

Structured Pedagogy Evidence Summary

This write-up summarizes an evidence review of Structured Pedagogy literature conducted in July 2024. The main objective was to compare effect sizes of different learning interventions in LMICs (Low and Middle Income Countries), to inform program design decisions for a new foundational learning NGO. It summarizes the evidence base around two questions: 1) How does Structured Pedagogy compare to other education interventions that aim to increase learning outcomes in both effect size and cost-effectiveness? 2) What accounts for the large variation in effect size among programs? In addition to the Structured Pedagogy literature, we also draw on studies of Targeted Instruction (often called Teaching at the Right Level, or TaRL), which shares many pedagogical elements with Structured Pedagogy. At the end of the write-up, we briefly explore the evidence studying the relationship between education and economic outcomes.

We conclude that a well-implemented Structured Pedagogy program can achieve a very large impact, with extremely high cost-effectiveness; evidence on this question is strong and encompasses many rigorous studies across many countries and regions. Effective teacher training and coaching appear to be critical drivers of lesson plan adoption and implementation quality by teachers, and should be central to the design of any new program. We also conclude that increases in learning are likely to have large positive impacts on future earnings of individuals, as well as eventual economic growth in regions that improve learning, though the evidence base on this question is weaker due to the long time lag between educational interventions, particularly in the early grades, and those students entering the workforce.

Introduction: What is Structured Pedagogy? How is it similar to and different from Targeted Instruction?

Structured Pedagogy is an approach intended to simplify the process of implementing effective, learner-centered instructional practices for teachers in LMICs by providing them with daily lesson plans. The lesson plans are designed in line with the evidence on how children learn math and literacy skills, including clear and direct instruction in daily bite-sized learning objectives, active engagement of learners in the lessons through fun practice activities and instructional routines, and use of phonics for teaching reading. Some Structured Pedagogy programs include highly scripted lesson plans, while other programs are much less prescriptive guides for teachers; typically the instruction style that they replace uses a highly teacher-centered “lecture and repeat” approach.

In addition to lesson plans, Structured Pedagogy programs typically include a package of: student textbooks aligned to the lesson plans, teacher training sessions to accompany the rollout of the approach, and teacher coaches who visit classrooms and provide advice to teachers on improving their implementation. Some programs also incorporate systems for

gathering and responding to data on student learning during the program's implementation. Most Structured Pedagogy programs target learners in grades 1-3 and assume a fairly uniform starting point for students, as they are generally being exposed to learning content for the first time.

Targeted Instruction, or TaRL, is an evidence-based reading and math catch up program typically delivered to students in grades 4-6. It's usually delivered outside of the school day, either during inter-term holidays or after school. Students are given a simple test at the start of the program, and are sorted temporarily by learning level rather than age or grade level, in order to receive instruction at their level. The instruction itself shares some similarities to the Structured Pedagogy approach in that it draws on the science of how children learn, and it includes both direct instruction in concepts and fun, learner-centered practice activities. Targeted Instruction activities may be somewhat less prescribed than Structured Pedagogy, although both may rely on similar types of instructional routines.

How does Structured Pedagogy compare to other education interventions in effect size and cost-effectiveness?

We reviewed three major meta-analyses of learning interventions in LMICs, and found that **all three cite Structured Pedagogy as among the 2-3 most high impact and/or cost-effective learning interventions ever studied globally**, alongside TaRL and sharing information with families about the benefits of education. This is likely why the Global Education Evidence Advisory Panel named Structured Pedagogy as among its "great buys."¹ **Likewise, the Copenhagen Consensus identifies Structured Pedagogy as among its 2 top education interventions, alongside TaRL, suggesting the tantalizing possibility of a 65:1 return on investment.**² Below, we summarize three meta-analyses:

Snijlsteit et al. 2015³ review 238 studies on education interventions in LMICs, 21 of them about structured pedagogy. Comparing across interventions, they determine that structured pedagogy is one of the most impactful ones. The range of impact estimates varies widely: language learning gains go from -0.14 SDs to 0.9 SDs, with a uniform distribution. The average impact is 0.23 SD. The meta-analysis doesn't explore which variables drove these differences in effect sizes, but they highlight the value of better understanding why outcomes vary so widely. We explore this question below under the implementation heading.

¹Global Education Evidence Advisory Panel, 2023. [Cost Effective Approaches to Improve Global Learning](#).

²Noam Angrist, Elisabetta Aurino, Harry Anthony Patrinos, George Psacharopoulos, Emiliana Vegas, Ralph Nordjo, and Brad Wong, 2023. [Improving Learning in Low- and Lower-Middle Income Countries](#). Journal of Benefit-Cost Analysis, 1-26. Cambridge University Press.

³Snijlsteit, B, Stevenson, J, Phillips, D, Vojtkova, M, Gallagher, E, Schmidt, T, Jobse, H, Geelen, M, Pastorello, M, and Evers, J, 2015. [Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review](#), 3ie Systematic Review 24. London: International Initiative for Impact Evaluation (3ie).

The Learning at Scale report⁴ was funded by the Gates Foundation to understand what are the components that lead to successful programs at scale that increase learning gains. They include eight programs, six of which are Structured Pedagogy and one of which is TaRL. All the programs show positive learning gains, the most impactful ones being **Tusome (Structured Pedagogy in Kenya), with learning gains of 0.71 SD** in Oral Reading Fluency (ORF), and SERI (Teaching at the Right Level in India) with learning gains of 1.3 SD in ORF. Some other programs also show highly promising results, such as Ghana Learning (Structured Pedagogy in Ghana), that reports 0.52 SD in ORF. Oral Reading Fluency is considered a critical precondition of reading with comprehension, as students need to be able to read quickly enough to focus their attention not on decoding words, but on engaging with the meaning of sentences and paragraphs as a whole.

A 2020 meta-analysis by Angrist et al compares 150 learning interventions using LAYS (Learning-Adjusted Years of Schooling) and concludes that Structured Pedagogy and Targeted Instruction are among the most impactful and cost-effective interventions.⁵ They show the full set of learning gains from all the studies (from 46 countries). The effect sizes for Structured Pedagogy are between -0.2 and 1.04 SD. Somewhat similarly, the effect sizes for Targeted Instruction are between 0 and 0.8 SD. Both interventions show a similar dispersion.

What Role Does Implementation Quality and Program Design Play in Effect Sizes?

Effect sizes are much larger in some Structured Pedagogy programs than others. RCTs exploring this question have identified several factors that drive effect size, including the importance of providing structured lesson plans and coaching, higher dosages of training and ongoing teacher coaching, and high quality implementation of coaching follow up. Below, we highlight some RCTs that explore these questions:

Northern Uganda Literacy Project (Uganda)

A randomized evaluation in Northern Uganda of the Northern Uganda Literacy Project (NULP), which aimed to increase reading skills in local languages, implemented by local education company Mango Tree,⁶ showed large learning gains. One permutation of the program was implemented with high quality and intensity (21 days of teacher training and 16 teacher coaching visits in the first year). **This achieved an improvement of 1.35 SD in local language literacy scores (at a cost of \$15 per year per student)** - among the largest effect sizes ever studied. NULP also experimented with another program permutation aimed at replicating an approach more similar to a government scale up: a lower-intensity version where they reduced

⁴ Jonathan Stern, Matthew Jukes, Benjamin Piper, Joseph DeStefano, Jessica Mejia, Peggy Dubeck, Bidemi Carrol, Rachel Jordan, Cosmus Gatuyu, Tabitha Nduku, Maitri Punjabi, Christine Harris Van Keuren, Fatima Tufail. 2021. [Learning at Scale: Interim Report](#). USA, RTI International.

⁵ Noam Angrist, David K. Evans, Deon Filmer, Rachel Glennerster, Halsey Rogers, Shwetlena Sabarwal. 2020. [How to Improve Education Outcomes Most Efficiently? A Comparison of 150 Interventions Using the Learning-Adjusted Years of Schooling Metric](#). World Bank Group.

⁶ Julie Buhl-Wiggers, Jason T. Kerwin, Ricardo Montero de la Piedra, Jeffrey Smith, and Rebecca Thornton, 2023. [Reading for Life: Lasting Impacts of a Literacy Intervention in Uganda](#).

the training days and the number of coach visits. **With the lower intensity program, they achieved an improvement of 0.73 SD in local language (at a cost of \$6 per year per student) - still a very large impact.** The scale-up of the light-touch intervention maintained impact. This study shows rigorous evidence that quality implementation of structured pedagogy with high dosage can lead to an extremely large impact, and that both a high-intensity and a lower intensity version of the program can achieve high cost-effectiveness

Additional papers on the NULP study explored variation in student and teacher performance. They found extremely high levels of variation among student performance. While this paper reports the average effect size of the program as a massive 1.4 SD increase in test scores, there is more than 1 SD of differential effect among students, suggesting that some students' performance is only mildly improved by the program, while many other students improved dramatically. It is not readily apparent which students will benefit the most, as some who most benefited were lower performers at baseline, while others were higher performers.⁷ While most studies only reported average effect sizes, these results highlight the value of maintaining student-level data for understanding and potentially addressing disparities in student learning outcomes.

Another NULP paper explored the question of variation in teacher performance, finding that the lowest performing teachers at baseline achieve much less increase in their teaching effectiveness as a result of the Structured Pedagogy intervention compared to average or higher performing teachers.⁸ This finding suggests that teachers need to have a basic level of motivation and/or competence in order to benefit from this sort of pedagogy-improving intervention - it's unlikely to help teachers who just aren't trying.

Tusome and PRIMR (Kenya)

Tusome was a nationwide, evidence-based structured pedagogy intervention in Kenya conducted across several years (from 2015 to 2021). The program is composed of four main "ingredients": teacher guides with daily lesson plans for English and Kiswahili, student textbooks matching the teacher guides, training to increase teacher capacity deployed through a cascade model, and teacher support through government coaches (CSOs) in public schools and program coaches in low-cost private schools. The effect sizes in Oral Reading Frequency between baseline and midline were around +0.7 standard deviations⁹. Between baseline and

⁷ Julie Buhl-Wiggers, Jason T. Kerwin, Juan S. Muñoz-Morales, Jeffrey A. Smith, and Rebecca Thornton, 2021. [Some Children Left Behind: Variations in the Effects of an Education Intervention](#). NBER Working Paper, Cambridge, MA, USA.

⁸ Julie Buhl-Wiggers, Jason T. Kerwin, Jeffrey A. Smith, and Rebecca Thornton, 2023. https://drive.google.com/file/d/1ImkKLxiY1e-2YsQEsM3VxEftJ4SmTVHZ/view?usp=drive_link Learning More About Teachers: Estimating Teacher Value-Added and Treatment Effects on Teacher Value-Added in Northern Uganda. NBER Working Paper, Cambridge, MA, USA.

⁹ Freudenberger, Davis, 2017. [Tusome External Evaluation - Midline Report](#). Management Systems International (MSI).

endline, evaluators estimate that the reading gains in english are equivalent to students attending one more year of schooling¹⁰.

Tusome evolved from a pilot program implemented by RTI in Kenya between 2011 and 2014, the Kenya Primary Mathematics and Reading (PRIMR) Initiative. It was a randomized evaluation that compared three treatment arms with different components and a control group: (1) teacher training and teacher support through coaches, (2) revised student textbooks + teacher training and teacher support through coaches and (3) structured teacher lesson plans + revised student textbooks + teacher training and teacher support through coaches¹¹. The evaluation found a positive impact on English and Kiswahili reading as well as on mathematics. The arm including all the program ingredients showed the largest learning gains. Despite its increased cost, it was also the most cost-effective option. Tusome built on these results and designed the program accordingly.

Tusome is a key piece of evidence showing that large-scale, nationwide literacy programs can have large impact on learning gains when implemented properly and consistently over time.

Monitoring Teachers and Changing Teaching Practice (South Africa)

Another randomized evaluation tested the impact of providing lesson plans + training or lesson plans + coaching in South Africa. Both appear to be more effective than control, but impact for lesson plans + training is 0.11 SD (not significant), and impact for lesson plans + coaching is 0.25 SD (significant). Impact is MUCH larger in urban schools (+0.42 SD in training, and +0.75 SD in coaching). The authors calculate a 0.41 SD increase in reading scores per \$100 spent.

Read Liberia (Liberia)

Read Liberia was an RCT in 90 government schools (45 treatment, 45 control) that tested learning gains of a traditional structured pedagogy USAID intervention (teacher and principal training, student textbooks, teacher guides and coaching). The trainings were carried out successfully, with high attendance and coach visits to schools. However, teacher adoption of the materials and pedagogy was very low. Despite this, they see a significant improvement in reading fluency: in control schools, the correct words per minute were 14.5, whereas in treatment schools, it was 29.7. This represents an improvement of 0.6 standard deviations. In Liberia, the MoE guidelines specify an ORF benchmark of 35 words per minute for the learners at the end of grade 2. In control schools, the students reaching the benchmark were around 13.8%, whereas in treatment, they were around 35.5%. The authors were uncertain how to account for these gains, as the implementation fidelity was extremely low, as is highlighted in the next section.

¹⁰ Keaveney, Fierros, Rigaux, Menendez, 2020. [Tusome External Evaluation Endline Report](#). NORC at the University of Chicago.

¹¹ Piper, Zuilkowski, Dubeck, Jepkemei, J. King, 2018. [Identifying the essential ingredients to literacy and numeracy improvement: Teacher professional development and coaching, student textbooks, and structured teachers' guides](#). Elsevier.

Cross-country studies of implementation quality

Structured Pedagogy

A 2024 qualitative observational study of large scale government implementation of Structured Pedagogy in Liberia (Read Liberia) and Uganda (LARA project) explored what happened in two implementations that did not achieve large learning gains. The study found that teachers shifted the topics they taught in response to the new curriculum, but that there were major gaps in how these teachers implemented the new approach¹². For example, some of the teachers providing phonics instruction did not appear to know all the sounds that letters are supposed to make, and thus were teaching incorrect sounds to students. Likewise, many teachers continued to use prior pedagogical approaches like lecture and repeat, showing particular reluctance to implement student-centered practice activities that were included in the lesson plans. These implementation issues highlight the critical importance of strong training and coaching practices to achieve the necessary knowledge and substantial behavior changes that the Structured Pedagogy lesson plans incorporate.

Targeted Instruction (meta-analysis)

An analysis of Targeted Instruction implementation quality draws a similar conclusion, suggesting that the effect size in pedagogy-focused interventions is heavily dependent on the quality of implementation, perhaps because substantial effort and behavior change is necessary from teachers in order to achieve the desired effects. A meta-analysis from Angrist and Meager combines the evidence from 5 Targeted Instruction RCTs. They find that there is high variation in learning gains (from 0.07 to 0.78 SDs), depending on both who is teaching and program implementation. They find that the programs that use teachers have an impact of 0.07 SDs, while the programs that use volunteers have an impact of 0.24 SDs, suggesting that volunteers implement better.

They also study questions about the role of effective implementation, which they define according to 2 variables: program 'takeup', measured by teacher attendance and presence of classroom materials; and "fidelity", which measures whether the teacher implemented the program fully, including assessing the students' learning level, grouping the students by their level, and then instructing them at their level. They report the effect size conditional of the intervention being effectively implemented, and the results are notably bigger: The effect of Treatment On the Treated (TOT) for teachers is 0.21 SDs, and Treatment On the Treated for volunteers is 0.76 SDs. The key takeaway here is that whether or not a program is fully implemented by the teacher or volunteer is a huge driver of impact in Targeted Instruction, similarly to Structured Pedagogy.

¹² Hoadley, Ursula, 2024. "[How do structured pedagogy programmes affect reading instruction in African early grade classrooms?](#)" International Journal of Educational Development, 107.

Does improved educational outcomes lead to individual and/or national economic outcomes?

Given the long time delay between primary education and adult income, the evidence base is somewhat limited. However, we can cite several studies that address this question. The most-cited meta-analysis studying the effect of skills on income suggests that 1 standard deviation increase in test scores leads to a 10-45% increase in individual income.¹³ Additionally, there are two high-quality RCTs showing that education access improves monthly income years later, by 24% in Ghana¹⁴ and by 8% in Colombia¹⁵.

Regarding additional societal benefits, several macroeconomic studies^{16,17,18} show a positive correlation between quality of education and economic growth. The authors argue that it is the quality of education that causes growth.

See [this evidence review](#) by Ambitious Impact on long term returns to education for a more comprehensive treatment of these questions.

Comparing Cost-Effectiveness to other income-generating interventions

An evidence review from Founders' Pledge estimates the cost-effectiveness implications of investments in learning improvements compared to the impacts of cash transfers. They estimate that an intervention such as Imagine Worldwide adaptive digital exercises in Malawi, which has an impact on test scores of approximately 0.2 standard deviations for \$15 per year, would be **11 times as cost effective as GiveDirectly's cash transfer program at improving wellbeing through higher income**¹⁹.

Conclusion

¹³ Hanushek, Eric A. & Schwerdt, Guido & Wiederhold, Simon & Woessmann, Ludger, 2017. "[Coping with change: International differences in the returns to skills](#)," Economics Letters, Elsevier, vol. 153(C), pages 15-19.

¹⁴ Duflo, Esther and Dupas, Pascaline and Kremer, Michael, [The Impact of Free Secondary Education: Experimental Evidence from Ghana](#). 2017. University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2021-73.

¹⁵ Bettinger, E.P., Baker, R. (2014). [The Effects of Student Coaching An Evaluation of a Randomized Experiment in Student Advising](#). Educational Evaluation and Policy Analysis, 36(1), 3-19.

¹⁶ Hanushek, Eric, A., and Dennis D. Kimko. 2000. "[Schooling, Labor-Force Quality, and the Growth of Nations](#)." American Economic Review, 90 (5): 1184-1208.

¹⁷ Hanushek, E.A., Woessmann, L. [Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation](#). J Econ Growth 17, 267–321 (2012).

<https://doi.org/10.1007/s10887-012-9081-x>.

¹⁸ Hanushek, E.A., & Woessmann, L. (2020). [Education, knowledge capital, and economic growth](#).

¹⁹ Vadim Albinsky. Effective Altruism Forum 2023. [Are education interventions as cost effective as the top health interventions? Five separate lines of evidence for the income effects of better education \[Founders Pledge\]](#).

Given the evidence above, we conclude that it is very likely that a well-implemented Structured Pedagogy program can achieve a very large impact, with extremely high cost-effectiveness, comparable to the most cost-effective foundational learning interventions in the world, including Targeted Instruction. Effective teacher training and coaching at sufficient dosage appear to be critical drivers of lesson plan adoption and implementation quality by teachers, and should be central to the design of any new program. Finally, we briefly discuss the economic impact of educational outcomes, and despite limited evidence, we believe it is likely that increased learning leads to higher income later in life, as well as eventual economic growth.