



Evaluation of Parental Engagement SMS Intervention in Lilongwe Rural

Pre-Analysis Plan

IDinsight

**Evaluation of Parental Engagement SMS Intervention in
Lilongwe Rural
Pre-Analysis Plan**

July 2022

Authors

Jack Thunde: Jack.Thunde@IDinsight.org

Valentina Brailovskaya: Valentina.Brailovskaya@IDinsight.org

Arthur Chibwana: Arthur.Chibwana@IDinsight.org

Leonard Oruko: Leonard.Oruko@IDinsight.org

Mtise Mwanza: Mtise.Mwanza@IDinsight.org

Acknowledgements

About IDinsight

IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers.

IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Dakar, Lusaka, Manila, Nairobi, New Delhi, Rabat, and Remote. Visit www.IDinsight.org and follow on Twitter [@IDinsight](https://twitter.com/IDinsight) to learn more.

Contents

1. Background Information	4
2. The Evaluation	6
2.1 Design Overview	6
2.2 Sampling and Randomization Procedure	7
2.3 Power Calculations	7
2.4 Exclusion criteria and sample selection	8
3. Data Collection and Timelines	9
3.1 In-person school visits	10
3.2 Phone survey	10
4. Outcomes	12
4.1 Primary Outcome	12
4.1.1 Attendance:	12
4.2 Secondary Outcomes	13
4.2.1 Child time spending:	13
4.2.2 Child-parent engagement:	13
4.2.3 Learning outcomes:	13
4.2.4 Children Motivation:	13
5. Analytical Model	14
6. Limitations & Mitigation Strategies	17
6.1 Threats to validity	17

1. Background Information

GiveDirectly is a nonprofit organisation that runs social cash transfer programs targeting the world's poorest households. The organisation believes people living in poverty deserve the dignity to choose how best to improve their lives, and money enables people to have this choice. Cash allows individuals to invest in what they need instead of relying on aid organisations and donors to choose for them¹. During the Covid-19 pandemic, GiveDirectly targeted over 4,000 vulnerable urban households in Mgoni area of Lilongwe, providing each household three monthly transfers of \$47 to help them combat the effects of the pandemic.

Movva is a tech organisation that offers smart communication for educational engagement. The organisation combines behavioural economics and artificial intelligence to improve educational indicators². Their interventions have increased family engagement in talking about school activities by 59%, impacting over 2 million students and families worldwide and during the Covid-19 pandemic reducing truancy by 16%³. The SMS-based intervention will consist of messages containing general information about the value of parental engagement in child education. It will be sent out to guardians of school-going children twice a week.

IDinsight is a mission-driven global advisory, data analytics, and research organisation that helps global development leaders maximise their social impact. We tailor a wide range of data and evidence tools, including randomised evaluations and machine learning, to help decision-makers design effective programs and rigorously test what works to support communities.

IDinsight, GiveDirectly and Movva are partnering to rigorously evaluate a Movva SMS-based intervention aimed at increasing parental engagement to improve schooling outcomes for children of GiveDirectly recipients in Lilongwe Rural. The objective of this study is to assess the impact of the SMSs on child attendance and enrollment and non-school activities like child labour and parental aspirations for their children.

Several relevant studies suggest that the intervention may be a promising solution to improve schooling outcomes in Malawi. SMS interventions for

¹ <https://www.givedirectly.org/about/>

² <https://movva.tech/en/about/>

³ <https://movva.tech/en/nudgebots-to-change-study-habits/>

parents have been shown to create literacy gains for preschool children (York et al. 2019, Doss et al. 2019) and help parents track school progress (Bergman 2019). Increased parental engagement in their children's education through games and positive reinforcement has been shown to improve economic, psychological, and socio-emotional outcomes many years later (Gertler et al. 2021, Walker et al. 2021). Angrist et al. 2021 demonstrate that SMS messages with learning activities improve learning by 0.12 SD in Botswana. The intervention used by Movva has been evaluated in Brazil and Cote D'Ivoire⁴. In Brazil, the authors found that general text messages that increase the salience of children's outcomes bring about increases in attendance, test scores, and promotion rates (Cunha et al., 2017). The largest impacts⁵ are found for SMSs sent three times a week without requiring parents to respond. Delivery times or consistency of delivery times do not have differential impacts.

IDinsight was asked to conduct a randomised control trial (RCT) in Lilongwe Rural to understand the causal impact of the SMS-based intervention on children's educational outcomes across all ages. The evaluation will focus on impacts on child attendance and enrolment. The results of the RCT will inform decisions on whether to scale up the intervention to more districts and will contribute to the literature on how nudges can be used to influence behaviour change.

⁴ The latter is not yet available to the public

⁵ 0.33 SD on math scores, 6 percentage points on attendance (control group 87%), 3.2 percentage points on promotion rate (control group 0.93)

2. The Evaluation

2.1 Design Overview

The primary purpose of the impact evaluation is to understand the impact of Movva's SMS-based intervention on children's school outcomes. IDinsight will estimate the impact of the SMS intervention through a randomised control trial (RCT) across ~2600 GiveDirectly beneficiary households with children of school-going age. The RCT will take place in Lilongwe Rural, TA Khongoni. The estimated intervention impact will be measured after comparing outcomes between two groups: 1) Treatment - those that will receive SMS interventions 2) Control - those that will not receive SMS interventions.

The primary and secondary research questions for this evaluation are described below.

Primary question

- What is the impact of Movva's SMS-based intervention on GiveDirectly beneficiary children's school attendance?

Secondary questions

- What is the impact of Movva's SMS-based intervention on children's motivation to go to school?
- What is the impact on children's time use?
- What is the impact on parental engagement with children?
- What is the impact on student learning outcomes?
- What is the impact on re-enrollment?⁶

The impacts of the intervention will be quantified via 1) phone surveys conducted with parents and children 2) in-person attendance checks and "progression score" based on test score data collected from school records.

⁶ We have considered including enrollment as a primary outcome as well, however, ultimately decided against it for the following reasons 1) enrollment is likely to be nearly universal; in the pilot data, we see that the only children who are not enrolled are 17 and 18 years old and it's unlikely that the messages will influence enrollment for those children 2) we are surveying children mid academic year, and while parents do make enrollment decisions term-by-term, dropouts are likely lower during this time compared to between years. 3) We believe that the primary channel for impact through the intensive margin, rather than extensive margin on education attainment.

2.2 Timeline

The estimated timeline for project and evaluation activities is outlined below in Figure 1.

Sampling and randomisation are scheduled for July 2022 and September 2022. The intervention will take place between August and December 2022. The data collection will take place in December 2022-January 2023. The phone surveys will take place between December 2022 and January 2023; the school visits will be completed in January 2023.

Figure 1: Impact evaluation Gantt chart

Activity	2022												2023				
	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
Study Design																	
Sampling/Randomization					x		x										
Intervention																	
Intervention: SMS messages are sent						x	x	x	x	x							
Endline																	
Preparation for endline										x							
Data Collection: School attendance spot check (all classrooms in ~30 schools, geography of 15 schools is less known)												x					
Data Collection: Phone survey (40 min/hh)											x	x					
Data Cleaning/Analysis											x	x	x				
Report Writing													x	x			

2.2 Sample selection

The study will consist of all GiveDirectly beneficiaries living in Lilongwe Rural (TA Khongoni) who have shown interest in taking part in Movva's SMS intervention and have children of school-going age. GiveDirectly beneficiaries who do not agree to take part in the SMS intervention or do not have school-age children will be excluded.

Givedirectly will provide a list of beneficiary families with at least 1 child of school-age to IDinsight and who have agreed to be a part of the study. It is estimated that there will be roughly 2600 unique households. This sample will be used for randomization.

In case the population is larger than 2600, IDinsight will randomly select 2600 households.

The sample or the study population is unlikely representative of the broader population, since GiveDirectly works with the most vulnerable households.

2.2 Sampling and Randomization

GD baseline data shared with IDinsight contains more than 1 member of the household and in some instances those members report different children. This may either be driven by either different names used for children or split care responsibilities (the former is more likely). Before randomization, we will randomly select 1 member for each household to participate in the study and survey the children reported by the randomly selected household member. The treatment will be administered to one randomly selected household member.

We will randomise all households provided by GiveDirectly into treatment and control groups using Stata v17. Randomisation will be stratified on key variables to be determined after the receipt of the final baseline dataset from GiveDirectly. The randomization will be completed in 2 waves: 1) July and 2) in September to accommodate GiveDirectly's enrollment schedule. The randomization will be stratified by 1) age group of a randomly selected child 2) sex of the randomly selected child 3) distance to school (top/bottom 50th percentile), 4) whether a parent ever reports talking to a teacher (proxy for parental engagement). The randomization strategies will be identical in both waves.

2.3 Exclusion criteria

All children both enrolled and not enrolled at baseline will be included in the study.

Prior to randomization, we will exclude children that are not listed to be in the list of schools that GiveDirectly shared. This means that the children attend schools outside of the area of interest and we will not be able to collect attendance data for them (preliminarily about 3% of all children, 114 in wave 1). In addition, we schools which contained ≤ 5 children per school (24 children in wave 1) and those that were not expected to be saturated in wave 2. Those schools are relatively far from the areas which GD is serving.

Prior to collecting endline data, we will exclude household/child pairs that have

not lived together for the past 6 months since the children would not be exposed to the intervention. We devised the following protocols to exclude children who have not lived with the parents in the past 6 months.

1. The enumerator will have a roster with all households & children prior to initiating the first contact. The order of the children that show up in the roster will be randomised within a household.
2. Enumerator will be instructed to ask a parent “In the past 6 months, have [child1] lived with you for at least 5 days in a typical week?”. If the answer is “yes”, the child of interest will be the first randomly selected child. If the answer is “no”, the enumerator will ask the same question about [child2]. If no children lived with the parent/guardian in the past 6 months, the household will be dropped from the study and not interviewed at endline.

We anticipate that less than 10% of households will be excluded by this exclusion criteria, according to the findings in the pilot.

We do not believe this exclusion strategy from the study will generate different exclusion rates between treatment and control groups into the study, since it is unlikely that the short SMS intervention would affect whether a child lives with parent in the short run, in the middle of the school year. We will check whether this exclusion criteria generates different non-response rates between treatment and control groups. In case of differential attrition, we will apply Lee bounds to see the impacts on the estimates.

2.3 Power Calculations

Power calculation results are presented below in Table 2 for our 3 outcomes of interest (attendance, parental engagement and chores). Summary statistics presented in the table below are from the pilot with the population of interest conducted in May 2022. The following inputs were included in the power calculations:

- Correlation between the outcome value and baseline predictors (e.g. outcome at endline, sex of the child, distance to school) is assumed to be 0.2 (assumed)
- Power of the study is 0.8
- Significance level is 0.1 (we choose 0.1 instead of 0.05 relying on the assumption that the donors are willing to scale up the program with a higher threshold for false positives)
- Sample size is adjusted for 15% non-response at endline
- Randomisation is at the household (individual level)

Please note that we are 1) not accounting for power gains due to stratification 2) potential higher correlation between baseline predictors and baseline outcomes.

It's likely that these calculations are conservative.

Table 2: Results of power calculations for outcomes of interest (both primary and secondary)

Outcome	Inputs					Sample size	
	Control Mean	Control SD	MDE	Percentage MDE	Correlation with BL vars	Number of students per group	Total Number of Students (2 treatment groups)
Child attended school yesterday	0.73	0.43	0.05	6%	0.20	1300	2600
Number of days attended (out of 5)	4.46	0.72	0.07	2%	0.20	1300	2600
Engagement (SD)	0.00	0.98	0.10		0.20	1300	2600
Chores (hours)	1.82	2.28	0.24	13%	0.20	1300	2600

With the sample size of about 2600, we should be able to detect impacts between 2-13% depending on an indicator. For subgroup analyses (e.g. comparing longer vs shorter intervention period) that equally split the sample, we are powered to detect impacts of about 9% compared to one of the treatment groups.

3. Data Collection

3.1 In-person school visits

Shortly after the end of the intervention, in-person visits to collect data on attendance and test scores will be conducted in schools. During visits, we will ask teachers to show our team ledgers with historical attendance data of GiveDirectly beneficiary children. We will also ask teachers to show our team Class Progress books for end of term test scores.

3.2 Phone survey

Close to the end of the intervention, all research participants will be contacted and data will be collected through phone surveys. The phone survey will be roughly 30 to 45 minutes with both caregivers and children being interviewed. As some of the surveys will take place during the school term, interviews will

be scheduled with caregivers to find a time in which both they and their child would be available. Once calls have been scheduled, SMS reminders will be sent to the participants the night before and the morning of the call. If it so happens that the participants are unavailable at the agreed upon time, calls will be rescheduled up to 7 times.

Interviews will start by obtaining verbal consent from the caregiver for their and their child's participation. Enumerators will then build rapport with caregivers by asking them questions about their education, their aspirations for their child, how their child spends their time, their school attendance and academic progression. Upon establishing rapport, enumerators will ask the caregiver to hand the phone over to the child and will begin the child interview. Once the child interview is completed, the child will be asked to hand the phone back to the caregiver for the rest of the interview. The survey will contain the following sections:

- Parents:
 - Education level of parents
 - Aspirations for children
 - Children's time spending (yesterday)
 - School attendance (past 1 week)
 - Academic progression
 - Engagement with child's school life
 - Parental Child Discipline practices
 -
- Children:
 - School attendance (yesterday)
 - General engagement with parent
 - Engagement of parent with school life
 - Time spending (yesterday)
 - School motivation
- Parents (only asked from the random sample in the control group to validate the correlation between parent/child reports):
 - General engagement with child
 - Receipt/understanding of SMS message

4. Outcomes

4.1 Primary Outcome

4.1.1 Attendance

We considered two ways of collecting information on attendance (1) school records which we believe are accurate⁷, but likely subject to missing data (2) children's reports which may be less accurate but we have greater assurance that the data will be more complete.

1. School records:

In our pilot study, we were unable to find most children's records within schools. Qualitative evidence suggests that it's driven by 1) lack of parental knowledge of which grade the children belong to 2) mismatch in names given to the research team by parents vs those that are used in schools. At endline, we will devise a number of strategies that will improve our ability to find child records (such as looking for children across grades), however, it's still possible that children will not be found in school records and attendance data will have many missing observations.

2. Children reports about attendance yesterday

Collecting information on child attendance may be subject to reporting bias since some of the children will be very young, however, we believe that children's report on attendance to be the most reliable for impact quantification given constraints. First, unlike the school attendance data, we expect the child report data to be more complete since we have been successful at interviewing children during piloting. Second, parental reports may not be a reliable source of this information since previous (unpublished) research suggests that the intervention changes parental awareness of children's lives, therefore, they will know more about their lives compared to control parents⁸.

Unfortunately due to project and intervention time constraints⁹, some of our

⁷ The school attendance data kept in ledges is reliable according to [Baird et al. 2011](#). During piloting, we have found the information to be well-organised and well-kept.

⁸ Movva research team has learned that parental reporting of their engagement with children and knowing about children's lives is affected by the intervention, making comparison of parental reports between treatment and control groups invalid.

⁹ The report needs to be finalized in March and intervention won't be completed until December.

phone survey data collection will fall on exam/holidays during which the school will be closed. If that is the case, we will ask the child to think of the last day school was opened and record whether he/she was present on the last day.

At the analysis stages, we will construct an index of the following 2 variables:

- 1) Whether a child reports attending school yesterday (or the last day when the school was opened).
- 2) Percentage of days (out of 5) child attended school in the week of interest (according to school records)

We will construct an Anderson index (Anderson 2008)¹⁰ out of these two variables to create a final variable for the analysis. If the child is not enrolled, attendance will be set to zero. If the observations are missing for one of the outcomes, only 1 variable will be used. The determination of whether the intervention is successful or not will be made based on the estimate on this variable. The intervention will be considered successful if the estimate has a pvalue smaller than 0.1

The supplementary tables (where the impact estimates on 2 outcomes will be examined separately) will be used to assess the magnitude of the effect.

4.2 Secondary Outcomes

4.2.1 Child time use

We will present the results on total hours that the child spent helping parents yesterday (according to children's reports).

In supplementary tables, we will split the results by the type of chores they helped with (household chores, farming activities, non-farming business activities). The disaggregation is done in order to understand whether exactly children are freed from household chores which should give a more complete picture of the intervention's channel of impacts. The 3 supplementary outcomes of interest will include:

1. total hours of HH chores
2. total hours of farming activities

¹⁰ <https://are.berkeley.edu/~mlanderson/pdf/Anderson%202008a.pdf>

4.2.2 Child-parent engagement

We will construct a single index composed of two themes of child-parent engagement:

- 1) general engagement (e.g. playing, asking about feelings, reading)
- 2) school engagement (e.g. asking about HW, helping with HW etc).

We choose to construct indexes rather than reporting impacts on raw scores on each question in the survey to decrease the total number of outcomes which reduces complexity of interpretation of the results. The raw questions (ranging from 1-4) represent the frequency with which parents engage with their children *according to children's reports*. To construct a single measure, we use Anderson index which finds the weighted average of the reports across questions, and then standardise it against the control group (subtract mean and divide by standard deviation). The impacts will be expressed in standard deviations compared to the control group.

In order to understand whether impacts are driven by general or school engagement, we will supplement the main table with 2 sub-indexes constructed the same way.

4.2.3 Learning outcomes

During school visits, we plan on collecting administrative data on a "progression score" based on all tests that children take in December 2022. This score is constructed by schools and is a summary of student performance across various subjects.

The tests based on which the score is constructed are different for different schools and grades and are likely not comparable. To account for those differences, we will standardise test scores within grade-school against the control students in the same grade-school¹¹.

It is possible that there will be a lot of missing data associated with not being able to locate children.

4.2.4 Children Motivation

We will collect data on whether a child is intrinsically or extrinsically motivated to go to school and do homework (score ranging from 1-5). We will construct an Anderson index from 2 questions, then, the index will be standardised against the control group (subtract mean and divide by standard deviation of the control group). The impacts will be expressed in standard deviations compared to the control group.

¹¹ Subtract mean & divide by standard deviation of the control group.

4.2.5 Children re-enrollment

Using the administrative data we will record whether a child is re-enrolled (i.e. present in records at all during the next term). This will also be supplemented by children's reports. Here, we will look at 2 outcomes:

- 1) Re-enrollment according to the administrative data (binary)
- 2) Re-enrollment according to children's reports (binary)

As with attendance, we will create an Anderson index which will be used in quantification of the impact. We will present results separately for each measure in supplementary tables to understand the magnitude of the effect.

5. Analytical Model

The average treatment effects (ATE) of Movva's SMS intervention on various outcomes will be estimated using the following Ordinary Least Squares model:

Base specification:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 Y_o + \beta_3 X_i + \delta_1 M_i + \alpha_i + \gamma_t + \varepsilon_i$$

Where:

- Y_i denotes the outcome variable for child i
- T_i denotes the treatment status of household j (1 for Treatment Group; 0 for Control Group)
- Y_o denotes the outcome variable for child i at baseline. In case if the baseline variable doesn't exist, this will be omitted (at baseline, we have collected attendance, number of hours a child did chores, whether a parent visited school, whether a parent talked to the teacher, whether a child is enrolled in school). These baseline variables will be used when studying these outcomes at endline. For other outcomes, no baseline values will be included.
- X_i baseline characteristics - These are included to improve precision of the impact estimate.
 - HH-level
 - Household farms
 - Household does casual labour
 - Number of school-age children in the household
 - Age of parent

- Gender of the parent
- Whether parent is married or not
- Whether parent ever visits school or talks to the teacher (engagement proxy)
- Child-level
 - Age of child
 - Ever repeated grade
 - Days missed in the past 1 week
 - Total hours of chores
- When X_i is missing for an observation, we include an indicator term for missingness M_{io} and replace X_i with its mean
- β_1 is the estimated treatment effect of Movva's SMS intervention compared to control
- α_i strata fixed effects
- γ_t Since the phone data collection will span roughly 2 months, we will include interview date fixed effects. If there is at least one day with lower than 30 interviews/day, we will use week fixed effects instead.
- ε_i error term

The above equation will be estimated using OLS, with Huber-White robust standard errors.

Learning Outcomes:

- When analysing learning outcomes, we will include grade-school fixed effects to account for differences in tests that children taken in different schools/grades. This will impact the number of observations that will go in the estimation since some will be lost due to no variation in treatment variable (i.e. by chance, we may not have treatment and control kids for all grades within a single school), suggesting that the estimation will be made on the sub-sample.

Conclusions & Inference:

We will declare that the intervention had a statistically significant impact on attendance (which is the definition of success) if the pvalue on the attendance index measure is smaller than 0.1, so we do not believe that there is a need for multiple hypotheses adjustments. There are five indexes in secondary outcomes based on which we will determine whether intervention had impacts on the intermediate outcomes and other outcomes of interest. We believe that the construction of indexes is sufficient for addressing potential for false positive results and will not apply further adjustments.

Subgroup analysis:

We will use Causal Forest Plots, following Athey et al. 2019 to understand whether the intervention had heterogeneous treatment effects on attendance. Some candidate variables include:

- Baseline attendance (above/below median)
- Gender of the child
- Grade level of the child (1-4, 5-8, 9-12)
- Gender x grade level of the child
- Distance to school (top/bottom 50%)
- Gender of the interviewed parent (targeting female vs male beneficiaries may have different impacts)
- Literate vs illiterate caregivers as SMS may have been less successful for parents who cannot read/write
- Whether parent reported to ever visit school at baseline (proxy to interest in school, parents with low interest may be affected more)
- Top/bottom 50% percentile for number of total hours that child helps with chores at baseline

6. Limitations & Mitigation Strategies

6.1 Technical concerns

Differential non-response

In case we see differential non-response between treatment and control groups, we will apply Lee bounds to examine the extent to which this contributes to the estimate.

Spillovers within classroom:

Previous research on the intervention showed that there could be possible within-classroom spillover effects driven by the fact that treated students may be more engaged in learning compared to control students. This higher engagement may improve the classroom environment and affect non-treated students. We planned on mitigating the issue by examining the saturation rate within a classroom at randomization stage and ensuring that no more than 20% of children within a classroom are treated. The data on classroom assignment was going to be derived from parental reports given to GiveDirectly. After piloting, we have learned that parents do not give accurate information on children's classroom assignment and we will not be able to ensure low saturation rate within the classroom at randomisation since the data on correct classroom assignment will be collected only at endline during in-person visits. Given the fact that there may be a lot of missing data from in-person visits, we may not be able to adequately understand whether spillovers contribute to the results.

Imperfect Takeup:

It's possible that parents may either refuse to receive messages or ignore them. Lower takeup will result in lower power to detect statistically significant differences.

To look at treatment on the treated effects (TOT) on attendance, we propose using administrative data from Movva to estimate the rate of treatment. We will define a household as treated if they successfully received 80% of messages (they didn't bounce back or the household didn't refuse participation.) However, we note that receipt of messages will not mean that the messages are actually read.

Validating reports given by children:

Our main impact specification relies on children's reports. Based on our understanding of the context/intervention, they are the most reliable compared to parental reports or administrative attendance data (due to potentially large numbers of missing records). In order to validate the truthfulness of children's reports we will compare attendance records to children's reporting for the sub-set of children for which the interview dates match dates for the attendance data collection. This will provide limited evidence to reliability of children's reports.

www.IDinsight.org
@IDinsight

IDinsight