How do transformative early learning gains shape learning and life in adolescence? Evidence from West Africa

Authors and institutional affiliation: Alex Eble, Ila Fazzio, Peter Boone, Aliu Balde, Baboucarr Bouy, Alpha Camara, Jenny Pei-Tseng Hsieh, Fatou Bin Jobe, Robin Lumsdaine, Biapte Nabunque[1]

Date of latest draft: July 5, 2025

Abstract

Our study asks two questions. One, do extremely large, intervention-driven early learning gains in foundational literacy and numeracy, achieved among primary school-aged children in extremely income-poor settings persist or fade out over time? Two, do these gains have any knock-on effects for families and communities? We answer these questions by following up with participants from studies in The Gambia and Guinea Bissau which showed large learning gains from bundled interventions targeting early grade reading and learning skills. Several years after the conclusion of these studies, we will return to find and interview the original trial participants enrolled in the studies and their families. We will also administer tests of learning, measuring foundational literacy and numeracy, higher-level academic skills, and "life skills" among these participants. We will compare these outcome variables across intervention and control villages.

Keywords: Foundational literacy and numeracy, learning, fade out, persistence, bundled interventions, poverty, West Africa

JEL codes: I25, O15, J24, I21, I30

1. Introduction

Despite great progress in increasing schooling access and completion in developing countries, there are still pockets of extreme poverty in which most children born today will grow to adulthood without ever acquiring foundational literacy and numeracy skills. These skills will prevent them from participating in most spheres of society and the economy, imposing a low ceiling on their future incomes and levels of prosperity, as well as those of their home communities.

[1] Eble (corresponding author, email: eble@tc.columbia.edu): Columbia University and NBER. Balde, Camara, Fazzio, and Jobe: Effective Intervention. Boone: Effective Intervention and Center for Economic Performance, London School of Economics. Lumsdaine: American University and NBER. Buoy: Ministry of Public Service, Administrative Reforms, Policy Coordination and Delivery, The Gambia. Hsieh: National Foundation for Educational Research. Nabunque: Effective Intervention

Efforts to raise learning in such populations have typically yielded only incremental learning gains, usually of a small fraction of a standard deviation (Ganimian and Murnane 2016; Glewwe and Muralidharan 2016; Evans and Yuan 2022). More recently, a series of studies have shown that larger gains are possible when multiple interventions known to work in isolation are bundled together (Mbiti et al. 2019; Kerwin and Thornton 2021; Gray-Lobe et al. 2022). The largest of these gains were documented in two studies, one in The Gambia and the other in Guinea Bissau (Eble et al. 2021; Fazzio et al. 2021). These two studies documented intervention-driven learning gains among primary-aged children a full order of magnitude larger than even some of the most effective interventions studied previously. Their results demonstrated that such bundled interventions can dramatically increase the likelihood of children in these areas ever reaching literacy or numeracy, from a baseline case of essentially zero chance to between 25 and 60 percent likelihood.

We will estimate the extent to which the learning gains measured in the original studies persist or fade out when the children are adolescents, whether this affects their likelihood of success and persistence in schooling, and whether this affects beliefs, aspirations, and non-academic life outcomes, such as family formation, measured in adolescence.

Our work follows a series of rigorous long-term follow-up studies of transformative development interventions in similar contexts. These studies have studied the effects of anthelmintic treatment for children (Baird et al. 2016), free secondary education (Duflo, Dupas, and Kremer 2021), and unconditional cash grants (Blattman, Fiala, and Martinez 2020), among others. We also hope to lay the ground for the creation of a larger evidence base on the long-term effects of transformative childhood interventions, particularly those with a focus on helping children of early primary school age attain foundational literacy and numeracy, in very low-income contexts. This would build on, and follow the model of evidence bases gathered on the longer-term impacts of educational interventions from other, wealthier settings (Campbell et al. 2002; Heckman et al. 2010; Chetty, Friedman, and Rockoff 2014) and help guide policy intervention in such settings. More specifically, we wish to understand and quantify the extent to which such gains in foundational literacy and numeracy persist and propagate in the community, as opposed to fading out or having children from other communities experience catch-up. This would help understand whether scarce policy resources should be focused on reaching large gains in early childhood, or whether those alone are insufficient for large later life gains, suggesting the need for scaffolding later on and possibly other, complementary subsequent interventions.

2. Research Design

Hypotheses and outcome variables

Our research is designed to test the broad hypothesis that the learning gains from the two studies persisted into adolescence and have yielded salutary effects on outcomes related to education, such as belief formation, life satisfaction, aspirations, and early fertility-related decisions. The main outcomes of interest can be divided into families of outcomes, as in (Kling, Liebman, and Katz 2007). In our study, we have the following four families: educational progression; learning; beliefs and aspirations; and family formation and gender norms.

The primary outcome for each family of outcomes is as follows:

- For *educational progression*, the primary outcome will be the highest grade attended. Secondary outcomes will include whether the trial participant is still in school, whether they were in school in the previous year, the following outcomes collected conditional on the participant being in school this year or the previous: {annual educational expenditure on the trial participant's education, location of school relative to trial participant's village of enumeration, type of school}, variables collected conditional on the trial participant not being in school this year: {reason left school, intention to go back to school, main daytime activity, whether they are employed, type of job, work hours, and work income}; time use, and the highest grade attended of the participant's next older and younger siblings.
- For *learning*, the primary outcome will be a composite score on a test comprising assessment of early grade reading and math skills and higher-grade learning. Secondary outcomes will include summary scores on of these two categories of skills plus on a third category of other skills, commonly known as "non-cognitive" or "life" skills; scores on subtasks within each category; and spillover of learning to the nearest older and younger sibling of the participant, as measured by ASER-style tests.
- For *beliefs and aspirations*, the primary outcome will be an index of the participant's future-oriented beliefs aspirations, created by binarizing their responses to the relevant beliefs and aspirations questions in the survey, harmonizing them so that positive values connote higher aspirations, and summing them. Secondary outcomes will include an analysis of the first principal component of this same index of outcomes as a y variable, individual responses on these participant aspirations questions, an index of caregiver aspirations for the participant, and individual responses on caregiver aspirations questions.
- For *family formation and gender norms*, the primary outcome will be an index of the participant's responses to relevant questions on family formation and gender norms, created by binarizing the responses, harmonizing them so that positive values connote higher aspirations, and summing them. Secondary outcomes will include an analysis of

the first principal component of this same index of outcomes as a y variable, as well as individual level responses to each question.

In all cases, outcomes will be measured at the level of the study participant.

Identification
Identification will come from the randomization in the original studies, and we will treat each participant according to their initial randomization status regardless of later movement or migration.

Intervention
The interventions being studied are described in greater detail in Eble et al. (2021) and Fazzio et al. (2021). The basic model motivating both interventions is to combine, or "bundle" multiple interventions, known to work in isolation, with the goal of having a synergistic effect whose overall impact on learning is far greater than the sum of the anticipated benefits from each intervention in isolation.

In The Gambia, the intervention was delivered via 12 hours per week of after-school lessons, combining several important components:

- Previously untrained "para teachers";
- Scripted lessons;
- Extensive monitoring of teachers (also known as "coaching") with the goal of increasing teacher skill and raising student learning;
- Extensive community sensitization; and
- A high level of per-child resources spent to ensure fidelity of the intervention.

In Guinea Bissau, the intervention was delivered by running schools in lieu of the government. Unlike in The Gambia, schools were run by individuals who had previously trained and worked as teachers, but that intervention also used the other components of the bundle in The Gambia. Both studies took place in small, remote rural areas of the country in question.

Original studies
The original studies were cluster-randomized control trials, with clusters of adjacent villages randomly assigned to groups by computer algorithm, stratified by region and degree of remoteness in the case of The Gambia, and by values of a composite variable that considered a weighted average of village's remoteness, size, and type of school available, and mother's education and fluency in Crioulo in Guinea Bissau. All individuals will retain the status assigned to them at randomization. Following the original studies, blinding was not feasible or necessary, and so will not be used. Because the majority of these outcomes are either objective

(performance on learning assessments) or not directly targeted by the educational intervention (aspirations and beliefs), we anticipate minimal risk of performance and expectancy biases.

Sample and statistical power

For our primary outcome, the unit will be learning at the individual trial participant level. We will travel to all clusters (55 control and 56 treatment in The Gambia; 33 control and 16 treatment in Guinea Bissau), and based on a pilot done in summer of 2023, we estimate that we will be able to reach at least 70 percent of participants in each study (original number of participants: 2,458 control and 2,060 intervention in The Gambia; 1,463 control and 649 intervention in Guinea Bissau).

Under these assumptions, and using an intra-cluster correlation coefficient of 0.1647 in The Gambia and 0.1493 in Guinea Bissau (calculated using the control group test scores at endline), we estimate that our MDE will be 0.2344 in The Gambia, and 0.3605 in Guinea Bissau. Given that the original treatment effects were 3.2 and 5.3 SD in the original studies, this ensures that we are powered to detect any difference that is as large or larger than a tenth of the original effect size.

3. Data

Our data will come from three main sources. The first source is the data collected during the course of the original studies. The second source is a set of surveys administered to study participants in late 2024 and early 2025. The third source is a learning assessment administered to study participants in 2025.

We will aim to collect data from all participants in both studies. Our approach involves two stages of data collection: first, an attempt to find all participants in the place in which they were enumerated at the start of the original studies. In cases where the participant is not present in the village, we will attempt to find a time to visit them during a "callback" where the village is later visited at a prespecified date.

We anticipate difficulty in reaching some trial participants if they migrate to other areas. We will follow best practice from other long-term follow-ups, such as that of Baird et al. (2016) in Kenya and Blattman, Fiala, and Martinez (2020) in Liberia, by including a second, "mop-up" attempt to find those not present in the village, looking for them instead in wherever they may have migrated to. This involves using contact information obtained from the surveys in the village of their next of kin, identifying whether they are in the country or not, and then for those who have not left the country, attempting to find them where they are and administer the survey to them there. We will administer phone surveys to those who have permanently migrated out-of-country, but will not be able to conduct learning assessments for them.

The overarching goal is to reach at least 80 percent of those initially enumerated in each study, and the extent of mop-up will be determined by success in the initial survey efforts. For reference, during an unannounced pilot we were able to find 50-60% of original study participants in their villages of enumeration, and obtained at least some contact information for essentially all of the 400 enumerated participants we attempted to find who were not present in the village on the day of visit.

We will measure learning on three fronts:

1. Early grade literacy and numeracy skills, measuring the persistence or gains of the effects we measured in the original studies;
2. Later grade math and reading skills, measuring the impact of the early grade gains on later learning; and
3. Practical skills related to productivity in employment, measuring the impact of the original interventions on adolescents' acquisition of skills to navigate and make informed decisions in daily tasks/challenges they may face at life or work.

4. Analysis

Our analysis will follow the research design of the original studies. We will test our hypotheses with linear regression following the research design in the original two studies, regressing the outcome variable on treatment status and the original stratification variables and nothing else, clustering standard errors at the cluster (Gambia) or tabanca (Guinea Bissau) level.

We will estimate the main parameter of interest, the coefficient on treatment status, separate by study country, with one set of regressions and estimates for Guinea Bissau and a separate set of regressions and estimates for The Gambia. To account for potential differentials in attrition across trial arms in both contexts, we will compute and report Lee Bounds.

Statistical model

$$y_i = \beta_0 + \beta_1 Treatment_i + \gamma X_i + \epsilon_i$$

We will estimate this equation, regressing the outcome *y* measured at the participant *i* level, on a constant, treatment status at randomization ($Treatment_i$), and fixed effects for the strata variables (X_i). We will cluster standard errors at the level of randomization (cluster in The Gambia and village in Guinea Bissau). We will estimate one version of this equation for each of the two countries *c*.

We divide our outcome variables into four families and specify a primary outcome for each family, with other variables clearly indicated as secondary. We will conduct no further pre-specified adjustment for multiple hypothesis testing.

We will conduct heterogeneity analysis by the dimensions investigated in the original studies: gender, baseline wealth, region, distance to main road[2], caregiver education level, whether there was a school in the village at baseline, and baseline caregiver aspirations. These variables are all based on data collected in the original two studies.

As in the original studies, we will present intent-to-treat estimates throughout. We will also estimate a version of treatment-on-the-treated estimates, estimating potential heterogeneity by takeup of the intervention, as measured by attendance in intervention classes during the original studies.

5. Interpreting Results

Our interpretation of the results is straightforward. The magnitude and significance of the estimates we generate will inform our assessment of whether the early learning gains that the two interventions yielded for participants, measured at the endline of the original studies when they were ages 8-11, persist into late adolescence (when they are 15-18 years old).

If the results remain large and positive, this would tell us that such large pushes in the early primary years would have a clear impact on later learning, and justify the placement of substantial resources into these early grade learning projects. If the results diminish in magnitude or disappear completely, it would tell us something equally important – that, alone, large investments in early grade skill acquisition in very low-income contexts alone are insufficient to spur later grade learning.

In either case, we would be careful not to over-interpret these results, as many results previously have shown benefits of childhood interventions on adult wages decades after learning gains wane, and, conversely, adolescent learning alone does not mean a path out of poverty in such contexts.

6. Bibliography

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. 2016. "Worms at Work: Long-Run Impacts of a Child Health Investment." *The Quarterly Journal of*

[2] We will analyze heterogeneity by distance to the main road only in The Gambia, following the heterogeneity analysis in the original papers.

Economics 131 (4): 1637–80.
Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2020. "The Long-Term Impacts of Grants on Poverty: Nine-Year Evidence from Uganda's Youth Opportunities Program." *American Economic Review: Insights* 2 (3): 287–304.
Campbell, Frances A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. 2002. "Early Childhood Education: Young Adult Outcomes From the Abecedarian Project." *Applied Developmental Science* 6 (1): 42–57. https://doi.org/10.1207/S1532480XADS0601_05.
Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–79.
Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2021. "The Impact of Free Secondary Education: Experimental Evidence from Ghana." National Bureau of Economic Research.
Eble, Alex, Chris Frost, Alpha Camara, Baboucarr Bouy, Momodou Bah, Maitri Sivaraman, Pei-Tseng Jenny Hsieh, Chitra Jayanty, Tony Brady, and Piotr Gawron. 2021. "How Much Can We Remedy Very Low Learning Levels in Rural Parts of Low-Income Countries? Impact and Generalizability of a Multi-Pronged Para-Teacher Intervention from a Cluster-Randomized Trial in the Gambia." *Journal of Development Economics* 148: 102539.
Evans, David K., and Fei Yuan. 2022. "How Big Are Effect Sizes in International Education Studies?" *Educational Evaluation and Policy Analysis* 44 (3): 532–40.
Fazzio, Ila, Alex Eble, Robin L. Lumsdaine, Peter Boone, Baboucarr Bouy, Jenny Hsieh, Chitra Jayanty, Simon Johnson, and Ana Filipa Silva. 2021. "Large Learning Gains in Pockets of Extreme Poverty: Experimental Evidence from Guinea Bissau." *Journal of Public Economics* 199 (104385).
Ganimian, Alejandro, and Richard Murnane. 2016. "Improving Education in Developing Countries: Lessons From Rigorous Impact Evaluations." *Review of Educational Research* 86 (3): 719–55.
Glewwe, Paul, and Karthik Muralidharan. 2016. "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications." In *Handbook of the Economics of Education*, 5:653–743. Amsterdam, Holland: Elsevier.
Gray-Lobe, Guthrie, Anthony Keats, Michael Kremer, Isaac Mbiti, and Owen W. Ozier. 2022. "Can Education Be Standardized? Evidence from Kenya." *Evidence from Kenya (June 5, 2022). University of Chicago, Becker Friedman Institute for Economics Working Paper*, no. 2022–68.
Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1–2): 114–28.
Kerwin, Jason T., and Rebecca L. Thornton. 2021. "Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures." *Review of Economics and Statistics* 103 (2): 251–64.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.
Mbiti, Isaac, Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani. 2019. "Inputs, Incentives, and Complementarities in Education:

Experimental Evidence from Tanzania." *The Quarterly Journal of Economics* 134 (3): 1627–73. https://doi.org/10.1093/qje/qjz010.

7. Administrative information

Funding: Funding was provided by Effective Intervention, a UK charity.

Institutional Review Board (ethics approval): IRB approval was obtained from Teachers College, Columbia University, the University of The Gambia Research Ethics Committee, and the Ministry of Education in Guinea Bissau.

Acknowledgments: We are grateful for both the intellectual contributions and hard work from Alieu Gibba, Abhishek Thakur, Filipa Silva, and others on our teams in The Gambia and Guinea Bissau that made this work possible. We are also grateful for the intellectual contributions of Lisa Kuhn, Isobel Beveridge, Elizabeth Martin, Molly Jay, and Charalambos Kollias, and of the National Foundation for Educational Research (UK) as an organization, in designing the assessment.