

A Mixed-Methods Longitudinal Study¹

CONCERN
worldwide

ENDING
EXTREME POVERTY
WHATEVER
IT TAKES

The Safe Learning Model



An image taken by a female class two student during a photovoice exercise
'If you are late, teachers will flog you' February 2020

Research Objectives

- Explore the socio-cultural context of children's lives and the impact on their learning
- Consider gender dynamics in children's lives, including in their learning
- Consider children's well-being and changes in this over time
- Explore the everyday practices of children (including attendance), their teachers and principals in primary schools
- Track children's progress in literacy over time
- Over time to assess the impact of the Safe Learning Model on children's wellbeing, literacy, gendered experiences, participation in education and the attitudes of teachers, principals, and community stakeholders.



Context of the Research

Developed by Concern Worldwide, the 'Safe Learning Model' adopts a holistic approach to the education of children in extreme poverty, bringing together interventions to prevent and respond to School-Related Gender-Based Violence (SRGBV) at multiple levels of the socio-ecological model within an education programme, in order to realise sustainable improvements in children's literacy, wellbeing and gender equality in schools and communities. This approach to learning aimed to counter the issues of SRGBV and gender stereotypes, which can keep children from learning and cause them to drop out of school. The model was piloted (2017/18) and implemented (2018-2021) in Tonkolili district of Sierra Leone based on the assumption that;

“children's educational progress will be enhanced when they live in communities that are underpinned by support for gender equality and children's wellbeing.”

The model consisted of three core components; Literacy Activities, School Level SRGBV Prevention & Response, and Community Level SRGBV Prevention & Response. This brief presents a description of the methodology of the three-year mixed-methods longitudinal study conducted by University College Dublin (UCD) School of Education to evaluate the effectiveness of this integrated model, providing an overview of the design, methods and analytical approach. Additional thematic briefs that include more detailed information on the Safe Learning Model and interpretation of the key findings as well as the overall executive summary of results are available².

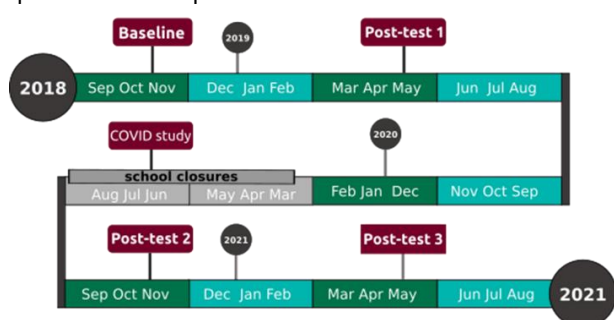
¹ This briefing is based on original research undertaken by Devine, D; Sugrue, C; Symonds, J; Sloan, S; Kearns, M; Bolotta, G; Samonova, E; Capistrano, D Smith, A and Gibbons, R; University College Dublin, School of Education.

² More information on the Safe Learning Model and study is available at <https://www.concern.net/knowledge-hub/education-safe-learning-model-research> or contact: Amy Folan, Senior Education Advisor, Concern Worldwide amy.folan@concern.net

Methodology

Study Design

The Safe Learning Model research was conducted in the Tonkolili district of Sierra Leone, from September 2018 to June 2021 with annual approval by the University College Dublin Human Research Ethics Committee. To inform the full study, a pilot study was conducted from September 2017 to May 2018. The pilot phase allowed Concern to develop effective methods of delivering the Safe Learning Model as an integrated programme. It also enabled the UCD research team, working in partnership with Concern, Ministry of Education, Science and Technology (MEST), Ministry of Social Welfare, Gender and Children’s Affairs (MSWGCA) and local research partner NestBuilders International (NBI), to explore the feasibility of evaluating the intervention, gain insight and understanding of children’s everyday lives in Sierra Leone, and develop research tools and protocols that can be used during a full-scale evaluation of the Safe Learning Model. The full study drew on data gathered using both quantitative and qualitative research methods. This mixed methods approach maximised potential strengths of both qualitative and quantitative methods, allowing the researchers to explore diverse perspectives and answer research questions that neither quantitative nor qualitative methods could answer alone. Outcomes were measured at baseline (November 2018). Post-test data collection was planned for three time points at the end of the school year (May/June) in 2019, 2020 and 2021. Post-test 1 was completed in May/June 2019 but post-test 2, planned for May/June 2020, was postponed and completed in November 2020 due to school closures arising from the COVID-19 pandemic. A COVID sub-study was conducted during this time to measure the impacts of the pandemic and school closures on children’s education and wellbeing. Post-test 3 took place in May/June 2021.



Post-test data collection was planned for three time points at the end of the school year (May/June) in 2019, 2020 and 2021. Post-test 1 was completed in May/June 2019 but post-test 2, planned for May/June 2020, was postponed and completed in November 2020 due to school closures arising from the COVID-19 pandemic. A COVID sub-study was conducted during this time to measure the impacts of the pandemic and school closures on children’s education and wellbeing. Post-test 3 took place in May/June 2021.

Quantitative Methods

The quantitative component of the study was designed as a cluster randomised controlled trial with four arms (a control group and three intervention groups). The control arm received a standard package of livelihoods and health services that were available within all participating communities irrespective of group allocation as part of the wider integrated programme. The intervention arms received different levels of the model, for further details of each component see the Safe Learning Model Briefing Document.

Control	Treatment 1 No SRGBV	Treatment 2 Light SRGBV	Treatment 3 Full SRGBV
25 Communities	25 Communities Literacy	25 Communities Literacy + School Level SRGBV	25 Communities Literacy + School Level SRGBV + Community Level SRGBV

Participants

Five Chiefdoms in Tonkolili were selected for implementation of model. Enrolment data from all primary schools in these Chiefdoms was collected and schools were screened for eligibility to participate, using the following criteria:

1. The school offered the full range of primary grades (classes 1-6)
2. The school had sufficient enrolment numbers (more than 15 children in Class 3 and on average at least 10 children in Classes 4-6)



Many communities had only one primary school, however in a small number of larger communities with multiple primary schools that met the inclusion criteria, one school was selected to avoid potential contamination between trial arms. In these communities, schools were eligible to be recruited if they were unapproved with at least one teacher on payroll. One hundred primary schools were recruited to participate in the research. This involved meeting with the school head teacher and community leaders who were asked to sign a memorandum of understanding.



Within each school, the target cohort were children entering Class 1 in September 2018, the first year of compulsory education in Sierra Leone. Prior to baseline data collection, information sessions for caregivers were held in each school in October 2018, where the research was discussed with caregivers in the local language (Krio, Limba, Temne or Kuranko). A list of names of children whose caregiver agreed for their child to participate was collated. Participants provided informed consent for their interviews to be audio-recorded and data to be used and published in the study.

Randomisation

To minimise the risk of contamination between trial arms, allocation to one of the treatment groups was at the school level. Schools were randomly allocated by the evaluation team to one of four arms, in May 2018 using the program Minim2 to create groups that were balanced across pre-specified characteristics, in this case:

- Chiefdom
- School status (approved or not approved)
- School size



Allocation based on Chiefdom was used to avoid one area having very few intervention schools and another too many. School status as approved or not approved was used because this impacts whether a school receives support from the Government including teacher salaries, teaching, and learning materials and subventions based on student population. While not all teachers in Government approved schools are paid and qualified, schools that are not approved do not receive any form of financial or in-kind support from the Government resulting in an increased number of unpaid and unqualified teachers and decreased access to teaching and learning materials, which in turn negatively impacts on the quality of teaching and learning in unapproved schools regardless of interventions being carried out. School size was also used to ensure there were a similar proportion of small, medium and large schools in each arm. The size of the school strongly impacts the viability of the school as well as the quality of teaching and learning. Small schools are often susceptible to closure, lower teacher attendance, with many classes combined resulting in multi-grade structures that are under-resourced. Subventions are typically based on per student ratios; however, they are inadequate to cover the basic needs of education at a per student level. Therefore, the larger the school population the more likely the school will be able to pool subvention funds to increase access to teaching and learning materials as well as water and hygiene facilities. School size was based on the total number of children enrolled in Classes 1-3, with schools categorised into three equally sized groups representing small (less than 105 children), medium (105 – 151 children) and large (more than 151 children) schools within the sample.

Measures



Literacy was measured using the Early Grade Reading Assessment (EGRA), one of the most widely used global literacy assessments available for use in the primary grades in low and middle-income countries, designed to measure foundational literacy skills through a collection of subtasks. The assessment consisted of six subtasks: letter identification, letter sounds, familiar words, invented words, reading passage and comprehension. The EGRA was orally administered with children in English (although the instructions were delivered in children's mother tongue to ensure understanding), by trained enumerators (blind to allocation status) on a one-to-one basis. Each assessment took approximately 20 minutes to complete. Children were timed in all tasks, with performance evaluated in terms of correct responses per minute.



Wellbeing was evaluated using subjective measures, where respondents' own perception of wellbeing was measured. Psychological wellbeing was estimated in relation to a child's general levels of 'feeling good' and 'doing well'. Measures of 'feeling good' and 'doing well' were conceptualised and developed during the pilot phase, accounting for four indicators of 'doing well' (autonomy, competence, resilience and helping) and four indicators of 'feeling good' (care, acceptance, safety, and happiness). The resulting instrument, the Child and Adolescent Personal and Social Assessment of Wellbeing (CAPSAW), is a 32-item measure of wellbeing across four dimensions: a global dimension, and dimensions of experience with peers, teachers, and family. Each dimension includes eight items capturing different aspects of 'feeling good' and 'doing well'. Physical wellbeing was estimated by accounting for student responses to questions on whether they had enough food, enough to drink, whether they feel healthy, and how often they miss school due to illness.



The research adopted a multidimensional construction of violence, taking into account three different dimensions: 'direct violence' that relates to experiences of physical violence (such as being beaten), and psychological violence (such as name calling and verbal threats); 'structural violence' that accounts for child suffering on account of a lack of resources (such as hunger and missing school due to work); and 'cultural violence' which refers to cultural norms that perpetuate both direct and structural violence in society (for instance gender inequality and high levels of poverty and wealth inequalities). Structured questionnaires were administered verbally with children on a one-to-one basis. Children were asked to indicate their experience of direct violence (being beaten or physically hurt, being whipped, or caned, or being called bad names, teased, or threatened verbally), including who had hurt them by responding 'yes' or 'no' to a set of options: nobody, someone from your family, a teacher, an adult in my village/town, a boy child or youth, and a girl child or youth.

Attrition

The study included a total of 3,118 children at baseline, however, only 1,403 children were present at each of the four waves of the data collection. While attrition from the sample, evaluated as a respondent not being present in all four waves, is high with over half the sample missing at least one wave of data collection, characteristics such as language, religion, socio-economic status (SES) and household make-up of the children in the full sample and in the sample present for all data collection points are very similar, with the only exception being that females are over-represented in the common sample (48%) relative to the full sample (43%). This indicates that, except for sex, attrition in the sample is predominantly random, meaning it is not necessary to re-weight the data. To evaluate if characteristics of the sample could predict attrition, four logistic regressions were run to calculate the odds ratios of the student missing at least one wave of data collection, with each model holding measures that did not change over time (eg: intervention group and sex), while also adjusting for the children's self-reported frequency of missing school at each wave. The results indicate that only when adjusting for missing school for work at baseline is sex statistically significant, suggesting that boys are likely missing school in subsequent waves owing to work commitments outside of school.

While all intervention effect models were run on the full sample in line with an Intention-To-Treat approach, models were also run on the sample that was present for all waves of data collection in order to evaluate if results were consistent with different sample specifications and results remained consistent.

Analytical Approach

Quantitative analysis of data was undertaken using an Intention-To-Treat (ITT) approach. This approach estimates the effect of the intervention based on the intervention assigned (initial randomisation), irrespective of whether the children in such groups received the full intervention (i.e., were present in school throughout the duration of the programme or at every wave of data collection).



To evaluate causal effects, longitudinal analysis of covariance was estimated where differences between the control and intervention groups that might occur through random fluctuations and measurement error, could be accounted for by adjusting for the baseline value, enabling a more accurate estimate the intervention effects. While this provides the statistical significance of an intervention, it does not readily inform of the practical significance of the result which can be provided by measuring the effect size. As a rule of thumb, an effect size of 0.2 can be interpreted as a small effect, a value of 0.5 as medium, and 0.8 as a large effect. In certain instances, the outcome variable is measuring two opposing points, either where the variable was measured in such a format (such as the Violence subscale), or where the variables are skewed to the extent as to invalidate assumptions of normal distribution, as with some of the EGRA subtasks. In such scenarios, logistic mixed effects models were employed and intervention effects were measured using odds ratios.

Lastly, longitudinal mixed effects models were run, utilising data from all four waves, to better understand the effect of the interventions over the full duration of the study. Three models were run for each outcome variable:

- Model 1: Month + Intervention received (baseline: Control)
- Model 2: Model 1 + Sex + Sex x Intervention
- Model 3: Model 2 + SES + SES x Intervention



Model 1 outlined the raw intervention effects observed, adjusted only for time. Model 2 added sex and an interaction between sex and the intervention group to identify if there were differences by sex in the efficacy of the intervention. Lastly, Model 3 adjusted for SES and an interaction between the SES of a student and the intervention group they were in, to evaluate whether this affected the efficacy of the intervention received.

Finally, to be consistent with findings presented in interim reports on the impacts of the intervention after 1 year and 2 years, the study also reported effects of the three intervention group combined, relative to the control group. However, certain caveats should be noted in interpreting findings for the combined intervention groups. Firstly, combining the intervention groups increases the sample size in these analyses, which increases the likelihood of results being statistically significant. Furthermore, each group had different intervention programmes, and it is not possible to evaluate which parts of the different intervention programmes are responsible for any intervention effects observed in the combined intervention groups. As such, in order to evaluate the intervention effects, it is best practice to look at the individual intervention group effects.

Qualitative Methods

Qualitative data was collected over the period of three years (2018-2021). In total, four waves of data collection were conducted, including an additional sub-study on the impact of the COVID-19 pandemic and school closures. The fieldwork took place in four case study communities that were selected based on the following criteria:

- Coverage of all arms of interventions and control schools
- Linguistic diversity
- Coverage of different types of schools (e.g., faith-based schools (Christian, Muslim), community schools, etc.)
- Proximity to Magburaka.

Interviews and Focus Groups

During year one of the study, researchers selected four case study families of class 1 children in each selected community (in total eight families of boys and eight families of girls - in year 2 one of the families in group one was replaced) and conducted in-depth semi-structured interviews with children, parents/caregivers, grandparents and an older sibling. During the next rounds of the fieldwork,

follow up interviews with case-study families were conducted, supplemented by focus group interviews with children in the case study classes and interviews with teachers and community leaders. A total of 376 semi-structured interviews and group discussions were conducted. Interviews and group discussions were transcribed in English, and transcripts were coded to identify the most recurrent themes. Codes were then clustered to generate a frame and to examine the relationships between different clusters of codes allowing for the identification of an intergenerational structure of themes. In total 9 themes, 300 major codes, and 28,000 sub-codes were generated, highlighting how issues relevant to the overall study such as education, gender, and wellbeing were differently interpreted and framed by children, parents, and elders.

Group 3- Full SRGBV	Group 2- Light SRGBV	Group 1- No SRGBV	Control
4 case study families	4 case study families	5 case study families	4 case study families
Interviews with school principal + teachers	Interviews with school principal + teachers	Interviews with school principal + teachers	Interviews with school principal + teachers
Group discussions with chief, SMC, elders, youth	Group discussions with chief, SMC, elders, youth	Group discussions with chief, SMC, elders, youth	Group discussions with chief, SMC, elders, youth
Group discussions with participants of school clubs/SEL	Group discussions with participants of school clubs/SEL	-	-
Structured observations in classrooms	Structured observations in classrooms	Structured observations in classrooms	Structured observations in classrooms
Photovoice with children	Photovoice with children	Photovoice with children	-

Systematic Classroom Observations

In years two and three, researchers conducted systematic observations of classroom practice using the Classroom Assessment Scoring System (CLASS) and Observational Research and Classroom Evaluation (ORACLE) tools. In 2020 observations were conducted in person, however in 2021, due to the COVID-19 pandemic and travel restrictions, lessons were video recorded and later analysed remotely. In total, over 42 hours of structured observations were conducted over the course of two years.

Photovoice



An additional participatory method called photovoice was used. Children were given digital cameras and asked to take pictures of the things they like and do not like in school. Afterwards pictures were discussed with each child individually providing deeper insights into everyday life from their perspectives and potentially introducing new topics that otherwise might have been overlooked or poorly understood from an adult viewpoint.



This briefing is based on original research undertaken by University College Dublin- School of Education. The Safe Learning Model intervention and research was funded by a grant from the Irish Government and implemented in partnership with the Sierra Leone Ministry of Basic and Senior Secondary Education, however the content within this briefing is entirely the responsibility of the authors and does not necessarily represent or reflect Irish Aid Policy.

CONCERN
worldwide

Amy Folan
Senior Education Advisor
Strategy, Advocacy and Learning Directorate
amy.folan@concern.net