HANDWASHING NUDGES IN ZAMBOANGA DEL NORTE: FINAL REPORT

This report is made possible by the generous support of UNICEF and the American people through the United States Agency for International Development (USAID). The contents of this report are the responsibility of IDinsight, and do not necessarily reflect the views UNICEF, USAID, or the United States Government.

July 6, 2020

USAID FROM THE AMERICAN PEOPLE

IDinsight DATA. DECISIONS. DEVELOPMENT.

unicef for every child

CONTENTS

Abbreviations

DepEd	(Philippines) Department of Education
EHCP	(Philippine) Essential Health Care Program
HWWS	Handwashing with water and soap
GHW	Group handwashing station
MOOE	Maintenance and other operating expenses
pp	Percentage points
RCT	Randomized-controlled trial
WASH	Water, sanitation and hygiene
WinS	WASH in Schools

Figures

Tables

EXECUTIVE SUMMARY

In 2019-2020, IDinsight evaluated a handwashing "nudges" intervention in Zamboanga del Norte elementary schools, with the support of UNICEF, USAID (through a WASHPaLS behavior change innovations grant), and in collaboration with the Philippines Department of Education. Nudges are environmental changes that "[alter] people's behavior in a predictable way."[1] The tested intervention consisted of the following nudges: painted footpath with spray-painted footprints from toilet stall to handwashing station, calendar of handwashing posters, "watching eye" sticker above the handwashing station, and arrow sticker pointing to a soap dish. This report details the results of the evaluation, which are intended to inform scale-up recommendations about the pilot intervention.

FINDINGS

To evaluate the intervention, IDinsight conducted an impact evaluation utilizing a cluster randomized-controlled trial (RCT), and a process evaluation utilizing structured pre-post nudge observations and qualitative key informant interviews. We originally randomized 25,966 students clustered in 132 schools (66 Treatment, 66 Control). Attrition due to safety and security risks posed by armed rebel groups in the region led our final sample to be comprised of 99 schools (49 Treatment, 50 Control), and 1158 handwashing observations. As attrition was anticipated and accounted for in our power calculations, and was neither *differential* nor *selective*, we do not think it biases impact estimates.

We found that:

- Handwashing with water and soap after toilet use rates among pupils in treatment schools were **17.3 percentage points (pp) or 148 percent** (p=0.010**) higher than in control schools
- Access to functional handwashing facilities with soap near toilets in treatment schools were **20.2 pp or 38%** (p<0.0001***) higher than in control schools
- The program simultaneously nudged students to wash hands with soap in classrooms that already had soap, and nudged teachers to provide soap where it was not already available.

Interviewed principals, teachers, and students all liked the nudges intervention, though some did not fully understand the purpose of the nudges or their design. This led to some unintended interactions with the nudges, such as teachers moving the posters calendar outside of the bathroom and into the classroom; however, as the impact findings showcased, the nudges were still able to fulfill their function as reflexive cues to trigger handwashing behavior. Several interviewees also emphasized that they needed improved water access in order to further improve pupil handwashing rates.

Ultimately, the process evaluation found the vast majority of nudges remained in good condition in the four-month period between baseline implementation and endline data collection. That the intervention largely lasted through the four-month implementation period is particularly noteworthy because it cost approximately P570 (USD 11.24[2]) to install in one section during the implementation pilot. If scaled up in the Philippines, the projected cost is a modest PHP 561 (USD 11.07) per classroom, or PHP 3930 (USD 77.51) per school.[3]

---

[1] Thaler, Richar H. & Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness* (2009). London: Penguin Books, 2009.

[2] The conversion rate here and throughout the report is PHP 50.7 = USD 1.

[3] This assumes schools have on average 6 sections and 1 group handwashing station, all of which are eligible for nudges. For more information on cost estimates, see section Findings 2.1.

RECOMMENDATIONS

Based on these findings, IDinsight recommends that UNICEF and DepEd scale up the nudges in schools across the Philippines.

Scale-up efforts should follow the following considerations:

- *The nudges should be part of several other programs supporting WASH in schools (WinS), such as those targeting WASH infrastructure, knowledge, social norms, and other determinants of pupil handwashing behavior .*
  Nudges are not a stand-alone solution. Their success is dependent on being surrounded by enabling conditions, like available water for handwashing and basic awareness of the importance of handwashing after using the toilet. However, in locations with these enabling conditions, the intervention should be included in the WinS budget as a low-cost way to improve pupil handwashing rates. However, in locations with these enabling conditions, the intervention should be included in the WinS budget as a low-cost way to improve pupil handwashing rates.

- *The nudges should only be installed in schools and sections that meet the minimum eligibility requirements.*
  At the school-level, the pupil-to-toilet ratio should be under 100, water should be available daily even if hours are limited, and there must be at least one functional handwashing station (indoors or outdoors). At the level of the handwashing station (indoors or outdoors), the handwashing station needs to have water and soap, with a nearby functional toilet, and a clear, unobstructed, paved path from toilet to the handwashing station. These eligibility requirements are individually listed in Table 10 of the report. If a school or section does not meet the minimum eligible requirement, it will not be possible to install the nudges, or the nudges will not unable to perform their function of encouraging handwashing at the handwashing station.

- *The nudges should be adapted to fit the local context, but should be scaled up as a package without major design changes*
  When installing the nudges to a new school or province, appropriate localizations such as materials, language (for the calendar posters), or colors used should be made them to make them more suitable; however, the intervention should be scaled up as a package, without major alterations to the individual nudges. For example, the calendar posters have been professionally designed to evoke emotion, rather than to provide detailed information about handwashing, and so their design should not be altered. The eye sticker was designed to create social pressure for students facing the sink to exhibit proper handwashing behavior, and so should not be changed into a face sticker. Both these changes may lessen the nudges' intended impact. The intervention was evaluated as a package, and such changes could have unpredictable and perhaps negative effects on its impact on handwashing.

- *Scale-up efforts should include materials or activities to build knowledge and buy-in of the nudges among principal and teachers*.
  As previously stated, interviews revealed principals and teachers lacked some understanding of the purpose of the nudges and the rationale behind their design. For example, many

interviewed teachers did not understand that the arrow was intended to point to the soap dish, or that the posters were intended to remind students about handwashing immediately after toilet use. Such gaps in understanding may have been why we observed in some classrooms that teachers had moved the soap dish from beneath the arrow sticker, or the posters from behind the bathroom door. During scale-up, such knowledge gaps could undermine school staff's ability to install nudges or to maintain the nudges present in their school, reducing the impact of the intervention. Additional orientation activities for school staff would rectify this potential issue.

- *The nudges may be included in plans to improve school design in the time of the COVID-19 pandemic, with appropriate minor design changes.*
  Handwashing with soap can reduce transmission rates of COVID-19, and the evaluation results demonstrate that the nudges can substantially increase handwashing with soap rates. The intervention can also be modified to remind pupils to follow social distancing guidelines when handwashing, though their impact on social distancing has not been tested.

INTERVENTION

1. BACKGROUND AND CONTEXT

In 2009, the Philippine Department of Education (DepEd) introduced the Philippine Essential Health Care Program (EHCP)[4]. The EHCP program addressed diseases prevalent among Filipino schoolchildren such as respiratory tract infections, parasitic infections, diarrhea, and tooth decay by promoting water, sanitation and hygiene (WASH) activities in school. In 2016, DepEd institutionalized the program through its WASH in Schools Policy. WinS works to support infrastructure, knowledge, and behavior change related hygiene and sanitation practices in public elementary schools.

However, IDinsight's previous research on the program HiFive for Hygiene and Sanitation ("HiFive") found that even when a functional handwashing station and soap were available nearby, pupils practiced handwashing with soap (HWWS) after toilet use in only 2.2% of instances. The HiFive program was a 6-week teacher-led handwashing campaign that focused on motivating improved independent handwashing behavior and triggering discussions on how to better maintain clean and functional handwashing facilities. It did so through a set of curricular tools aimed at integrating behavioral motivations in lessons on hygiene. Evaluations took place in Camarines Norte and Puerto Princesa. In Phase I of HiFive, we conducted an impact evaluation to estimate the causal effect of the intervention on HWWS rates among elementary school children in the Philippines. In Phase II, we conducted a process evaluation to assess the program's implementation. Results found that HiFive improved handwashing rates by only a modest 3.7 pp, prompting IDinsight to develop the Handwashing Nudges intervention and evaluation in partnership with UNICEF and DepEd to more significantly increase pupil handwashing rates.

This intervention was tested in 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, the highest in the Philippines.[5] Additionally, the incident rate of stunting among children below 5 in Zamboanga del Norte was 41.4 percent in 2015.[6] Zamboanga del Norte was identified as one of the priority programme areas under UNICEF's Country Programme of Cooperation (2019-2023) with the Government of the Philippines.

2. INTERVENTION OVERVIEW

The handwashing nudges intervention is a package of "behavioral nudges" that aims to encourage pupils to approach the handwashing station after toilet use and wash their hands with soap. According to Thaler and Sunstein, a behavioral nudge "alters people's behavior in a predictable way" by changing how choices are presented to them.[7] Our handwashing nudges were comprised of:

---

[4] The program was designed as part of a partnership between DepEd, LGUs, GIZ and GlaxoSmithKline, and was funded by UNICEF and other partner organizations.

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

[6] Food and Nutrition Research Institute. "2015 Updating of the Nutritional Status of Filipino Children and Other Population Groups Anthropometric Survey." (2016). Taguig City: Department of Science and Technology.

[7] Thaler, Richard H. & Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness* (2009). London: Penguin Books, 2009.

- Painted footpath with spray-painted footprints[8] from toilet stall to handwashing area
- Posters with simple visual handwashing messages in toilet stall
- "Watching eye" sticker above water source of handwashing area
- Arrow sticker pointing to designated soap location[9] by handwashing area

Photos of the nudge design can be seen in the Appendix.

2.1. DESIGN PROCESS

We designed, piloted, and refined the nudges in this intervention. Our design process was comprised of three key steps:

1. **Review of existing literature**. We sought to understand what has worked to increase handwashing rates in studies in other geographies, and brainstormed how to adapt previously-studied nudges to the Philippines context.

2. **Scoping visits to Zamboanga del Norte elementary schools**. Based on our observations, we refined the list of nudge possibilities to ensure the nudges were well-suited to our context.

3. **Pilot installations.** We tested different installation protocols and materials, in order to specify efficient protocols for many different classroom layouts. We also solicited feedback from teachers, principals, and students about the piloted nudges. We finalized the nudge design based on insights from the pilots.

Below, we detail each nudge and their theorized behavioral change pathway.

2.2. BEHAVIORAL CHANGE FRAMEWORK

The conceptual framework for the handwashing nudges intervention is based on the psychological determinants of handwashing discussed in Aunger et al. 2010[10], and combines this model with insights from behavioral economics' choice-architecture approach. Our framework also fits into the broader Integrated Behavioral Model for Water, Sanitation, and Hygiene.[11]

Broadly speaking, the nudges work to combat forgetfulness and build habits. Nudges can help to change behavior when the desired behavior is not restricted by lack of knowledge or lack of enabling conditions, but rather by *behavioral barriers.* In our evaluation of the HiFive intervention, we found that Filipino pupils generally had knowledge of proper handwashing behavior and access to functional toilets and handwashing station. The primary barriers to handwashing following toilet use for pupils in our study was forgetfulness and lack of habit formation, both behavioral barriers. The evidence on behavioral change also supports the use of nudges to address these two specific barriers.[12]

---

[8] If the floor could not be painted, we installed either only spray-painted footprints or footprint stickers.
[9] This was often a soap dish, but could also be a tiered organizer or plastic container. If the handwashing station had no designated soap location, we provided a soap dish.
[10] Aunger, Robert, et al. "Three kinds of psychological determinants for handwashing behaviour in Kenya." Social science & medicine 70.3 (2010): 383-391.
[11] 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.
[12] Aunger (2010).

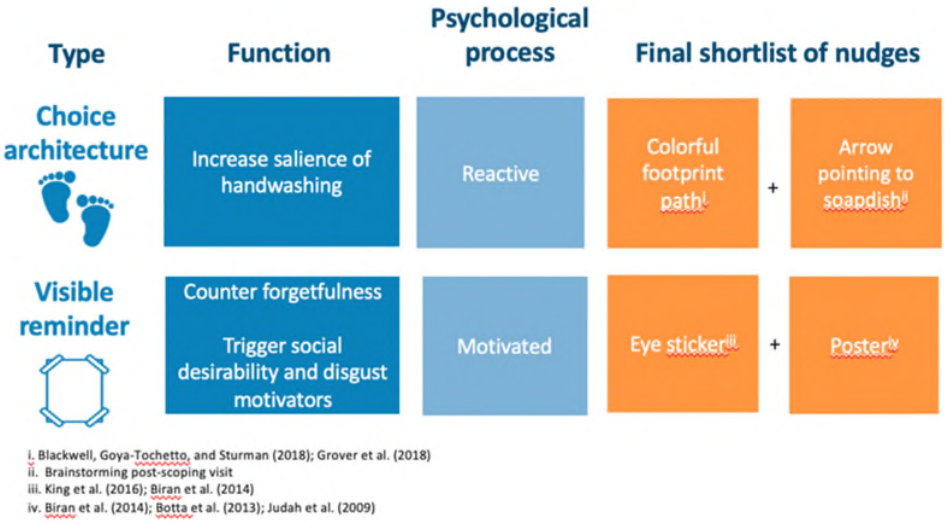
We divide the set of nudges into two types:

1. **Choice architecture** (painted footpath, arrow sticker): These are small changes to the environment that provide subconscious cues in the environment where a behavior takes place to make that behavior more likely. The painted footpath and arrow sticker both increase the salience of handwashing. They encourage children to approach the handwashing station, and to wash their hands with soap, respectively. Such contextual cues are theorized to trigger reactive processes, which are responses "triggered automatically by a particular kind of stimuli"[13] that help to form habits.

2. **Visible reminders** (posters, eyes sticker): These are visual cues that remind students about handwashing, directly counteracting forgetfulness and present bias. Present bias is the tendency of people to value pay-offs closer to the present, and is a documented phenomenon in human decision-making.[14] The poster and eye sticker also aim to trigger motivators like disgust (the desire to perform behavior to feel clean) and social affiliation (the desire to behave in similarly to one's peers).[15]

The intervention's conceptual framework, as well as each nudge's theorized pathway to behavioral change, is summarized in Figure 1. In the figure, we also cite the original source of each nudge idea.

Figure 1 Behavioral Change Framework



[13] Aunger, Robert, et al. "Three kinds of psychological determinants for handwashing behaviour in Kenya." Social science & medicine 70.3 (2010): 383–391.
[14] O'Donoghue, T. and Rabin, M., 2015. Present bias: Lessons learned and to be learned. American Economic Review, 105(5), pp.273–79.
[15] 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.

METHODOLOGY

1. OVERVIEW AND RESEARCH QUESTIONS

Results of this evaluation are intended to inform scale-up decisions by UNICEF and DepEd. The evaluation is divided into two parts: the impact evaluation and the process evaluation. There were approximately four months in between nudge installation and data collection.

Impact evaluation

The impact evaluation aims to evaluate the causal impact of the behavioral nudge intervention. To do so, IDinsight conducted a cluster randomized controlled trial (RCT). To estimate causal impact of the behavioral nudges intervention, we measured the difference in outcomes of interest between the treatment and control groups. We assigned treatment status to schools using a stratified randomization procedure, to ensure that treatment and control groups were balanced in terms of observable and unobservable characteristics that could affect outcomes of interest.

Process evaluation

The process evaluation aims to assess whether the intervention was implemented and functioned as intended, as well as to identify ways to modify or improve the intervention moving forward. This was done through structured observations during nudge installation and data collection, and key informant interviews with principals, teachers, and students.

Our research questions, as well as their associated data collection activities, are listed in Table 1.

Table 1 Research Questions

Research Question (RQ)	Details	Data Collection Activity
RQ 1 *Impact*	**Practice of HWWS after toilet use** Does the behavioral nudge intervention increase the rate of independent handwashing with soap after toilet use by pupils at school?	Handwashing Observations
RQ2 *Impact*	**Handwashing facilities** Does the behavioral nudge intervention increase student access to handwashing facilities with soap and water?	Facilities Observations
RQ3 *Process*	**Implementation of intervention** Were the nudges installed as intended, and did they function as intended?	Facility Observations Interviews

2. EVALUATION DESIGN

2.1. SCHOOL ELIGIBILITY

We determined eligibility for the study at the school-level.[16] To be eligible for the intervention and study, schools needed to meet a set of minimum inclusion criteria. This was intended to ensure that schools participating have the required infrastructure in place for the intervention to be effective.

We used the SY 2018-2019 data from DepEd's WinS Online Monitoring System (OMS) to assess the eligibility of public elementary schools. DepEd division offices collect this data via self-assessment by schools. Observations from the scoping visits conducted for the study aligned with the data from OMS, which indicates this data is reliable.

The minimum inclusion criteria included:

(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 does not have any planned WASH programming in the upcoming 2019-20 school year beyond standard WinS activities

To these infrastructure criteria, we added an additional criterion:

(v) School must be situated in a district deemed safe by DepEd.[17]

Based on all five criteria, 210 out of 634 (33%) schools were eligible for random selection.

2.2. SCHOOL SAMPLING STRATEGY

We assigned schools 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 outcomes 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. Based on how many criteria schools met on a series of WASH indicators, schools received WinS IQ scores ranging from 0 to 9.

We also determined whether a school was "nudge enabler" or "non-nudge enabler", based on school-level variables from the OMS data that we hypothesized would be conducive to the nudges intervention. Specifically, these features were having no Information, Education, and Communications (IEC) campaign materials on hygiene in toilets or handwashing areas; and having lights in toilets. We determined 114 schools had such features.[18]

---

[16] We determined eligibility for each data collection activity at the level of the section or group handwashing station. For more information, see Data Collection Activities on sampling.

[17] 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.

[18] We do not use this nudge-enabler variable in our analysis of treatment heterogeneity. At the time, it was part of an effort to determine which variables can approximate schools that are more conducive to nudges. However, with more context about WinS policy, we realized indicators chosen may not approximate this well. For example, we presumed having no IEC

First, we randomly sampled 100 of these 114 "nudge enabler" 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). To achieve our final sample of 132 schools, 32 schools were randomly drawn from the 96 non-nudge enabler schools remaining in the eligible population.[19] 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.

Sampling weights corrected for the oversampling of "nudge enabler" schools. The final results are generalizable to all Zamboanga del Norte schools that meet the school eligibility criteria.

2.3. FINAL SAMPLE

We initially randomly selected 132 DepEd schools to comprise the study sample, assigning 66 schools to the treatment group, and 66 schools to the control group.

Based on our sample size calculations[20], a sample size of 100 schools would have enabled us to detect a difference of at least 7 percentage points or greater in handwashing rates between treatment and control schools. Due to concerns that many schools would attrite from the study, we selected an additional 32 schools to ensure the study would have enough statistical power.

Ultimately, for our main finding on the practice of HWWS after toilet use (RQ1), our study sample was 99 schools (49 treatment, 50 control). Our sample size for RQ2 was 110 schools (54 treatment, and 56 control).[21] Schools attrited primarily due to safety concerns for enumerator visits; while school areas were deemed safe by DepEd during the initial sampling (before installation), safety conditions changed by the time we carried out data collection. In some cases, safety conditions changed over the course of data collection, due to a dynamic security situation in the area. Other reasons included classroom-level eligibility, reluctance to participate in the study, and repeated unsuccessful visits. Schools were unsuccessfully visited at least three times before being deemed ineligible. Figure 2 shows the sample size throughout the course of the study.

However, neither the rate of attrition nor the composition of the remaining sample in terms of baseline school-level indicators differed between treatment and control schools (see Table 2). As such, attrition was unlikely to affect the internal validity of our study. This point is further detailed in Technical Considerations and Limitations.

---

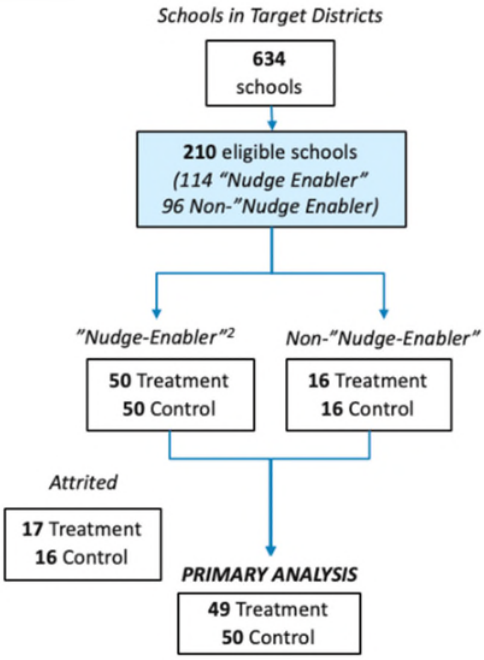
posters would make the poster reminders more surprising, but IEC posters may actually correlate with other WinS programs like additional WASH infrastructure support, which would be conducive to nudges. We do not include this "nudge-enabler" distinction in our analysis, but control for it as part of strata fixed effects in our regression.

[19] We randomly select a sampling fraction of 37/74 per WinS strata.

[20] See Appendix for details.

[21] The discrepancy between the sample sizes is due to different eligibility criteria for handwashing observation and facilities observations. In some cases, we conducted facilities observations but could not conduct handwashing observations (for example, if students were absent or no sections were eligibility for handwashing observations but there were available sections or GHW stations). In others, we could conduct handwashing observations but not facilities observation (for example, if we had to leave abruptly due to safety concerns). Please Data Collection Activities for details.

Figure 2 Sample Size and Randomization Flow Chart



Our baseline sample of 66 schools per treatment arm was well-balanced along key observable school-level characteristics, meaning that there were no statistically significant differences in the key characteristics between treatment and control schools.[22] The p-value from a joint test of orthogonality[23] of all the balance check variables is 0.85. After attrition, our final sample (49 Treatment, 50 Control) remains well-balanced along key characteristics, with a p-value on joint test of orthogonality of 0.94. This balance check is shown in Table 2. Note that both Treatment and Control groups skew towards schools with higher WinS index scores, indicating they have better WASH facilities than the average school in Zamboanga del Norte.

Table 2 Balance Check of Final Sample Along Key School-Level Characteristics (n=99 [49 T, 50 C])

Variable	Mean(C)	Mean(T)	Diff	P-val
Regular availability of soap	0.78	0.80	0.02	0.85

[22] Please find the mean comparisons between treatment and control schools and associate p-value of the difference in means of our original sample of 132 in the Appendix.

[23] Joint tests for orthogonality is a complementary approach to testing for balance by presenting bivariate t-tests for each covariate. It runs the following regression: Treat = a + b1*X1 + b2*X2 + b3*X3 + ….+b20*X20 +u
And then tests the joint hypothesis b1=b2=b3=…=b20=0 using an F-test.

Maintenance and Other Operating Expenses (MOOE) funds for soap	0.66	0.69	0.04	0.72
LGU funds for soap	0.1	0.08	-0.06	0.75
No IEC poster in toilet	0.56	0.61	0.05	0.60
No IEC poster about handwashing in toilets	0.56	0.67	0.12	0.25
Toilet has lights	0.88	0.78	-0.18	0.17
Has GHW facility	0.62	0.67	0.06	0.58
Has section handwashing facility	0.6	0.63	0.03	0.74
Low WinS Index Score (5 and below)	0.12	0.12	0.01	0.97
Medium WinS Index Score (6-7)	0.48	0.49	0.01	0.92
High WinS Index Score (8-9)	0.4	0.39	-0.01	0.90

T=Treatment; C=Control

3. DATA COLLECTION

IDinsight conducted three data collection activities, as shown below in Table 3.

Table 3 Data Collection Activities

RQ	Activity	Approach	Sample Size
1	Handwashing after toilet use observations	Trained enumerators observed for at least 2 hours in one section of each randomly sampled grade level (1-6) in 99 schools	**99 schools (49 T, 50 C)** 298 sections[24] 1158 HW opportunities

[24] In the Philippines, school staff often use "classrooms" to refer to the physical space where classes take place, and "sections" to refer to the group of students (the class) that occupies the classroom. This is because several sections sometimes share one classroom, either simultaneously, or at different times of the day. In this report, we use sections and classrooms synonymously, as we observed both the physical space and the students who occupy it during our evaluation.

2-3	Handwashing facilities observations	Trained enumerators observed all pupil handwashing facilities (section and group) near toilets	**114 schools**[25] **(54 T, 56 C)** 545 sections 44 GHW stations
3	Key informant interviews	Selected enumerators who received additional qualitative research training conducted semi-structured qualitative interviews in T schools	**6 T schools** 36 interviews[26] 4 principals 16 teachers 16 students

We had different sampling procedures for each data collection activity, and determined eligibility for each activity at the level of the section or group handwashing station.

Sampling: Handwashing after Toilet Use

To be eligible for handwashing after toilet use observations, sections in both control and treatment schools had to be both *available, eligible for nudge installations,* and *observable.* We assessed eligibility for treatment schools during both baseline and endline data collection, and control schools at endline data collection.[27]

For a section to be *available*, enumerators had to be able to visit the section when students were present. A section could be unavailable if it was locked, no students are present, the teacher refused to allow the enumerator to stay, etc. To be *eligible for nudge installations*, the section had to (1) have a functional toilet; (2) have a functional handwashing station; and (3) have a clear, unobstructed path from toilet to the handwashing station. This was the same eligibility criteria the installation team used to determine which treatment section would receive nudges.[28] To be *observable*, enumerators had to be able to sit in a location that gave them clear view of the toilet stall door and handwashing area without disrupting classroom activities. For example, a section that had curtains around their handwashing area would not be observable, and therefore would be ineligible for the handwashing observations.

We randomly selected one section within each grade level from grades 1-6. If the selected section was not eligible for observations, we moved on and observed in the next randomly-selected section. If all sections within a grade were not eligible, we observed an additional section in the next grade.

[25] The discrepancy between the sample size for handwashing observations and handwashing facilities observations is due to different eligibility criteria for handwashing observation and facilities observations, which is detailed in the next section, and safety and security concerns (rebel groups are active in areas of Zamboanga del Norte.

[26] While we conducted 42 interviews (6 principals, 18 teachers, 18 students) as planned, we analyzed only 36 interviews, as enumerators failed to save 6 interview audio files.

[27] This approximates eligibility at baseline, as we did not conduct baseline data collection. However, this was unlikely to bias our estimates. We explore implications for our findings further in the Appendix.

[28] However, there was a subtle difference between nudge eligibility between installation and data collection. During installation, if toilets or handwashing stations appeared to be *temporarily out-of-order* (for example, if there was no stored water by the sink or if the toilet was clogged), installers would consider the section nudge eligible and install nudges. However, during data collection, we considered sections with toilets or handwashing stations temporarily out-of-order at the time enumerators visited as *ineligible* for nudges. Enumerators did not have access to observations made at installation, and did not consider whether the section had nudges in making their determination on the three criteria for nudge eligibility.

Sampling: Facilities observation
To be eligible for handwashing after toilet use observations, sections and group handwashing stations had to only be *available* to facilities observations. A section had to be unlocked, so that enumerators could enter and observe the toilet and handwashing area. A group handwashing station had to be approachable and at least one toilet stall unlocked, so that enumerators could approach near it and make observations. For example, if the group handwashing station was behind some construction that pupils were not allowed to pass, the group handwashing station would be considered unavailable, and ineligible for facilities observations.

Sampling: Key Informant Interviews
We purposely selected two large, two medium, and two small treatment schools[29] from different districts in Zamboanga del Norte for the interviews. At each school, we interviewed the principal, three teachers, and three pupils. This added up to 42 interviews in total: six with principals, 18 with teachers, and 18 with pupils.

During data collection, we made two visits for the interviews. During the initial visit, when enumerators also conducted handwashing observations and facilities observations, enumerators obtained permission from the principal to pass out parental consent forms for pupil interviews in classrooms. Pupils brought the forms home for their parents' signatures, and then returned them to their respective teachers. Enumerators returned on a different day to collect the returned parental consent forms, and to conduct interviews with all stakeholders. Principal and teacher consent forms were presented and signed immediately before their interviews began.

We randomly chose one teacher in grades 1-2, one teacher in grades 3-4, and one teacher in grades 4-5. This was to ensure we captured opinions of teachers of both older and younger pupils. For pupil interviews, we randomly selected three sections, one each from grades 4-6.[30] Pupils from grades 1-3 were not interviewed because they would likely have been too young to fully understand the questions. We provided parental consent forms to all pupils in each selected section, and the pupils who returned consent forms constituted our pupil interview sample frame for the section. In each chosen section, we randomly selected one pupil.

For teachers and pupils, we resampled if our selected respondent was unwilling to participate. For principals, we sought to determine the principal would agree to be interviewed before we made our second visit. However, we did not have to resample schools chosen for the key informant interview.

4. ANALYTICAL FRAMEWORK

4.1. IMPACT EVALUATION

All findings represent intent-to-treat effects, as compliance (that is, nudges were installed in sections or group handwashing (GHW) stations that were eligible for nudges) was very high in the study. The intent-to-treat treatment estimate is the policy relevant estimator as it evaluates the effectiveness of

[29] We define large schools as having more than 10 sections in grades 1-6, medium as having 5-10 sections in grades 1-6, and small schools as having less than 5 sections in grades 1-6.
[30] If a section in one grade level was not available, we would select another section in grades 1-6.

the intervention as it is implemented. The Pre-Analysis Plan was registered publicly in Clinical Trials. Analyses and data visualizations were conducted in Stata/IC 15.1.

Analytical model

The effect of the treatment on the outcome variables for RQ 1 is estimated by the following weighted least squares Linear Probability Model:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \overrightarrow{\beta_2}\vec{X}_{ij} + \overrightarrow{\beta_3}\vec{S}_j + \varepsilon_{ij} \qquad (1)$$

Where:

- Y_{ij} denotes the outcome variable for pupil *I* in school *j*, classified as a binary variable
- T_j denotes the treatment variable (binary variable for whether school *j* received the nudge intervention)
- $\vec{X}_{ij}$ represents a vector of pupil, school, and class level co-variates[31]
- $\vec{S}_j$ represents strata fixed effects
- 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 or vector of coffecients determined by the regression model (β_1 is coefficient of interest)

Heterogeneity in the treatment effect was explored for different groups in the sample. As specified in the pre-analysis plan, sub-group analysis was done on pupil gender, pupil grade level (grades 1-3 in comparison to grades 4-6), and WinS Index Scores (as a measure of level of WASH infrastructure). For sub-group analysis, a single regression on the full sample with an interaction term between the treatment term and the sub-sample was used.

A similar model was used for RQ2. However, the unit of analysis was at the facility level. Class and school level covariates were still used, but no pupil covariates were.

β_1 is the coefficient of interest, denoting the estimate of treatment impact. For the outcome of handwashing with water and soap for example, a coefficient of .10 can be interpreted as the following: "Nudges increased the probability that the pupil washed their hands with soap and water after toilet use by 10 percentage points." Throughout the report, in addition to describing the treatment effect in the text, we also show predicted means $\widehat{Y_{ij}}$ from equation 1 for treatment and control schools, and their associated 95% confidence intervals.

Sample Weights

Sample weights were included to account for the uneven sampling probability. The purpose of including sampling weights is to recover the population estimates from a sample where every

[31] These were district, grade, gender, class size, total pupils, and strata fixed effects. The school and class characteristics are not at the *ij* level but rather only the *j* level because they are characteristics of the school and not the individual pupil. We omitted the nudge conduciveness index, which is a deviation from the Pre-Analysis Plan. This is discussed further in the Appendix.

observation does not have an equal chance of being selected, such that estimates are generalizable to nudge-eligible schools in ZDN. Specifically, since a pre-determined number of classes were sampled in each grade within a school,[32] pupils at schools with a greater number of sections had a lower probability of being selected. Additionally, we oversampled "nudge-enabler" schools, relative to their proportion in the overall sample of eligible schools. Sample weights equal to the inverse of the probability of a school being selected out of the "nudge-enabler" schools in the eligible school population was applied. For handwashing outcomes, this was multiplied by the inverse probability of a classroom being selected for observation to account for the uneven sampling probability of classrooms in a grade.

Multiple Interference Correction

As a result of the large number of hypotheses being tested, the analysis for this study adjusted for multiple inference using the Holm-Bonferroni procedure.[33] This ensured that statistically significant findings are more likely to reflect actual change rather than a false positive. For all regression results presented in the report where multiple inference correction was used, the corrected p-value is indicated. We conducted multiple interference corrections for our analysis of the presence of functional handwashing facility near toilet facility, presence of functional handwashing facility with soap near toilet, and functioning handwashing outlet with or without soap per pupil. The research question on facility access is the only one with three outcomes inferred from one family (set) of observed variables. However, we did not adjust in our subgroup analyses, because none of the subgroup analysis treatment effects for handwashing rates were statistically significant.

An asterisk (*) is added when the p-value less than 0.10. A p-value less than 0.10 indicates that the difference was statistically significant at the 10% significance level. Two asterisks (**) are added when the p-value is less than 0.05. Similarly, this indicates that the difference is significant at the 5% level, the conventional threshold at which the evidence is strong enough to reject the null hypothesis that the treatment had no impact for the entire population. Three asterisks (***) indicates the strongest level of statistical significance at a p-value of less than 0.01. This indicates that the difference is significant at the 1% level.

Data Visualization

Data visualizations in this report showing predicted population means (for the overall sample or a specific treatment group) from the linear probability regression in equation 1. They include a 95% confidence intervals around the means to highlight the precision of each estimate. These are represented using error bars. The confidence interval is also indicated in brackets directly following estimated population means described throughout the report. Estimated population means at the 95% confidence level can be interpreted as follows: there is a 95% chance that the true population mean is between the numbers indicated in the brackets.[34]

---

[32] For classroom observations, one class in each grade in each school was sampled independent of the total number of classes in the grade. For the pupil survey, 8 total pupils in each grade from up to 2 classes were sampled.

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

[34] As a more technical definition, a 95% confidence interval means that if the data generation process were run repeatedly, and the confidence interval were to be re-calculated each time, 95% of the time the true parameter would lie within the interval.

4.2. PROCESS EVALUATION

Nudges Presence and Condition

Analysis and visualizations were conducted in Stata. The sample of the analysis includes only those that meet all of the following criteria:

- School is in treatment[35]
- School is one in which facilities observations were conducted
- Facility was available for facility observations at both baseline and endline[36]
- For each type of nudge, facilities for which baseline records shows the relevant nudge was installed

Key Informant Interviews

Interviews were recorded, transcribed, and coded on Excel utilizing the principles of Framework Analysis.[37] Framework Analysis is a variant of thematic content analysis that draws out significant units of meaning to describe and interpret what is happening in a particular setting from multiple perspectives, primarily but not exclusively around an a priori issue.[38] In this case, we utilized it to describe and interpret what happened in the section or at the school from the perspective of each key stakeholder, primarily around their interaction and reaction to the nudges.

[35] Note, however, that enumerators observed no presence of nudges in any control schools.

[36] We assume that we could not observe at some installed sections at endline because of random reasons, such as a facility being locked due to a teacher's absence. In the Appendix, we detail a conservative estimate in which we assume that all installed facilities not visited at endline no longer had nudges. Only five footpaths installed at baseline were not revisited and recorded at endline.

[37] This was the project protocol laid out to enumerators. In the course of the project, enumerators conducted 42 interviews, but failed to save the audio files for six interviews, so we were unable to analyze these six interviews. Ultimately, we completed data analysis for 32 interviews.

[38] Ritchie, J. and Spencer, L. (1994) Qualitative Data Analysis for Applied Policy Research.

FINDINGS

1. IMPACT FINDINGS

1.1. HANDWASHING RATES

For this evaluation, observed hand-washing after toilet use is the primary outcome of interest. As HWWS is the most effective method of handwashing, we report HWWS rates for our analysis on the full sample as well as subgroups. For our full sample, we also report handwashing with only water rates and handwashing with at least water rates. Handwashing with water reduces bacterial contamination, but is significantly less effective than handwashing with water and soap.[39]

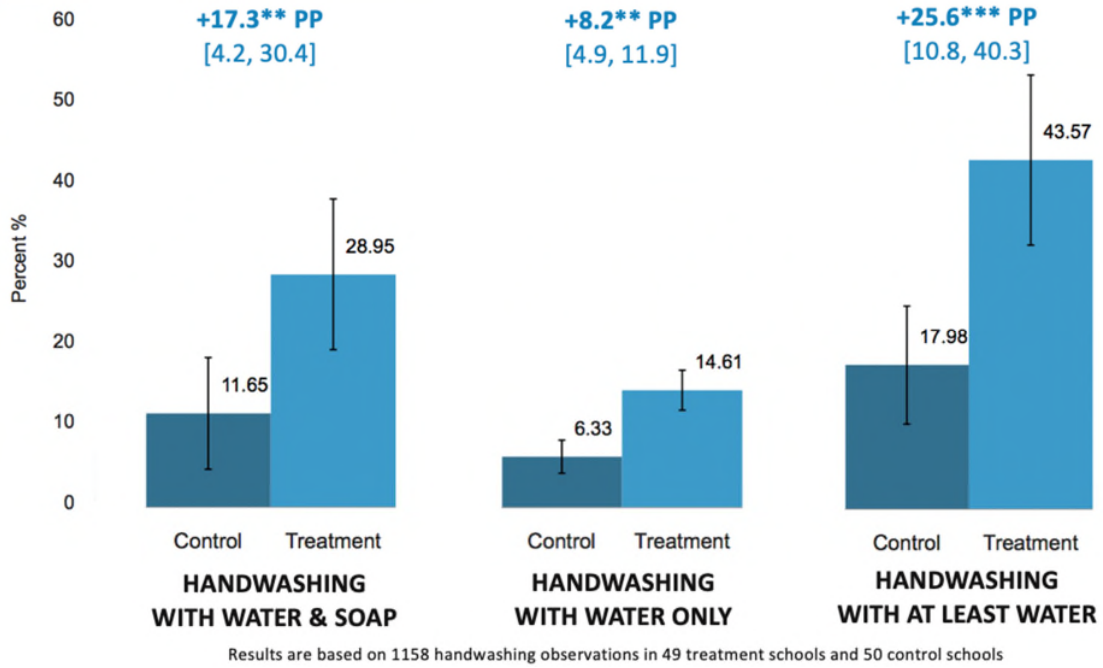
Nudges increased HWWS rates by 17.3 pp [4.2, 30.4] (p=0.010**), or 148 percent. Among pupils in grades 1-6, the HWWS rate was 11.7% [4.7, 18.6] in control schools, and 29% [20.0, 38.3] in treatment schools. Rates in control schools were higher than rates found previously in Camarines Norte and Puerto Princesa.[40] This may have been because our sample schools had better infrastructure than the schools in the HiFive Phase I and Phase II evaluations, due to stricter eligibility criteria for receiving nudges. This may also have been due to the influence of the COVID-19 epidemic. Data collection took place at the beginning of the pandemic, when teachers and pupils may have had heightened conscientiousness of handwashing. This would have led to increased HWWS rates in both control and treatment schools, but would not have affected our treatment effect size, as the impact of COVID-19 on handwashing rates would have been the same across the two groups. We discuss this further in Technical Considerations and Limitations.

Water-only handwashing in treatment schools was also 8.2 pp (p=0.000***) higher than in control schools, where it was 6.33%. Combining the nudges' impact on handwashing with water-and soap and with water-only, this led to an increase of 25.6 pp (p=0.001***) in handwashing with at least water. All results are shown in Figure 3.

[39] Burton M, Cobb E, Donachie P, Judah G, Curtis V, Schmidt WP. The effect of handwashing with water or soap on bacterial contamination of hands. International Journal of Environmental Research and Public Health 2011; 8: 97-104.

[40] In our HiFive Phase I evaluation, among pupils from grades 1-6 in control schools, observed HWWS rates after toilet use was an estimated average rate of 2.2% [0.8%, 3.6%]. In our HiFive Phase II pre-post evaluation, during baseline data collection, we observed HWWS rates after toilet use of 2.45% among grades 1-6 pupils.

Figure 3 Observed Rates of Handwashing After Toilet Use

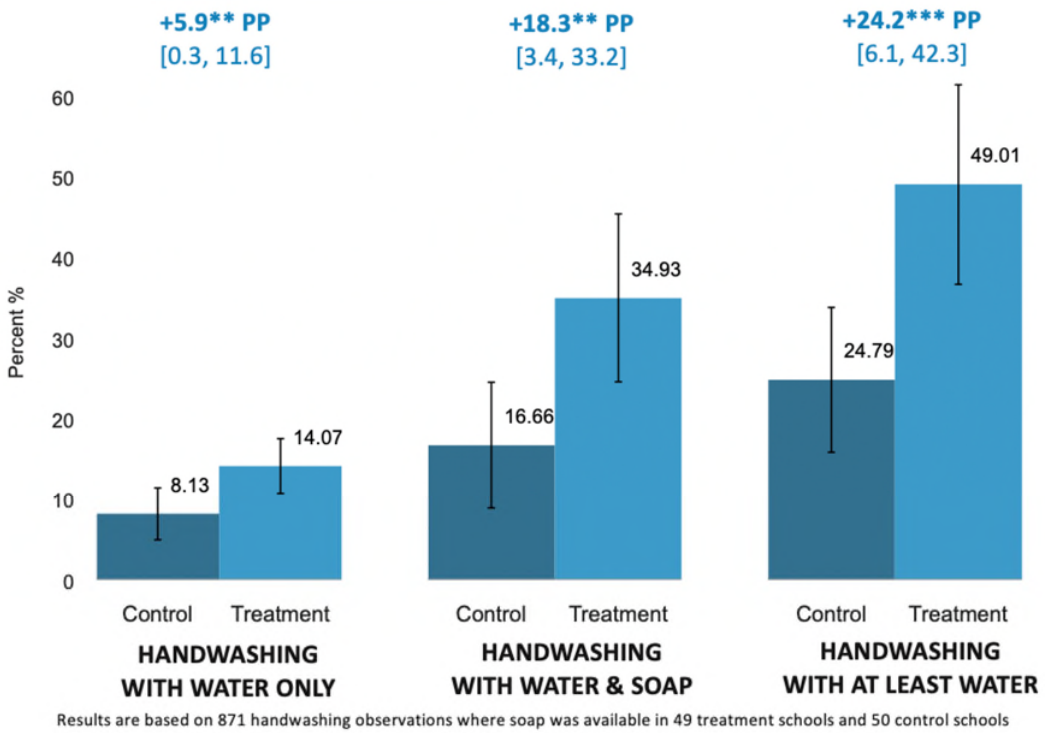


As the treatment effect on handwashing with water and soap is twice greater in magnitude than handwashing with only water, this suggests that when nudges influence pupils to approach the handwashing station, pupils are typically choosing to wash their hands with water *and* soap, rather than simply water. A multi-factorial clustered randomized trial would be needed to more fully explore the effects of the nudges on handwashing with water only vs. handwashing with water and soap.

However, this introduces the question of whether the increase in water-only handwashing is because pupils *chose* to wash hands with water though soap was available, or *had* to wash hands with water because soap was unavailable. An additional analysis considering only handwashing observations at sections with soap available found that there was still a small increase in water-only handwashing, of 5.9 pp [0.3, 11.6] (p= 0.038**) (Figure 4). This provides suggestive evidence[41] that for a small subset of students, the nudges were effective at bringing them to the handwashing station, but did not translate to the desired behavior of handwashing with water and soap, even when soap was present. The majority influenced by the nudges chose to wash hands with water and soap.

[41] This analysis provides suggestive, not causal evidence because soap availability was only measured at endline, and was impacted by the intervention, as discussed later in this section.

Figure 4 Observed Rates of Handwashing After Toilet Use in Sections with Soap



Subgroup Analysis

Subgroup analysis on HWWS rates was conducted on three characteristics: pupil gender, pupil grade, and the level of WASH facilities at the school. The intention was to explore whether particular groups responded more positively to the nudges intervention than others.

The intervention had no differential handwashing impact on grade groups or genders. Pupils in grades 1-3 and pupils in grades 4-6 responded similarly to the nudges: effect sizes were 17.8pp and 16.2pp respectively, p-value of difference= 0.847). Boys and girls also responded similarly: effect sizes were 17.1pp and 17.5pp respectively, p-value of difference=0.926).

Figure 5 Handwashing with Water and Soap by Grade Group

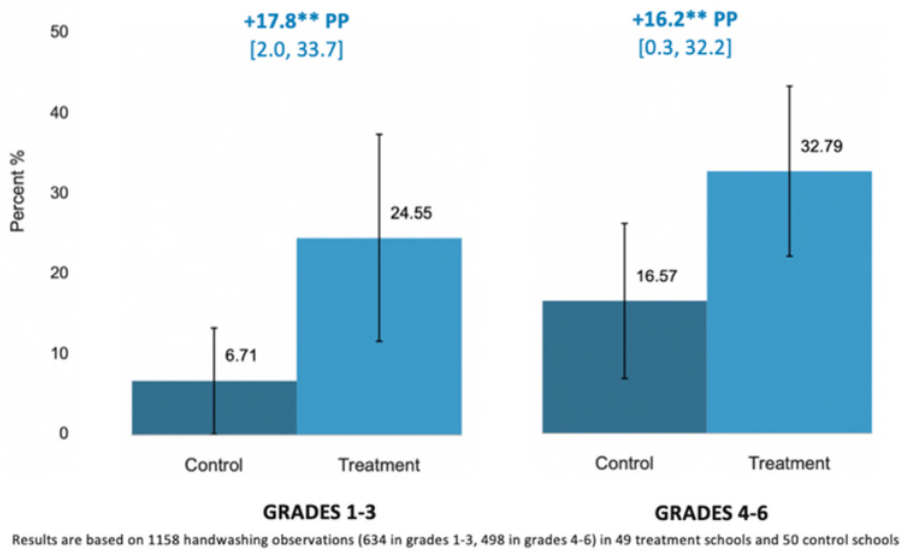
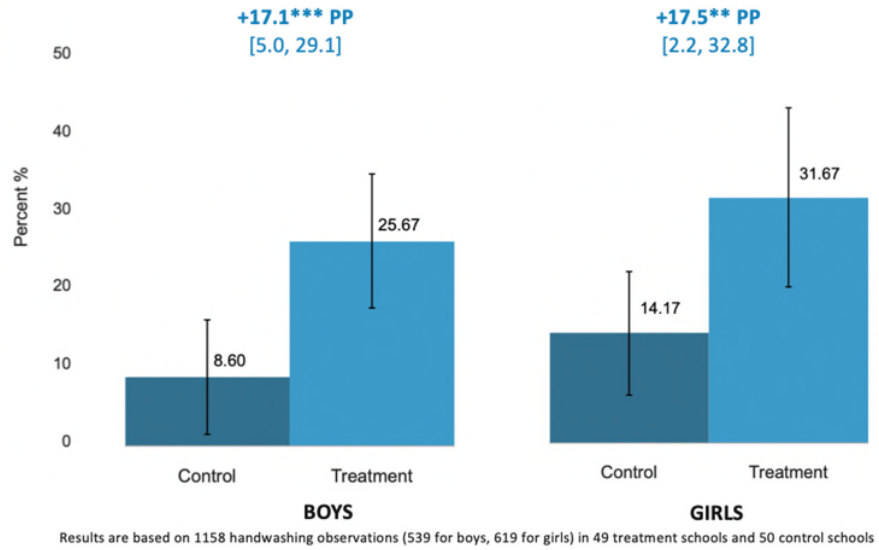


Figure 6 Handwashing with Water and Soap by Gender



To compare schools with worse WASH facilities with schools with better WASH facilities, we utilized schools' WinS IQ scores,[42] an index constructed based on selected WASH-related indicators from OMS data. We conducted two analysis: the first assessed heterogeneous effects using the continuous WinS index score (0-9), as was pre-specified. The second analysis was exploratory and classified WinS index

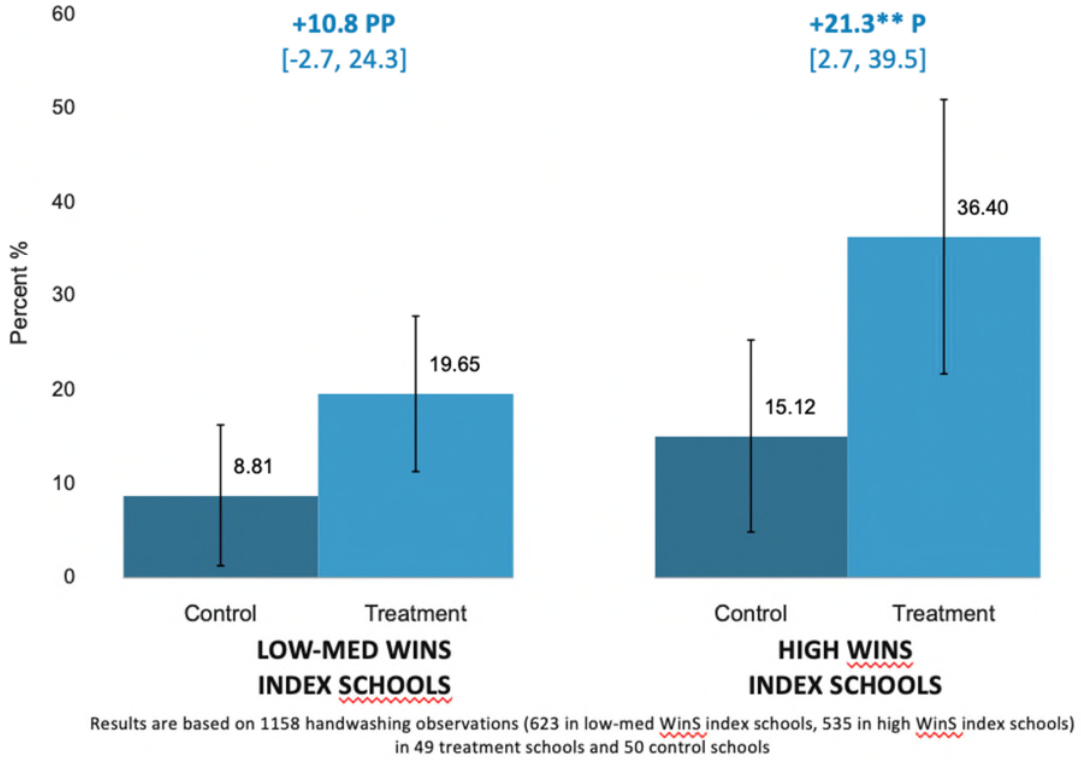
[42] WinS IQ scores were generated using data from OMS, including schools' water availability, soap availability, and toilet to pupil ratio.

as a categorical variable, comparing those schools which were considered High WinS index at baseline to those that were Low or Medium.[43] Impact estimates on HWWS for both groups are shown in Figure 7 and in Table 4.

There is suggestive evidence that nudges may be more effective with better facilities. However, the results are inconclusive, due to low sample size.

In both subgroups, HWWS was higher in the treatment group than in the control. Among "Low-Med" WinS Index schools, nudges increased HWWS by 10.7 pp, but the 95% confidence interval around this difference ranges from -2.7, a modest decrease, to 24.3, a large increase. In "High" WinS Index schools, the estimates of treatment impact are more precise: results suggest a 21.1 pp increase, with the difference ranging from a small increase of 2.7 to a large increase of 39.5 (Figure 7). The impact estimate for High WinS Index Schools was twice as large as for Low-Med WinS Index School; however, the 10 pp difference in treatment effects between WinS categories was not statistically significant at conventional levels (p=0.373) (Table 4).

Figure 7 Handwashing with Water and Soap by WinS Index Scores



[43] WinS IQ scores of 3-5 were determined at baseline and labeled "Low", 6-7 "Medium", and 8-9 "High." Since only a small number of schools fell into the "low" WinS score category (12 schools of out 99), treatment impact among "low" and "medium" schools was compared with the impact among "high" schools. This is preferable to estimating impact across the three clusters, which we did not do due to the small size of "low" WinS score clusters.

However, it is reasonable to expect schools with less WASH infrastructure to derive less benefit from the nudges, as there may be other barriers to handwashing at these schools. For example, the school may have lower social norms around handwashing, or lack supporting infrastructure like regular water access or soap availability.[44]

Table 4 Results of Subgroup Analysis on Handwashing with Water and Soap

Impact compared by	Coefficient	P-value
Pupil gender	0.00	0.926
Grade Group	-0.02	0.847
School WinS Index Score (continuous)	.01	0.89
School WinS Index Score (high vs low-med)	0.10	0.373

1.2. ACCESS TO HANDWASHING FACILITIES

It is important to understand whether the nudge intervention increased access to functional handwashing facilities[45] with soap near toilets, as access to such a facility enables pupils to wash their hands with water and soap. In other words, it is a precondition for our RQ1 outcome, HWWS after toilet use.

As was described in the Methodology section, we selected schools which tended to have better WASH infrastructure in place, including better soap availability. Seventy-five percent of schools that met our infrastructure eligibility requirements had regular soap availability, whereas only 47% of schools that did not meet our eligibility requirements had regular soap availability.[46] In the context of this study, we did not provide soap as part of the nudge package. This decision was made for several reasons:

First and foremost, we wanted to test the hypothesis that the nudges could increase soap access by changing behavior. Based on OMS data at baseline, 68 percent of schools in our sample have MOOE funds set aside for soap, suggesting most schools did not face financial restrictions to providing soap. Although in previous empirical literature, nudges primarily served to encourage student behavior change rather than increase access to soap and water availability[47], we hypothesized that the nudges may encourage teachers to replenish soap and water supply. Schools rely on staff to put out the soap — sometimes on a daily basis, as soap left out overnight, especially outdoors, are at risk of being stolen. Several of the nudges could be said to serve a reminder function; footpath and eye sticker draws attention to the handwashing station, and the arrow sticker carves out a visual space for soap (or lack thereof).

[44] In fact, stakeholders in our qualitative interviewers specifically mentioned that inconstant water access limited the benefit pupils derived from the nudges.

[45] This is defined as a handwashing station usable by students with water.

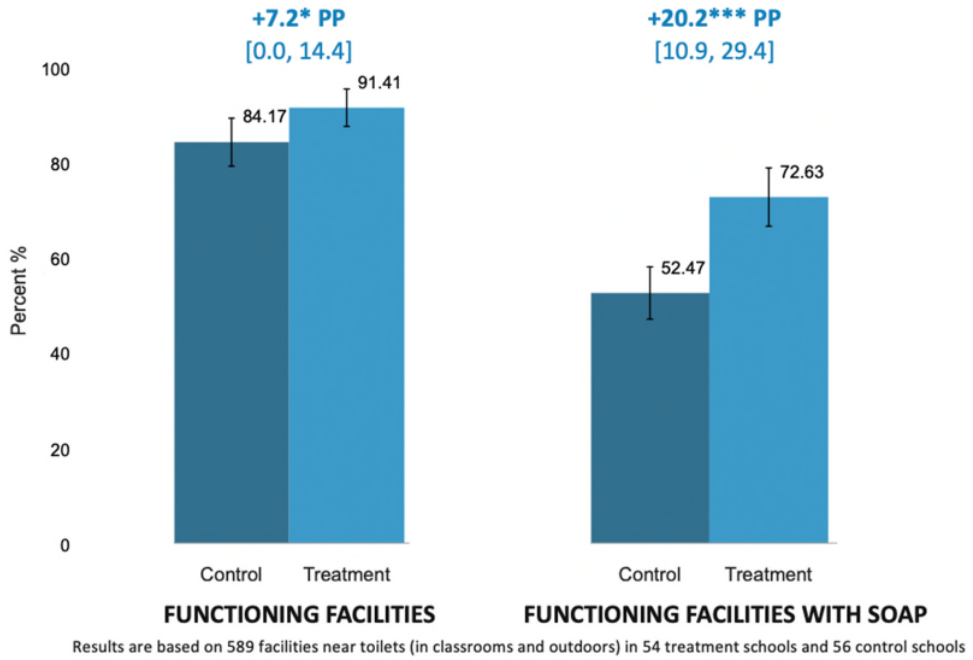
[46] These figures are based on OMS indicators. See the Appendix for a comparison between schools that met our infrastructure eligibility requirements (our sampling frame) and schools that did not.

[47] In one study, soap and water were already uniformly available (Blackwell et al 2018). In another, it was provided as part of the intervention (Dreibelbis et al 2016).

Additionally, we wanted to mirror scale-up conditions, where soap would likely not be provided, and to ensure implementation fidelity. In Zamboanga del Norte public schools, generally funds for soap are already disbursed as part of general maintenance funds (MOOE); principals and teachers are responsible for purchasing and putting out soap by the handwashing station. Due to the four months between installation and observation, had we decided to make soap provision part of the intervention, it would have been difficult for us to ascertain that soap was delivered and made available at all treatment schools for the entire implementation period.

The nudges increased access to functional handwashing facilities with soap near toilets by 20.2 pp [0.11, 29.4] (p=0.000***), or 38 percent. Nudges also slightly increased access to functional handwashing facilities near toilets, by 7.2 pp (p=0.1*). Both results are shown in Figure 8.

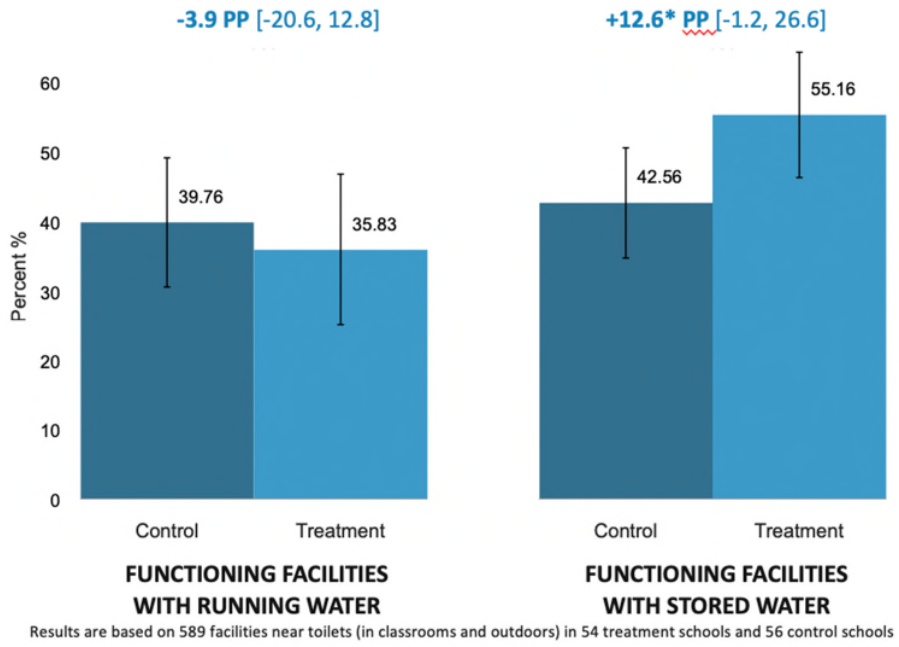
Figure 8 Access to Functional Handwashing Facilities Near Toilets



These findings are particularly significant because the eligibility criteria for the study already restricts the sample to those with better WASH infrastructure than the overall population of schools in Zamboanga del Norte. According to OMS data, 68 percent of our schools had MOOE funds for soap; yet during our facility observations at endline, just over half of toilets (52.5% [50.0, 58.0]) in control schools actually had a nearby functional handwashing facilities with soap available. This suggested there is room for improvement in soap distribution, which is exactly what we found in the intervention schools. Without requiring more funds for soap, the nudges were able to increase pupil access to soap, likely by reminding teachers, who are primarily responsible for maintaining facilities day-to-day, to supply soap more consistently at handwashing stations.

To get insight into how nudges improve access to functional facilities in general, treatment effects were compared between two handwashing station types: "faucet" and "stored water."[48] At faucet facilities, students turn on running water using a faucet. At stored water facilities, students use buckets to scoop water from a stored water container next to the sink basin to wash their hands. As can be seen in Figure 9, the slightly higher access to functional facilities in treatment schools compared to control schools was driven entirely by its higher access to facilities *with stored water,* which increased by 12.6 pp (p=0.072*). There was no effect on access to facilities with faucets (p=0.643). This suggests the nudges were also encouraging teachers to provide or replenish stored water.

Figure 9 Access to Handwashing Facilities Near Toilet By Type



1.3. MEDIATION ANALYSIS

As the findings above show, nudges significantly increased pupil HWWS rates. However, nudges also increased pupil access to functional facilities with soap near toilets; in other words, it also increased the availability of soap and water. The nudges were intended to lead to *behavioral change*. Therefore, we conducted mediation analysis to understand whether the increased pupil HWWS rates was indeed driven by students' behavioral change, or by increased opportunity to wash hands with soap due to increased availability of soap and functioning handwashing facilities.

Controlling for soap access[49], treatment effect size reduces slightly, from 17.3 (p=0.010**) to 16.4 (p=0.013**), but remains positive and statistically significant (Table 5). This suggests that although

[48] These are considered exploratory analysis, as they were not pre-specified in our PAP.

[49] We did not and could not control for access to a functional handwashing station, as all handwashing observations took place in classrooms with a functional handwashing station. Having a functional handwashing station was a section eligibility requirement for handwashing observations.

access to soap was an important enabling condition, handwashing rates increased primarily because of *student behavioral change attributable to the nudges.* In other words, while teachers putting soap and stored water out more often did lead some students to wash hands more, the nudges also directly influenced students to choose to wash their hands more often.

Table 5 Mediation Analysis Regression Results

	(1) Main Specification	(2) Including Soap Availability
Treatment	**0.173****	**0.164****
	(0.07)	**(0.07)**
Soap available		0.08**
		0.033
Female	0.029	0.028
	(0.03)	(0.03)
Number of pupils in class	0.005	0.005
	(0.00)	(0.00)
School Population	0.001	0.001
	(0.00)	(0.00)
Grade Dummies	Yes	Yes
District Dummies	Yes	Yes
Strata Dummies	Yes	Yes
Adj. R-squared	0.428	0.433
Observations	1158	1158

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2. PROCESS FINDINGS

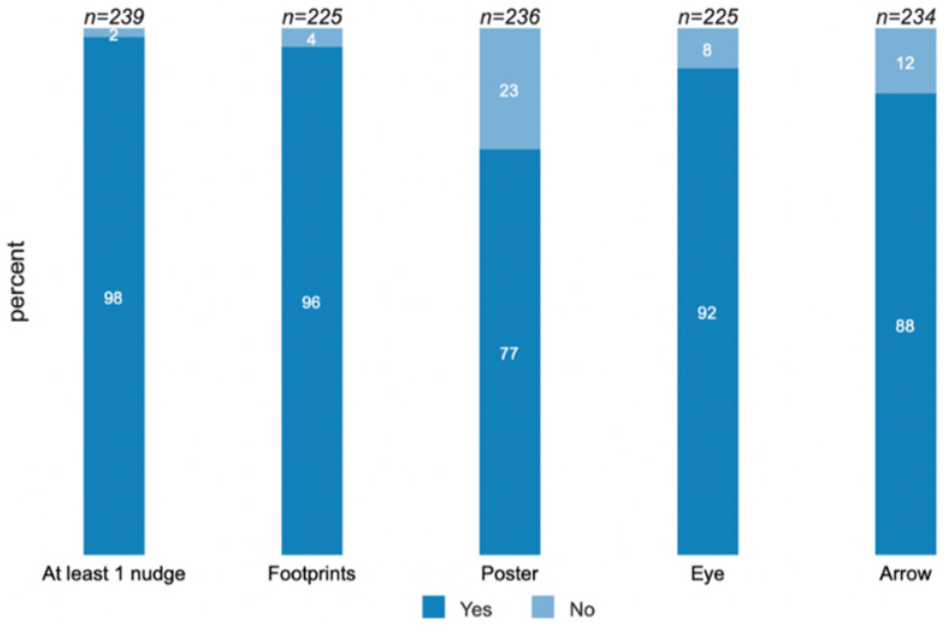
The process evaluation is intended to inform our recommendations by shedding light on how well the nudges lasted through the four-month period between implementation and observation, how staff and pupils engaged with the nudges, and what perspectives the stakeholders had about the nudges. Structured observations in treatment schools and one-on-one key informant qualitative interviews comprise process evaluation data sources. The first section presents findings on the state of the installed nudges at endline. The second section presents key informant reflections on the nudges.

2.1. CONDITION OF NUDGES IN IMPLEMENTATION

There was a four-month gap between nudge installation in October 2019 and endline data collection in February 2020. Despite this significant amount of time, among the sections where nudges were installed, over 98 percent still had at least one nudge visible at endline. For each individual nudge

category, more than 75 percent installed remained at endline, as can be seen in Figure 10.[50] The calendar of posters was the nudge with the greatest percentage of missing installations. Out of the sections where it was installed, 23 percent lacked the posters calendar in their toilet stalls at endline.[51]

Figure 10 Presence of Installed Nudges at Endline



Enumerators marked the condition of each nudge observed at endline as either "good", "fair," or "bad." The definition provided to enumerators for each label is presented in Table 6 below. Overall, the vast majority of the nudges that remained at endline were in good condition, as seen in Figure 11. The footpath was the only nudge with significant variation in its condition, with 11 percent of installed sections having footpath in "bad" condition, and 23 percent in "good" condition. As the descriptions for "bad" and "fair" indicate, the deterioration was primarily due to flaking or fading paint. For the poster calendar, eye sticker, and arrow sticker, there was relatively little variation in the condition of the nudge. The reasons why nudges went missing or deteriorated in condition are listed in Table 7, as reported by interviewed teachers and principals.

[50] For this analysis, we consider only 1. For each nudge, this analysis considers only those at: 1) schools that have not attrited; 2) schools in treatment; 3) facilities which we observed at both nudge installation and endline data collection; 4) facilities for which baseline records show the given nudge was installed.

[51] Enumerators marked the poster calendar "not present" if it was not in the toilet stall, but also recorded if the posters calendar was present anywhere else in the section. There were thirteen sections in which the posters calendar was moved from inside the toilet stall, where it was originally installed, to another location in the classroom. If these sections are included in the analysis, the percentage of "missing" installation decreases to 17.5%, which is still higher than the other nudge types. In our conservative estimate of "present" nudges (see Appendix), 6% of the installed footpath, an increase from 4% in the main analysis, was not present at endline.

Figure 11 Condition of Nudges Present at Endline

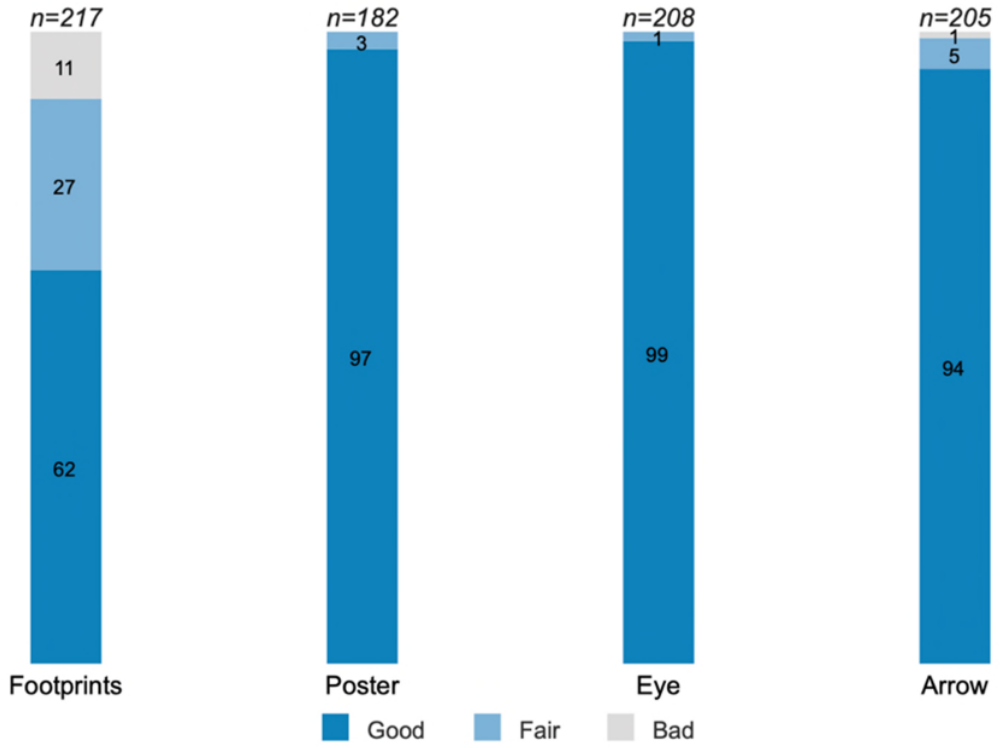


Table 6 Definitions of Nudge Condition

Nudge	Good	Fair	Bad
Footprints (for all designs) [52]	Footpath and footprints are still almost entirely visible; path may have flaking at edges but is clear Most to all spray-painted footprints are still present and visible; path guide is clear Most of the foot stickers are still present and visible; path is clear	A few to some footprints have faded, and more than footpath edges have faded; footpath has some gaps but path still visible A few to some footprints have faded; path guide has gaps due to paint fading but start and end point still clear	Most of the footprints have faded; footpath has big gaps and/or start and end points unclear Most of the footprints have faded; path guide has big gaps and/or start and end point unclear Most of the foot stickers have faded;

[52] The primary design was a painted footpath with spray-painted footprints. If this could not be installed, the section received either only spray-painted footprints, or spray-painted stickers. The definitions provided for these alternate designs are also shown in the table.

		A few to some foot stickers are gone; path is still clear though it may have gaps	path has big gaps or is unclear
Poster	The poster's main image is fully untouched, but the poster may have very minor flaws such as a curved body or bent corners	The poster's main image may have some creases or water stains, but all image and text are still fully legible.	Some of the image and text are damaged and not legible
Eye	Good - the eye sticker is entirely or mostly intact	Fair - the eye sticker may be partially torn or damaged but is still readily identifiable as an eye	Bad - the eye sticker is so torn or damaged it cannot be easily identified as an eye
Arrow	Good - the arrow sticker is entirely or mostly intact	Fair - the arrow sticker may be partially torn or damaged but is still readily identifiable as an arrow	Bad - the arrow sticker is so torn or damaged it cannot be easily identified as an arrow

Table 7 Reported Reason for Nudge Condition

Nudge	Reason for "Not Present"	Reason for "Fair" or "Bad" Condition
Footprints	o N/A	o Flaking paint
Poster	o Fallen off o Gone	o Dropped in water
Eye	o Fallen off o Gone	o Torn
Arrow	o Fallen Off o Gone	o Torn

These findings on the condition of the nudges at endline are particularly noteworthy because the nudges were relatively inexpensive to install. In our study, the nudges cost PHP 3420 or USD 67.50 to install per school, with an average of 5 sections and 0.4 GHW station installations per school.[53] We estimate that if scaled up in the Philippines, the nudges would cost PHP 3930 or USD 77.51 per school. Assuming each school would have 6 sections and 1 GHW station, and all facilities are nudge eligible, this means nudges would cost PHP 561 or USD $11.07 per handwashing station.[54] For a breakdown of the inputs for our cost estimates, please see the Appendix.

[53] We installed at 299 sections and 20 GHW stations in 57 schools.
[54] For this estimate, we assume fixed costs are shared equally across sections and group handwashing stations, and that the school does not need to pay for labor (by, for example, relying on parent volunteers).

2.2 INSIGHTS FROM STAKEHOLDERS

Stakeholder perceptions

According to interviews, both pupils and teachers had limited interaction with the nudges.[55] Pupils were reported to be using the footpath, though this was both to walk from the toilet stall to the handwashing station, and to play on the footpath. Some pupils were also flipping the calendar of posters to look at the posters, according to teachers. In general, teachers stated that pupils showed curiosity about the nudges after installation, but that pupil enthusiasm dimmed after some time, so that the nudges are now simply part of the classroom. Teachers interacted with the nudges primarily in maintaining them: changing the poster displayed monthly;[56] sweeping up flaked paint from the footpath; and touching the stickers.[57]

Broadly, most interviewed principals, teachers, and students were enthusiastic about the nudges.[58] Students mostly liked the nudges because they were colorful and attractive in the classroom. When asked about whether or not they liked the footprints, one student stated, ”I like [the footprints], because it looks good.” Principals and teachers liked the nudges because they appreciated the nudges' purpose of improving student hygiene, and saw no disadvantage to having the nudges by handwashing stations. One anonymous principal stated, “I am happy if there are new things in my school, especially if they are lasting, useable and useful.“

Interviewees understood some nudges better than others. This is to be expected, particularly among teachers and pupils. The purpose of some nudges (such as the footpath) are more self-explanatory than others (such as the stickers). Prior to installations, we conducted an orientation on the nudge intervention for principals of treatment schools, but not for teachers or students. This was because we needed principals' buy-in for the installation of nudges, but wanted to limited the amount of information about nudges given to teachers or students. Effective nudges are designed to work even if they are not understood, meaning that even without orientation, teachers and students should still theoretically respond to them.

Most principals, teachers, and pupils understood the purpose of the footpath. Most interviewed pupils were also able to explain the message of a poster in the calendar when shown. This is important, as the poster is the only nudge with a specific message that pupils must comprehend. Pupils may have understood these nudges because teachers had correctly explained it to them. Many of the teachers interviewed stated that they explained the nudges after installation to students, often because the

[55] Most nudges were installed in sections, where teachers and pupils spend most of their time. In the schools we conducted at, no principals reported interacting with nudges, either in sections or at group handwashing stations.

[56] Among those who said the calendar of posters was occasionally flipped, while a few teachers allowed pupils to flip the calendar, most flipped the calendar themselves. One teacher even said she did not trust a student to do flip the posters calendar without damaging it.

[57] One respondent noted they got the stickers laminated and then re-installed them. Conversely, another stated they stored away the stickers after they fell off.

[58] There were some outlier responses, primarily from respondent who did not know or understand the nudges. One interviewed teacher, whose classroom did not have any nudges, at first said they did not support the nudges. However, after the interviewer showed the nudges and explained their purpose, the teacher stated they now understand and supported them. Another suggested they would remove the arrow sticker from their classroom, but then specified it was because they did not know what purpose the arrow had. When asked whether they liked or disliked specific nudges, several students also provided extremely short responses (silence or a nod) without providing a rationale. This is likely due to the students' shyness, rather than reluctance to provide a negative response, as these students had similar demeanors in the remaining duration of the interview.

students asked. "[My students] liked it and they asked, 'Ma'am! What is it for?' So we taught them and demonstrated how to use it, especially the footprints," one teacher said.

Conversely, principals, teachers, and pupils most commonly could not explain or incorrectly explained the purpose of both the eye sticker and the arrow sticker.[59] However, according to the behavioral change model, lack of understanding should not impede impact the purpose of the nudges on pupils. This is because, aside from the calendar of posters, the other nudges' pathways to impact do not rely on conscious knowledge about their purpose or design. Their pathways are either subconscious or reflexive. For example, while the eye sticker's purpose was not well-understood, when pupils were asked how the eye sticker made them feel, several said that the eye sticker made them feel "watched." This is the exact intention of the nudge's design, which is meant to activate the motivator of social desirability by making pupils feel watched.

Suggestions for Improvement

The interviewees had several suggestions for improving the nudges. These are presented in Table 8. Many of these suggestions are based upon a trade-off between functionality and cost (for example, lamination of posters or stickers), or are matters of personal preference (color of the nudges). Many of these adjustments can be decisions made by individual schools in future installation efforts. There are only two suggestions that should not be implemented: adding additional facial features to the eye sticker, and additional sticker next to the arrow to signal purpose of the arrow sticker. This is because these changes significantly alter the nudges' design, and in ways that could impede the intervention's behavioral change model. The "watching eye" sticker would likely not make pupils feel as "watched" if it was changed to a smiley face. As these two specific changes may *negatively* affect the impact that the nudges have on handwashing and other outcomes of interest, we do not recommend adopting them.

Table 8 Suggestions for Nudge Design Changes from Principals and Teachers

Nudge	Suggested Changes
Footpath	o Use different (more long-lasting) paint o Paint in various different colors
Poster	o Change binding to make easier to flip and hang o Laminate posters to be waterproof o Enlarge posters and change colors
Eye	o Laminate sticker to be waterproof o Add other facial features* o Change color
Arrow	o Laminate sticker to be waterproof o Add additional sticker to signal purpose* o Change color

[59] Common alternative explanations were that the arrow or eye sticker were pointing to the faucet, or to the sink.

* *not recommended*

Several school staff also specifically requested additional support to improve water access at their school. As one principal stated, "[I hope UNICEF] can launch a program helping every school to have enough source of water. We cannot fully maximize [UNICEF]'s programs due to deficiency of water supply." For this principal and other respondents, more stable water access would increase the impact the nudges have on pupil handwashing rates.[60]

3. TECHNICAL CONSIDERATIONS AND LIMITATIONS

Timing and COVID-19

Endline data collection began in February 2020, at the beginning of the COVID-19 pandemic in the Philippines. This timing likely increased rates of handwashing in both control and treatment schools. While data collection occurred just before widespread school cancellations and movement restrictions, teachers and students may have already been showing heightened vigilance over handwashing. Shortly before we arrived, DepEd Zamboanga Del Norte division sent a memo about COVID-19 and the importance of handwashing to all schools. In key informant interviews, several respondents mentioned COVID-19 when discussing hygiene messages they had recently heard or delivered. However, since our evaluation design randomized treatment and control groups, any boost to base handwashing rates due to COVID can reasonably be expected to be comparable across the two groups. Therefore, this is unlikely to bias our estimates of impact.

Attrition

If schools in our sample are lost to follow up for non-random reasons, then results from our final sample may be biased. Attrition is unlikely to be a source of bias in our estimates. The primary causes of attrition are school-wide ineligibility for nudge installation or handwashing observations (determined using the same criteria between treatment and control schools), and safety and security concerns.[61] These causes of attrition are unlikely to differentially impact treatment and control schools. This is supported by the comparable rates of attrition for treatment and control. From the original sample of 132 schools, 17 treatment schools and 16 control schools attrited. Finally, there is no evidence of selective attrition. The remaining sample (49 treatment schools, 50 control schools) is well-balanced along baseline school-level variables (see Table 2).

Observer Bias

Teachers and students may behave differently when being observed than they would otherwise, particularly because handwashing is a socially-desirable behavior. To minimize this risk, enumerators:

- Did not conduct section facility observations at lunch
- Did not visit treatment schools if they were part of October installation team, and visited treatment and control schools on different days if they were not part of the October team[62]

[60] This aligns with our subgroup analysis of schools by WinS Index categories, which suggests that nudges are more effective in schools with better WASH facilities.

[61] There were reported active rebel presence in the surrounding areas of or en-route to some schools within our sample.

[62] Some enumerators may have been better at concealing the true purpose of the study than others. By assigning each enumerator randomly to control or treatment schools each day, we balanced out the effect of such differences across control and treatment groups. This ensured our impact estimate remains unbiased.

- Did not reveal to teachers and pupils in observed classes that they were observing handwashing, stating only that they intended to observe "normal classroom behavior."
- Asked questions and recorded information about the classroom unrelated to hygiene, as "red herring" questions
- Removed the UNICEF logo from enumerator name tag and refrained from mentioning UNICEF throughout the school visit.[63]

We conducted a robustness check to see whether estimates are sensitive to dropping the two schools we suspect observation effects were present (i.e. those we observed before changing the protocol around UNICEF's logo). We dropped 93 handwashing observations where teachers appeared to know the study's purpose (unreliable) or explicitly reminded pupils about handwashing (reminder), and re-ran analysis of the impact of the nudges on HWWS. These represent overly conservative estimates of impact, since reminders were likely not influenced by our presence.[64] The result is a more conservative impact estimate of 14.1 pp (p=0.029**), which is lower than the estimate of 17.3 pp from our main finding. However, the effect size remains large and statistically significant. For full details, see the Appendix.

External Validity

Evaluation findings are representative of Zamboanga Del Norte (ZND) schools that meet the school-level eligibility requirements, as laid out in Methodology 2.1. Any generalization of results to other schools in Zamboanga del Norte or in the Philippines should be approached carefully. In particular, schools in the sample may differ from other schools in two distinct ways. First, UNICEF has identified Zamboanga del Norte DepEd division as a priority target district for WinS programming, so its schools have benefitted from more focused WinS support from UNICEF than other districts. Second, schools identified as eligible for the intervention and included in the evaluation sample met minimum inclusion criteria related to their current WASH conditions and implemented programs. Therefore, sample schools are likely to have better hygiene and sanitation infrastructure than the average school, both in the province and nationally.

[63] During the first day of data collection, principals and teachers in two treatment schools discovered that the enumerators were observing handwashing. An investigation discovered that this was because UNICEF was highly associated with WinS and handwashing in Zamboanga del Norte, and because treatment schools remembered they had received the handwashing nudges in October, as part of a UNICEF project. They therefore associated the data collection team's visit in February with the installation team's visit in October, and with handwashing. After protocols changed to eliminate the mention of UNICEF, enumerators did not report that any additional schools discovered the study's purpose.

[64] There was no indication that teachers gave these reminders because of our presence. Teachers were often reminding pupils to wash their hands in the course of organizing classroom activities. Nonetheless, we mark and drop the relevant observations to obtain a more conservative estimate.

DISCUSSION AND RECOMMENDATIONS

Summary of Findings

The results of the evaluation found that the handwashing nudges intervention led to substantial, statistically significant increases in several outcomes of interest. Compared to control schools, in treatment schools handwashing with soap rates were 17.3 pp (p= 0.010**) higher, and handwashing with at least water rates were 25.6 pp (p= 0.001***) higher. Mediation analysis suggests that student behavioral change drove improvements in handwashing rates, which aligns with our Theory of Change model for the intervention. The nudges also increased access to functional facilities near toilets by 7.2 pp (p=0.1*), and to functional facilities with soap near toilets by 20.2 pp (p<0.0001***). As all outcomes were measured four months after implementation, these findings suggest persistence of the nudges' impact.

To our knowledge, this is the first published, fully-powered randomized control trial of a school-based behavioral nudge campaign to improve children's handwashing. In designing this study, we drew from previous work pioneering the use of nudges in schools. In comparison to similar studies, our effect size of 17.3 pp on the handwashing with water and soap rate is moderate. It is higher than the results of a study using arrows pointing to handwashing stations in the US in adult bathrooms[65], but lower than the impact estimates of several studies testing a variety of nudges in other resource-poor school settings[66], though these studies relied on within-subject comparisons. Figure 12 shows how our impact estimates compare to these studies. Table 9 lists the location, context, and specific intervention that the study tested. While the HiFive program is not a handwashing nudges study, we include it in both the graph and the table as a benchmark, given it also targeted pupil handwashing in Philippines public elementary schools, and is a precursor to our current study. As seen in the graph, the nudges intervention's impact estimate is more than two times greater than that of the HiFive program.

Table 9 Summaries of Comparison Studies

Study	Details
Blackwell (2018)[1]	*Location*: United States *Context*: Adult bathrooms *Intervention:* smiley face stickers by sink and footpath with arrows[67]
Grover (2018)[2]	*Location*: Bangladesh *Context:* Schools *Intervention:* paved and painted footpath, infrastructure upgrades[68]

[65] Blackwell C, Goya-Tocchetto D, Sturman Z. Nudges in the restroom: How hand-washing can be impacted by environmental cues. Journal of Behavioral Economics for Policy 2018; 2(2): 41-47

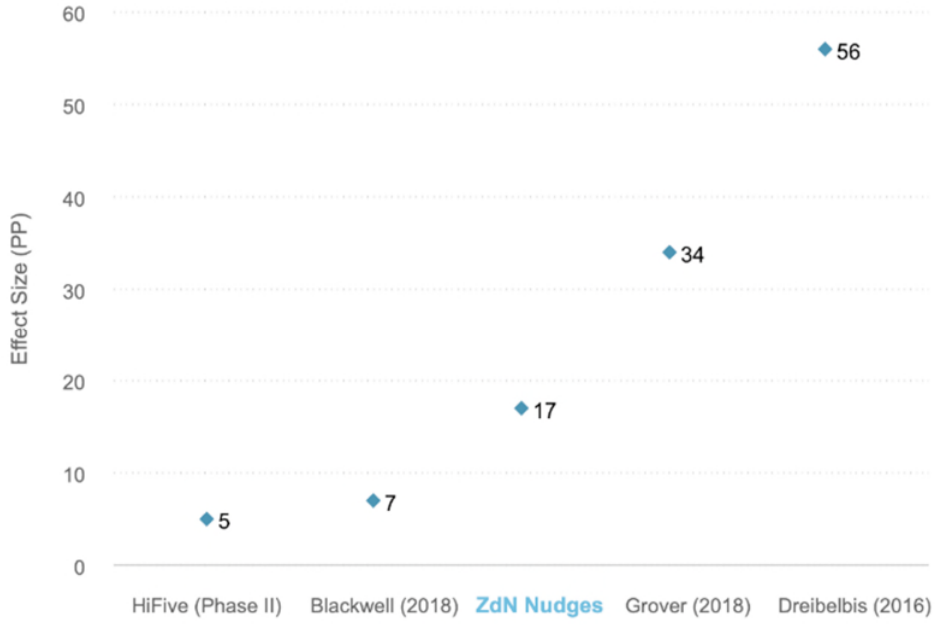
[66] Dreibelbis, R., et al., 2016; Grover, E., et al., 2018. "Comparing the behavioural impact of a nudge-based handwashing intervention to high-intensity hygiene education: a cluster-randomised trial in rural Bangladesh." BMC Public Health 13, pp. 1050; Grover, E., et al. "Comparing the behavioural impact of a nudge-based handwashing intervention to high-intensity hygiene education: a cluster-randomised trial in rural Bangladesh." BMC Public Health 13 (2018): 1050.

[67] The treatment effect presented here is specifically for the *arrows footpath*, as this estimate was statistically significant.

[68] The treatment presented here is for the study arm that implemented footpath and infrastructure upgrades simultaneously.

Dreibelbis (2016)[3]	*Location:* Bangladesh *Context: S*chools *Intervention*: Handprints by sink and footpath with footprints

Figure 12 Comparison of Treatment Effects Across Handwashing Studies



Our process evaluation findings show that in general, all key stakeholders—principals, teachers, and students—enjoyed and supported the nudges. However, several school staff emphasized at the end of their interviews that their school needed more consistent water access, because it was difficult to improve student hygiene practices without it. Some staff explicitly tied this request to the nudges, arguing lack of water access limited the nudges' impact on HWWS. Others simply made this request, likely because they believed the interviews would be relayed to UNICEF, who supports DepEd with various WASH initiatives. Nonetheless, their statements reinforced that pupils' behavior is not the only barrier to handwashing. Staff also commonly did not know or incorrectly stated the purpose of the eye and arrow stickers.

Recommendation for the Philippines

As our cost analysis shows, we project it would cost an estimated PHP3930 (USD 77.51) for one school to install the nudges in five classrooms and one group handwashing station. Given the modest costs and positive impact results, we recommend that UNICEF and DepEd scale up the nudges in eligible schools across the Philippines.

However, there are several considerations that should be considered in scale-up efforts:

- The nudges should be considered as one element of an overall WinS programme, which should also include interventions to improve WASH infrastructure, knowledge, social norms, and other determinants of pupil handwashing behavior.
- The nudges should only be installed in schools that meet the minimum eligibility requirements (Table 10)
- The nudges should be adapted to fit the local context (e.g. translating posters into the local language, or choosing different paint colors that stand out against the classroom floor), but should be scaled up as a package (including all the nudges) without major design changes
- Scale-up efforts should include materials or activities to build knowledge and buy-in of the nudges among principal and teachers
- The nudges should be installed to improve school design during the COVID-19 pandemic, and may be adapted at group handwashing stations to encourage adherence to best practices for preventing disease transmission.

We discuss these considerations in detail below.

The nudges should be considered as one element of an overall WinS programme, which should also include interventions to improve WASH infrastructure, knowledge, social norms, and other determinants of pupil handwashing behavior.
The reason that the nudges should serve as a complement to other WinS interventions is because nudges are not a stand-alone solution. Their success is dependent on availability of handwashing stations and basic awareness of the importance of handwashing after using the toilet. Therefore, they should work alongside interventions that support WASH infrastructure and raising WASH awareness in schools. In our studies, approximately two-thirds of schools in Zamboanga del Norte were deemed ineligible for the nudges (that is, they did not meet minimum WASH infrastructure criteria). These schools require additional infrastructure support to be ready for the nudges. Even in schools that meet minimum eligibility requirements, subgroup analysis suggests that the nudges may be more effective in schools with better WASH infrastructure. While we did not have a knowledge eligibility requirement, we also presumed pupils at school in our sample schools within our sample were aware of proper handwashing behavior, based on our HiFive findings and the presence of UNICEF's WinS-related activities in Zamboanga Del Norte. Our qualitative interviews with students validated this assumption. However, this may not be true for all schools or provinces in the Philippines. Therefore, UNICEF and DepEd should support a comprehensive/multi-faceted WinS program, to ensure they are improving all aspects of WASH conditions in all schools.

The nudges should only be installed in schools and sections that meet the minimum eligibility requirements
We list these minimal eligibility requirements, drawn from the study's eligibility criteria, in Table 10. These requirements must be met because the nudges require enabling WASH conditions such as functional access to a toilet and a handwashing facility, water and soap availability, or a paved path between the toilet and handwashing station. If a school or section lacks these conditions, the nudges cannot perform their function as visual reminders or contextual cues to encourage handwashing with soap.

Table 10 Recommended Installation Criteria for Scale-Up

At the School	At Sections or GHW Stations
o Pupil-to-toilet ratio of under 100 pupils o Water available at least some hours of the school day daily o At least one functional individual or group handwashing station near toilets	o Functional handwashing station (with either running water or stored water regularly available) o Functional toilet stall near handwashing station o Clear, unobstructed, paved path from the toilet stall to the handwashing station o Soap regularly available at handwashing stations[69]

The nudges should be adapted to fit the local context, but should be scaled up as a package (including all the nudges) without major design changes.
We recommend that each school or province that wishes to install the nudges appropriately localize the nudges. We will provide a forthcoming "How-To" for those who wish to do, in order to help guide stakeholders through testing the nudges in their local context, and making appropriate changes such as: translating the language and visualizations in the posters; refining installation protocols; constructing nudges out of different materials based on local availability and pricing of materials. Another possible change would be separating the posters from each other, to reduce the risk that pupils with dirty hands are flipping through the posters in the toilet. In this scenario, teachers would need to change the posters every month (rather than flipping from one poster to the next.)

However, we do not recommend making major changes to the intervention package. This is because the positive findings documented in this report are for the current nudge design package, as implemented in our pilot in Zamboanga del Norte. Making major changes to the design of the nudges or selectively implementing only some of the nudges could change the impact of the intervention in unpredictable ways, potentially reducing its impact.

Scale-up efforts should include materials or activities to build knowledge and buy-in of the nudges among principal and teachers.
As previously stated, our qualitative interviews indicate there may be some gaps in staff understanding of the nudges' purpose and design. This led to some unintended consequences, such as teachers moving soap dishes from beneath the arrow sticker, putting away fallen nudge stickers, or removing posters from the bathroom and placing them inside the classroom. However, such knowledge gaps and their resulting actions appeared to have had minimal negative effect on our pilot. Results show the intervention had positive impact, and most staff still liked the nudges. However, we provided the nudges to participating schools free-of-charge. In a more decentralized plan for scale-up, it will be important for principals and teachers to deeply understand and support the nudges, as they will be responsible for overseeing the planning, budgeting, installation, and maintenance of the nudges. The condition of the nudges after a four-month period suggests schools may need to touch-up installed nudges roughly every six months. A greater understanding of the nudges' purpose may

[69] While this was not part of our eligibility criteria for sections or GHW stations, 79 percent of schools in our final sample had regular soap availability, and 68 percent had MOOE funds dedicated to soap. This suggests that it is important for schools hosting the nudges to have soap provisions.

also stop teachers from taking negative actions such as those stated above, which might reduce intervention impact.

The nudges should be installed to improve school design during the COVID-19 pandemic, and may be adapted at group handwashing stations to encourage adherence to best practices for preventing disease transmission.

Handwashing with soap reduces infectious disease transmission rates, particularly the spread of COVID-19, studies show.[70] We recommend UNICEF and DepEd implement the nudges in schools as part of larger COVID-19 response efforts. The nudges may also be adapted at group handwashing stations to re-enforce social distancing and other best practices. Ideas include: putting eye and arrow stickers on every other faucet to encourage pupils to use faucets spaced adequately far apart; coloring footprints that are six feet apart from each other on the footpath a different color, so students using the footpath know how far to stand apart while waiting to use the handwashing station; putting a handwashing sign at the school entrance and creating an additional footpath from the school gates to the closest group handwashing station, to remind pupils to wash their hands when first entering the school; and appending messaging related to COVID-19 to the current posters, while maintaining their appeal to emotional motivators like disgust. Though the impact of the nudges on adherence to social distancing and other best practices have not been evaluated, such changes to nudge design would be relatively low-risk.

Recommendations for Other Contexts

Our study provides additional support to a growing number of studies[71] that suggest that nudges can be effective in a variety of low-middle income country contexts. However, our design and evaluation experience in Zamboanga del Norte suggest for nudges to be successful in a new context, it is important that:

1) *The nudges are context-appropriate*—Nudges that were considered during the nudge design phase were determined not to be appropriate in Zamboanga del Norte.[72]
2) *They are installed in a setting with enabling conditions*—Because nudges are designed to primarily target behavioral change, they are likely to be less effective in locations with minimal WASH infrastructure and awareness.
3) *Determine a plan for ensuring nudge sustainability*—If the nudges do not last, they are less likely to have long-lasting effects. School staff and perhaps the wider school community must continue to monitor and restore the nudges (e.g. repaint path and footprints; replace torn/lost eye and arrow stickers; etc.) as necessary.

[70] Pogrebna, Ganna & Kharlamov, Alexander. (2020). The Impact of Cross-Cultural Differences in Handwashing Patterns on the COVID-19 Outbreak Magnitude. 10.13140/RG.2.2.23764.96649.

[71] These studies include Grover (2018) and Dreibelbis (2016). See Table 9 for details on the studies.

[72] We considered but ultimately did not implement various handwashing nudges drawn from the literature. For example, we considered placing mirrors by the handwashing station. A handwashing behavioral-change resource mentioned mirrors as a potential way to attract children to the handwashing station, as the mirror would link handwashing to the intrinsically attractive behavior of checking their reflections in the mirror (Neal, D., Vujcic, J., Hernandez, O., & Wood, W. (2015). *The science of habit: creating disruptive and sticky behavior change in handwashing behavior.* Washington DC, USA. USAID/WASHplus Project. Neal, David, Jelena Vujcic, Orlando Hern). However, during field scoping visits in the nudge design phase, we found out that many classrooms already had mirrors near their handwashing station in Zamboanga del Norte. We thus did not incorporate mirrors as a nudge into our intervention.

Further Research Agenda

There are some potential areas for additional research. First, while these handwashing nudges have been replicated in several contexts, it may still be useful to replicate them in other geographies. New replication studies could add to the literature in several ways.

First, if structured as a multi-armed RCT and/or with a large enough sample size, new studies could explore *which* nudges primarily drive the positive impact on handwashing rates. While we have our hypotheses about which nudge is most effective, our study was not intended to evaluate each individual nudge, only the whole package intervention.

Second, they may help generate additional nudge ideas. We know from our design process that nudges effective in previous studies were not appropriate for ours, and imagine the opposite may be true in new contexts. Some of these nudges might address more specific barriers to handwashing. Our nudges were designed to bring students from the toilet stall door to the handwashing station. However, results indicate there was a small percentage of students whom the nudges led to the handwashing station, but who did not wash their hands with soap even though soap was available. Perhaps replication studies might choose to specifically convert such students, who are currently washing their hands with water, to wash their hands with water *and* soap.

Third, future studies may wish to measure persistence effects of handwashing nudges. To do so, these studies may measure handwashing multiple times by utilizing alternative data collection approaches such real-time sensor data or liquid soap volume (though these are more expensive than real-time handwashing observations).