Pre-analysis plan [DRAFT] | January 20, 2020

INTRODUCTION

IDinsight is conducting a cluster randomized controlled tiral (RCT) to assess the impact of school-based handwashing nudges on primary school students' observed handwashing rates in the Philippines.

This pre-analysis plan documents the key research questions the evaluation seeks to answer and specifies the analysis that will be perfomed for each question. The first version of this plan was written after installation of nudges in schools when we collected implementation data and before endline data collection on outcomes of interest.

1.1 PROJECT CONTEXT

School children in the Philippines suffer from a high burden of preventable diseases, with hygiene deficiencies identified as a common cause. Handwashing with soap is considered to be one of the most effective measures to reduce respiratory tract infections and diarrhea and prevent the transmission of infectious diseases.

In 2009, the Philippine Department of Education (DepEd) formally introduced the Philippine Essential Health Care Program (EHCP)[1] to address high prevalence of childhood diseases including diarrhea, respiratory tract infections, parasitic infections and dental caries. Funded by UNICEF in partnership with other organizations, the EHCP focuses on promoting daily group handwashing (GHW) with soap, group tooth-brushing with toothpaste, and biannual de-worming in public elementary schools. The EHCP was institutionalized through DepEd's WASH in Schools Policy (WinS) in 2016.

Despite this policy and its associated programs, IDinsight's research[2] in Philippines public schools has found that independent handwashing with soap (iHWWS) rates are below 10%, even when adequate facilities are available. The research also has found that targeted behavior-change campaigns conducted by teachers have not effectively led to meaningful changes in handwashing behavior, despite strong implementation. Therefore, there is an ongoing need for effective programming which instills the habit of regular handwashing by pupils in schools.

1.2 INTERVENTION

1.2.1. INTERVENTION OVERVIEW

The intervention will install behavioral "nudges" in the handwashing environment to influence pupils' handwashing behavior. The nudges are designed to address the primary barriers to handwashing

[1] The program was designed as part of a partnership between DepEd, LGUs, GIZ and GlaxoSmithKline.

[2] We previously conducted the HiFive for Hygiene and Sanitation ("HiFive") Phase I and Phase II evaluations. HiFive was a behavior-change intervention that focuses on motivating improved independent handwashing behavior and triggering discussions on how to better maintain clean and functional handwashing facilities. Phase I was an impact evaluation to estimate the causal effect of the intervention on iHWWS rates among elementary school children in the Philippines. Phase II was a process evaluation to assess the program's implementation. HiFive as a handwashing intervention was a precursor to the nudges intervention, and our experience with HiFive prompted the development of the nudges study and evaluation.

among children in the Philippines, which our previous studies have found to be forgetfulness and more broadly, a lack of habit formation.[3]

There are two intervention points:

1. Reminders to children to wash their hands immediately after they use the toilet. This counteracts forgetfulness and present-bias.
2. Relatively low-cost changes to the physical environment to provide subconscious cues that encourage children to approach the handwashing station. This increases the salience of handwashing and therefore make the behavior more likely.

Based on smaller-scale prior studies[4][5], we have identified several promising nudges. These were chosen and adapted to be culturally appropriate, contextually feasible, cost-effective, and scalable for implementation sites in the Philippines:

1. *Visible signage* crafted to trigger behavioral motivators for handwashing
2. *Colorful pathway* overlaid with footprints leading from the toilet to the handwashing station
3. *Sticker of a pair of eyes* placed above the handwashing station
4. *Soap dish with an arrow* beside it on the sink or counter to remind children to wash their hands with soap

1.2.2. BEHAVIORAL CHANGE THEORY & HYPOTHESIS OF CHANGE

The intervention aims to increase handwashing practice by utilizing behavioral nudges to address lack of habit formation and forgetfulness, two psychological factors associated with behavioral change. Our previous studies found that these were the primary barriers to handwashing following toilet use among children in Camarines del Norte province and Puerto Princess province in the Philippines. The case for focusing on habit formation and forgetfulness is also reinforced by research from two relevant fields, neurobiology and behavioral economics.

Neurobiological studies find that children's brains are underdeveloped in regions associated with attention and future-oriented behaviors.[6] Insights from the field of behavioral economics suggest that attention is a limited resource,[7] and we can expect children to be distracted by classroom activities as they leave the toilet.[8] Furthermore, people tend to be 'present-biased' when making decisions, meaning a future good (i.e. remaining disease-free) is often valued disproportionately less than the present cost or inconvenience of taking an action (i.e. washing one's hands). These psychological factors challenge the ability to form healthy habits.

---

[3] Our prior work observing Philippine pupils' handwashing behavior at critical times reveal handwashing rates to be as low as 2.2%; foremost among reasons youth did not wash their hands is "I forgot".

[4] Blackwell, Calvin, Daniela Goya-Tocchetto, and Zack Sturman. "Nudges in the restroom: How handwashing can be impacted by environmental cues." *Journal of Behavioral Economics for Policy* 2.2 (2018): 41-47

[5] Grover, Elise, et al. "Comparing the behavioural impact of a nudge-based handwashing intervention to high-intensity hygiene education: a cluster-randomised trial in rural Bangladesh." *Tropical Medicine and International Health* 23.1 (2018): 10-25.

[6] Yurgelun-Todd D. Emotional and cognitive changes during adolescence. Current opinion in neurobiology. 2007;17(2):251-7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17383865. Accessed June 11, 2011.

[7] Kahneman, Daniel. "Maps of bounded rationality: Psychology for behavioral economics." American economic review 93.5 (2003): 1449-1475.

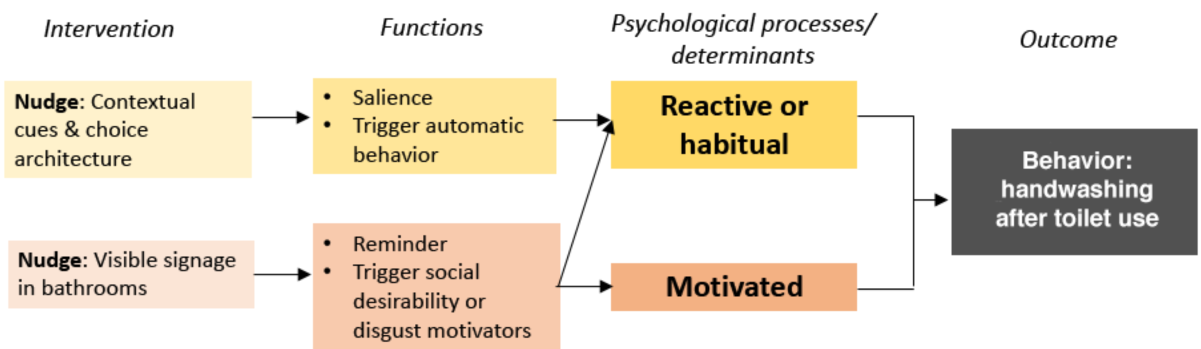
[8] Most DepEd schools have toilets and handwashing facilities within each classroom, so this intervention is designed for those types of facilities.

Our conceptual framework for inciting behavior change among pupils is shown in Figure 1. It is based on the psychological determinants of handwashing discussed in Aunger et al. 2010[9], but combines this model with insights from behavioral economics' nudge theory to influence pupil's decisions to wash hands. Our framework also fits into the broader Integrated Behavioral Model for Water, Sanitation, and Hygiene (IMB-WASH).[10]

Our two intervention points map onto Aunger et al.'s model[11] for behavioral response, specifically the two distinct psychological processes the authors highlight for behavioral change:

- *Reactive*: These are "automatic responses to the presence of a specific cue, such as an object, person, message, or time of day" that serve to reinforce the development of habits over time.
- *Motivated*: These are trigged by the conscious desire to satisfy a need, for example to reduce feelings of disgust, increase social affiliation or conform to a social norm.

Figure 1: Behavioral Change Framework



The first two nudges (footprints and arrow pointing to soapdish) are based on the choice-architecture approach. As described by Thaler and Sunstein[12], a "nudge" is a deliberate design choice that changes an individual's context and physical environment to encourage the desired behavior, rather than altering conscious decision-making processes. In this case, our desired behavior is handwashing. The nudges above function to increase salience and to trigger automatic behaviors (i.e. following the footprints by default or feeling like you are being watched and therefore want to carry out the desirable behavior), and therefore operate through the *reactive* channel to change behavior.

The remaining nudges (visible bathroom signs and soap dish with arrow) at its most basic level serve as a reminder function. We also intend to craft the signs' specific messaging content to trigger *emotional motivators* for handwashing, such as disgust and social affiliation.[13]

---

[9] Aunger, Robert, et al. "Three kinds of psychological determinants for handwashing behaviour in Kenya." Social science & medicine 70.3 (2010): 383-391.
[10] Dreibelbis, Robert, et al. "The integrated behavioural model for water, sanitation, and hygiene: a systematic review of behavioural models and a framework for designing and evaluating behaviour change interventions in infrastructure-restricted settings." BMC public health 13.1 (2013): 1015.
[11] Please see Section 1.3 Evidence Review for more information on this behavioral change model from Aunger and colleagues.
[12] Thaler, Richar H. & Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness* (2009). London: Penguin Books, 2009.
[13] Biran, Adam, et al. "Effect of a behaviour-change intervention on handwashing with soap in India (SuperAmma): a cluster-randomised trial." The Lancet Global Health 2.3 (2014): e145-e154.

2. EVALUATION OVERVIEW

2.1 OVERALL DESIGN

To evaluate the impact of behavioral nudge intervention, IDinsight will conduct a cluster randomized controlled trial (RCT). Across Zamboanga del Norte province, IDinsight will randomly select 132 Department of Education (DepEd) schools to comprise the study sample. We will assign 66 schools to the treatment group, and 66 schools to the control group. Treatment schools will receive the behavioral nudge intervention, while control group will not. Both treatment and control schools will receive DepEd's national WASH in Schools (WinS) policy. WinS promotes "correct hygiene and sanitation practices among school children and a clean environment in and around schools" and includes initiatives related to infrastructure, knowledge, and behavior change, but does not include other programs with an explicit focus on handwashing promotion.
We will assign the behavioral nudges intervention to schools using a stratified randomization procedure, to ensure that treatment and control groups are balanced for both observable and unobservable characteristics that could affect outcomes of interest. This allows us to isolate the impact of nudges from other factors that may be influencing outcomes of interest at schools. The evaluation will estimate the causal impact of the behavioral nudges intervention by measuring the difference in outcomes of interest between the treatment and control groups collected during data collection.

2.2 RESEARCH QUESTIONS

The research will consist of two parts with associated research questions:

1. An impact evaluation to estimate the causal impact of the behavioral nudges on iHWWS among pupils from the two groups.
    a. **Research question 1 (RQ1): Practice of iHWWS after toilet use—**Does the behavioral nudge intervention increase the rate of independent handwashing with soap after toilet use by pupils at school?
    b. **Research question 2 (RQ2): Handwashing facilities—**Does the behavioral nudge intervention increase student access to handwashing facilities with soap and water?

2.3 STUDY SETTING & TARGET POPULATION

This study targets elementary school students in Zamboanga del Norte. This is a province situated in the region of Mindanao in the Philippines. In 2018, it had a poverty incidence rate of 51.6 percent[14], the highest in the Philippines. UNICEF has designated Zamboanga del Norte a target district for DepEd's national WASH in Schools (WinS) programming for 2020. WinS promotes "correct hygiene and sanitation practices among school children and a clean environment in and around schools" and includes initiatives related to infrastructure, knowledge, and behavior change, but does not include other programs with an explicit focus on handwashing promotion. All schools in the study will have WinS policy implemented.

[14] The Poverty and Human Development Statistics Division. "First Semester 2018 Official Poverty Statistics of the Philippines" (2019). Quezon City: Philippine Statistics Authority.

Within Zamboanga del Norte, DepEd has identified all school districts within Zone 1-3, and several in Zone 4, as sites for the study. The zones are geography-based organizational clusters that DepEd uses to administrate its schools. DepEd excluded the remaining school districts within the school division due to safety and accessibility concerns. We also excluded schools that were receiving additional WASH-related programming besides WinS. The remaining schools were available for the study, and a subset of them were selected as the sampling frame (see Section 2.5 Recruitment Strategy).

Most DepEd schools have toilets and handwashing facilities within each classroom, so this intervention is designed for those types of facilities. However, for classrooms that do not have in-classroom handwashing facilities, the intervention will also be implemented at group handwashing facilities that are outside but close to these classrooms within the schools. We will specifically target pupils from two sets of grades at the sample schools: pupils from grade 1-3, and pupils from grade 4-6. UNICEF and DepEd had previously noted the importance of understanding the potentially different impact of handwashing interventions on pupil iHWWS in these grade groups, in order to inform future implementation plans.

2.4 OUTCOME MEASUREMENT

Accurately measuring handwashing practices poses an important challenge in assessing the effectiveness of handwashing programs, and each method presents difficulties related to cost, practicality or validity. Common techniques used for measuring handwashing practices include: inspection of hand cleanliness or microbial contamination of hands[15], sensor in soap dispenser[16] or embedded in soap bars[17], observed handwashing in households[18], and structured observation of handwashing with soap[19].

Kaltenthaler et al.[20] demonstrate various microbiological methods for assessing handwashing practice in hygiene behavior studies. However, a Water & Sanitation Program study[21] finds large differences between levels of hand contamination at random and at critical times, suggesting that microbiological testing of the same subject’s hands can yield variable results. Hence, large sample sizes are needed to overcome the variability noted in the results of hand microbiology testing, rapidly increasing the cost

[15] Greene, Leslie E., et al. "Impact of a school-based hygiene promotion and sanitation intervention on pupil hand contamination in Western Kenya: a cluster randomized trial." *The American journal of tropical medicine and hygiene* 87.3 (2012): 385-393.

[16] Hussam, Reshmaan, et al. "Habit Formation and Rational Addiction: A Field Experiment in Handwashing." (2016).

[17] Ram, Pavani K., et al. "Is structured observation a valid technique to measure handwashing behavior? Use of acceleration sensors embedded in soap to assess reactivity to structured observation." *The American journal of tropical medicine and hygiene* 83.5 (2010): 1070-1076.

[18] Biran, Adam, et al. "Effect of a behaviour-change intervention on handwashing with soap in India (SuperAmma): a cluster-randomised trial." The Lancet Global Health 2.3 (2014): e145-e154.

[19] Galiani, Sebastian, Paul J. Gertler, and Alexandra Orsola-Vidal. "Promoting handwashing behavior in Peru: The effect of large-scale mass-media and community level interventions." (2012).

[20] Kaltenthaler, E. C., and J. V. Pinfold. "Microbiological methods for assessing handwashing practice in hygiene behaviour studies." The Journal of tropical medicine and hygiene 98.2 (1995): 101-106.

[21] Ram, Pavani. "Practical guidance for measuring handwashing behavior: 2013 update." Water & Sanitation Program: Working Paper (2013).

of observation[22]. Many methods using sensor technologies do not provide respondent-specific information, do not inform about rates of handwashing with soap at critical times, and may require expensive specialized hardware[23]. Self-reported handwashing is generally an efficient measure and has been shown to be associated with health outcome; however it has also been repeatedly shown to overestimate handwashing behavior due to the social desirability of handwashing[24].

While structured observation can be expensive and time-consuming and can pose a risk of influencing behavior through “observer effects”[25], it is regarded to be among the most reliable methods of measurement in many contexts[26]. Being able to observe many pupils at school on a given day can make the process less time- and cost-intensive.

Based on the assessment of the advantages and challenges described above for each of the available methods to measure handwashing, and on the specific evidence needs, constraints and context of the study, the evaluation will use structured observation after toilet use among pupils at sample schools. This will be combined with observations of the presence of functioning handwashing facilities at schools, and of the availability of water and soap at these facilities. The study design takes places in schools expected to have these handwashing enablers. These observations will be used to understand pupil opportunity to practice handwashing with soap at school, and to corroborate this assumption. This approach was also chosen because it is what IDinsight used in two previous handwashing studies in similar Philippines schools, meaning that we have had the opportunity to test and refine its feasibility in this context.

2.5 SAMPLE SIZE CALCULATIONS

This study is powered to detect at least 7 percentage points or greater in handwashing rates between treatment and control schools. The minimum detectable effect size (MDES) of 7 percentage points was determined as a policy-relevant benchmark based on IDinsight’s conversations with our partner, UNICEF, about school-based handwashing campaigns in the Philippines. In our previous study, handwashing rates in Philippines schools were estimated at around 2 percent. As in our previous studies, a scale-up recommendation to DepEd will require that the program demonstrate improvements to handwashing that move rates beyond the threshold of 15 percent. Therefore, assuming conservatively that the baseline control group handwashing rates among the schools in our sample is 10 percent and the intra-cluster correlation is 0.1[27], a sample size of 100 schools in total with MDES of 7 pp provides sufficient power to inform the policy decision above with some buffer.

---

22 Ibid.

23 Ram, Pavani K., et al. "Is structured observation a valid technique to measure handwashing behavior? Use of acceleration sensors embedded in soap to assess reactivity to structured observation." *The American journal of tropical medicine and hygiene* 83.5 (2010): 1070-1076.

24 Ibid.

25 Halder, Amal K., et al. "Observed hand cleanliness and other measures of handwashing behavior in rural Bangladesh." BMC Public Health 10.1 (2010): 545.

26 Ram, Pavani. "Practical guidance for measuring handwashing behavior: 2013 update." Water & Sanitation Program: Working Paper (2013).

27 ICC=0.07 from previous HiFive study. We use an ICC of 0.1 as a more conservative estimate, in the event that handwashing outcomes are more correlated within schools in ZDN compared to the provinces (Puerto Princesa and Camarines del Norte) where our previous HiFive took place.

Provided the true effect of the program is indeed as small as 7 percentage points, a sample size of at least 100 schools with 26 handwashing observations per school enables the experiment to be well-powered, or sensitive enough, to detect a difference of this magnitude or greater with an 80 percent chance of statistical significance.

To buffer against attrition due to safety risks (i.e. not being able to observe a school because the area is deemed unsafe), we expanded the sample size to 132 in total, with 66 schools per arm.

- **Pupils sampled per school:** 26 (13 pupils each from grade 1-3 and grade 4-6)
- **Alpha:** 0.05
- **Power:** 0.8
- **Intra-cluster Correlation Coefficient (ICC)**: 0.1 (conservative estimate, based on previous work in Philippines schools)
- **Baseline correlation:** 0 (baseline will not be conducted)
- **Average outcome in control group**: 0.1 (this is a conservative overestimate of baseline handwashing rates based on previous work in Philippines schools)

2.6 RECRUITMENT OF PARTICIPANTS (SAMPLING)

School eligibility

Intervention status will be randomly assigned to 132 schools across the Schools Division of Zamboanga del Norte (ZDN) Province, resulting in 66 schools per arm.

To be eligible for the intervention and study, schools will need to meet a set of minimum inclusion criteria. This is intended to ensure that schools participating have the required infrastructure in place for the intervention to be effective. The minimum inclusion criteria will include:

(i) Water for handwashing was available at the school for at least certain days of the week
(ii) The overall pupil to toilet ratio equaled 100 or lower
(iii) School has at least one individual or group handwashing station
(iv) School is situated in a district deemed safe by DepEd[28]
(v) School does not have any planned WASH programming in the upcoming 2019-20 school year beyond standard WinS activities

We used data from DepEd's WinS Online Monitoring System (OMS) tsto assess the eligibility of public elementary schools. DepEd division offices collect this data via self-assessment by schools. Based on the alignment between initial observations from our scoping and piloting trip for the study and the OMS data, we believe this data is reliable. We can further confirm these assumptions with data from Research Question 2 (RQ2).

Based on the above, 210 out of 634 (33%) schools were eligible for random selection.

Randomization procedure

[28]Within Zamboanga del Norte, DepEd has identified all school districts within Zone 1-3 and several in Zone 4, as sites for the study. DepEd excluded the remaining school districts within the division due to safety and accessibility concerns.

Schools were assigned to treatment or control group based on a two-step stratified randomization procedure. Stratification ensures the schools are balanced across relevant characteristics that may affect our outcome of interest. In our case, our strata were determined by the "WinS implementation quality index score" (WinS IQ). Each school was assigned a WinS IQ score using data from the OMS. These criteria capture whether schools have infrastructure conducive to our outcomes of interest. Scores ranged from 0 to 9 based on how many of the criteria they met, and was based on a series of nine sanitation, hygiene, and overall WASH indicators.[29]

We also determined variables in the OMS data that may indicate that the school has basic features which are conducive to the nudge intervention, or "nudge enabling" features. Specifically:

- No Information, Education, and Communications (IEC) campaign materials on hygiene in toilets or handwashing areas
- Toilets have lights

We categorized schools as having "nudge enabling" features based on whether a combination of these basic criteria were satisfied[30]. We determined 114 schools had such features.

First, we randomly sampled 100 of these 114 "nudge enabled" schools. We randomly assigned 50 schools to treatment group and 50 schools to control group, stratifying by three categories of WINS quality index scores (Low 2-5, Medium 6-7, and High 8-9). Within our sample of 132 schools, we oversampled "nudge enabled" schools in an attempt to hold constant bathroom environments across schools in our evaluation sample. This was done out of the concern that substantial differences in bathroom environments at baseline (for instance, some more cluttered than others) may dilute the effect of nudges.

To achieve our final sample of 132 schools, 32 supplementary schools were randomly drawn from the 96 non-nudge enabler schools remaining in the eligible population. Again, half were randomly assigned to treatment group and half were randomly assigned to control group within each WiNS index strata to ensure balance on this key variable. Note, this was done in order to introduce variation in baseline handwashing environment, since we hypothesize this is an important mediator for the impact of nudges. This procedure aims to ensure that the nudges are installed in more environments likely to be conductive to them, as such environments would also be targeted in any scale-up decision.

However, we will not be comparing treatment effects in nudge-enabled and non-nudge-enabled schools. We constructed the "nudge-enabled" indicator only to identify schools more likely to have non-cluttered facilities, in order that we can select for a higher probability of non-cluttered classrooms during sampling, as such classrooms would likely be more conducive to the nudges. This indicator was created using available school-level administrative data. However, in our evaluation, we are ultimately interested in the effect of "clutteredness" at the handwashing facility-level (rather than the school-level) on the impact of the nudges. Therefore, rather than using our "nudge-enabled" constructed

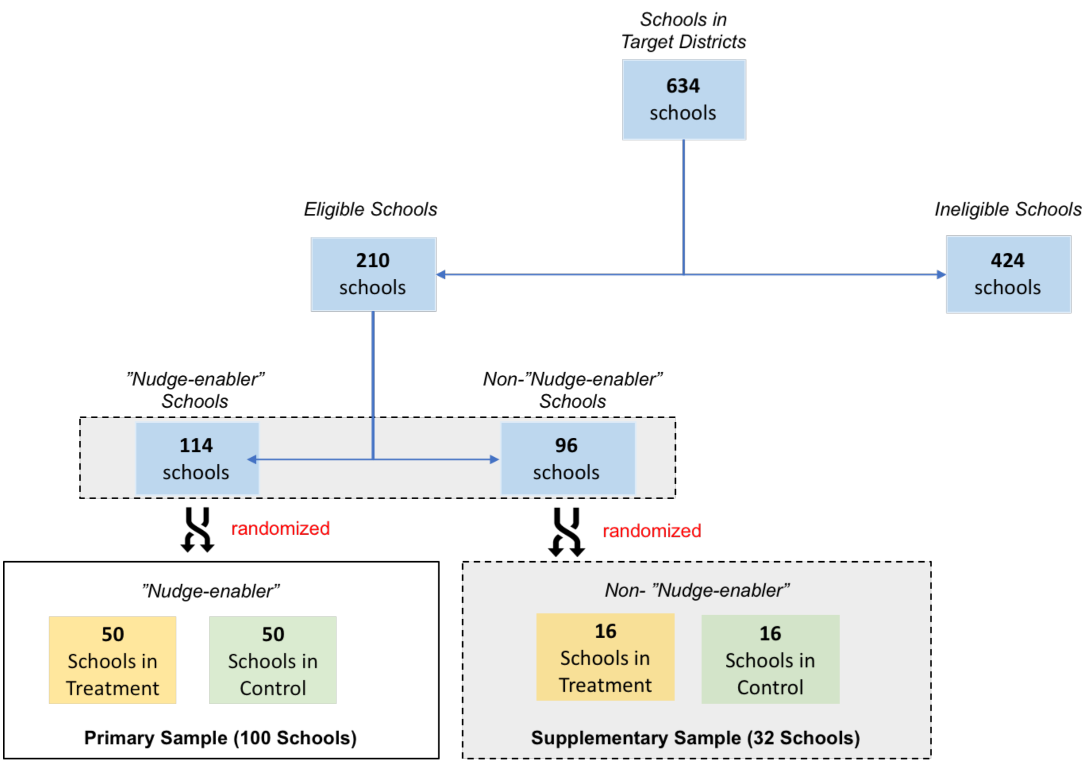
[29] IDinsight used the same methodology of developing WiNS IQs for school selection in the first round of the HiFive assessment. Findings in both HiFive assessments (Phase I and Phase II) confirmed that handwashing facilities were not the primary barrier to handwashing behavior, validating this selection process.

[30] We consider schools to be "enablers" based on three tiers: first, if no_iec_handwashing, no_iec_toilet, and toilet_has_light are satisfied, then if both no_iec_handwashing or no_iec_toilet are satisfied, then if at least 2 of the above three criteria are satisfied.

indicator for analysis, during the evaluation, we will gather handwashing facility-level data on “clutteredness,” and conduct analysis on its effect on the impact of the nudges directly.

Figure 2 shows the flow of randomization.

Figure 2: Flow of Sample & Randomization



2.7 STAKEHOLDER ENGAGEMENT

This study is led by IDinsight in close partnership with the WASH section of UNICEF Philippines. The project’s stakeholders and roles are as follows:

- **UNICEF** WASH team will advise on the intervention, convene and build consensus among stakeholders, and fund intervention design, implementation, and elements of the study
- **Zamboanga del Norte DepEd** will support the intervention design and implementation, provide routine monitoring and support to schools during implementation process, and advise IDinsight on safety and security matters
- **IDinsight** will design the intervention, manage the intervention implementation, and evaluate the program
- **WASHPaLS** will fund the majority of the evaluation and provide technical oversight

IDinsight will solicit and incorporate both DepEd and UNICEF’s contextual knowledge of Zamboanga del Norte schools and the region to strengthen the nudge design and evaluation design. The intervention will be tested as part of DepEd’s national WinS policy, and will serve as part of UNICEF’s strategic focus on improved hygiene and handwashing in schools. IDinsight will present evaluation results and discussion implications with UNICEF, DepEd, and WASHPaLS at the conclusion of the study. IDinsight will also work with all parties to disseminate the study process and results to add to the global WASH knowledge and practice, especially around iHWWS.

3. DATA COLLECTION

3.1 KEY OUTCOMES

This section describes the key outcomes and unit of analysis. Each research question focuses on different indicators based on the data collected. The pre-analysis plan in Appendix 1 describes the analytic model and metrics for each research question in more detail. Some indicators for the process evaluations will be based on data collected during nudge installation. The remaining indicators will be based on observed data collected during the end-line.

The primary research question (RQ1) will specifically target two groups of pupils at sample schools: pupils from grade 1-3, and pupils from grade 4-6. UNICEF and DepEd are interested in the potentially different impact of the intervention on pupil iHWWS in both of these grade groups, in order to inform in which grades to implement the nudges in the future. The impact may differ between these two groups due to different receptiveness to motivators for handwashing behavior targeted by the nudges, or different pre-existing pupil behaviors and perceptions around handwashing. The two age groups were identified in collaboration with UNICEF and are in line with the distinction often made in schools between "lower grades" (1-3) and "higher grades" (4-6). We will report estimates for the pooled sample, but will also conduct sub-group analysis to determine whether the impact of the key outcomes for this research question was statistically significantly different in the two samples.

In addition, within each sample, sub-group analysis will be conducted to compare the average impact of the behavioral nudges by:

- pupil gender
- enumerator gender
- WinS implementation quality index score
- whether the handwashing environment is conducive to nudges, as determined by a constructed index on homogeneity of the toilet and handwashing environment (see Table)
- class size

More information on the analytic models used for each research question, sub-group analysis and co-variates included is provided in the pre-analysis plan in Section 6. An overview of the key indicators, unit and sample for each research question is presented in Table 1 below.

Table 1: Key Outcomes by Research Question

RQ.	Outcome category	Outcome	Unit (sample)
RQ1 *(impact)*	**Independent handwashing after toilet use**	Pupil washed hands *with soap and water* after toilet use	Pupil (grade 1-3 & grade 4-6)
		Pupil washed hands *with water* after toilet use	Pupil (grade 1-3 & grade 4-6)
RQ2 *(impact)*	**(a) Availability of handwashing facilities with water and soap**	Availability of functioning handwashing facilities	School
		Availability of functioning handwashing facilities with soap	
		Presence of functional handwashing facility with soap near toilet facility	Toilet facility
		Presence of functional handwashing facility near toilet facility	
RQ3 *(process)*	**Implementation of nudges**	Presence of nudges at intended location in handwashing facility or toilet facility	Toilet facility
		Condition of nudges at handwashing facility or toilet facility	

3.2 DATA COLLECTION ACTIVITIES

Table 2 outlines our plans for data collection activities and timing. We plan to observe handwashing behavior 3-4 months after implementation to measure the medium-term effects. Drawing on our experience assessing HiFive[31] in 2017 and 2018, we will measure handwashing through in-person observations by trained enumerators. Structured observation is regarded to be among the most reliable methods of measurement in many contexts.[32] Enumerators will visit all schools in the sample, and record observations using SurveyCTO. To minimize observer effects, they will inform principals, teachers, and students only that they are observing general classroom activities. We will detail additional procedures for enumerators during training and in our Data Collection Guidelines document, to be created prior to the endline observations.

Data Collection tools for immediately after implementation and 4 months post-implementation will be developed in the coming months. We will pilot these tools at a subset of schools that were excluded from our sample because they will be receiving additional WASH programming from another DepEd partner during the duration of our study. These schools also hosted piloting of our nudge installation protocols, and will therefore be appropriate for testing the portion of our data collection tools related to nudge implementation (RQ3).

IDinsight will ensure data quality by following robust protocols, including:

- Enumerator training to ensure enumerators make classroom observations correctly and consistently
- Spot-checks on enumerator by IDinsight staff to ensure correct and consistent procedure and data entry
- Frequent data back-checks by IDinsight to spot and correct any data collection flaws in real time
- Use of Survey CTO software for data collection through all survey instruments, as SurveyCTO's internal checks and clear interface reduce risk of enumerator error
- Application of various quality control checks during initial data analysis, including checks for missing data and data entry checks

[31] Please see Footnote 2 for more information about the HiFive intervention and HiFive Phase I and Phase II evaluation.
[32] Please see Section 1.3 Evidence Review.

Table 2: Overview of Data Collection Activities

Assessment Component	Data Collection Activities	Time relative to implementation	Description
1. iHWWS rates *(impact evaluation, RQ1)*	iHHWS after toilet use observations in classroom *(treatment & control)*	4 months post-implementation	Enumerators will observe whether pupils wash their hands with soap, wash without soap, or do not wash, after they have used classroom toilets. This will be used to estimate the causal impact of the nudges intervention on iHHWS rates.
2. Homogeneity in toilet and handwashing environment *(impact evaluation, RQ1)*	Structured facilities observations for both treatment and control *(treatment & control)*	4 months post-implementation	Structured observations of each facility will assess factors such as: presence of existing WASH-related informational posters, whether the handwashing area is used for other non-WASH purposes (e.g. storage, dish-washing area), whether the handwashing station has objects unrelated to handwashing[33], how many primary colors are present in health corner where handwashing station is present, and other factors related to visual clutter to be finetuned after piloting. These will be used to construct a measure of whether the facility is conducive to nudges, to be used in subgroup analyses.
2. Access to soap and water *(impact evaluation, RQ2)*	Structured facilities observations *(Treatment & control)*	4 months post-implementation	Structured observations of whether classroom handwashing facilities including soap and water. These will be used to estimate the causal impact of the nudges intervention on pupil access to soap.
3. Nudges implementation *(process evaluation, RQ3)*	Structured facilities observations *(treatment)*	During implementation visit but after nudge installations are completed	Structured observations of facility after implementation, to assess presence and condition of nudges in toilet and/or handwashing station. These will be used to assess if installation proceeded as intended.
4. Nudges implementation *(process evaluation, RQ3)*	Structured facilities observations *(treatment)*	4 months post-implementation	Structured observations of presence and condition of nudges in toilets and/or handwashing facilities. These will be used to assess if installed nudges functioned and lasted as expected.

33 A concern is that handwashing areas both across and within schools may differ in terms of how conducive they are to nudges; more cluttered and busy environments may dilute the impact of nudges.

4. TECHNICAL RISKS AND MITIGATION STRATEGIES

"Clutteredness" in school handwashing environment

Schools in our sample may differ in terms of baseline toilet and handwashing facility environments. A potential concern is that among handwashing areas that are cluttered and busy, all nudges by the handwashing area (footpath, arrow sticker, eye sticker) would be less likely to grab attention by producing the necessary element of surprise; therefore, any potential impact of the intervention may be diluted. This is particularly likely if the bathrooms already have existing WASH-related informational posters, are crowded with other features (e.g. cabinets or items for storage), or used for other purposes such as tooth-brushing or cleaning dishes. Since we hypothesize that baseline bathroom "clutteredness" factors like these mediate the impact of nudges, substantial differences in the toilet/handwashing environment present the risk of reducing the impact of nudges overall. If the average school is very cluttered, we may not have a large enough sample size to estimate the differential effects of the nudges in cluttered schools and uncluttered schools.

We address this risk in two ways:

1) During school sampling, we prioritized selecting schools with "nudge enabling" handwashing environments, based on characteristics available in the OMS data which we hypothesized could reflect conduciveness to nudges. This is described in Section 2.5 Recruitment of Participants (Sampling).

2) Given that OMS data is at the school-level and there may be difference within a school in terms of handwashing facilities, during our endline data collection, we will also collect detailed information on toilet and handwashing facility environment. Using this data, we will explicitly model for heterogeneity of "nudge enabling" environments in our analysis of the impact of nudges.

Observer effects and social desirability bias

Measuring handwashing behavior is challenging, since it is particularly prone to over-reporting and difficult to observe directly, as described in Section 1.3 Evidence Review. Based on this review, this evaluation will use a combination of indicators that are cost- and time-effective, and meet the practical needs and constraints for the evaluation, in order to measure handwashing practices among pupils. The research team carefully designed and piloted the observation protocols to minimize influence of observation on handwashing behavior, and to mitigate potential reporting biases. These protocols have also been successfully refined and used in two previous IDinsight studies to observe handwashing in DepEd schools in the Philippines.

For indicators based on observed data, the risk of observer effects will be mitigated by not revealing to teachers and pupils in observed classes the specific information that the observer will collect throughout the observation period. While DepEd will be informed about the overall aim of the study, and permission will be obtained from teachers for observing their class, enumerates will state to teachers that the reason for observation is to observe "normal classroom activities," with no mention of handwashing, sanitation, or WinS programming. Further, it is important to note that if there is any behavior change as a result of observation despite the precautions put in place to minimize this, given the nature of random assignment, it is assumed that any resulting change in behavior would, on average, be similar across the treatment and control schools. Therefore, the impact estimate will still provide an unbiased estimate of the impact of the behavioral nudges. These measures are intended to mitigate biases in data collected on handwashing behavior.

Treatment compliance
Within our study, treatment-compliant schools refer to schools assigned to treatment that receive the nudges and maintain them over the course of the intervention, and schools assigned to control that do not implement the nudges. This intervention does not require schools or principles to take any action other than providing nudge installation teams permission to visit and implement the nudges, which are presented to them as desirable classroom improvements. Therefore, non-compliance among treatment schools is not likely. While it is possible that teachers may remove nudges after implementation, we conducted focus groups with teachers during the intervention design phase to ensure teachers did not have strong objections to the presence of the intervention nudges in their classrooms. We will also conduct orientation with all treatment school principals to receive their support for nudge installation.

Non-compliance among control schools is also unlikely. While it is possible that there may be replication of the nudges in control schools, which are cheap and scalable environmental changes that are easily replicable, we will minimize this risk through our implementation protocols. We will limit communication about the nudge installations, and the exact specification of the nudges, to relevant individuals to treatment schools, primarily principals. We will also ensure DepEd does not announce the nudge installation plans to control schools.

As such, our estimates will be intent-to-treat (ITT) estimates. Nonetheless, we will also record at endline whether any treatment schools removed nudges or any control schools introduced nudges. In the event we find non-compliance, we will also estimate treatment-on-the-treated effects.

Spillovers
Potential spill-overs are expected to be minor, and would likely result in the treatment estimate being a lower bound of the impact of the program. The nudges intervention was randomly assigned at the school-level. Distances between schools within Zamboanga del Norte vary. Spillovers would occur when students from schools of different treatment statuses interact. For example, students from treatment schools who interact often with friends and family from control schools with lower handwashing rates may be influenced to was their hands less often than treatment students without such friends. Students may also transfer between schools of different assignment statuses during the study, such that a student from a treatment school may transfer into a control school and influence their new classmates to wash their hands more often. However, pupils are expected to attend schools closest to their homes, and to have friends who are also from their *barangay* (neighborhood) and who would attend the same school. Therefore, spillovers are unlikely to have a meaningful effect on the treatment estimates.

Representativeness and External Validity
We intend to generate evidence on the effectiveness of behavioral nudges to inform national scale-up plans in the Philippines, and to add to the global body of research on handwashing and behavioral change interventions.

The target population for the nudges intervention includes pupils in grades 1 to 6 in selected elementary schools in Zamboanga del Norte. It is important to note two ways in which these schools may differ from other schools in the Philippines. First, these schools are located in DepEd school divisions in which UNICEF has designated a priority target district for WinS programming. These areas are chosen because of high rates of poverty as well as particularly willing partners in the associated DepEd Division offices. Second, schools that were identified as eligible for the intervention, and were included in the evaluation sample, met minimum inclusion criteria related to the current WASH conditions and programs being implemented at these schools, as described in Section 2.5 Recruitment of Participants (Sampling). It is important to keep in mind the characteristics of the schools participating, when considering how the evaluation findings may predict the impact of the intervention in other schools in the Philippines. However, it is also important to note that the intervention is intended for, and meant to be scaled up in, environments similar to those meeting our minimum inclusion criterial; that is, classrooms with minimum handwashing enablers such as facilities, water, and soap.

For outcomes based on pupil classroom observations, all grade 1-6 pupils at sample schools will be eligible for the evaluation sample frame, with the exception of pupils in classrooms where the only handwashing facility available to pupils when in class is not observable (only inside the toilet and not visible to surveyors), or there is no toilet for pupils to use inside or directly outside the classroom.

5. ETHICAL CONSIDERATIONS

ETHICAL CLEARANCE

Given the nature of the evaluation activities and the relatively low sensitivity of the survey questions, the study poses little risk to participants. IDinsight will apply for ethical clearance for the research through a research ethics committee accredited by the Philippine Health Research Ethics Board (PHREB). We will apply for ethical clearance for the study in October 2019, after we have finalized the evaluation design and prepared all the necessary documentation.

RECRUITMENT STRATEGY

DepEd national and division staff were consulted in the design of the nudges and evaluation. DepEd divisions and treatment schools in Zamboanga del Norte will be provided information of the details of the nudges and installation prior to implementation. This will be done either in-person or via phone by either IDinsight staff or supporting members of DepEd.

No additional implementation activities will be conducted and no supplies will be distributed to schools or pupils as part of the evaluation. In addition, WinS programs have previously been implemented in these school divisions with support from UNICEF, and all participating schools (both treatment and control) may thus be familiar with both WinS programs and UNICEF.

INFORMED CONSENT & ASSENT

Prior to visiting schools, letters from DepEd divisions will be obtained, and will be presented to principals at all schools visited to describe the research aims and activities, and confirm DepEd authorization. Upon first arrival at each school, permission will be obtained from the principal prior to collecting any observation or survey data.

Once principal permission is granted, teachers of classes selected for observation or pupil surveys will be sensitized about the study and the data collection activities.

CONFIDENTIALITY

All pupil observations will be conducted at the school. The observations will be conducted inside the pupils' classroom or at the group handwashing station, and we will not collect uniquely identifiable information about pupils.

Data will be stored electronically. All devices used for data collection will be password protected and utilize software that includes built-in security features, such as Secure Sockets Layer (SSL) technology, as well as encryption to prevent access to data in the event devices are lost or stolen. All respondents will be given a unique ID code, and Personal Identifiable Information will be saved in separate files from the survey responses and stored on a secure, encrypted IDinsight server.

6. DATA ANALYSIS

The pre-analysis plan outlines the analytic models that will be used to analyze the data collected as part of the behavioral nudges evaluation, the indicators that will be analyzed, and other technical considerations.

The primary research questions the evaluation aims to answer is:

RQ. 1 **Practice of iHWWS after toilet use**: What is the causal impact of the behavioral nudge intervention on the rate of independent handwashing with soap after toilet use by pupils at school?

In addition, the evaluation aims to answer the following secondary research questions:

RQ. 2 **Handwashing facilities**: What is the causal impact of the behavioral nudge intervention on pupil access to handwashing facilities with soap and water?

For each research question, the following topics are discussed in this pre-analysis plan:

a) Research question
b) Indicators, unit of analysis, and sample
c) Analytic model
d) Sub-group analysis
e) Presentation of results

RESEARCH QUESTION 1

1.a. Research question 1: When implemented with the WinS policy, does the *nudges* intervention increase the prevalence of independent handwashing with soap at critical times by pupils at school, relative to when the WinS policy alone is implemented?

1.b. Indicators, unit of analysis, and sample: Table 3 below provides a summary of the indicators for this research question, and the unit of analysis and sample corresponding to each indicator. These indicators are based on observed data.

Note on the sample: Observation data will be explored for the pooled sample, and separately for pupils from grades 1-3 and from grades 4-6. The intention with splitting the sample is to identify whether there is a differential impact of the nudges intervention on these two grade groups.[34] This can inform decisions by UNICEF and DepEd regarding the grade(s) in which the behavioral nudges are more effective and should be implemented. The impact may differ between these two groups due to different receptiveness to motivators for handwashing behavior targeted by the nudges, or different pre-existing pupil behaviors and perceptions around handwashing.

[34] The approach to test the difference in impact of the program for these two sub-samples will be separate regressions for each sub-sample and then a test comparing the treatment effects across the two regressions. This will be performed instead of a single regression on the pooled sample with an interaction term between the treatment term and the sub-sample for the sake of simplicity, as the former approach would require multiple interaction terms for all co-variates, as well as three-way interactions when exploring other sub-groups.

Note on sample weights: Sample weights will be included to account for the fact that the sampling probability of classes for a given grade is different across schools. Since a constant number of classes will be sampled for each grade in each school[35], schools with more classes for a given grade will have a lower probability of being selected. This will be accounted for by using weights that represent the inverse probability of class selection. By doing so, each pupil observed or surveyed will have the same probability of inclusion into the sample.

Table 3 below provides a summary of the indicators for this research question, and the unit of analysis and sample corresponding to each indicator. Table lists the co-variates we will use in addition to the indicators in our analytical models to answer research question 1. We will report estimates for our outcome indicators both with and without controls.

[35] For classroom observations, one class in each grade in each school will be sampled independent of the total number of classes in the grade.

Table 3: Outcome Indicators for Research Question 1

Outcome category	Outcome	Metric	Unit	Sample	Analytic model[36] (LPM – linear probability model; OLS – ordinary least squares)
Independent handwashing after toilet use	Pupil washed hands *with soap and water* after toilet use	**Y.1** – binary variable for whether pupil was **observed** to wash hands *with water and soap* after toilet use	Pupil	Separately for: Grade 1-3 pupils Grade 4-6 pupils	**LPM**: Y = Y.1
	Pupil washed hands *with water* after toilet use	**Y.1** – binary variable for whether pupil was **observed** to wash hands *with water* after toilet use at school	Pupil	Separately for: Grade 1-3 pupils Grade 4-6 pupils	**LPM**: Y = Y.1

[36] Unless otherwise stated, the independent variable of interest is treatment assignment.

Table 4: Control Variables for Research Question 1

Unit	Co-variate	Metric	Hypothesized relationship to outcome
School	District	**X.1** – categorical variable where each value category represents the school district where the school is located	School district is associated with geographic location, which may be associated with varying levels of rurality, access to water, socio-economic status of teachers and pupils, or support received from DepEd division office.
	Number of pupils enrolled (pre-intervention)	**X.2** – continuous variable to represent the number of pupils enrolled at the elementary school prior to intervention[37]	Given that school budget is partially determined by number of pupils enrolled, and that larger schools tend to be located in urban or semi-urban areas, school resources and the school pupil and teacher population may vary depending on school size. This will be estimated using information reported by principals during end-line school visits or using the data reported in DepEd's Online Monitoring System (OMS), depending on how reliably information reported by principals can be collected.
	WinS implementation quality index score (pre-intervention)	**X.3** – index variable between 0 and 11 composed of 11 binary indicators related to WASH conditions at the school, as reported in OMS data by principals prior to the intervention[38]	Higher score indicates a higher level of reported quality of WinS implementation, and may thus be associated with a higher level of handwashing infrastructure and supplies, as well as more consistent implementation of WinS program activities involving pupils, such as group handwashing and tooth-brushing.
	Nudge-enabled	**X.4**–binary variable for whether school is "nudge-enabled" or not "nudge-enabled"	Nudge-enabled schools are those with combinations of features that indicate they may be more conducive to nudge intervention than schools without those features. Please see Section 2.5 Recruitment of Participants (Sampling) for details of these nudge-enabling features. Because of the presence of these features, the nudge intervention may have greater impact on nudge-enabled schools than schools that are not.
Class	Number of pupils in the classroom	**X.5** – continuous variable where value represents the number of pupils present in the classroom during the observation period	Classroom size may be associated with school resources and teacher quality or attention, which may influence pupil opportunity and incentives to wash hands with soap. Further, if handwashing with soap after toilet use is considered to be a norm (or habit) that the nudges remind pupils of, and if the class is considered as the primary 'social unit' for this norm, the number of pupils in this social unit may influence the likelihood of the norm arising and persisting.

[37] This variable may be converted to an ordinal variable to mitigate random error due to imprecision in the number of pupils reported.
[38] See Table 7 in the *Sampling* section of the evaluation design document for a description of 11 indicators an additional information on the online monitoring system (OMS)

	Nudge-conducive	**X.6**- Binary variable indicating whether the handwashing area location near/in the classroom is "nudge conducive"	Nudge-conducive handwashing areas are those with combinations of features that indicate they may be more conducive to nudge intervention than handwashing areas without those features. Please see Section 3.2 Data Collection Activities for details of these nudge-conducive features. The nudge intervention may have greater impact for nudge-conducive handwashing areas than non-nudge conducive areas.
Unit	**Co-variate**	**Metric**	**Hypothesized relationship to outcome**
Pupil	Gender	**X.7** – binary variable for whether pupil observed is female	Traditional gender roles tend to involve more tasks that may require handwashing for girls than boys, such as cooking or cleaning, and may thus facilitate building handwashing habits among girls.
	Grade	**X.8** – categorical variable where each value category represents the grade of the pupil observed/surveyed	The behavioral nudges targeted at changing pupil behavior may be more engaging and/or effectively convey intended messages to pupils of lower or higher grade.

1.c. Analytic model

Specification for LPM: The effect of the treatment on the outcome variables listed in **Error! Reference source not found.** for which the analytic model indicated is the Linear Probability Model will be measured by conducting the following regression:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \overrightarrow{\beta_2}\vec{X}_{2ij} + \varepsilon_{ij}$$

Where:

- Y_{ij} denotes the outcome variable for pupil *i* in school *j*, classified as a binary variable, as specified in Table 3.
- T_j denotes the treatment variable (binary variable for whether school *j* received the nudge intervention)
- $\vec{X}_{2ij}$ represents a vector of pupil, school, and class level co-variates as specified in Table
- ε_{ij} denotes the pupil error term *i*, clustered at the school-level to reflect the fact that the treatment assignment was at the school level
- β_n denotes the coefficients determined by the regression model (β_1 is coefficient of interest)

1.d. Sub-group analysis: the impact of the nudges intervention on the outcomes of iHWWS after toilet use will be estimated using the analytic model described in sub-section 1.c, for the pooled sample and each of the two samples corresponding to grade groups (grades 1-3 and grades 4-6). The impact for these two grade groups will be compared to assess whether the effect of the treatment is different across grade groups.

In addition, within each sample, sub-group analysis will be conducted to compare the average impact of the nudges intervention by:

- pupil gender
- WinS implementation quality index score
- number of pupils
- nudge conduciveness of handwashing area

1.d.i. Treatment impact by grade group: The impact of the nudges intervention on the outcomes of iHWWS after toilet use will be estimated separately for pupils grade 1-3 and grade 4-6. To test whether the impact of the intervention on these key outcomes was statistically significantly different in the two samples, the impact estimates in the two models corresponding to the two different samples will be compared by testing the equality of the regression coefficients.[39]

1.d.iii. Treatment impact by pupil gender: In order to compare the average effect of the intervention on the outcomes of independent handwashing after toilet use between girl and boy pupils, sub-group analysis will be conducted using the following regression:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 F_i + \beta_3 (T_j \times F_i) + \overrightarrow{\beta_4}\vec{X}_{4ij} + \varepsilon_{ij}$$

Where:

- F_i is a dummy variable denoting whether pupil *i* is female
- $\vec{X}_{4ij}$ represents a vector of pupil, school, and class level co-variates as specified in Table 2, excluding pupil gender
- All other regression terms are defined as they were in the previous regression specification and β_3 is the coefficient of interest.

[39] Observations in which the student grade was unknown will be excluded.

1.d.iv. Treatment impact by WinS implementation quality index score: In order to compare the average effect of the nudges intervention on the outcome of independent handwashing after toilet use across different levels of WinS implementation quality, sub-group analysis will be conducted using the following regression:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 W_j + \beta_3 (T_j \times W_j) + \overrightarrow{\beta_4} \vec{X}_{4ij} + \varepsilon_{ij}$$

Where:

- W_j is an ordinal variable between 0 and 11 denoting the WinS implementation quality index score for school j
- $\vec{X}_{4ij}$ represents a vector of pupil, school, and class level co-variates as specified in Table 2, excluding WinS implementation quality index score
- All other regression terms are defined as they were in the previous regression specification and β_3 is the coefficient of interest.

1.d.v. Treatment impact by nudge conduciveness: In order to compare the average effect of the nudges intervention on the outcome of independent handwashing after toilet use depending on the context's conduciveness to nudges, sub-group analysis will be conducted using the following regression:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 W_j + \beta_3 (T_j \times N_i) + \overrightarrow{\beta_4} \vec{X}_{4ij} + \varepsilon_{ij}$$

Where:

- N_i will be a binary variable (to be constructed after seeing the distribution) =1 if the pupil was exposed to a nudge-conducive handwashing environment at the time of toiet use and =0 if not
- $\vec{X}_{4ij}$ represents a vector of pupil, school, and class level co-variates as specified in Table 2, excluding WinS implementation quality index score
- All other regression terms are defined as they were in the previous regression specification and β_3 is the coefficient of interest.

1.e. Presentation of results: a regression table will be included for each specification to describe the correlation between key indicators of interest and co-variates. Data visualizations describing the level of the outcomes variables in the treatment and control group may also be included.

RESEARCH QUESTION 2

2.a Research question 2: What is the causal impact of the behavioral nudge intervention on pupil access to handwashing facilities with soap and water?

2.b. Indicators, unit of analysis and sample: Table 5 below provides a summary of the indicators for this research question, and the unit of analysis and sample corresponding to each indicator. These indicators are based on enumerator observations of toilet and handwashing facilities at the schools visited.

Table 5: Outcome indicators for Research Question 2

Outcome category	Outcome	Metric	Unit	Sample	Analytic model[40] (LPM – linear probability model; OLS – ordinary least squares)
(a) Availability of handwashing facilities with water and soap	Availability of functioning handwashing facilities *with soap*	**Y.1** – continuous variable where value represents the number of functioning handwashing facilities *with soap* at the school	School	All T&C schools	**OLS** with Y=Y.1
	Presence of functional handwashing facility *with soap* near toilet facility	**Y.1** – binary variable for whether there is a functioning handwashing facility with water *and soap* inside classroom for classroom toilet facilities and within 10 meters of functioning toilet facility for block toilet facilities	Toilet facility	All toilet facilities at T&C school	**LPM** with Y=Y.1
	Presence of functional handwashing facility near toilet facility	**Y.1** – binary variable for whether there is a functioning handwashing facility with water inside classroom for classroom toilet facilities and within 10 meters of functioning toilet facility for block toilet facilities			**LPM** with Y=Y.1

[40] Unless otherwise stated, the independent variable of interest is treatment assignment.

The analytic model used to answer research question 2. will also include the following co-variates, as they were defined in the analytic model for research question 1 (see Table 4):

- **School-level variables**: school district; WinS implementation quality index score (pre-intervention)

In addition, for outcomes for which the unit indicated in Table 5 is the toilet facility and the analytic model is LPM, the following will be included as co-variates:

- **School-level variable**: number of pupils enrolled at the school (pre-intervention)
- **Facility-level variable:** Location of the toilet facility observed, defined as a categorical variable where each value category represents which of following location the classroom facility is inside a classroom from grade 1-6, or outside a classroom.

2.c. Analytic model

2.c.i. ***Specification for LPM:*** The effect of treatment on the outcome variables listed in Table 5 for which the analytic model indicated is the Linear Probability Model will be measured by conducting the following regression:

$$Y_{kj} = \beta_0 + \beta_1 T_j + \overrightarrow{\beta_2} \vec{X}_{2kj} + \varepsilon_{kj}$$

Where:

- Y_{kj} denotes the outcome variable for facility *k* in school *j*, classified as a binary variable, as specified in Table
- T_j denotes the treatment variable (binary variable for whether school *j* received the nudges intervention)
- $\vec{X}_{2kj}$ represents a vector of school and facility level co-variates as specified in sub-section 2.b
- ε_{kj} denotes the facility error term *k*, clustered at the school-level to reflect the fact that the treatment assignment was at the school level
- β_n denotes the coefficients determined by the regression model (β_1 is coefficient of interest)

2.c.ii. ***Specification for OLS analytic model***: The effect of treatment on the outcome variables listed in Table for which the analytic model indicated is OLS will be measured by conducting the following regression:

$$Y_j = \beta_0 + \beta_1 T_j + \overrightarrow{\beta_2} \vec{X}_{2j} + \varepsilon_j$$

Where:

- Y_j denotes the outcome variable for school *j*, classified as a binary variable, as specified in Table
- T_j denotes the treatment variable (binary variable for whether school *j* received the nudges intervention)
- $\vec{X}_{2j}$ represents a vector of school level co-variates as specified in sub-section 2.b
- ε_j denotes the school error term
- β_n denotes the coefficients determined by the regression model (β_1 is coefficient of interest)

2.d. Sub-group analysis

2.d.i. *Treatment impact by toilet facility location*: In order to compare the average effect of the intervention on the outcome of presence of a functioning handwashing facility with soap near school toilets between classroom toilets and out-of-classroom toilets, sub-group analysis will be conducted using the following regression:

$$Y_{kj} = \beta_0 + \beta_1 T_j + \beta_2 L_k + \beta_3 (T_j \times L_k) \overrightarrow{+ \beta_4} \vec{X}_{4j} + \varepsilon_{kj}$$

Where:

- Y_{kj} denotes the outcome variable for toilet facility *k* in school *j*, classified as a binary variable, as specified in **Error! Reference source not found.**
- T_j denotes the treatment variable (binary variable for whether school *j* received the nudges intervention)
- L_k is a dummy variable denoting whether toilet facility *k* is located inside a classroom
- $\vec{X}_{4j}$ represents a vector of school level co-variates as specified in sub-section 2.b, excluding location of toilet facility
- ε_{kj} denotes the facility error term *k*, clustered at the school-level to reflect the fact that the treatment assignment was at the school level
- β_n denotes the coefficients determined by the regression model (β_3 is coefficient of interest)

2.d.ii. *Treatment impact by WinS implementation quality index score*: In order to compare the average effect of the nudges intervention on the outcome of presence of a functioning handwashing facility with soap near school toilets across different levels of WinS implementation quality, sub-group analysis will be conducted using the following regression:

$$Y_{kj} = \beta_0 + \beta_1 T_j + \beta_2 W_j + \beta_3 (T_j \times W_j) + \overrightarrow{\beta_4} \vec{X}_{4kj} + \varepsilon_j$$

Where:

- W_j is an ordinal variable between 0 and 11 denoting the WinS implementation quality index score
- $\vec{X}_{4kj}$ represents a vector of school and facility level co-variates as specified in sub-section 2.b, excluding WinS implementation quality index score
- All other regression terms are defined as they were in the previous regression specification and β_3 is the coefficient of interest

2.e. Presentation of results: a regression table will be included for each specification to describe the correlation between key indicators of interest and co-variates. Data visualizations describing the level of the outcomes variables in the treatment and control group may also be included.

MULTIPLE HYPOTHESIS TESTING CORRECTION

Subgroup analyses and analyzing multiple indicators of the same category of outcomes increases the number of hypotheses being tested. As such, standard statistical significance levels would likely result in finding significant coefficients by chance alone (i.e., false positives). To correct for this increased likelihood of false positives, the analysis will be adjusted for multiple inference using the

Holm-Bonferroni procedure[41]. The correction will impose a more conservative threshold for statistical significance. Normal, uncorrected p-values will also be reported.

Table 6 below summarize the outcomes described for each research question above, and indicates the number of hypotheses being tested with each "family" of outcomes. The analysis will correct p-values based on the number of hypotheses being tested for each family of outcomes.

[41] Abdi, Hervé. "Holm's sequential Bonferroni procedure." *Encyclopedia of research design* 1, no. 8 (2010): 1-8.

Table 6: Number of Hypotheses Tested For Each Family of Outcomes

Research Question	Outcome Group	Indicator Type	Indicator	Sample	Number of hypotheses tested within each family of outcomes for p-value correction
1	Independent Handwashing With Water and Soap After Toilet Use	Main	Independent HWWS after toilet use	All grades	n/a because main effect
		Main	Independent HWWS after toilet use among grades 1-3	Grades 1-3	
		Main	Independent HWWS after toilet use among grades 4-6	Grades 4-6	
		Sub-group	Gender	All grades	4
			Wins implementation quality	All grades	
			Number of pupils	All grades	
			Nudge conduciveness of handwashing area	All grades	
	Independent Handwashing With At Least Water After Toilet Use	Main	Independent HW with water after toilet use	All grades	n/a because main effect
2	Availability of Handwashing Facilities	Main	Availability of functioning handwashing facilities *with soap*	All HW facilities	4
			Availability of functioning handwashing facilities	All HW facilities	
			Presence of functional handwashing facility *with soap* near toilet facility	All HW facilities	
			Presence of functional handwashing facility near toilet facility	All HW facilities	
	Availability of Toilets	Main	Functioning toilet per pupil	All schools	n/a because separate outcome family

LIMITATIONS AND CORRECTIONS TO THE ANALYSIS

Outliers and Missing Values*:* Missing values can take the form of non-response (e.g. uncompleted surveys), partial response (e.g. "Don't know" responses), or errors in the data. Minimal missing data is expected since the digital survey forms include constraints that require data to be entered before the enumerator can proceed with data collection. If co-variates have missing values for some observations, a dummy variable will be generated and included in the analytic model, and missing values for co-variates will be replaced with 0[42].

Cross-tabulations: cross-tabulations of variables may be conducted to provide additional insights on the relationship between variables within the sample of schools surveyed.

Additional Hypothesis Testing: Additional analyses may be conducted following the conclusion of data collection if additional questions of interest arise. Any analyses that are not specified in this pre-analysis plan will be indicated and justified as such.

[42] Alternatively, and depending on the number missing values across co-variates, multiple imputation may be used to account for missing values in the analysis.

ANNEX C: GENDER ANALYSIS PLAN

Table 1 Behavioral Nudges Gender Analysis Plan

GENDER ANALYSIS QUESTIONS *What do we need to know? What are possible gender gaps? What are possible unintended consequences?*	ANSWERS *What are the answers to the gender analysis questions?*	WHY IT MATTERS *Is there a gender issue/gap that may affect grant implementation? Why does it present a barrier or opportunity for the research? How critical to the success of the research grant is it to address it?*	GRANT RESPONSE *Describe how the gender gap will be addressed in the design and implementation of the research grant (i.e. to close the gap, address the barrier or accelerate on the opportunity). Be specific.*	MONITORING *How will we monitor or measure this?*
1. **How do independent handwashing rates (with and without soap) between boys and girls compare?**	Traditional gender roles tend to involve more tasks that may require hand-washing for girls than boys, such as cooking or cleaning. Thus, they may facilitate building hand-washing habits among girls, and lead girls to have a higher independent handwashing rates (with or without soap) than boys.	This is both a challenge and an opportunity for the intervention. If a gender gap in independent hand-washing is discovered, implementers can incorporate new gender-targeted measures into our intervention design. These measures can ensure that the nudges are appropriately targeting both genders.	The intervention can alter or add measures that specifically target the gender found to have lower handwashing rates. For example, if boys are found to have lower handwashing rates, implementers can put up posters designed to appeal more to boys than to girls. These measures could help close the possible gender gap in handwashing.	The evaluation will compare the handwashing rates of boys and girls in control schools, in order to determine if there is an existing gender gap in independent handwashing rates.
2. **Will behavioral nudges based on disgust and social affiliation have greater impact on independent handwashing rates of girls than boys?**	Due to gendered differences in desirable traits in children, girls may be more sensitive than boys to messages related to cleanliness and conformity. Therefore, behavioral nudges based on disgust and social affiliation	This is both a challenge and an opportunity for the intervention. As designed, the intervention may have a greater impact on girls than boys. If so, this would have the positive effect of leading to greater handwashing rates	Post designs will feature both boys and girls carrying out the same activities (handwashing alone and with peers) to ensure that there are no gendered differences in the message that they are sending. If there is a gendered difference in the impact of the intervention, and	The evaluation will compare the difference in handwashing rates between control group and treatment group by gender. That is, it will compare the nudges' impact on the

GENDER ANALYSIS QUESTIONS *What do we need to know? What are possible gender gaps? What are possible unintended consequences?*	ANSWERS *What are the answers to the gender analysis questions?*	WHY IT MATTERS *Is there a gender issue/gap that may affect grant implementation? Why does it present a barrier or opportunity for the research? How critical to the success of the research grant is it to address it?*	GRANT RESPONSE *Describe how the gender gap will be addressed in the design and implementation of the research grant (i.e. to close the gap, address the barrier or accelerate on the opportunity). Be specific.*	MONITORING *How will we monitor or measure this?*
	may have a greater impact on handwashing rates for girls than boys.	overall, but the negative effect of creating or widening a gender gap in handwashing rates. Since nearly all of the bathrooms in our sample schools are shared by boys and girls, there is no opportunity to test different messages for boys and girls.	we suspect it is because boys and girls have different levels of sensitivity to specific emotional appeals, the intervention can be adapted to appeal equally to boys and girls in future iterations. For example, half the posters can make appeals more effective on boys, and the other half can make appeals more effective on girls.	handwashing rates of girls and of boys. This will determine if the nudges affect boys and girls in a differentiated way.
3. Does enumerator gender affect handwashing behavior at all, or among boys and girls differentially?	It is possible that due to gendered differences in how children perceive adults, they may be more likely to wash their hands when a female enumerator is observing the classroom than a male enumerator, since they may associate female adults with reinforcing hygienic behavior. It is also possible that a female enumerator may have a greater effect on handwashing rates for girls	If there is a gendered element to how students respond to female and male enumerators, this may present a challenge for the study's observations, because it could affect the effect size observed.	To address this issue, we will do our best to hire a gender-balanced team of enumerators to conduct observation and send split male and female observers equally between treatment and control schools.	In addition to comparing the difference in handwashing rates between girls and boys, the evaluation will compare the difference between overall and gender-specific handwashing rates based on enumerator gender in control schools, to isolate the effect of enumerator gender from the nudge treatment.

GENDER ANALYSIS QUESTIONS *What do we need to know? What are possible gender gaps? What are possible unintended consequences?*	ANSWERS *What are the answers to the gender analysis questions?*	WHY IT MATTERS *Is there a gender issue/gap that may affect grant implementation? Why does it present a barrier or opportunity for the research? How critical to the success of the research grant is it to address it?*	GRANT RESPONSE *Describe how the gender gap will be addressed in the design and implementation of the research grant (i.e. to close the gap, address the barrier or accelerate on the opportunity). Be specific.*	MONITORING *How will we monitor or measure this?*
	than boys, as gender norms lead girls to perform more handwashing-related chores (cleaning).			