OVERVIEW - IMPACT EVALUATION 2025

WHAT

- In 2024, an evaluation of STiR's programmes in Indonesia and Uganda was carried out by Deloitte, supported by Professor Nishith Prakash of Northeastern University.
- The study evaluated STiR's programmes through a mixed-methods approach, collecting data directly from students to assess learning outcomes and social-emotional learning. The study also estimates impact on teacher and official behaviours and mindsets.
- It employed a quasi-experimental design to compare randomly selected treatment and control groups, using statistical techniques to control for variables that might influence student learning (eg student demographics, family characteristics, and economic factors) in order to estimate the causal impact of STiR's programming. Additionally, a statistical robustness check was conducted to further strengthen causal inference and ensure the reliability of the findings.

WHY

- The quasi-experimental design was chosen to ensure a balance of rigour and pragmatism.
- While not as rigorous as a randomised controlled trial (RCT) this design allowed us to assess our pre-existing programmes. We randomly selected schools at the district level, amongst treatment (districts where the STiR programme exists) and control (where the programme does not).
- We also recognised the need to produce some quantitative evidence quickly. A criticism of STiR has been our lack of data on student learning, and this approach allowed us to address that in a shorter time frame than an RCT.
- Deloitte were selected as the partner after an open tender process, assessed according to pre-defined criteria by a committee within STiR.

FINDINGS

This section highlights key findings at the student and teacher level. A summary of key statistics across all research questions can be found here. The full Indonesia and Uganda reports can be accessed here and here respectively.

Foundational learning

- The evaluation found a statistically significant impact of STiR's programming on literacy and numeracy in Uganda, and literacy in Indonesia.
- It also found relatively minor impact on numeracy in Indonesia, but this was not statistically significant – indicating that the observed difference could be due to chance and may need further research to estimate the causal effect of STiR's intervention.

Impacts as expressed as both standard deviation and percentage effect sizes can be seen in the following table:

Country	Outcome	Statistically Significant	Impact (in % change)	Effect Size (in terms of standard deviation change)
Indonesia	Numeracy	No	4.4	0.06
Indonesia	Literacy	Yes	16.4	0.16
Uganda	Numeracy	Yes	14.2	0.17
Uganda	Literacy	Yes	15.6	0.13

Social Emotional Learning

- The evaluation found limited impact on social emotional learning in Indonesia and Uganda.
- Overall, STiR's interventions appear to enhance peer collaboration among students in Indonesia and conflict resolution skills among students in Uganda. However, the impact on empathy and emotional response is less clear, indicating a potential area for further exploration and programme refinement.

Impact on teachers

- In Uganda, the evaluation found that teachers were more likely to be engaging in effective professional development behaviours, such as peer learning, coaching and observation.
- The evaluation found no correlation between self-reported intrinsic motivation and student learning.
- However, it found correlations between the predictors of intrinsic motivation (a sense of autonomy, mastery and purpose) and changes in teaching behaviour and student learning.

INTERPRETATION

- From our perspective, the impact on foundational learning is very positive. While the effect sizes are in the 'modest' range for educational interventions, they are high for programmes that run at scale (we reach over 100,000 teachers in Uganda and Indonesia combined) and at low cost, given our overall cost per child per year of under $1.
- We hypothesise that our programme amplifies the priorities of the system. In both countries, there is a greater systemic focus on foundational skills as opposed to SEL, which would explain why there is greater impact on the former.
- The findings on motivation appear counterintuitive, but are consistent with wider research on motivation, that is that self-reported motivation often does not correlate with positive outcomes, whilst the presence of autonomy, mastery and purpose often does.
- The wider field however is unclear on what are the 'best bets' in terms of practical things we can do to promote autonomy, mastery and purpose.
- We think this calls for a more practically-informed 'motivation science 'that can help us cut through some of the conceptual complexity surrounding motivation to focus on helping ourselves and others identify the most useful things we can do to build motivation at scale.
- We believe the practical mechanisms of our programme – including peer learning and observation, supporting teachers to make sense of new material, and providing support for teachers to make decisions about how to integrate learning into their classrooms – provide a strong foundation to support others to integrate the principles of motivation science into their work.