-set, and plausible interaction effect sizes about half the main ATE (i.e., ~0.05–0.10 SD), power for detecting heterogeneity is limited. We pre-emptively note that null findings should be interpreted as non-rejection of zero, not as evidence of homogeneity.
9. **Causal interpretation of moderators.** Even with joint balance on the moderator set, the CATE conditional on the X-set is not a *manipulation* of X. If household SES is correlated with unobserved drivers of treatment response, our estimates describe *who benefits*, not *why* they benefit. We will avoid causal language for the moderator effects themselves.
10. **One-sided tests.** Tests are pre-specified one-sided in the H1-implied direction. Findings in the *opposite* direction (i.e., *larger* effects among lower-SES households, the standard compensatory pattern) will be reported as such, with two-sided p-values reported alongside the pre-specified one-sided tests for transparency.
11. **The other-subjects null effect.** The motivation for this amendment is the absence of a detectable *average* spillover to non-targeted subjects. If the BLP test in Procedure 2 fails to reject the null of no heterogeneity, the null average effect cannot be attributed to averaging across heterogeneous responses, and the paper will report that interpretation accordingly.

## A1.10 Implementation Notes

Code for both procedures will be archived in `7_Script/Analysis/` under filenames `27_het_linear_interactions.R` (Procedure 1) and `28_het_causal_forest_CLAN.R` (Procedure 2), referenced by git commit hash at the time of registration.

## Appendix A1.A — Joint Balance Test on Wave-6 SES Moderators

Conducted prior to filing this amendment. Sample: $N = 1{,}690$ Wave-6 non-attriters. Specification: `y ~ T + ρ_b + δ_d`, cluster-robust SE on `emis_id`. Joint $F$-test against the full set of 17 candidate Wave-6 covariates: randomization-inference $p = 0.97$ ($B = 1{,}000$ within-strata school-level permutations). Individual-level: 16 of 17 balanced at $p > 0.10$; the single marginal failure is *Has mother* (control 0.89, treatment 0.84, $p < 0.05$). Full results in `5_output/Tables/balance_post_treatment.tex`.

---

## References (Additions)

*Athey, S., Tibshirani, J., and Wager, S.* (2019). "Generalized random forests." *Annals of Statistics* 47(2), 1148–1178.

*Chernozhukov, V., Demirer, M., Duflo, E., and Fernández-Val, I.* (2018). "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments." NBER Working Paper 24678.

*Cunha, F. and Heckman, J. J.* (2007). "The Technology of Skill Formation." *American Economic Review* 97(2), 31–47.

*Heckman, J. J.* (2006). "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312(5782), 1900–1902.