Impact of Graduation on Child Development in Uganda — Wave 2 (Ancillary/Extension)
Registry Identifier: NCT05531812 (IPA-2022-UG)

Pre-Analysis Plan
October 1, 2025
Principal researchers: Peter Rockers, Lasse Brune, Nathanael Goldberg, Dean Karlan, Doug Parkerson, Chris Udry

Purpose

This document outlines the pre-analysis plan for analysis of a second wave of child assessment data collected from children in households previously enrolled in the Graduating to Resilience (G2R) project in Uganda (AEARCTR-0004080). The first wave of child assessments in 2022 were pre-registered in clinicaltrials.gov (NCT05531812). We now expect to analyze wave 1 and wave 2 data together and thus details of the wave 1 pre-analysis plan are described in this document along with the pre-analysis plan for wave 2.

Randomization

The G2R trial had two stages of randomization. In the first stage, villages were randomly assigned to treatment (T) and control (C) villages. Villages were stratified into host and refugee communities and randomized within strata in this first stage of randomization. The second stage of randomization assigned households in treatment villages to one of three types of treatment arms (T1, T2, or T3) or within-village control (C1). The only difference between T1 and T2 was the receipt of coaching sessions in either an individual (T1) or group (T2) format, while T3 households received individual coaching but no asset transfer. Households selected for interviews in the control villages were spillover controls (C2).

Wave 1. In the first wave of data collection for this study (2022), households from host and refugee communities with eligible children (26-40 months of age) from arms T1, T2, and C1 were visited. Data from T1 and T2 will be combined as a single treatment group in the analysis. The results of the G2R trial indicate that impacts on household assets, consumption, income, and food security were similar in T1 and T2. While potential spillovers within villages may limit our ability to cleanly identify the impact of Graduation, the results of the main G2R trial suggest that such spillovers were small and that household-level impacts on income and consumption comparing T1 and T2 to C1 were significant, which presents the opportunity to examine whether these impacts confer benefits to children.

Wave 2. In the second wave of data collection (2025), households from host communities only with eligible children (0-9 years of age) from arms T1 and C2 will be visited. The use of pure controls (C2) in wave 2 will give a cleaner identification of the impact of the Graduation program as compared to wave 1.

Outcomes

Wave 1. The primary outcome is domain-specific child development scores measured using the Bayley Scales for Infant and Toddler Development version III (BSID-III). The BSID-III assesses five domains of skill development: cognitive, language, motor, social-emotional, and adaptive behavior. Each domain will be examined as a separate primary outcome. Domains include a set of age-appropriate tasks that the child is asked to complete. For each domain, raw data will be converted to z-scores by standardizing within the study population. Secondary outcomes are

domain-specific child development scores measured using the Caregiver-Reported Early Development Instruments (CREDI). Like the BSID-III, the CREDI assesses cognitive, language, motor, and social-emotional skills. While the BSID-III relies on direct assessments of children, the CREDI is based on caregiver reporting on child skills. Raw item data will be converted to norm-referenced z-scores for each domain using the CREDI scoring procedure.

Wave 2. Primary outcomes differ depending on the age of the child. For all study children regardless of age, anthropometric measures (height and weight) will be collected and converted to z-scores based on WHO Child Growth Standards and hemoglobin levels will be captured using HemoCue devices and categorized according to anemia status (<11.5 g/dL). For children 0 to 3 years of age, child development will be measured using the CREDI. For children 4 to 9 years of age, fluid intelligence, school readiness, and executive function (EF) skills will be measured. Fluid intelligence will be measured using the Raven's Progressive Matrices (RPM). The RPM is a validated test that has been used for many years in diverse settings around the world. School readiness will be measured using the *International Development* and *Early Learning* Assessment (IDELA). The IDELA was created by Save the Children and partners and includes items across the four domains—motor, literacy, numeracy, and social emotional. It has been tested and validated in several LMICs. Working memory, a core EF skill, will be measured using the forward digit span task. We will measure inhibitory control and cognitive flexibility using the Head Toes Knees Shoulders task. Secondary outcomes at wave 2 include child diet diversity and caregiver-child interactions.

Analysis

Wave 1. In our main analysis, we will estimate average treatment effects of Graduation on outcomes using an intention-to-treat approach. We will combine T1 and T2 groups into a single treatment group for our analysis and compare to within-village control (C1). We will also examine heterogeneity in treatment effects by the duration of program exposure *in utero*. The relative contributions of poverty-related adversity experienced *in utero* versus post-birth on child development is an important research question. We will also estimate impacts within refugee and host strata. We will fit a set of unadjusted and adjusted linear regression models to estimate average treatment effects. All models will include variables used for stratification in the original randomization procedure. Adjusted models will include the following additional covariates to improve precision: child age and gender; caregiver age and education level; and measures of household consumption and food insecurity measured at baseline of the G2R trial.

Wave 2. In our main analysis, we will estimate average treatment effects of Graduation on outcomes using an intention-to-treat approach. We will adjust standard errors in all models to account for clustering in the study design. We will fit a set of unadjusted and adjusted linear regression models to estimate average treatment effects. All models will include variables used for stratification in the original randomization procedure. Adjusted models will include the following additional covariates to improve precision: child age and gender; caregiver age and education level; and measures of household consumption and food insecurity measured at baseline of the G2R trial.