Impact of a multi-level intervention, Sundara Grama, on latrine use and safe disposal of child faeces in rural Odisha, India

Bethany A. Caruso, Department of Environmental Health, Emory University

Gloria D. Sclar, Department of Environmental Health, Emory University

Parimita Routray, Independent Consultant

Corey Nagel, College of Nursing, University of Arkansas

Fiona Majori, London School of Hygiene and Tropical Medicine

Steven Sola, Department of Environmental Health and Department of Epidemiology, Emory University,

Will Koehne, Department of Epidemiology, Emory University

Renee DeShay, Department of Behavioral Sciences and Health Education, Emory University

Shivika Udaipuria, Department of Epidemiology, Emory University

Rebekah Williams, Department of Behavioral Sciences and Health Education

Thomas Clasen, Department of Environmental Health, Emory University

Grantee Final Report

Accepted by 3ie: August 2019

Note to readers

This impact evaluation has been submitted in partial fulfilment of the requirements of grant TW14.1006 issued under Promoting Latrine Use in Rural India Evidence Programme. This version is being published online as it was received. A copy-edited and formatted version will be available in the 3ie Impact Evaluation Report Series in the near future.

All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Bethany A. Caruso at bcaruso@emory.edu.

The 3ie technical quality assurance team comprises Neeta Goel, Radhika Menon, Sayak Khatua, Shaon Lahiri, Anmol Narain, Rosaine Yegbemey, Marie Gaarder, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder.

Funding for this impact evaluation was provided by the Bill & Melinda Gates Foundation. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: Caruso, BA, Sclar, GD, Routray, P, Majorin, F, Sola, S, Koehne, W, DeShay, R, Udaipuria, S, Williams, R, Nagel, C and Clasen, T, 2019. Impacts of a multi-level intervention, Sundara Grama, on latrine use and safe disposal of child faeces in rural Odisha, India, 3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation (3ie)

Acknowledgements

We are especially grateful to the participants in all of the communities in Puri, Odisha with whom we have worked; we are ver cognizant of the tremendous amount of time each participant provided to enable this research to be completed.

This work would not have been possible without the care, diligence, thoughtfulness, critical thinking, and wisdom of our dedicated research team: Puspanjali Barik, Rajani Barik, Subhashree Chatterjee, Anjali Dakshinaray, Snahaspada Das, Sujata Das, Munmun Dasmohapatra, Archana Harichandan, Sadasiva Kothia, Prasant Kumar Mahapatra, Alfred Mohanty, Ramesh Mohanty, Parbati Nayak, Rajashree Nayak, Satyajit Parija, Subhrakanta Pattanayak, Bagdatta Ray, Sunita Ray, Manaswini Rout, Sandhyarani Rout, Sonam Sahoo, Indrajit Samal, Aparna Singh, Swarnalata Swain, and Santosh Kumar Tripathy. We also thank Belen Torondel for her feedback on the original proposal and during the formative research phase.

Many thanks to our partner, Rural Welfare Institute (RWI), notably Mr. Prabhakar Nanda, for implementing the intervention and learning with our team as we made intervention adaptations.

We extend our thanks to the International Initiative for Impact Evaluation (3ie), our funder, and the Research Institute of Compassionate Economics (r.i.c.e.), for valuable insights and feedback throughout the research. We are particularly appreciative of the thoughtful support of Neeta Goel, Radhika Menon, and Shaon Lahiri (3ie) and Sangita Vyas (r.i.c.e.).

Summary

Globally, an estimated 892 million people, approximately 12% of the global population, defecates in the open. Ending open defecation by 2030 is the aim of Sustainable Development Goal target 6.2, importantly shifting focus to sanitation behavior from just sanitation access, as was the focus in the Millennium Development Goal era. In India, where an estimated 60% of those practicing open defecation reside, the government has also shifted focus to prioritize ending open defecation over increasing coverage alone. Research has demonstrated that access to sanitation enables but does not guarantee use. Therefore, there remains a need to understand barriers to latrine use among household members that own latrines and to create and evaluate interventions to address those barriers to increase use.

Emory University undertook formative research to understand specific barriers of latrine use in rural Odisha, India—including the validation of previously identified barriers to use—and to use findings to design a theoretically-informed intervention to increase latrine use and safe disposal of child feces. The resultant multi-level intervention, called *Sundara Grama,* included community-level activities that were designed to reach latrine owners and not-latrine owners alike.

Community-level activities included:

- A music and humor filled **Palla**, which is a folk dance performance common in Odisha, that communicated messages about latrine use, health, child feces disposal, and the importance of overall village cleanliness;
- A **Transect Walk** that toured the village and marked piles of feces with colored powder along the way;
- A **Community Meeting** to discuss the village state and create a plan for assuring its cleanliness;
- The **recognition of latrine-using households**, specifically those whose members all use the latrine all the time, with a banner hung in front of their house to indicate their latrine use behavior has been confirmed and recognized by community members at the community meeting;
- **A Village Map Painting** of all households, with special recognition of those using the latrines at all times and a description of the community action plan decided in the meeting.

Household-level activities included:

- A **targeted visit for latrine owners** specifically, reiterated messages from the other activities and elicited commitment from the household members to use the latrine to keep the village clean and beautiful.
- **Latrine repairs** were carried out to provide minor repairs to those latrines that were not functional and to doors to all latrines that did not have one or had one that was broken.

Finally, a **Mother's Group Meeting** was created for mothers and caregivers of children under age five, regardless of their household latrine status to provide action knowledge and hardware to enable the save disposal of child feces.

We conducted a cluster-randomized trial to determine if the intervention had an impact on latrine use and safe child feces disposal behavior. We engaged 72 villages to evaluate the intervention, with 66 involved in the trial (33 intervention, 33 control) and 6 engaged in simultaneous qualitative inquiry (3 intervention, 3 control). In the trial villages, a census was conducted from February 2017 to April 2018 to identify the latrine status of all households and determine latrine use behaviors of all members in latrine owning households, including disposal of child feces. The Palla, Transect Walk, Community Meetings, Household Recognitions and Visits, and Mother's Group activities were carried out June-July 2018, Wall Paintings were done in September 2018, and Latrine Repairs were completed in July-November 2018. Endline data collection took place from November 2018-February 2019, again targeting all households in the engaged villages to determine latrine ownership and querying use among appropriate households.

Intervention delivery was observed in all villages to understand if each of the activities were carried out as designed and reached target participants. Qualitative research was conducted in a subset of non-trial villages and after endline to elicit perceptions of the interventions. Questions about the intervention were asked during endline data collection to assess awareness of and participation in activities.

Latrine use increased in both intervention and control communities. We found an increase in reported latrine use among individuals age 5 and older of 6.4% (95% CI 2.0%-10.7%, p=0.004) in the intervention group at endline after accounting for the increase in latrine use observed in the control group and an increase in reported safe child feces disposal of 20.4% (95% CI 11.7%-29.2%, p<.001) in the intervention group at endline after accounting for the increase in safe disposal of child feces observed in the control group. No difference was observed between intervention and control groups in the proportion of households that did not have a latrine at baseline and the proportion of households that had one at endline.

Overall, the intervention activities were well received, particularly the Palla, which participants found entertaining and funny. With the exception of the Mother's Group, which reached an estimated 96% of latrine owning households with children under age 5, reach of all other activities could have been improved. Poor recruitment, specifically of community members in hamlets or other parts of the village, may be a cause. Women indicated barriers to attending activities, particularly the Palla and Community Meeting, where men were in attendance, as well as the Mother's Group if their families restricted them from leaving the house. Activities were delivered with fair to good fidelity overall, suggesting that improvements could result in greater increases in behavior change.

Our result demonstrates that theory-informed interventions designed to change behavior can be impactful. Latrine use behavior is changing in the research area overall, but increased 6.3% more in the intervention area. Importantly, our intervention also increased reported safe child feces disposal by over 20%. Safe faces disposal practices were largely not practiced in our research area before our intervention, primarily because the importance of safe disposal was not understood. Additional investment in refining this and like interventions are warranted to bring these efforts to scale, particularly as safe child feces disposal has yet to be an investment

and communication priority in government campaigns to date and costs needed for safe management of child feces disposal programs, like ours, do not need to be extensive to enable change.

Moving forward, policy makers should leverage this and like programs to not only continue to influence behavior change, but to sustain changes already made. Increased investment to develop and evaluate evidence-based interventions specifically targeting behaviors is warranted. In turn, researchers need to engage target populations, apply theory to intervention design, and conduct rigorous process evaluations to inform future adaptation and scale-up.

Contents

Figures and tables

Appendix Figures and Tables

Abbreviations and Acronyms

3ie	International Initiative for Impact Evaluation
BCD	Behavior Centered Design
CM	Community Meeting
DID	Difference in Difference
GoI	Government of India
FGD	Focus Group Discussion
HH	Household
HV	Household Visit
IDI	In Depth Interview
MG	Mother's Group
OD	Open defecation
ODF	Open Defecation Free
PE	Process Evaluation
RA	Research Assistant
RWI	Rural Welfare Institute
SBM	Swachh Bharat Mission
TW	Transect Walk
WASH	Water, Sanitation, Hygiene

1. Introduction

Globally, an estimated 892 million people, approximately 12% of the global population, defecates in the open. Ending open defecation by 2030 is the aim of Sustainable Development Goal target 6.2, importantly shifting focus to sanitation behavior from just sanitation access, as was the focus in the Millennium Development Goal era. Research has demonstrated that access to sanitation enables but does not guarantee use. In an ad hoc analysis as part of a systematic review, Garn et al (2017) found that each 10% increase in coverage led to a 5.8% increase in use, revealing that open defecation is still practiced by a considerable proportion of those who own latrines. Therefore, concerted effort is needed to increase latrine use among latrine owners if the benefits of sanitation, including impacts on diarrhea, active trachoma, schistosomiasis, height-for-age, and hookworm, *A. lumbricoides, S. stercoralis,* intestinal protozoa infections, and well-being are to be realized (Freeman et al., 2016, Freeman et al., 2017, Sclar et al., Sclar et al., 2018).

In India, where an estimated 60% of those practicing open defecation reside, the government has also shifted focus to prioritize ending open defecation over increasing coverage alone. Since the 1980s, the Government of India (GoI) has implemented a series of missions and campaigns, focusing particular attention to rural parts of the country. In 1986, the government of India launched the Central Rural Sanitation Program (CRSP), the first large-scale country wide sanitation program in India, through which rural households below the poverty line (BPL) were provided subsidies for building toilets. In 1999, CRSP became the Total Sanitation Campaign (TSC), with new emphasis placed on community mobilization and information, education, and communication (IEC) activities. Subsidies (now called 'incentives') were still provided exclusively to BPL households and changed in amounts over the course of the 12-year campaign. The Nirmal Gram Puraskar (NGP) was set up in 2003, financially rewarding Gram Panchayats that attained 100% sanitation coverage. In 2012, the TSC became the Nirmal Bharat Abhiyam (NBA) and extended the financial incentives, which were increased, to non-BPL households that fit certain criteria. In 2014 the Swachh Bharat Mission (SBM) was launched in both urban (SBM) and rural areas (SBM-Gramin) with the goal of ending open defecation by Gandhi's 150th Birthday on October 20th, 2019. The SBM also includes IEC activities and subsidies. As a result of these successive campaigns, many villages have experienced multiple government campaigns and with different subsidy amounts.

The Government of India has been publically tracking latrine construction efforts and the declaration of open defecation free (ODF) villages, districts, and states on the SBM Dashboard (Ministry of Drinking Water and Sanitation) which reports that 92,541,952 latrines have been built and 556,441 villages have been declared ODF since the start of the campaign[a]. The Research Institute for Compassionate Economics (r.i.c.e.) found a reduction in open defecation, from 70% in 2014 to between 40-50% in 2018 in the northern states of Bihar, Madhya Pradesh, Rajasthan, and Uttar Pradesh among rural households surveyed at the start of SBM in 2014 and again in 2018.(Gupta et al., 2019) They report that the proportion of latrine owners practicing open defecation in 2014 and 2018 remained the same, suggesting that the reduction in open defecation was driven by latrine coverage increases alone. Further, the data also suggest that not all coverage increases amounted to use and that those who previously owned latrines and did not use them continue to defecate in the open. Therefore, there remains a need to understand barriers to latrine use among household members that own latrines and to create and evaluate interventions to address those barriers to increase use.

[a] SBM Dashboard (https://sbm.gov.in/sbmdashboard/) accessed on 8 April 2019 at 3:07 pm EST. Numbers are accurate at that date and time.

Several studies, by our team and others, have identified barriers to latrine adoption in rural India. Perhaps chief among these, especially in Hindu populations, are deep cultural attitudes of purity and pollution, where OD is viewed as more healthy and pure than the use of household latrines (Coffey et al., 2015, Routray et al., 2015). Other barriers revolve around deficiencies in the functionality or acceptability of Government of India subsidized latrines, which often relate to issues of privacy and ability to practice cleansing rituals (poor design, incomplete construction, low wall height, no roof or door, lack of water, susceptibility to clogging, lack of light, bad odor) (Coffey et al., 2015, Routray et al., 2015, Banda et al., 2007, Barnard et al., 2013). Some of the barriers involve practical issues (burden to clean, inconvenient location for men working in fields) or habits (time to socialize, exercise, accustomed to OD) that favor OD over use of household latrines (Coffey et al., 2015, Routray et al., 2015, Banda et al., 2007, Barnard et al., 2013). Tied to purity is a fear of pit latrines filling up too quickly, requiring a household member to manually empty the pit - an act of ritual pollution (Coffey et al., 2015, Torondel, 2015). Uncertainty about how quickly pits fill and how to safely empty and dispose of pit contents is another barrier to use.

Additionally, latrines are often perceived to be for women. Women are often the primary users of latrines in a household (Coffey et al., 2015, Routray et al., 2015, Barnard et al., 2013, Clasen et al., 2014). Privacy, modesty, and safety of daughter-in-laws are key motivating factors for men to build a household latrine (Coffey et al., 2015, Routray et al., 2015, O'Reilly and Louis, 2014). Women also have their own motivations for latrine use. Women report having greater privacy, handling their menstruation more comfortably, saving time, and not needing to wait for a family member to accompany them (Routray et al., 2015, Caruso et al., 2017, Hirve et al., 2015, Hulland et al., 2015, Sahoo et al., 2015). Conversely, latrines can reinforce women's isolation in the household, limit their mobility outside the home, and are not always perceived to be a better option than OD (Coffey et al., 2015, O'Reilly and Louis, 2014, Caruso et al., 2017).

From January to April 2017, Emory University undertook formative research to understand specific barriers of latrine use in rural Odisha, India—including the validation of previously identified barriers to use—and to use findings to design a theoretically-informed, multi-level intervention to increase latrine use and safe disposal of child feces. At present, according to the SBM dashboard, only 84% of Odisha state has access to a household latrine, the second lowest latrine coverage in the country.[b]

This report presents results of a cluster randomized trial conducted to evaluate the intervention designed. In the following sections, we describe the intervention, including the proposed theory of change, and our strategy for evaluating intervention delivery; the study design, including sample size, sample selection, tools, data collection procedures, and ethics; and research findings, including process and impact evaluation and cost analysis. We then present a discussion of findings and conclusions.

[b] Ibid.

2. Intervention

2.1. Description

2.1.1. Setting

Formative research, intervention design, intervention piloting, and the evaluation were carried out in rural Puri District, Odisha State, India.

At the time of intervention implementation, Odisha state had the lowest toilet coverage of any state (56.29%), just below the state of Bihar (57.35%) and the fourth lowest number of open defecation free villages (21.87%) (Ministry of Drinking Water and Sanitation). The sanitation context of rural Puri district is better than the overall context of the state, with 36.8% of households using an improved sanitation facility in rural Puri and 23.0% overall in rural Odisha (International Institute for Population Sciences (IIPS) and ICF, 2017a).

According to the National Family Health Survey conducted in 2015-2016, the majority of people in Odisha live in rural areas (83%) and more than 46% of the population are recognized by the government as living below the poverty line (BPL) (International Institute for Population Sciences (IIPS) and ICF, 2017b). The predominant religion is Hindu (95%), and 20% of households belong to scheduled castes and 23% to scheduled tribes (International Institute for Population Sciences (IIPS) and ICF, 2017b).

In Odisha state, 84% of rural households have electricity, 88% have an improved drinking water source, and 23% have an improved sanitation facility, and 11% use clean fuel for cooking (International Institute for Population Sciences (IIPS) and ICF, 2017b). Ten per cent of rural children in Odisha under age 5 had diarrhea in the 2 weeks preceding the 2015-2016 NFHS-4 survey, and among those 69% received oral rehydration salts, and 69% were taken to a health facility (International Institute for Population Sciences (IIPS) and ICF, 2017b). Thirty-five per cent of rural children under age five are stunted. Among rural women, 65% are literate and 65% over age six ever attended school (International Institute for Population Sciences (IIPS) and ICF, 2017b).

The water, sanitation, electricity, and cooking context of rural Puri district is better than the overall rural context of Odisha state: 94% of rural households have electricity, 94% have an improved drinking water source, 36.8% of households have an improved sanitation facility, and 14% use clean fuel for cooking (International Institute for Population Sciences (IIPS) and ICF, 2017a). Seven per cent of rural children under age 5 had diarrhea in the 2 weeks preceding the 2015-2016 NFHS-4 survey, and 17% per cent of rural children under age five are stunted. Among rural women in Puri, 83% are literate and 79% over age six ever attended school.

2.1.2. Intervention Design

Formative research was carried out from January-April 2017 to inform the design and piloting of a theoretically-informed intervention that aimed to increase latrine use (including the safe disposal of child feces) among latrine-owning households at an average cost of 20USD per household, a policy-relevant stipulation required by 3ie. As part of this formative research, we carried out an intervention design process that was informed by the Behavior Centered Design (BCD) steps (Aunger and Curtis, 2016), the Intervention Mapping approach (Bartholomew et al., 2011), causal analysis using problem trees (Starr and Fornoff, 2016), and theory of change creation (Starr and Fornoff, 2016, Aunger and Curtis, 2016, De Silva et al., 2014). We identified behavioral factors and subfactors that determine latrine use and safe disposal of feces using an expanded version of the BCD Checklist from (Aunger and Curtis, 2016) that included relevant and distinct

components of both the COM-B model of behavior (Michie et al., 2011) and the RANAS model (Mosler, 2012).

Through this process we identified the following behavioral barriers as responsible for non-use of latrines for defecation or disposal: i) non-functional latrines, ii) lack of practical knowledge regarding latrine use, iii) preference for open defecation, iv) latrine use not prioritized/valued, v) unsuitable latrine design, and vi) inaccessible water. Latrine design and water access were deemed beyond our ability to address given the limitations of funding and time, thus we focused on the remaining four barriers. We do acknowledge that some households or household members will not be able to use their latrines because of latrine design and water issues, despite potentially being motivated to use their latrines by our intervention activities. For some, these barriers may be too great to overcome. Water is needed to flush the type of latrines available in Puri and while Puri is not water scarce, ability to access water can be variable by village (variable water sources), household (variable proximity to sources), and individual (variable physical ability to get, carry, and use water). Latrine design may also be quite variable. Based on our formative work, we know the common latrine design is challenging for people with disabilities and older village residents who cannot easily get inside the latrine and squat.

In regards to safe disposal of child feces specifically, we determined the primary barrier to be a lack of awareness and action knowledge regarding safe practice for disposal of feces. (See Section 2.3 for a discussion of the theory of change).

2.1.3. Description of Intervention Package

Based on the barriers identified in the formative research phase, we developed a multi-level intervention to address the primary barriers to latrine use (non-functional latrines, lack of practical knowledge regarding latrine use, preference for open defecation, latrine use not prioritized/valued) and safe disposal (lack of action knowledge). The overall motto of the intervention was 'moro swacha, sustha, sundara grama' or 'my clean, healthy, beautiful village'. Recognizing status as an important behavioral driver during the formative work, we intentionally focused on the reputation of the village, among other villages, as a driver. This motto was to be repeated across all activities, along with the name of the intervention, *Sundara Grama,* or 'beautiful village.'

The intervention activities were to be delivered at multiple levels, including the village-level, the household-level and at a 'sub-group' level for all mothers/caregivers of children under age five. Each intervention activity is **briefly** described below by level of delivery as designed and intended to be delivered. A more detailed description of each intervention activity can be found in the *Sundara Grama Intervention Manual*.

Village-level activities:

- *Pre-intervention community visits:* Community mobilizers from Rural Welfare Institute (RWI) were to make preliminary visits to each village to build rapport with key village stakeholders, foster support for the intervention, plan intervention logistics (e.g. location and date for activities), and learn about the social dynamics of each village before any intervention activities were to take place.

- *Palla performance:* A 'palla,' a traditional folk art performance, was to be the first activity to take place in each village. Pallas were performed by local troops hired, trained and managed by RWI. Songs and skits aimed to engage village members around the health and non-health benefits (i.e. comfort, privacy) of latrine use, as well as increase action

knowledge around the practices of latrine use, pit emptying, and safe disposal of child feces.

- *Colored powder transect walk:* After the palla performance, ideally within the week, community mobilizers were to conduct a surprise transect walk in each village, intended to enable village members to recognize the amount of fecal contamination in their village due to open defecation, and therefore re-evaluate the condition of their village environment and likely generate a sense of disgust and possibly shame. By starting first thing in the morning, the walk was designed to be carried out when all or most household members were still in the household to enable participation by all. The use of bells and other noise makers were to be employed to draw attention to the activity and summon participants. All walks were to weave through the village streets and eventually end up at known defecation sites (which were to be previously identified by community mobilizers during the pre-intervention community visits). At the first sight of human feces, community mobilizers were to use brightly colored powder, traditionally used for the Hindu spring festival known as 'Holi,' to mark piles of feces. Community mobilizers were then to distribute bags of powder to participants and encourage them to mark all piles of feces seen during the walk. After the feces marking activity, the walk ended at a water source to have a reflection discussion among participants and have a handwashing demonstration, for which participants were provided soap.

- *Community meetings:* Community meetings, one for women and one for men, were to be facilitated by community mobilizers and were meant to help participants decide upon a set of action steps to achieve the goal of a "swacha, sustha, sundara grama" or 'clean, healthy, beautiful village'. Community members were to be encouraged to identify actions they felt could achieve this goal (whether cleaning the village pond or roads, or other actions that emerged). If not brought up as part of the action planning process, community mobilizers were to suggest including latrine use by all members of the household at all times. During the meeting, those in attendance were also to be asked to identify 'positive deviant' households where all the members always used their latrine for defecation at all times. This information would later be used to formally recognize these 'positive deviant' households and make them known to other community members for their behavior.

- *Positive deviant household recognition:* Households identified as positive deviants by participants in the community meetings were to be provided with a banner to display (at their discretion) in front of their house to publically recognize that all members of their household use the latrine all the time—as confirmed by members of the village—thus praising their contribution to achieving a 'clean, healthy, beautiful village.'

- *Village wall painting:* As a final activity, community mobilizers were to hand draw maps of each village, depicting major features in the village, like roads and temples, as well as all households. This draft map was to be copied by local artisans in the form of a mural painted in a location agreed upon by participants in the community meeting. The mural was to distinctly identify the positive deviant households to serve as a reminder to all village households and to motivate all households to have all members use their latrines all the time. To the side of the map, the wall painting also displayed the action steps decided upon in the community meeting to remind the community of the decided upon actions to achieve the goal of achieving a 'clean, healthy, beautiful village.' The mural's display of the action steps also enabled members not present at the meeting to be aware of what was discussed and decided upon.

Group level activity:

- *Mothers meetings:* Community mobilizers were to hold 'mothers meetings', which were open to all mothers, as well as other caregivers, of children under age five, with the aim of providing participants with information about health risks of unsafe child feces management, and the necessary action knowledge and hardware (i.e. plastic scoops and potties) to enable the practice of safe child feces disposal. Participants were to be provided information on how to use the hardware provided, including how to properly dispose of feces and clean and store the hardware. As meetings were open to all mothers/caregivers of children under age 5 in the village, regardless of latrine ownership, the community mobilizers were to emphasize that those who owned a latrine should dispose of feces and any water used to wash hardware or cloths into the latrine, while those without latrines should bury feces and contaminated washing water.

Household-level activities:

- *Household visits:* A community mobilizer was to make individualized visits to all latrine-owning households to reflect with members on the intervention activities to date and reiterate key messages, with the understanding that messages would be repetitive for some household members and new to other members who had not been able to attend or hear about any of the intervention activities previously carried out. Visits were not forced; household members could refuse the visit or aspects of the visit. Community mobilizers also were to lead household members in a voluntary pledge to working towards the village goal of achieving a 'clean, healthy, beautiful village', which was to include latrine use by all members of the household at all times. At the end of the visit, each household was to be given a poster, which they could refuse, with key messages about latrine use to serve as a reminder of their commitment.

- *Latrine assessment and repairs:* Acknowledging that a key barrier to latrine use is lacking access to a *functional* latrine, India-based Emory team members were to assess latrines with representatives from local contracting partners who repair latrines to jointly identify repairs needed. Latrine-owning households that were in need of minor repairs and that were deemed eligible for assistance were then to have their latrines fixed.

2.1.4. Key Partners Involved in the Delivery of the Intervention

Rural Welfare Institute (RWI), a grassroots NGO based in Nimapara District and led by director Mr. Prabhakar Nanda, was engaged as the implementing partner for delivery of the *Sundara Grama* intervention. The RWI implementing team consisted of 4 supervisors (1 woman, 3 men) and 16 community mobilizers (12 women, 4 men). The majority of community mobilizers were in their 20s and had completed +3 schooling (bachelor's degree). Some had previous work experience in social services and community development while for others this was their first official job position. The RWI mobilizer team was trained by Emory team members on all of the intervention activities over the course of 12 training days, which included both in-house and field practice. The RWI team was responsible for making initial visits to the community (to build rapport with community stakeholders and organize activity logistics) and leading the following activities: transect walk, community meetings, positive deviant identification and recognition, mother's group meetings, household visits, and mapping households for the wall paintings. The RWI team split into 4 sub-teams comprised of 1 supervisor and 4 mobilizers. Each sub-team implemented all of the intervention activities across the 8 to 10 villages to which they were assigned.

Additional partners were engaged to complete the palla performances, wall paintings, and latrine repairs. Palla performances were conducted by two palla groups with 5 to 6 members each. Community wall paintings were completed by two local artisan groups with 4 to 5 members each.

Latrine repairs were completed by two local contracting groups, Gopabandhu Seva Parisad (GSP) (assigned 20 villages in Pipli and Delang blocks) and Jageswari Jubak Sangha (JJS) (assigned 16 villages in Pipli and Nimapara blocks). An assessor from each group was paid to go to each selected household and complete a full assessment of the household's latrine. About 815 households were selected for latrine assessment across the 36 intervention villages. The two assessors completed their work over the course of 3 months. The two contracting groups completed latrine repairs in 457 households across the 36 intervention villages (33 trial and 3 qualitative). Households that planned to destroy their latrine, required an entirely new superstructure or pit, or used their latrine as storage and did not remove the storage items during the assessment were not eligible for repairs. The types of repairs included about 233 slab repairs, 153 l-pipe repairs, 43 pan repairs, 363 door repairs, 130 pit repairs, and 117 floor repairs. The average cost of latrine repairs per household, including material and labor, was 922.47Rs.

Control villages were not provided with intervention activities after endline.

2.1.5. Changes to the Intervention During the Course of the Study

Minor changes were made to intervention activities based on piloting in non-trial villages (e.g. slight revisions to activity scripts/facilitator guides) prior to delivering activities to intervention communities engaged in the trial. A minor change that could have resulted in a negative impact if not addressed, was the color of the powder used during the transect walk activity. In one pilot village, red colored powder was used and participants informed the RWI mobilizers it was not an appropriate color because of its connection to different religious practices. RWI and Emory team members discussed this lesson learned and decided both the colors red and orange (also noted as a religious/sacred color) would not be used during the transect walk. To ensure a non-religious color was used, Emory team added a question to the transect walk process evaluation survey about the color of the powder.

During intervention delivery, the household visit activity was revised due to resource constraints. Specifically, the household visit activity was originally designed to be a 45-minute visit that included 4 key activities: (1) reflection on village goal and action steps, (2) demonstration of how feces spread (using glass of water, mustard paste to represent feces, and piece of thread to represent a fly's leg), (3) discussion of individuals' latrine use with barrier planning, and (4) household commitment and distribution of a reminder poster. These activities incorporated a variety of behavior change techniques. However, RWI staff could only be employed for a set amount of time in order to meet the 20 USD per household cost limit to the intervention. Since the other *Sundara Grama* activities took longer than expected to complete, there was not enough time for RWI staff to complete a 45-minutes visit for every household with a latrine. In order to resolve this issue, we cut out activities #2 and #3 so that the visit only took about 10 minutes to complete and only included a quick reflection and then the household commitment with reminder poster. Unfortunately, this means some important behavior change techniques, such as barrier planning, had to be cut from the visit. However, with this modification, the household visit activity was able to be completed by RWI staff across all intervention villages. This revision to the household visit was made *before* the household visit activity commenced and RWI staff received a 'refresher training' on the activity to ensure they were properly trained on the changes. As such, *all intervention households* should have received the same version of the household visit activity.

Any deviations in delivery of the intervention from the original implementation plan will be reported in section *4.1 Intervention Implementation Delivery.*

2.1.6. Program Participants
The intervention targeted all members of households that owned at least one latrine across all engaged villages.

Six additional non-trial villages, three of which received the intervention, were engaged in qualitative activities. These villages were in the same blocks within Puri as the other villages, and are likely to be similar demographically, though they were not censused to enable reporting of specific characteristics. From a pre-enrollment village mapping phase, we know these 6 non-trial villages ranged in size from 90 to 174 households (mean: 130 households) and had a latrine coverage of 49.43% to 79.25% (mean: 60.33% latrine coverage).

2.2. Theory of Change

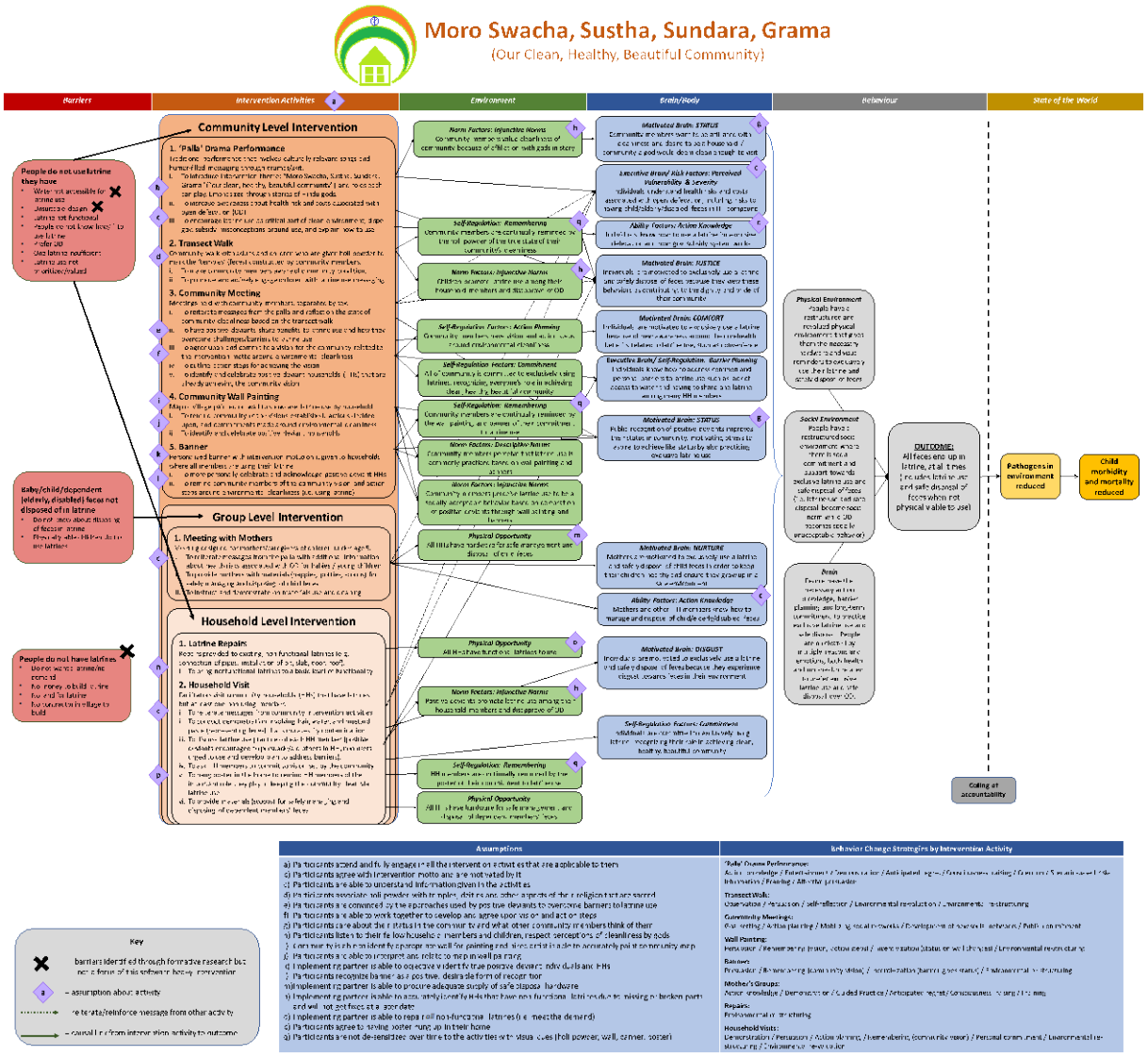
The theory of change (TOC) outlines all identified barriers, proposed intervention activities, environmental and psychosocial determinants targeted to achieve consistent latrine use, and lists of assumptions and behavior change strategies employed (See Figure 1). Overall, we surmised that latrine use would increase by improving the ***physical environment***, specifically by increasing access to *functional* latrines, by improving the ***social environment*** (norms) of latrine use and encouraging rejection of OD, and by targeting key ***'brain' or psycho-social determinants*** like personal-level motivators (namely status, justice, comfort, disgust, nurture), ability to practice latrine use and safe feces disposal, risk perceptions associated with OD, and 'self-regulation' or ability to sustain the behaviors.

The components in the theory of change are explained below ***by barrier***, providing details of how the intervention activities addressed the identified barriers, and discussing the behavioral techniques utilized. A final paragraph also describes the key ***motivators*** for latrines use identified through the formative research.

2.2.1. Behavioral barriers and Intervention Activities
To inform out intervention design process and the selection of intervention activities during the formative research phase (January- April 2017), we held a meeting in a rural village with community members and a partner organization Bhabagrahi Kala Niketan (BKN). The purpose of this meeting was to work with community members to build a problem tree to visually identify all possible barriers to latrine use and safe disposal of feces, and use the problem tree to identify possible solutions to address barriers (solution tree) (Snowdon et al., 2008).

The problem and solutions trees helped us to identify the behavioral factors that we needed to target through our intervention activities. Once the behavioral factors were identified, we used the intervention mapping approach (Bartholomew et al., 2011) to identify various behavior change techniques that could influence the specific behavioral factors identified (Michie et al., 2011, Bartholomew et al., 2011, Mosler, 2012). For example, we learned that mothers and caregivers of young children did not know about safe feces disposal practices or how to perform these behaviors. The behavioral factors we needed to target for these barriers included Psychological Capability and Ability (Michie et al., 2011, Mosler, 2012).

Figure 1: Theory of Change



Behavior change techniques recommended for targeting these factors included providing action knowledge (Mosler, 2012) and modelling, or providing an example for people to imitate (Bartholomew et al., 2011). We decided the action knowledge should be communicated to all community members, so decided to include it in a community-wide event, but that modelling would be best in a more intimate setting with mothers/caregivers to enable discussion and even practice. We therefore created activities around these behavioral strategies.

Below, the barriers to latrine use and safe disposal of feces are note along with further information about the behavioral factors and strategies utilized in the activities.

Non-functional latrine: Household latrine repairs create a physical environment that enables latrine use.

Lack of practical knowledge regarding latrine use: Palla performances use demonstration and action knowledge strategies to teach community members how to use a latrine.

Preference for OD: Palla performances use various strategies, like scenario-based risk-information communication, so individuals understand the health risks and costs associated with OD and unsafe disposal. Pallas also counter commonly cited benefits of OD with benefits of latrine use as a form of affective persuasion and a way to influence the cost-benefit perception around latrine use. Transect walks use environmental re-valuation and self-reflection to generate disapproval of OD and awareness that one's environment is dirty and requires change.

Latrine use not prioritized/valued: Palla performances, community meetings, wall paintings, and banners use a mix of affective persuasion, mobilizing of social networks, incentivisation, and environmental re-structuring to establish a status motive and injunctive norm around latrine use. The community meeting in particular utilizes goal setting and action planning to help establish a collectively-generated prioritization of latrine use. Remembering, pledging, and environmental re-structuring are also used through the community meeting, wall painting, and household visits to create a descriptive norm and social commitment towards latrine use.

Do not know about safe disposal: Mothers meetings provide action knowledge on the different ways to safely dispose of child feces depending on the age of the child and their defecation habits (i.e. infants vs. toddlers vs. young children), demonstrate these safe disposal practices and provide guided practice, and create an enabling environment by providing different types of safe disposal hardware (i.e. potties and scoops).

Cleanliness, beauty, and status as motivators: Importantly, all of the intervention activities reiterated and emphasized the intervention motto of a 'clean, healthy, beautiful village.' This motto was developed based on formative research findings that found cleanliness, beauty, and status in particular to be important motives for behavior in the study context. As such, the motto acts as an affective persuasion strategy that is consistently employed throughout the intervention. The motto aims to shift community members' views on latrine use from being a behavior that is not prioritized/valued to a behavior that has clear influential motivations. The motto is especially highlighted through targeted discussion in the community meeting activity, and the motivators (cleanliness, beauty, status) specifically informed one of the central skits in the palla about the goddess Laxmi.

2.3. Intervention Monitoring plan

2.3.1. Process Evaluation Indicators

We conducted a mixed-methods process evaluation to monitor delivery of all intervention activities, informed by the guidance of Saunders, Evans, and Joshi (Saunders et al., 2005). Specifically, we aimed to understand intervention fidelity, dose, reach, recruitment, satisfaction, and context. Each of these components is defined in Table 1 below (from Saunders, 2005).

We created a process evaluation manual that outlined how each intervention activity would be assessed along the key components noted. Specifically, for each activity, we created a table that noted each process evaluation component, the relevant process evaluation questions that would address that component, the data sources and tools needed to answer the questions identified,

when data collection would need to take place to gather accurate information, and what the data analysis plan would be.

Table 1. Process Evaluation Components to be Evaluated as Part of *Sundara Grama* Delivery *(Definitions from Saunders et al 2005)*

Component	Purpose
Reach (Participation rate)	The extent to which the intervention reaches the intended population
Fidelity	The extent to which the program is implemented as planned (e.g. in manner consistent with its design)
Dose Delivered (Completeness)	The extent to which program components are delivered
Recruitment	Procedures used to recruit and engage participants
Satisfaction	The extent to which the program is received by the target group, including satisfaction and enjoyment of intervention or program activities
Context	Aspects of the environment that may influence intervention implementation or outcomes

2.3.2. Process Evaluation Data Collection Tools and Data Sources

The process evaluation data collection tools were both qualitative and quantitative, and are summarized in Table 2 below, along with information on their intended purpose. The majority of process evaluation data was collected at the time the intervention activities took place, qualitative process evaluation data was collected after intervention delivery in a subset of non-trial villages and post-endine in trial villages, and some process evaluation questions were included in the endline survey.

Process evaluation data collected at the time intervention activities took place assessed reach, fidelity and dose. Qualitative activities, specifically in-depth interviews and focus group discussions, were used to assess the recruitment, satisfaction, and context components. These tools are mentioned in Table 2 and described in section *3.4.2. Data Description*.

Process evaluation questions asked as part of the endline survey assessed various components, depending on the activity, including reach, dose, and satisfaction. Specifically, all households, regardless of latrine or intervention status, were asked about awareness and attendance at intervention activities. For intervention communities, this enables further understanding of intervention reach. In control communities, questions about the intervention enable understanding of spillover (see section *3.4.2. Data Description* for more information on 'Section I' of the endline tool).

2.3.3. Analysis of Process Evaluation Data

Assessment of Intervention Reach

Intervention reach was assessed during intervention delivery and at endline.

For the Palla, Transect Walk, and Community Meeting, the Emory Process Evaluation (PE) team members used a tally counter device to count the total number of village members in attendance

at a specific time point during the activity (Palla = number of audience members at 30 minutes into the performance, Transect Walk = number of individuals participating in the walk upon reaching the OD sites, Community Meeting = number of individuals participating in the meeting during activity #3 out of 7). The PE team members 'tallied' the total attendees as well as sub-groups by sex and age (total adult men, total adult women, total boys ages < 18, and total girls ages <18). For the Mother's Group Meeting, community mobilizers filled out a roster sheet of participants who attended. Using village data collected at baseline, we estimated the approximate proportion of target attendees in attendance at each of these activities. For the Palla and Transect Walk, we determined the proportion of the whole village in attendance, for the Community Meetings, we assessed the number of adults over age 18 in attendance, and for the Mother's Group meeting, we assessed the number of caregivers from households with children under age 5 in attendance. For each activity, Reach scores were determined for each village. Reach scores range from 1-10; a score of 0 equates to ~1-10% participation, a 1 equates to 11-20% participation, etc.

For latrine repairs, we assessed reach during the endline survey. Respondents in all households that were supposed to get latrine repairs (as determined by our assessment in July – September 2019) were asked if they received latrine repairs.

Table 2: Process Evaluation Tools for Assessing *Sundara Grama*		
Tool Name	**Purpose**	**Data Collection Logistics**
Palla Activity Checklist	*To Assess:* *Reach:* To record estimates of the number of participants (by sex and age group) in attendance at each performance; *Fidelity:* To determine if palla was delivered according to design; *Dose:* To determine if all intended skits and messages were delivered.	Target: All Palla performances Timing: During palla performance Administration: Emory
Transect Walk (TW) Activity Checklist	*To Assess:* *Reach:* To record estimates of the number of participants (by sex and age group) in attendance during TW; *Fidelity:* To determine if TW was delivered according to design; *Dose:* To determine if each TW component, such as marking of feces and specific messages, was delivered.	Target: All Transect Walks Timing: During Transect Walk Administration: Emory
Community Meetings (CM) Activity Checklist	*To Assess:* *Reach:* To record estimates of the number of participants (by sex and age group) in attendance during the CM; *Fidelity:* To determine if the CM was delivered according to design; *Dose:* To determine if each CM component, such as creation of action steps and group commitment, was delivered.	Target: All Community Meetings Timing: During Community Meetings Administration: Emory
Mother's Group (MG) Activity Checklist	*To Assess:* *Reach:* To record estimates of the number of participants (by sex and age) in attendance during the MG; *Fidelity:* To determine if MG was delivered as designed; *Dose:* To determine if each MG component, such as demonstrations and hardware distribution, was delivered.	Target: All Mother's Group meetings Timing: During Mother's Group meetings Administration: Emory
Household Visit (HV) Logsheet	*To Assess:* *Reach:* To record the number of participants (by sex) in attendance during the HV; *Fidelity:* To determine if the HV was delivered according to design; *Dose:* To determine if each HV component, such as personal pledge and distribution of poster, was delivered.	Target: All Household Visits Timing: During Household Visits Administration: Community mobilizer from RWI conducting visit; Emory staff did oversight to make sure visits occurring, sheets being filled
Community Wall Painting Logsheet	*To Assess:* *Fidelity:* To determine if the Community Wall Paintings were created as planned.	Target: All Wall Paintings Timing: Once Community Wall Painting complete Administration: Emory
Community Member's Perceptions of Intervention activities (FGD)	*To Assess:* *Recruitment:* To determine if community members had challenges/issues attending activities *Satisfaction: To assess perceptions of intervention activities (likes, dislikes, etc.)* *Context: To understand contextual factors that may have impacted delivery, attendance, participant perceptions of intervention*	Target: Community members in sub-study villages Timing: Once Community-level intervention activities complete Administration: Emory

Table 2: Process Evaluation Tools for Assessing *Sundara Grama* (continued)		
Tool Name	**Purpose**	**Data Collection Logistics**
Mother/caregiver Perceptions of Child feces disposal messages and directed activities (Mother's group meeting) (IDI)	*To Assess:* *Recruitment:* To determine if mothers/caregivers had challenges/issues attending activities *Satisfaction: To assess perceptions of mother's group activity (likes, dislikes, etc.)* *Context: To understand contextual factors that may have impacted delivery, attendance, participant perceptions of the mother's group activity*	Target: Mothers of children age 5 and under in sub-study villages who attended the mother's group meeting Timing: Once Mother's group activity complete Administration: Emory
Community Mobilizer Activity Feedback – IDI Guides (one per activity)	*To Assess:* Community Mobilizer's perception of the different intervention activities and experience of implementation.	Target: 4 Community Mobilizers per activity Timing: Once majority of focal activities complete Administration: Emory
Community Mobilizer Implementation Feedback - Focus Group Discussion Guide	*To Assess:* Community Mobilizer's experiences of implementation.	Target: 5 to 8 Community Mobilizers in 3 FGDs Timing: Once majority of focal activities complete Administration: Emory
Community Member's Perceptions of Intervention activities (Post-endine) (FGD)	*To Assess:* *Recruitment:* To determine if community members had challenges/issues attending activities *Satisfaction: To assess perceptions of intervention activities (likes, dislikes, etc.)* *Context: To understand contextual factors that may have impacted delivery, attendance, participant perceptions of intervention*	Target: Men and Women (separate discussions) in trial intervention villages Timing: Post-endline Administration: Emory
Community Member's Perceptions of Intervention activities (Post-endine) (IDIs)	*To Assess:* *Recruitment:* To determine if community members had challenges/issues attending activities *Satisfaction: To assess perceptions of intervention activities (likes, dislikes, etc.)* *Context: To understand contextual factors that may have impacted delivery, attendance, participant perceptions of intervention*	Target: Men and Women in trial intervention villages, who did and did not exhibit change in latrine use Timing: Post-endline Administration: Emory
Mother/caregiver Perceptions of Child feces disposal messages and directed activities (Mother's group meeting) (IDIs)	*To Assess:* *Recruitment:* To determine if mothers/caregivers had challenges/issues attending activities *Satisfaction: To assess perceptions of mother's group activity (likes, dislikes, etc.)* *Context: To understand contextual factors that may have impacted delivery, attendance, participant perceptions of the mother's group activity*	Target: Mothers/caregivers of children age 5 and under in trial intervention villages who attended the mother's group meeting who did and did not exhibit change in child feces disposal behavior Timing: Post-endline Administration: Emory

At endline, each household, regardless of latrine ownership, was asked if they recalled or had any member attend the various community level activities (Palla, Transect Walk, Community Meeting), or had seen the wall painting. All households with children under age five, regardless of latrine status, were asked if a member from their household attended the Mother's Group Meeting. All households with latrines were asked if a community mobilizer visited their household.

Assessment of Intervention Fidelity and Dose.
We created a combined 'Fidelity/Dose' score for each of the community level activities (Palla, Transect Walk, Community Meeting) and the Mother's Group from relevant indicators in the activity-specific tools. Each activity has a maximum possible score, which is based on the number of components that the activity should include to have been delivered with fidelity and to be considered complete (dose). Scoring criteria for each of the activities are outlined in Appendices 2-5.

Common Fidelity components assessed across the various activities include: attendance by a key stakeholder, length of activity, appropriate pre-activity preparations made, delivery of activity components in correct order. For Dose, each activity had key 'events' that needed to take place for the activity to be considered 'complete'. For example, Palla included various songs, skits, and stories, as well as opening and closing remarks. Each event within the activity was assessed with at least one and sometimes several questions by the observer to indicate if each event within the activity was carried out completely as planned.

We determined fidelity/dose of the Household Visit activity by assessing whether or not all three components of the activity were completed by the community mobilizer: reflection of household practices, commitment, and poster hanging.

We determined fidelity/dose of the Wall Paintings by reviewing photos of each completed painting to determine if all three key components were included: map of entire village with houses, clear identification of 'positive deviant' households on the map, and notation of the specific action steps determined in the village's community meeting.

For latrine repairs, we assessed dose during the endline survey. We asked respondents in all households that were supposed to get latrine repairs and who indicated that they got repairs, what repairs they received.

Assessment of Intervention Recruitment, Satisfaction, and Context.
We translated qualitative data, collected within the non-trial villages engaged in the substudy and post endline, and notes written by the research team when they observed intervention delivery or recorded community members' perceptions when sharing trial findings back to intervention villages. Data were thematically analyzed to generate a preliminary understanding of recruitment, satisfaction, and context elements across activities.

For latrine repairs, we assessed satisfaction during the endline survey. We asked respondents in all households that were supposed to get latrine repairs and who indicated that they got repairs, how satisfied they were with the repairs they received.

2.3.4. Measures Taken to Ensure Data Quality
We anticipated that observation of the activities could influence delivery by the community mobilizers as well as reception of the activities by community members. Thus, we elected to collect monitoring data *for all activities.* In addition to ensuring that the monitoring exercise had a

uniform influence across intervention villages, the detailed data collected enables identification of systematic issues, across activities, villages, and community mobilizer teams.

All members of the Process Evaluation team were trained and had the opportunity to pilot the tools and practice data collection in the villages where the intervention was piloted. During the piloting, team members reported challenges using mobile phones for data collection. As such, all tools were converted to paper to enable the team members to easily move through the tool and take notes, which was particularly important if the activity components were delivered out of order or if an unforeseen event occurred that needed to be captured (e.g. interruption, weather, the need to move locations, etc.).

3. Evaluation Questions, Design, Methods, Sampling & Data Collection

3.1. Primary and Secondary Evaluation Questions

The objective of our evaluation is to evaluate a multi-level, theoretically informed intervention designed to increase latrine use for defecation among all members of households that currently own a latrine. Our primary outcome for the study is latrine use. Secondary outcomes are latrine coverage and latrine use determinants.

The main evaluation questions to be tested and associated hypotheses are as follows:

Main research question: Is latrine use among people who own a latrine in communities that received the intervention significantly different at endline than among people who own a latrine in communities that did not receive the intervention?

- H_1: Latrine use among people who own a latrine in communities that received the intervention will be significantly higher compared to latrine-owning households in control communities.

Secondary research question 1: Is latrine construction by endline among people who do not own a latrine in communities that received the intervention significantly different than among people who do not own a latrine in communities that did not receive the intervention?

- H_2: Latrine construction among non-latrine owners in communities that received the intervention will be significantly higher compared to controls.

Secondary research question 2: Are behavioral determinant scores (i.e. scores for social norms, abilities, physical opportunity, risk perception, motivation, and self-regulation) significantly different at end line among owners of latrines in intervention villages compared to owners of latrines in control villages?

- H_3: Latrine use behavioral determinant scores are significantly higher at endline among latrine owners in intervention villages compared to latrine owners in control villages.

Secondary research question 3: Are behavioral determinant scores (i.e. scores for social norms, abilities, physical opportunity, risk perception, motivation, and self-regulation) associated with latrine use?

- H_4: High latrine use behavioral determinant scores are significantly associated with latrine use.

3.2. Evaluation Design and Methods

3.2.1. Identification Strategy

We conducted a cluster randomized controlled trial (CRT) in rural Puri district, Odisha, India to determine if those who own latrines in villages that received a multi-level behavior change intervention use their latrines significantly more at endline than those in control villages. A cluster-randomized design was selected because the intervention was to be delivered at the village level. To assess immediate perceptions of the intervention and potential spillover, qualitative research was conducted in six non-trial villages, three of which received the intervention, within one month of intervention delivery. Additional qualitative research was conducted within weeks of endline data collection completion in four trial intervention villages, specifically in villages and among individuals for whom endline data collection indicated either a change in behavior or no change at all.

See Figure 2 for a flow diagram of the study design and Figure 3 for a timeline of study activities. See Appendix 1 for our pre-analysis plan.

Figure 2: Trial Flow Diagram

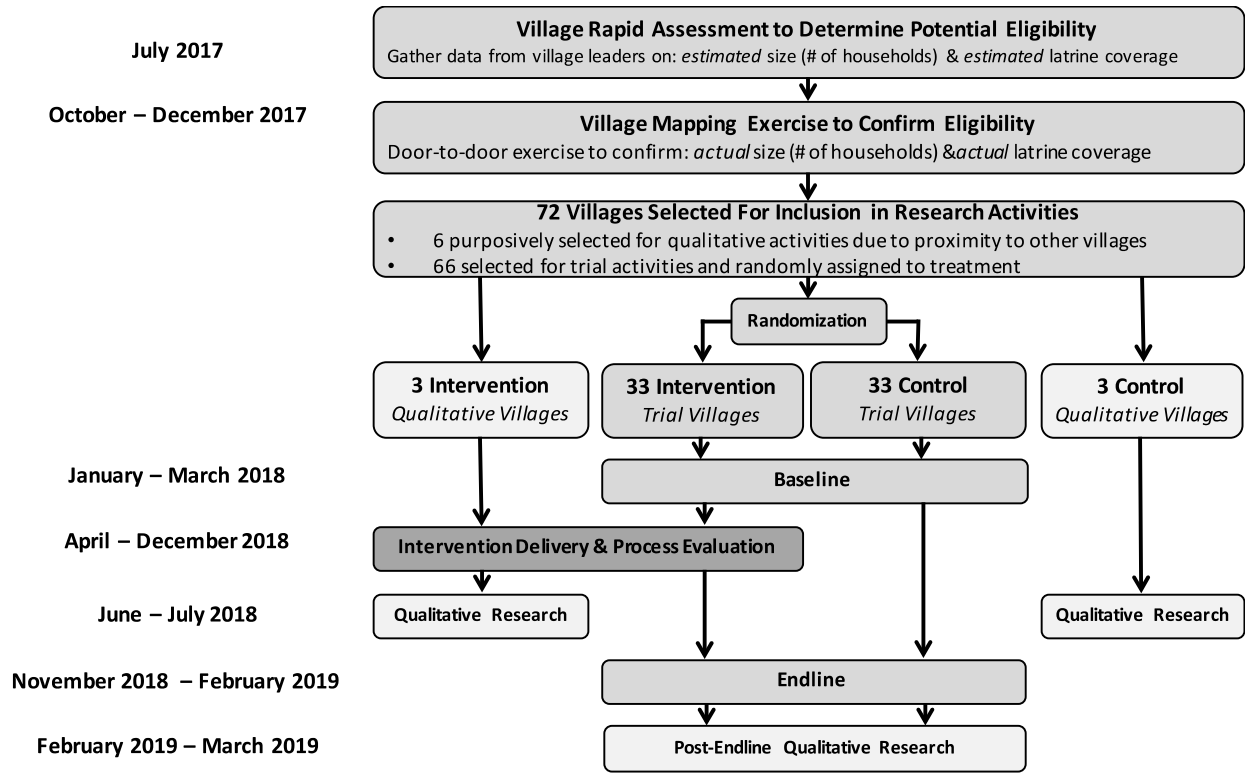
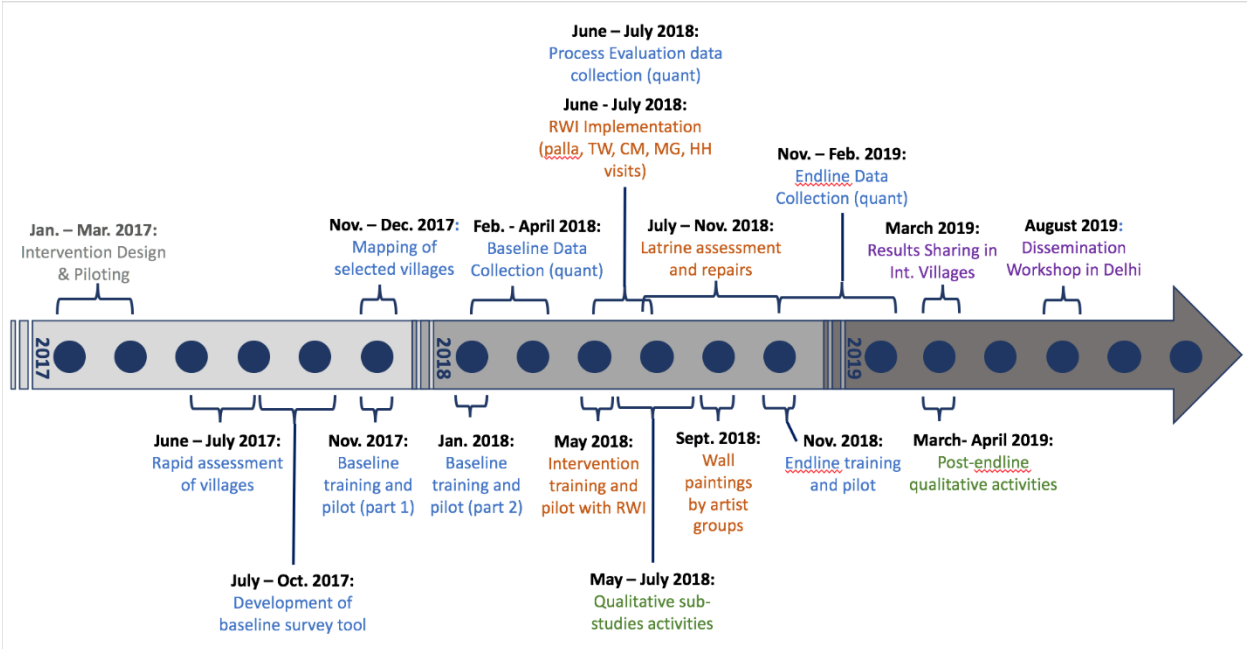


Figure 3: Trial Timeline

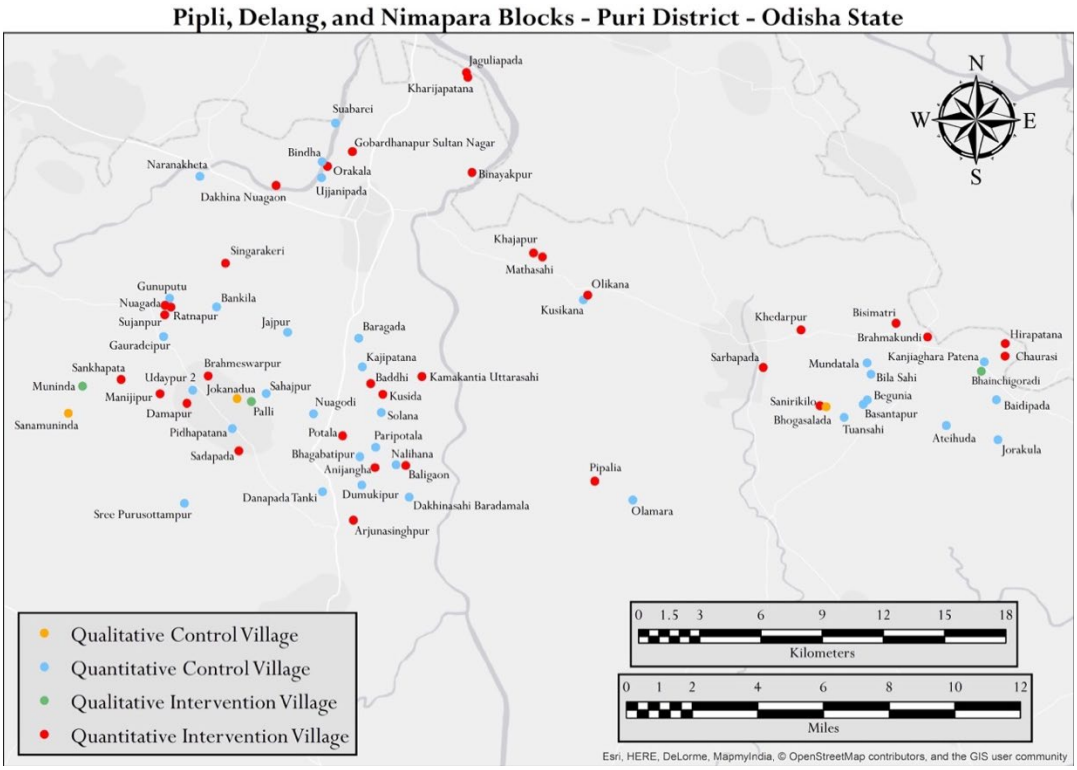


3.2.2. Sample Size

We identified 72 eligible villages to engage in research activities. Sixty-six villages are included in the CRT (33 treatment and 33 control). The remaining six villages were engaged in qualitative activities exclusively (3 treatment, 3 control).

Trial cluster sample size. We used the primary outcome of reported latrine use at the last defecation to determine a sample size of 33 villages per arm for the trial (See Figure 4 for a map of study villages). We used a simulation approach that accounted for baseline assessments of latrine use and adjusted for both within-person and within cluster correlations (Arnold et al., 2011). Latrine use data collected from 2011-2013 during a sanitation trial in Odisha was the source of simulation parameters (Clasen et al., 2014). Specifically, we assumed baseline latrine use of 45%, a village-level ICC of 0.10, and a within-person correlation from baseline to follow-up of 0.60. We conducted a rapid assessment of villages in the study area in 2017 to estimate latrine coverage. As no trials have previously investigated the impact of a behavioral intervention designed to increase latrine use among household that already owned a latrine, we identified a 10% increase in use (from 45% to 55%) as a minimum intervention effect, under the assumption that the theory-informed behavior change intervention would have a greater effect on latrine use than interventions that have increased latrine coverage alone. Garn et al 2017 conducted a post hoc regression analysis and found an increase in latrine use of 5.8% for every 10% increase in latrine coverage (Garn et al., 2017). Finally, we assumed villages would have an average of 292 eligible participants for whom latrine use could be assessed (cluster size coefficient of variation = 0.35), 10% loss to follow-up, 80% power, and 0.05 significance level. Sample size estimates were also checked using the *clustersampsi* add-on package in Stata v.14 (StataCorp, College Station, Texas, USA).

Figure 4: Map of Study Villages



Comparing the estimates that informed our initial sample size calculations to data collected from households surveyed during baseline, we observed slightly lower average village size (98 households estimated vs. 88 households actual), higher average latrine coverage (68% estimated vs. 75% actual), and lower average number of persons with latrines per village (292 estimated vs. 256 actual). In addition to updating our estimate based on our baseline enrolment numbers, we used the baseline data to calculate the village-level ICC, which was equivalent to our estimated ICC based on prior work in Odisha (0.10 estimated vs. 0.103 actual). However, the proportion of persons reporting using the latrine at baseline was higher than our estimate (0.45 estimated vs. 0.60 actual).

Based on our baseline enrolment, unequal cluster sizes with a CV = 0.37, baseline latrine use of 60%, and 10% loss to follow-up from baseline to follow-up, **we are powered to detect an absolute difference of 10% as planned (required N=7936 individuals per arm)**.

Appendix 6 displays the estimated parameters that informed our initial sample size calculations and the actual values from our baseline data collection.

Trial household sample size. In selected villages, all households that owned latrines were eligible for inclusion to assess latrine use of all household members at baseline and endline. Additionally, a subset of 20 households per village were randomly selected to have a household member complete an assessment of latrine use determinants at baseline and endline, with the aim of having 50% of respondents be female and 50% be male. Finally, at baseline, all latrine-owning households with children age five and under were asked to have a primary caregiver complete an assessment of child feces management practices and answer questions on behavioral determinants. At endline, the same assessment was asked to be completed by caregivers in all households with children age five and under, regardless of latrine ownership.

Qualitative sample size.
The study included two rounds of qualitative research – one round right after implementation of activities by RWI (we refer to this qualitative work as the 'Sub-study qualitative') and a second round right after endline data collection was completed (we refer to this qualitative work as the 'Post-endline qualitative').

Sub-study qualitative
A subset of six non-trial villages was engaged in qualitative research that took place after implementation in three treatment villages and three control villages. We purposively selected 3 villages from each study arm with one village from each study block (Pipli, Delang, Nimapara) to provide perspective from each block context. The Sub-study had four research aims:

1. To explore community members' perceptions of the *Sundara Grama* intervention (Noted in Section 2.3 'Intervention Monitoring Plan' and described in Table 2, this includes the 'satisfaction' element of the process evaluation, (what they liked, did not like, if/how it influenced them, etc.), as well as issues related to 'recruitment' and 'context')
2. To examine spillover effects in control villages (if men and women in control villages heard anything about the intervention activities, what they may have heard/learned related to latrine use, etc.)
3. To explore mothers' perceptions of the child feces disposal components of the *Sundara Grama* intervention (Noted in Section 2.3 'Intervention Monitoring Plan', this includes the 'satisfaction' element of the process evaluation, (what they liked, did not

like, if/how it influenced them, etc.), as well as issues related to 'recruitment' and 'context')
4. To explore what implementers (RWI staff) thought of the intervention (what went well, what did not go well, what would they improve, etc.).

The original sample size plan to meet aims #1 and #2 was to conduct 12-18 in-depth interviews (IDIs) with men and women and 6 sex-segregated focus group discussion (FGDs) with 5 to 8 participants in each. The original sample size plan to meet aim #3 was to conduct 12-18 IDIs with caregivers and 2 to 3 FGDs in control villages to understand norms around child feces disposal practices and to explore spillover of the mother's group meeting (which also met aim #2). The FGDs for aims #1 and #2 included questions that explored the child feces disposal components of the intervention (e.g. mother's group meeting activity) and in this way, also addressed aim #3. We aimed to engage several participants from each village by conducting at least 1 FGD and several IDIs in each village. The sample sizes were also determined based on an estimate of how many activities/participants were required to fully explore the qualitative research aim and reach saturation of themes across the six villages.

The sample size to meet aim #4 was 12 to 16 IDIs with RWI mobilizers and 2 to 5 FGDs with RWI mobilizers and supervisors. The sample size was determined based on the idea of conducting an interview with each mobilizer to gather individual perspectives and several group discussions to capture shared and differing perspectives on the challenges and success of implementation and again, reach saturation of themes.

For more information on the qualitative sub-study sample, please refer to section *3.4.1 Sample Selection.*

Post-endline qualitative

A subset of four intervention trial villages was engaged in the qualitative research that took place after endline data collection. The aim of this research was to understand what aspects of the *Sundara Grama* intervention may be of greatest value (i.e. led to behavior change), what aspects are not effective (i.e. did not influence behavior change), and what factors at the individual and community level we did not target effectively or may not be alterable. The sample size planned was 8 sex-segregated FGDs (2 per village) and 20 IDIs (5 per village) with one participant from each of the 5 household types we aimed to gather perspective from (see description of household types in 3.4.1 sample selection). Again, the sample size was determined in order to ensure the research aim was explored in each village (i.e. 2 FGDs per village, 5 IDIs per village) and to reach saturation of themes.

3.2.3. Randomization of Trial Villages

We assigned villages to control or intervention status using stratified randomization. While there are many potential criteria (program under which latrines were provided, proportion constructed with household funds versus government subsidy, etc.), we grouped eligible households into four strata based on village size and latrine coverage (median splits). Randomization was conducted within strata using the Stata '*randomize*' command with the *'block*' option and a defined seed to ensure reproducibility.

3.3. Ethics

The Institutional Review Board at Emory University in Atlanta Georgia (00098293) and the Ethics Review Committee at Xavier Institute of Management in Bhubaneswar, Odisha, India reviewed and approved study protocols. Trained research assistants read to all participants a consent

document in Odia, the local language, that described the study and objectives prior to any quantitative and/or qualitative data collection activities and then asked for the participants' verbal consent before continuing. Verbal consent was approved as not all study participants are able to read, particularly women, and therefore may not feel confident signing a document. All consent documents indicated that participants could cease participation or skip questions at their discretion, and research assistants were trained to respect participants' responses and decisions. Only members of the research team had access to participants' names or other identifying information, with the exception of those participating in focus group discussions (FGDs). In FGDs, we reminded participants that the information they shared would be known to the other participants, and while we requested all participants to not repeat shared information outside of the discussion, we warned that we could not control whether or not people shared their information and to participate at their discretion.

All data collectors had prior experience in conducting research and obtaining informed consent from participants. Prior to baseline all data collectors were re-trained on the informed consent process and the particular information sheet and consent for this study.

After endline was completed, Research Assistants (RAs) visited all intervention communities to share aggregated findings. Specifically, RAs held meetings at pre-arranged times open to all community members. At meetings, RAs went through broad results, sharing information specific to the village, including the number of households and individuals, the number and percent of households that have latrines, including a breakdown by condition (complete, in construction, abandoned) and type (single pit pour flush or flush to tank). The RAs then presented data on latrine use before and after the intervention in that particular village and in aggregate across all intervention villages to enable comparison and discussion. Specifically, RAs shared the total number of individuals that use latrines (including breakdown by sex and age), the number of households with all members using the latrine, and the number of households reporting disposing of child feces into the latrine. Community members were then able to ask questions about the findings and share opinions about the data, including their perception of its accuracy.

3.4. Sampling and Data Collection

3.4.1. Sample Selection

Trial Village and Household Selection.

We sought villages that were not declared open defecation free and, to power the study, had a mean size of 100 households (ideally between 50-150 households) and a mean latrine coverage of 60%.

To identify eligible villages, we first carried out a rapid assessment of villages from June to July 2017 in three blocks in Puri district (Delang, Pipli, and Nimapada). We focused on these blocks due to the ability of our team's and partner's ability to access them for evaluation and intervention delivery. We carried out this exercise, which involved talking with village leaders to gather information, assuming village sizes and coverage are always changing and that any records available would be dated, and that visiting and talking with village leaders would be able to provide the most up-to-date village information. We visited a total of 282 villages for the rapid assessment. This list of 282 villages was used as our sampling frame. From this list, we identified all villages that potentially could be eligible for inclusion. In November and December 2017, these villages were re-visited and mapped to both 1) confirm size and coverage numbers originally estimated and, if included in the trial, 2) later enable within-village randomization for administration of specific sub-sections of the baseline survey.

A total of 130 villages were mapped, including 115 presumed eligible from the rapid assessment. An additional 15 villages not involved in the rapid assessment were added in order to find enough villages to fit our criteria. Fifty-eight villages were excluded based on village size (N = 9, too large), low coverage (N = 38), being declared ODF free prior to baseline (N = 9), or having a large number of hamlets or hamlets located far from the central part of the village (N = 2).

The remaining 72 were selected to be included in research activities. We first purposively selected the 6 villages to engage in the qualitative research (further explained in following section). The remaining 66 selected were engaged in the trial and have a mean size of 97 households per village (range: 47-132) and a mean latrine coverage of 63% (range: 49%-95%). See Table 3 for target and achieved sample sizes for trial.

Trial Household Selection. All households with a latrine (regardless of functionality) were eligible for inclusion. Within households, data regarding the last defection event was sought for all permanent household members. This resulted in a baseline enrolment of 3972 households with latrines (Control=2045, Intervention=1927), for a total baseline sample of 16880 persons over the age of 5 years (Control=8654, Intervention=8226).

Table 3: Target and Achieved Sample for Trial			
	Target	**Actual, baseline**	**Actual, endline**
Villages	66	66	66
Intervention	33	33	33
Control	33	33	33
Households	All HHs that own latrines in each village	3978 (4251 Eligible)	4280 (4484 Eligible)
Intervention	All HHs that own latrines in intervention villages	1928 (2077 Eligible)	2100 (2170 Eligible)
Control	All HHs that own latrines in control villages	2050 (2174 Eligible)	2180 (2314 Eligible)
Individuals	All Individuals from HHs that own latrines in each village	13,406	13,406
Intervention	7936	6862	6862
Control	7936	6544	6544

Sub-Study Qualitative Village and Participant Selection.
The six qualitative villages were purposively selected from the pool of 72 eligible villages. They represent three pairs of villages from three distinct blocks that are within 1 kilometer of each other. Because of this proximity, we felt they would not be ideal for the trial due to the potential for spillover.

In each pair, one village was identified to receive the intervention and one to serve as a control. The proximity, while not ideal for trial activities, allows for qualitative assessment of whether or not spillover occurred in the control villages. One of the selected villages was unintentionally engaged in piloting activities. As such, a new village needed to be selected and only two of the three pairs were within 1 kilometer of one another as planned, enabling a realistic assessment of spillover in two of the engaged control villages. At the time of the mapping exercise, these six villages had a mean size of 102 households per village (range: 42-154) and a mean latrine coverage of 65% (range: 51%-94%), roughly matching the characteristics of the trial villages.

Target participants for the FGDs (to meet aims #1 and #2 described above in section *3.2.2 Sample size*) were men and women aged 18 years or older who reported attending several of the *Sundara Grama* intervention activities. The qualitative research team asked for support from village *Anganwadi* (rural child care center) workers in helping recruit target participants in advance since *Anganwadi* workers are often familiar with all households residing in the village as a result of their work. If the Angwandi worker was not able to recruit enough participants, then participants were recruited on the day of the focus groups through convenience sampling via going from house to house and snowball sampling.

Target participants for the IDIs among caregivers (to meet aim #3 described above) were women who had attended the mother's group meeting. Since the process evaluation data showed that a wide age range of women attended the mother's group meeting activity, women in different life stages from young wives, to mature wives, to grandmothers were purposefully targeted. Recruitment for the IDIs was primarily conducted the day before the interviews. Research assistants again engaged *Anganwadi* workers to help identify and recruit mothers to participate. After locating the initial contacts suggested by the *Anganwadi,* the research assistants conducted snowball sampling from initial contacts or other members of the community to seek out more participants.

Target participants for aim #4 were RWI community mobilizers and supervisors. Our team directly engaged with RWI staff and did not require any specific recruitment strategy. See Table 4 for target and achieved sample sizes for sub-study qualitative research.

Table 4: Target and Achieved Sample for Sub-study Qualitative Research

Activity	Target	Actual
IDIs	36 to 52	44
IDIs with Community members in intervention villages	6 to 9 (2 to 3 per village)	0*
IDIs with Community members in control villages	6 to 9 (2 to 3 per village)	0*
IDIs with Women who attended the Mother's Group Meeting in intervention villages	12 to 18 (4 to 6 per village)	24 9 women (aged 20-29) 9 women (aged 30-39) 1 woman (aged 40-49) 2 women (aged 50-59) 3 women (aged 60-69)
IDIs with RWI community mobilizers	12 to 16	20
FGDs	10 to 14	15
FGD with Mother's with children under age five in control villages[+]	2 to 3 (across 3 villages)	4 (across 3 villages)
FGD with Community member's in intervention villages	3 (across 3 villages)	4 (2 with men and 2 with women, across 3 villages)
FGDs with Community members in control villages	3 (1 across 3 villages)	4 (2 with men and 2 with women, across 3 villages)
FGDs with RWI community mobilizers and supervisors	2 to 5	3

*We ultimately decided FGDs were the more appropriate qualitative method for exploring aims #1 and #2 since the goal was to understand a *broad range of views and shared perspectives* on the intervention and whether spillover took place. In-depth interviews are more appropriate for gathering perspectives on personal experiences. As a result, we focused data collection efforts on conducting FGDs.

Post-Endline Qualitative Village and Participant Selection.
Four villages were selected for post-endline qualitative research. The villages represented all 3 study blocks, geographic spread, variety in percent change in latrine use (low to high percent change), and had intervention activities implemented by one of each of the four RWI supervisor teams.

Five in-depth interviews and two focus group discussions were planned for each village. Endline data was used to identify participants for each of the activities.

For in-depth interviews, in each village we sought to engage a household member from each of the five household types:

- Household with change in latrine use:
  None or only a few of members reporting use at baseline and all reporting use at endline;
- Household with *no* change in latrine use:
  None or few of members reporting use at baseline and no change in use at endline;
- Household with change in child feces disposal behavior:
  Household did not report practicing safe child feces disposal at baseline but reported practicing safe disposal at endline;
- Household with *no* change in child feces disposal behavior:
  Household did not report practicing safe child feces disposal at baseline or endline;
- Household with a new latrine:
  Household did not have a latrine at baseline but had a latrine at endline that they constructed themselves.

For all household types, we used endline data to ensure the household had reported attending 3 or more of the intervention activities. In regards to target respondent, for the safe disposal of child feces IDIs, the target respondent was the mother of the youngest child. For the other IDIs, the target respondent was any adult household member, either male or female, who ideally attended some of the intervention activities. Importantly, while endline data was used to target participants, in some cases those interviewed reported different behavior during the interviews than at baseline. In Table 5, the number of actual respondents engaged is based on what was reported in the IDIs, not at endline.

We ultimately elected to not conduct any IDIs with households that had new latrines at endline, contrary to our original plan. Endline data revealed that latrine construction in intervention and control households was comparable. Thus, we elected to focus efforts on learning more about latrine use and feces disposal behaviors.

For FGDs, we aimed to have one FGD with women and one with men in each identified village, with a target of 6-8 participants per discussion. Target participants were those who attended intervention activities and thus, could offer insights and feedback.

Table 5: Target and Achieved Sample for Post-Endline Qualitative		
Activity	**Target**	**Actual**
IDIs	20 (5 across 4 villages)	20
HH with change in latrine use	4 (1 per village)	9
HH with no change in latrine use	4 (1 per village)	3
HH with change in CFD practice	4 (1 per village)	5
HH with no change in CFD practice	4 (1 per village)	3
HH with new latrine at endline	4 (1 per village)	0
FGDs	8 (2 per village; 6-8 participants each)	6
Women only	4 (1 per village)	3
Men only	4 (1 per village)	3

3.4.2. Data Description

Trial Data Collection Tools

The quantitative baseline and endline instrument includes the 8 sections (A-H) that are described below. The endline instrument included a few additional questions in some of the sections and one entirely new section (section I) that asked about the intervention activities. Note that if a household was surveyed at baseline, then some sections or parts of a given section were skipped at endline as the information was not necessary to collect again as it was reasonable to expect no change in response (e.g. parts of section B, D, and E). For each section, the target respondent and objective is described.

Section A – Determination of Census Eligibility:

Respondents: All households in each community identified to be a part of the trial were eligible to complete section A, regardless of latrine status, if a member of the household who was over age 18 was home and willing to participate. Female respondents were prioritized.

Objective: The objective of this section was to determine if the household owned their own latrine, which is eligibility criteria for the remaining sections of the baseline tool. Basic demographic information (household size, caste) were also recorded to identify trends in ownership versus non ownership in the village and among all villages collectively.

Endline additions: For households that reported having more or fewer latrines compared to the number of latrines reported by the household at baseline, we included a few additional questions that asked about why a new latrine had been built or a latrine had been destroyed.

Section B – Household SES & Non-Sanitation Questions:

Respondents: All willing and consented participants in latrine owning households.

Objective: The objective of this section was to collect expanded demographic information from the participating household (religion, SES indicators, occupation, education, etc.).

Section C – Latrine Use of Household Members:

Respondents: All willing and consented participants in latrine owning households.

Objective: The objective of this section was to determine latrine use for defecation among all household members, including feces management practices among children under age five and those who are immobile.

Section D –Household Water Sources & WASH Facilities:

Respondents: All willing and consented participants in latrine owning households.

Objective: The objective of this section was to gain information about the household's WASH facilities, including water source, ownership of a bathing area, support (monetary or material) for latrine construction, and if the household member was aware of any latrine promotion activities that took place in the community to date.

Section E – Household Sanitation (Latrine History & Sludge Management):

Respondents: All willing and consented participants in latrine owning households.

Objective: The objective of this section was to gain more information about the household's latrine, including condition, if it is currently in use, and pit emptying needs and practices.

Section F – Behavioral Determinants - Latrine Use:

Respondents: Using data from a mapping exercise conducted prior to survey administration, households were randomly selected to complete this section of the instrument. Households were randomized using a random number generator in Microsoft Excel and were then randomly assigned. No more than 20 households were to complete per village. If a male household member was available, he was asked to complete with the aim of having 10 females and 10 males complete this section per village.

Objective: The objective of this section was to assess various factors, including attitudes, norms, risk perceptions, and motivations for using and/or not using their household latrine.

Section G – Behavioral Determinants – Child Feces Disposal:

Respondents: At baseline: all willing and consented participants in latrine owning households with children under age 5, ideally the primary or secondary caregiver. At endline: all willing and consented participants in all households in engaged villages with children under age 5 *regardless of latrine ownership*, ideally the primary or secondary caregiver.

Objective: The objective of this section was to assess current child feces disposal practices and various factors, including attitudes, norms, risk perceptions, and motivations that influence child feces management. Information was solicited from *non-latrine owning households at endline* in both intervention and control villages as the intervention was open to all caregivers of children age 5 and under, regardless of latrine ownership.

Section H – Latrine Spot Checks:

Respondents: No respondents. Observational data collected from all households that had a member consent to the baseline survey and agree to have observations conducted.

Objective: The objective of this section was to assess current status of latrines.

Endline additions: One additional question was added to observe whether or not the pit needed to be emptied.

Section I – Intervention Activities

Respondents: All willing and consented participants in latrine owning households.

Objective: For intervention households, the objective of this section was to determine which intervention activities any member of the household attended, if any, what the respondent remembered about the activity (if the respondent themselves attended), and whether or not the respondent heard or talked about the activity with others. For control households, the objective of this section was to assess spillover. Specifically, participants were asked if they had heard of any of the intervention activities.

Sub-Study Qualitative Data Collection Tools

The Sub-study qualitative research aimed to understand village member's perceptions of the various activities conducted in intervention villages, latrine use and safe child feces management behaviors of community members, and potential spillover in control villages. Sub-study tools are described below in Table 6.

Table 6: Sub-Study Qualitative Data Collection Tools

Tool	**Purpose**
IDI with Women who attended the Mothers Group Meeting in intervention villages	*To Assess:* Mothers' child feces management practices, their perceptions of the Mother's Group meetings, including the education delivery and informational value, and their perceptions and usage or non-usage of the potty and scoop received.
FGD with Mothers with children under age five in control villages	*To Assess:* Mothers' child feces management practices and spillover from intervention villages.
FGD with Community members in intervention villages	*To Assess:* Community members' perceptions of the intervention, including their opinions of each activity, how activities could be improved, and if the intervention impacted behavior in the village.
FGDs with community members in control villages	*To Assess:* Community members' sanitation behavior, the history of sanitation programs in the village, and if and what they had heard of the intervention occurring in adjacent villages.

Post-Endline Qualitative Data Collection Tools

The Post-endline qualitative data tools were designed to gain additional information about behavior and understand if, how, and why the intervention may or may not have had an influence. The tools are described in Table 7.

Table 7: Post-Endline Qualitative Data Collection Tools	
Tool	**Purpose**
IDI with adults from latrine-owning households in intervention villages	*To Assess:* Women and men's perceptions of the intervention, and how the intervention may or may not have influenced personal and household latrine use behavior.
IDIs with mothers/caregivers of children age 5 and under who do and do not safely dispose of child feces	*To Assess:* Participant's perceptions of the mother's group meetings (and other intervention components), and how the mother's group and broader intervention may or may not have influenced child feces management behavior,
FGDs with community members	*To Assess:* Community members' perceptions of the intervention, and if and how the intervention may or may not have influenced latrine use in the community.

3.4.3. Quality Control

The following quality controls were conducted on the baseline and endline survey data (i.e. quantitative data):

- *Field visits:* The field manager and Dr. Parimita Routray, Indian PI, made several field visits at the start of baseline data collection to ensure the data collection protocol was being properly followed by the team and to oversee data collection of some of the new enumerators. Dr. Parimita Routray conducted field visits at the start of endline data collection, as she did at baseline, to ensure protocol was being followed for the same purposes.
- *Accompaniments:* At the beginning of baseline and endline data collection, for the first few weeks, two trained field supervisors observed the data collection of enumerators, identified those who were struggling on certain sections of the survey (be it the behavioral determinants section, the latrine observation or some other specific part) and provided additional support and training. During endline data collection, field supervisors especially observed and provided additional support to the two new enumerators who were not part of baseline data collection.
- *Data entry and data checks:* At baseline and endline a "Progress Database" was maintained that recorded the number of surveys collected each day and in what village. The field manager and Dr. Parimita Routray also confirmed in the database each day that all surveys collected on the mobile phones were uploaded to our server. The research team then conducted checks on the uploaded data by checking that the number of surveys uploaded for a given village matched the number of surveys recorded in the "Progress Database." At baseline, the research team also examining the length of time each survey took (checking for surveys conducted in less than 15 minutes, which would flag a potential issue during delivery).
- *Data processing:* At baseline, duplicate households within each village were identified. The supervisor's tracking sheet was referred to for each duplicated household to determine the correct household identification number. Households were deleted if no information was found in the supervisor's tracking sheet, or because they were true duplicates. New household identification numbers were created for households based off

of the records from the supervisor's tracking sheet. This data processing also took place at endline.

For qualitative data, names and identifying information was excluded from transcripts to ensure anonymity. Additionally, attempts were made to have sufficient sample sizes to enable triangulation of findings.

3.4.4. Specifications

To evaluate the impact of the intervention on latrine use, we used GEE with robust standard errors to estimate a marginal (population average) model with the general form

$$g\left(E\left[Y_{ij}|x_{ij}\right]\right) = \boldsymbol{x'_{ij}}\boldsymbol{\beta}_{ij}$$

where g(.) is the link function**,** Y_{ij} is the outcome of interest for the j_{th} observation in the i_{th} cluster, $\boldsymbol{x'_{ij}}$ is a vector of covariates**, and** $\boldsymbol{\beta}_{ij}$ is a vector of regression coefficients. We specified an exchangeable correlation matrix as the most plausible and parsimonious choice of working correlation structure, noting that GEE with robust estimation yields valid estimates of model coefficients and standard errors when the correlation structure is mis-specified. We used the difference-in-differences method (as noted in our pre-analysis plan) rather than simply controlling for the baseline outcome in order to estimate the treatment effect. This resulted in the model specification

$$\begin{aligned}\text{post} - \text{intervention latrine use} = \beta_0 + \beta_1\left(\text{treatment}_{\text{assignment}}\right) + \beta_2(\text{time}) + \beta_3(\text{time} * \\ \text{treatment}) + \beta_4(\text{age}) + \beta_5(\text{sex}) + \beta_6(\text{male educational attainment}) + \\ \beta_7(\text{female educational attainment}) + \beta_8(\text{household size}) + \beta_9(\text{household SES})\end{aligned}$$

where the coefficient β_3(time*treatment) is the estimate of the causal effect.

GEE was chosen to account for the clustered nature of the data arising from the cluster-randomized design. Of the two most widely used approaches to modelling the correlation structure in clustered designs, GEE and multilevel modelling, we choose GEE because 1) it estimates the population-averaged, or marginal effects of the intervention which are more relevant in the context of a large-scale global health intervention, and 2) is more robust to misspecification of the true underlying cluster-correlation structure than multilevel models, which can be subject to substantial bias when the model is mis-specified(Hubbard et al., 2010). That said, as a sensitivity analysis we did fit all models as multilevel models (adjusting for village and household clustering) and did not observe any notable changes in model results.

Initially, we attempted to fit log-binomial difference-in-differences (DID) models with GEE as specified in our pre-analysis plan, but these models failed to converge (a problem frequently noted in the statistical literature). Instead, we fit linear probability models which, in contrast to non-linear DID models, allow for a straightforward interpretation of the DID coefficient. We used robust standard errors given the issues with heteroskedasticity implicit in this approach and examined the marginal probabilities of the outcome for each time*treatment combination in order to ensure that predicted probabilities fell within the unit interval [0,1]. Lastly, we conducted a sensitivity analysis by fitting equivalent logit DID models with GEE. Both the predicted probabilities and significance of model coefficients from those models did not differ substantively from the linear probability models. Our pre-analysis plan is in Appendix 1.

3.4.5. Strategies to Avoid Bias and Address Spillover Effects
Strategies to avoid bias in quantitative data collection:
Our research team took several steps to ameliorate various potential sources of bias.

Our outcome measure, which was collectively agreed upon by 3ie, r.i.c.e., and all four research teams engaged in assessing behavioral interventions on latrine use in India, was deliberately designed to avoid social desirability bias or respondents answering questions about latrine use as they think they should answer them to depict themselves in a postivie light. In an attempt to reduce social desirability bias, we had enumerators first tell the respondent that "I have seen that some people defecate in the open, and some people use the latrine. Now I want to ask about where you and your family members defecate." By adding this opening phrase, the aim was to indicate that the enumerator was neutral and to reduce perception of there being a 'correct' answer. In addition, based on feedback from our enumerators, the way in which the latrine use question was administered may have also helped to mitigate response bias. Instead of having each household member come to the enumerator and report their defecation practice, the primary survey respondent was able to report on their behalf. This approach could have helped to mitigate social desirability bias.

We recognize that bias, in the form of measurement error, could have taken place if a respondent reported on their family members' defecation practices, but did not report accurately simply because they did not have accurate information to report. In an attempt to mitigate this form of bias, we aimed to engage female household members to answer questions about family member latrine use, assuming they would have the best knowledge of their family member's defecation practices, particularly children.

It is possible, that the multiple rounds of surveying on latrine use/open defecation practices in a relatively short time frame could have led to courtesy bias, or a want to respond in a way that was not offensive to the enumerator team. Trial households experienced a baseline survey, an endline survey, and possibly a measurement team baseline survey, measurement team endline survey, and post-endline qualitative activities. This means a household could have been approached up to five times in the course of 1.5 years about latrine use.

Social desirability bias, courtesy bias, and measurement error all could have resulted in, and may explain, the reported increase in latrine use across intervention and control communities. However, we carried out the same research activities in intervention and control communities so expect equal influence on both arms, if at all.

Finally, by the time we were conducting the post-endline qualitative research, we did identify survey fatigue, which occurs when respondents are tired of taking surveys and decline to participate. Specifically, potential participants or their family members expressed frustration that our team was continuing to approach their household to engage in a post-endline qualitative research activity and refused to participate.

In regards to the process evaluation quantitative data, we avoided self-evaluation bias by having the majority of process evaluation data collected by a separate evaluator team rather than members of the implementer team. Based on results from the process evaluation data (reported below in section 4), it is clear that the enumerators were not biased towards answering that all activities were implemented with full fidelity.

Strategies to avoid bias in qualitative data collection:
We avoided social desirability bias in qualitative data collection by developing IDI and FGD guides that included exploratory, open-ended questions that were neutral in tone, did not pose leading questions, and did not suggest a right or wrong answer. In addition, during the consent process, research assistants specifically explained to participants that there were no right or wrong answers. Lastly, research assistants were trained to be effective interviewers and facilitators that do not guide the participant to certain responses but conduct the activities in a neutral disposition to elicit the participant's true experience.

Strategies to address spillover effects:
We attempted to mitigate spillover by mapping our selected villages (using GPS coordinates taken at the center of the village, roughly) and checking to make sure there was at least ~1km "buffer" between all villages before randomization into treatment arm.

We assessed spillover in the Sub-study qualitative research. We purposefully selected two "pairs" of intervention and control villages that were in close proximity (less than 1km) to understand spillover effects. We found that spillover was variable as it largely depended on the relationships that households from different villages had with each other. Spillover only took place in one of the intervention and control village pairs where control qualitative participants largely knew about only the palla and transect walk activities.

We also assessed spillover quantitatively by asking control households in the endline survey about whether or not activities had taken place in their village. We did not find any evidence to suggest spillover was an issue (see Table 8 in section *4.1.1. Intervention Reach)*.

4. Findings

4.1. Intervention Implementation Fidelity

Below we report on intervention reach, fidelity, dose, recruitment, and context.

4.1.1. Intervention Reach

Intervention reach was assessed during intervention delivery and at endline.

At endline, 85% of households surveyed from the intervention households reported that there had been an intervention of any kind in their village since April 2018, compared to just 3% of households in the control villages. Those who reported activities to have taken place, were then asked to report what the activities were without being offered response options. Respondents in the intervention communities who reported that activities to encourage latrine use took place in their village since April 2018 identified the Palla performance most often (94%), followed by the Transect Walk (46%), Community Meeting (42%) and the Mother's Group (35%) and Household Visits (35%). Only 14% of respondents reported to have seen the Wall Paintings, and 11% reported latrine repairs to have taken place. Among all respondents from intervention households, 80% reported the Palla, 39% reported the Transect Walk, 36% reported the Community Meeting, 30% reported the Mother's Group, 12 reported the Wall Painting, and 9% reported latrine repairs.

Table 8: Endline Participant Reporting of Intervention Activities
(N=5780 Households, including those without latrines)

	Intervention (n=2820)		Control (n=2960)	
	number	%	number	%
Have activities to encourage latrine use taken place in village since April 2018?				
Yes	2395	85%	81	3%
No	425	15%	2879	97%
Activities reported to have taken place (by those who reported that activities did occur) *(list not read to participant)*				
	(n=2395)		(n=81)	
Palla performance	2252	94%	29	36%
Transect walk	1106	46%	3	4%
Community Meeting	1007	42%	27	33%
Household visits / poster	843	35%	2	2%
Wall painting	337	14%	0	0%
Mother's Group	840	35%	1	1%
Latrine repairs / latrine assessments	267	11%	0	0%
Visit from government official	5	0%	12	15%
Media campaign (TV, newspaper, radio)	1	0%	1	1%
Other	11	0%	9	11%

Of the 81 respondents in the control villages who said a latrine use intervention activity took place since April 2018 (3% of respondents), the most commonly reported activities included (without prompting) the Palla (36%), Community Meetings (33%), and visits from government officials (15%) to have taken place. Among *all* respondents from households in the control villages, only 1% reported a Palla to have taken place, 1% reported a community meeting to have taken place, and <0.5% reported a visit from a government official. It is likely that the 1% of participants in the control community that reported Palla and Community Meeting activities heard about these events occurring in intervention villages.

From data collected during intervention delivery, we gave each village Reach scores based on estimated attendance (see section 2.3 for details). The Reach score for the Mother's Group was the highest (8.70 – i.e. 87% of targeted participants attended), distantly followed by the Palla (2.82), Community Meeting (1.27), and Transect Walk (1.03) (See Table 9).

As reported in Table 10, at endline, 66% of survey respondents reported having heard the intervention motto. The most commonly attended intervention activity by at least one member of intervention households was the Palla (66%), followed by the Community Meetings (39%), Transect Walk (27%), and Mother's Group Meeting (26%). However, Mother's Group attendance was 96% among households with children under age 5 (. Sixty-four percent of respondents indicated that community mobilizer conducted a household visit in their household, and only 13% indicated that they had seen the wall painting. Of those that saw the wall painting, 23% were male and 10% were female; we expect that restrictions on women's mobility likely explains why a greater proportion of men saw the wall painting.

At endline we asked all households that were selected to receive repairs whether or not they had any latrine repairs completed. Of the 403 intervention households selected to receive repairs (19% of the latrine owning households at baseline), 272 (67.5%) reported to have received the repairs, which is equivalent to 13% of all latrine owning households in the intervention villages.

Table 9: Reach and Fidelity/Dose Scores of Community Activities from Observations of Each Intervention Activity across all 36 villages

	Reach Score*		Fidelity/Dose Score			
	Mean	SD	Max Possible	Mean	SD	%
Palla	2.82	0.94	14.00	10.68	1.91	76%
Transect Walk	1.03	0.30	11.00	8.25	1.54	75%
Community Meeting	1.27	0.45	20.00	13.89	1.34	69%
Mother's Group	8.70	2.82	16.00	12.97	1.52	81%

*Max reach score is 10 for all activities.

Table 10: Reach of Community Activities as Reported by Respondents from Intervention Communities in the Endline Survey (N=2820 Households, including those without latrines)

	N	%
Recall Hearing Intervention Motto	1860	66%
Someone in Household attended:		
Palla	1853	66%
Transect Walk	775	27%
Community Meeting	1088	39%
Mother's Group (among all households)	728	26%
Mother's Group (among households with children under age 5)	728	96%
Household Visit conducted		
Yes	1814	64%
Refused Visit	21	1%
Respondent has seen Wall Painting	358	13%

4.1.2. Intervention Fidelity/Dose

Intervention Fidelity/Dose was assessed during intervention delivery by Emory Process Evaluation team members. Fidelity/Dose scores are presented in Table 9.

For the Palla, the mean Fidelity/Dose Score was 10.68 out of 14, indicating that 76% of all components were delivered as planned. Across the villages, 35% of pallas left out messaging about the importance of male latrine use, 35% did not discuss the need for men to be role models in their household, 42% forgot messaging about safely disposing of child feces into a latrine as a means of keeping children healthy, 50% did not mention that Odisha has one of the highest rates of open defecation, 62% did not indicate that work time would not be lost to sickness if people use latrines, and 65% did not discuss that money can be saved from illness-related costs if all use the latrine.

We engaged two different troupes to perform the Pallas. We explored whether or not their faithfulness to the script / messages varied. Specifically, we ran a t-test on the overall scores and found a significant difference in the scores for the troops (95% CI -3.763 to -1.581; p-value <0.001) . The mean score of Troupe 1 was 9.38 and the mean score of Troupe 2 was 12.053. It seems Troupe 1 most often missed messaging comparing the perceived benefits of open defecation compared to the benefits of latrine use, as well as messaging about pit emptying. Overall, Troupe 2 performed almost every Palla message every time while Troupe 1 varied greatly.

For the Transect Walk, the mean Fidelity/Dose Score was 8.25 out of 11, indicating that 75% of all components were delivered as planned. In 31% of villages, community mobilizers did not close the meeting with a statement of encouragement about the community's ability to become defecation free, like other villages in Odisha.

For the Community Meeting, there were supposed to be two community meetings per village, one for women and one for men, to create environments that enabled female attendance and participation. Due to resource constraints (time, funds allowable for intervention activities) in all villages there was only one meeting, open to both men and women. In total, there were 1245 participants in community meetings, including 598 women (48%), 403 men (32%), 118 boys under

age 18 (9%), 134 girls under age (18 11%). There was an average of 34 attendees per meeting (11 men, 16 women, 3 boys, 4 girls).

The mean Fidelity/Dose Score for the community meeting was 13.89 out of 20, indicating that 69% of all components were delivered as planned. There were key behavior change technique (BCT) elements that were not completed in a majority of community meetings. Specifically, while nearly all (97%) of meetings discussed challenges to latrine use, only 53% followed with a discussion of solutions to the challenges, potentially having a detrimental effect. Additionally, an activity intended to inspire the village to be a 'model' village had numerous components missing.

For the Mother's Group Meeting, the mean Fidelity/Dose Score was 12.97 out of 16, indicating that 81% of all components were delivered as planned. Observation revealed that there was a high rate of participation in discussions, however community mobilizers did not engage participants well in opportunities to practice using the hardware, and only 75% of meetings included a planned action planning segment. Finally, messages about how to handle child feces if households did not own a latrine were missed. Notably, only 78% of meetings discussed the importance of burying water used to wash materials/hardware tainted with child's feces if the household does not own a latrine and only 61% discussed the need to bury child feces if household did not have a latrine.

A desk review of household visit log sheets revealed that the three components of the visit were carried out in nearly all engaged households. All wall paintings had the three necessary elements: village map with houses, clearly identified positive deviant households, and the action steps decided upon by the village during the community meetings.

For latrine repairs, we asked households that reported having repairs done, what they had repaired. Doors were the most common repair (57%), followed by fixes to flooring (21%), pit lining (19%), slab covers (15%), and pipe connections (14%) (Table 11).

Table 11: Types of Repairs Reported at Endline (n=272 repairs)

Repair Type	Number of Repairs	% Repaired Latrines
Door	155	57%
Roof	0	0%
Walls	2	1%
Slab cover	40	15%
Pan	13	5%
Pipe connection	38	14%
Pit lining	52	19%
Flooring	57	21%
Parapet	1	0%
Other	8	3%

*Multiple repairs possible per household

4.1.3. Intervention Recruitment and Context

The qualitative research in the non-trial villages and at endline, and the post-endline community meetings revealed issues with attending the activities, specifically for the community level activities, and about the potential for ration cards to be taken away if people were caught practicing open defection.

For the Palla, Transect Walk, and Community Meetings, some participants indicated that they did not know about the activities. Those that did not know about the activities typically indicated that community mobilizers did not go to 'their side' of the village to tell them about the activities. In these cases, there is a clear issue of recruitment and community mobilizers should have made sure that all parts of the village were aware of the activities.

In other instances, recruitment efforts were made, but village dynamics impeded participation. Specifically, some participants indicated that they were not able to attend activities. Whether because of caste or 'tension', divisions in villages existed. If an activity was in a location associated with one 'side' of the village, non-occupants felt they could not attend. When we piloted in early 2017, we also had an incident where half of a village did not attend because of reported 'tensions'. Anticipating social dynamics within the villages themselves, community mobilizers were instructed to make visits to the villages and meet with various stakeholders prior to starting activities to identify if duplicate activities should occur in different parts of the village. In four villages additional Palla performances were scheduled to accommodate different parts of the villages. We did not follow-up or ask what the specific dynamics of what some of these 'tensions' were.

Women also reported barriers to attending activities, primarily related to social context. Some women indicated that community level activities, like the Palla, were scheduled when they had household work to complete, which prevented their participation. Other women indicated that they could not attend activities where their father-in-law or brother-in-laws were in attendance. Anticipating that women may not attend or participate community meetings, the intervention was designed to have sex-segregated community meetings to enable attendance, however only one meetings was completed per village due to constraints on time and funds.

Surprisingly, women also faced barriers to attendance to the Mother's Group Meeting. The Mother's Group Meetings were designed for mothers and caregivers of children under age 5. Often, the women in attendance were mothers-in-law. While we did not collect data on the roles of the women in attendance (i.e. mother, mother-in-law) we found that 36% of attendees were older than age 45, which we assume (with limitations) were mothers-in-law. Families would not permit their daughters-in-law to leave the house. When interviewed, daughters-in-law who had a mother-in-law attend shared the hardware provided with their daughters-in-law, but often failed to pass on any of the relevant information.

Finally, both the qualitative data from non-trial villages that received the intervention and notes from the post-trial sharing meetings revealed that people heard rumors that ration cards could be taken if people were caught practiving open defecation. In the two non-trial qualitative villages, ration cards were discussed differently. In one, participants reported that someone was seen practiving open defecation during the transect walk and a male told her she could have her ration card taken away. It remains unclear who the male was. In another village, taking ration cards was brought up by participants as a potential strategy to enforce latrine use in their community, though it is unclear whether they came up with this idea or had heard it being a tactic used elsewhere. During the post-eneline sharing meetings, participants from four intervention villages also noted that they heard ration cards could be canceled if people were caught open defecating. In one

village a participant noted tha this rumor seemed to have stopped people from practicing open defecation for a few days. Overall, as we did not collect systematic data on this topic, we do not have a sense of how wide spread these rumors about ration cards being canceled were, by whom the messages were delivered, when (before, after, or even during intervention activities) these messages were delivered and heard, who heard these messages, whether or not they were believed, or the extent they may have influenced behavior. Still, this information is important contextual information to share and could be investigated further.

4.1.4. Intervention Satisfaction

Participant satisfaction varied. In reference to the intervention overall, several participants in qualitative interviews and discussions, and those who commented during the community meetings, indicated that the repetition of messages through the various activities was important, and many suggested repeating some of the activities, particularly the Palla.

Overall, participants enjoyed the Palla activity, commenting on how entertaining and funny it was, though some shared that they did not think it provided them with any additional information that they did not know before.

Participants thought the transect walk was memorable, though some noted that it induced sentiments of shame, disgust, and even fear.

Perceptions of the community meetings were variable. Some village members were positive and indicated that the meeting inspired their community to clean up parts of their village. Others indicated that their village would never be able to change or work together because of pre-existing tensions.

In regard to the Mother's Group, participants and those who received materials from other members in their household who attended had varied opinions of the supplies. Several women discussed problems with the potties: they could only be used for kids of a certain age, kids fought over them, kids could not defecate well on them in a sitting position, they were hard to clean, their kids did not like them, they broke, and their children fought over them. The potties seemed to only have worked for some households. Scoops were described as more useful as they could be used for picking up feces when children defecated in the open and some women used them to have children defecate directly on. Regardless of methods use, women discussed the time consuming nature of cleaning the hardware, raising questions about whether or not this was done as expected. Some women used paper to cover the potty and scoop, but then just tossed feces-soiled paper in their trash piles.

Households that reported to have received repairs to their latrine reported their satisfaction with repairs via the endline survey. Of the 113 respondents who provided answers, 75% indicated they were satisfied or completely satisfied (Table 12).

Table 12: Satisfaction with Latrine Repairs (n=113 respondents)

	n	%
Completely satisfied	36	32%
Satisfied	49	43%
Somewhat satisfied	23	20%
Not at all satisfied	5	4%

4.2 Impact analysis

4.2.1. Descriptive Statistics

Trial Participants. The trial study population includes 36 gram panchayats and 66 villages across intervention and control communities. At baseline, 5864 households and 20370 individuals were represented and 85.7% of primary respondents were female (Table 13). Among those with at least one latrine at baseline, the predominant religion was Hindu (95.5%) while only 2.9% of households identified as Muslim (Table 14). The majority of households reported their caste/tribe as 'Other' (43.3%) since most households report their surname as their caste and do not identify with one of the government categories. As such, we marked these responses as "Other" but also recorded the surname provided. These surnames were then re-categorized into one of the government categories based on official government documents[c]. After re-categorizing, 41.4% of households were identified as Other Backward Caste, 35.7% General Caste, 12.1% Scheduled Caste, 1.0% Scheduled Tribe, 8.1% Other, and 1.6% not applicable. Three quarters (75.1%) of households received some sort of government funding, be it BPL, Antodaya, and/or ration card. For education, 38.9% of male head of households completed secondary education or higher with 10.6% having never attended school. In contrast, 20.6% of female head of households completed secondary education or higher with 29.7% having never attended school.

Table 13. Characteristics of the Full Trial Sample (Census Population)

Census population (all households surveyed)

	Intervention	Control	Total
Villages	33	33	66
Households censused	2846	3018	5864
Population censused	9922	10448	20370
Households per village	86.2 (25.4)	91.5 (24.6)	
Population censused per village [Mean (SD)]	300.7 (102.8)	316.6 (104.9)	
Population censused per study arm (%)			
Female†	4482 (49.8)	4770 (50.4)	9252
Male†	4514 (50.2)	4687 (49.6)	9201
Under 5 years	531 (5.9)	543 (5.7)	1074
Sex, primary respondent (n = 5864) (%)			
Female	2439 (85.7)	2581 (85.5)	5020
Male	407 (14.3)	437 (14.5)	844
Caste / tribe of household (n = 5864) (%)			
Brahmin	293 (10.3)	180 (6.0)	473
General	675 (23.7)	713 (23.6)	1388
Scheduled caste	254 (8.9)	235 (7.8)	489
Other backward caste	419 (14.7)	508 (16.8)	927
Scheduled tribe	28 (1.0)	12 (0.4)	40
Other	1174 (41.3)	1366 (45.3)	2540
Don't know	3 (0.1)	4 (0.1)	7

[c] We referred to documentation provided by India's Ministry of Social Justice and Empowerment to identify caste categories: http://socialjustice.nic.in/UserView/index?mid=76750.

Table 14. Characteristics of the Baseline Sample (Eligible Households with at Least 1 Latrine) (3978 Households)

Baseline study population (eligible households with at least 1 latrine)

	Intervention	Control	Total
Households eligible	1928	2050	3978
Religion of household (n = 3977) (%)			
Hindu	1854 (96.2)	1944 (94.9)	3798
Hindu / Muslim	2 (0.1)	0 (0)	2
Hindu / other	0 (0)	2 (0.1)	2
Muslim	42 (2.2)	72 (3.5)	114
Christian	1 (0.1)	0 (0)	1
Buddhist / Neo-Buddhist	1 (0.1)	1 (<0.1)	2
No religion	6 (0.3)	9 (0.4)	15
Other	22 (1.1)	21 (1.0)	43
Government subsidies (n = 3977) (%)			
BPL	143 (7.4)	181 (8.8)	324
BPL / Antodaya	10 (0.5)	11 (0.5)	21
BPL / Antodaya / Ration card	60 (3.1)	56 (2.7)	116
BPL / Ration card	319 (16.5)	299 (14.6)	618
Antodaya	121 (6.3)	108 (5.3)	229
Antodaya / Ration card	29 (1.5)	38 (1.9)	67
Ration card	753 (39.1)	857 (41.8)	1610
None	482 (25)	489 (23.9)	971
Don't know	11 (0.6)	10 (0.5)	21
Education of male head of household (n = 3820)			
Anganwadi	76 (4.1)	66 (3.4)	142
Primary	426 (23.0)	446 (22.7)	872
Upper primary	344 (18.6)	360 (18.3)	704
Secondary	527 (28.4)	570 (29.0)	1097
Senior secondary	77 (4.2)	84 (4.3)	161
Graduate / post-graduate	112 (6.0)	116 (5.9)	228
Never attended	184 (9.9)	219 (11.1)	403
Don't know	108 (5.8)	105 (5.3)	213
Education of female head of household (n = 3925)			
Anganwadi	60 (3.2)	65 (3.2)	125
Primary	560 (29.4)	562 (27.8)	1122
Upper primary	306 (16.1)	330 (16.3)	636
Secondary	301 (15.8)	350 (17.3)	651
Senior secondary	40 (2.1)	51 (2.5)	91
Graduate / post-graduate	40 (2.1)	25 (1.2)	65
Never attended	561 (29.5)	605 (29.9)	1166
Don't know	35 (1.8)	34 (1.7)	69

†All ages

On average, at baseline 63.7% of households in intervention villages and 67.3% of households in control villages reported owning at least one latrine. The majority of latrines were flush to pit (73.1%), fully constructed (87%), did not require repairs (68.6%), and were reported to be used for defecation (75.9%) (Appendix 11). In the baseline study sample, 77.3% of households reported they had received money and/or materials to construct their latrine, 65.1% of households received assistance from Swachh Bharat Mission (SBM), and 25.5% reported their latrine construction was entirely self-financed. Only 4.7% of households reported their latrine pit had ever filled. Among these, 73.5% reported the pit was then emptied. Based on surveyor observation, 41.1% of households had a handwashing station in or near their latrine and 68.7% had a water source within 30 feet of their latrine (See Appendix 11).

Sub-Study Participants. One hundred and fifty-two adults participated in focus groups to gain perceptions of the intervention (in intervention communities) and spillover (in control communities). Approximately half (47%) were women, 43% were from control villages, 66% reported to own a latrine, and of those 87% reported to use the latrine at all times (See Appendix 12). Sixty-one women participated in in-depth interviews (n=24, all from intervention villages) and focus group discussions (n=37, all from control villages) to understand their perceptions of the mother's group activity (in intervention communities) and spillover (in control communities). Fifty-nine per cent reported owning a latrine, and of those 42% reported to use the latrine at all times (See Appendix 12).

4.2.2. Balance Tables

Table 15 shows the balance between control and intervention arms at baseline. Latrine use (C: 61.7%; I: 60.4%) and safe child feces disposal (C: 3.4%; I: 6.0%) were nearly identical. Half of the population in the sample is female, the mean age is 20, mean household size is 4.7, and most people are Other Backward Cate followed by General Caste (See Table 15). Household water and sanitation-related characteristics of the full study population at baseline, including latrine type, latrine funding source, and information about pit emptying can be found in Appendix 11.

4.2.3. Research Analyses

Primary Outcome

Our primary research question for this evaluation was: Is latrine use among people who own a latrine in communities that received the intervention significantly different at endline than among people who own a latrine in communities that did not receive the intervention?

We hypothesized that latrine use and safe disposal of child feces among people in households that own a latrine in villages that received the intervention would be significantly higher compared to latrine owning households in control villages.

Our intention to treat (ITT) analysis revealed an increase in reported latrine use among individuals age 5 and older of 6.4% (95% CI 2.0%-10.7%, p=0.004) in the intervention group at endline after accounting for the increase in latrine use observed in the control group. (See Table 16).

There was an increase in latrine use in both intervention and control villages at endline compared to baseline. Latrine use increased in control villages by 14% (from 62% to 75%) and in intervention villages by 20% (from 60% to 80%). Proportion of increase among men and women were comparable across both arms. Men had slightly higher increase in both control (14%) and intervention villages (21%) compared to women (C: 13%, I: 20%), but more women (C: 80%, I: 84%) than men (C: 70%, I: 77%) were reported to use latrines in either arm at endline (See Figure 5). Proportion of use increase ranged from 17%-22% in the intervention villages and from 11%-

16% in the control villages across age categories. The smallest increase in the intervention villages (17%) was among 20-29 and 30-39 year olds, who had the greatest proportion of use at baseline. The greatest increase in the intervention villages (23%) was among 50-59 year olds, who had a low proportion of use at baseline (53%). Those age 60 and over had the lowest proportion of use in intervention villages, and had a moderate (20%) increase in use. This age set continues to have the lowest proportion of latrine users overall (See Figure 6).

Figure 5: Proportion of Individuals Over Age 5 from Latrine-Owning Households Who Used a Latrine at Last Defecation Event, by Sex

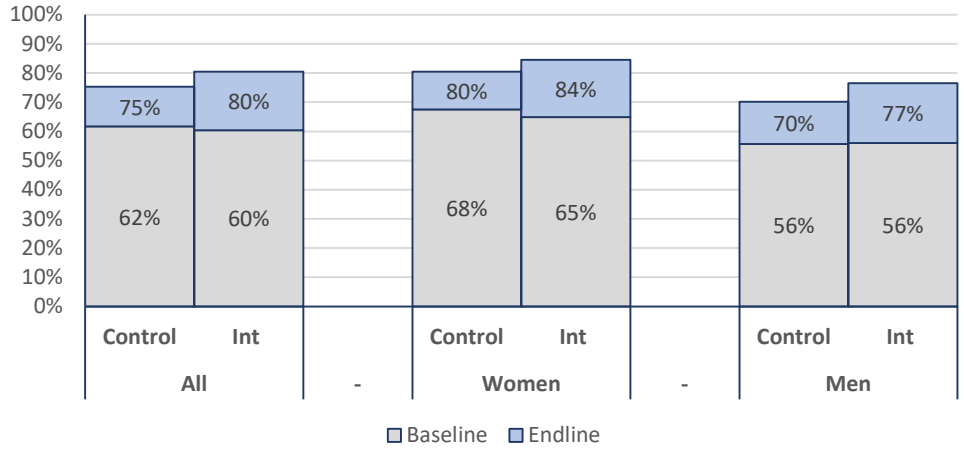


Figure 6: Proportion of Individuals Over Age 5 from Latrine-Owning Households Who Used a Latrine at Last Defecation Event, by Age Category

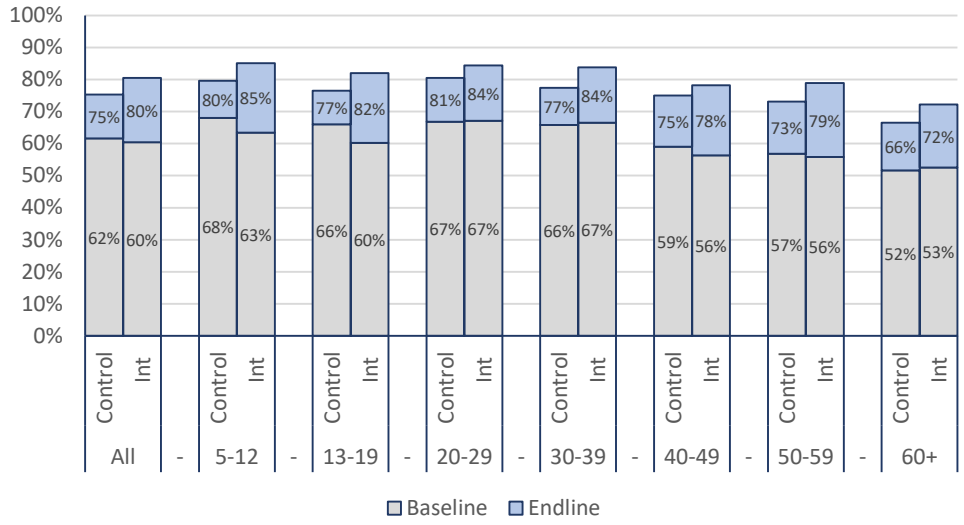


Table 15: Baseline Balance: Individual (N=13,812) and Household (N=3305) Characteristics

	Control		Intervention		
	Mean or n	SD or %	Mean or n	SD or %	Std. Dif.
Individual					
Latrine Use					
Yes	4231	61.7	3954	60.4	0.03
No	2631	38.3	2590	39.6	
Safe Child (u5) Feces Disposal					
Yes	7	3.4	12	6.0	0.13
No	200	96.6	187	94.0	
Female	3477	50.7	3263	49.9	0.01
Male	3385	49.3	3281	50.1	
Age	36.1	20.4	36.2	20.38	-0.01
Age Category					
0-4	207	2.9	199	3.0	0.07
5-12	726	10.3	719	10.7	
13-19	844	11.9	699	10.4	
20-29	1216	17.2	1171	17.4	
30-39	992	14.0	1061	15.7	
40-49	1043	14.8	983	14.6	
50-59	901	12.7	835	12.4	
60+	1140	16.1	1076	16.0	
Household					
Household size	4.7	2.0	4.7	1.9	
Caste					
General Caste	554	33.1	598	38.4	0.18
Scheduled Caste	182	10.9	207	13.3	
Other Backward Caste	764	45.7	589	37.8	
Scheduled Tribe	10	0.6	19	1.2	
Not Applicable	28	1.7	21	0.3	
Other	134	8.0	124	8.0	
SES Quintile					
Quintile 1	323	19.0	296	18.5	0.08
Quintile 2	363	21.3	300	18.7	
Quintile 3	345	20.2	316	19.7	
Quintile 4	344	20.2	343	21.4	
Quintile 5	329	19.3	346	21.6	
Education of Male HH Head (grade)	7.5	4.1	7.5	4.2	0.00
Education of Female HH Head	5.2	4.2	5.2	4.2	0.00

Table 16: Effect of Intervention on Latrine Use
(adjusted for clustering, intention-to-treat analysis)

	Effect Size	SE	95% CI%			*P*
Difference in Difference	0.06	0.02	0.02	-	0.11	0.00
Intervention Arm	-0.01	0.03	-0.08	-	0.05	0.67
Baseline Latrine Use	0.14	0.01	0.11	-	0.16	0.00
Number of Household Members	-0.01	0.00	-0.01	-	0.00	0.03
SES						
Quintile 1	*Ref*					
Quintile 2	0.09	0.02	0.04	-	0.13	0.00
Quintile 3	0.15	0.02	0.11	-	0.20	0.00
Quintile 4	0.24	0.02	0.20	-	0.28	0.00
Quintile 5	0.33	0.02	0.28	-	0.37	0.00
Sex						
Male	-0.09	0.01	-0.11	-	-0.08	0.00
Female	*Ref*					
Education of Male HH Head (grade)	0.01	0.00	0.01	-	0.01	0.00
Education of Female HH Head	0.00	0.00	0.00	-	0.01	0.00
Age Category						
5-12	0.04	0.01	0.02	-	0.06	0.00
13-19	0.04	0.01	0.01	-	0.06	0.00
20-29	0.04	0.01	0.02	-	0.06	0.00
30-39	*Ref*					
40-49	-0.03	0.01	-0.05	-	0.00	0.03
50-59	-0.05	0.01	-0.07	-	-0.02	0.00
60+	-0.09	0.01	-0.11	-	-0.07	0.00
Intercept	0.45	0.04	0.37	-	0.52	0.00

Our intention to treat (ITT) analysis revealed an increase in reported safe disposal of child feces of 20.4% (95% CI 11.7%-29.2%, p<.001) in the intervention group at endline after accounting for the increase in safe disposal of child feces observed in the control group (See Table 17).

Table 17: Effect of Intervention on Safe Disposal of Child Feces
(adjusted for clustering, intention-to-treat analysis)

	Effect Size	SE	95% CI%	*P*
Difference in Difference	0.20	0.04	0.12 - 0.29	0.00
Intervention Arm	0.03	0.02	-0.02 - 0.07	0.23
Baseline Safe Disposal	0.07	0.02	0.04 - 0.11	0.00
Number of Household Members	-0.01	0.01	-0.02 - 0.00	0.05
SES				
Quintile 1	*Ref*			
Quintile 2	0.07	0.04	-0.01 - 0.15	0.67
Quintile 3	0.07	0.04	0.00 - 0.14	0.54
Quintile 4	0.12	0.04	0.05 - 0.19	0.00
Quintile 5	0.17	0.04	0.09 - 0.25	0.00
Sex				
Male	-0.01	0.02	-0.05 - 0.04	0.73
Female	*Ref*			
Education of Male HH Head (grade)	0.00	0.00	0.00 - 0.01	0.34
Education of Female HH Head	-0.01	0.00	-0.01 - 0.00	0.11
Intercept	0.01	0.04	-0.07 - 0.09	0.76

Secondary Outcomes:

1. Latrine Construction:
We hypothesized that latrine construction among non-latrine owners in villages that received the intervention would be significantly higher compared to those residing in control villages.

We found no difference in the proportion of households that did not have a latrine at baseline and had one at endline (6.77% in control; 6.96% in intervention; p=.903).

2. Behavioral Determinants:
We hypothesized that latrine use behavioral determinant scores would be significantly higher at endline among latrine owners in intervention villages compared to latrine owners in control villages. We aimed to have 20 randomly selected respondents from latrine owning households in both intervention and control communities, with 10 being female and 10 being male at both waves (1320 at each wave, 660 per arm). A total of 1251 households completed behavioral determinants surveys from at both waves (C: 625, I: 626). However, of those, only 810 (65%) households had the same respondent at baseline and endline (C:398, I: 412).

We created scores from items representing each of the six behavioral determinants: ability (5 items), motivation (4 items), physical opportunity (2 items), risk perception (3 items), self regulation (5 items), and social norms (7 items). All items for ability, motivation, physical opportunity, risk perception, and social norms had four potential response options to indicate level of agreement: completely agree (1), mildly agree (2), mildly disagree (3), completely disagree (4). Self-regulation items also had four possible responses for each item. One item assessed level of agreement as previously described while the remaining four items assessed level of intention (1 item), level of commitment (1 item), and level of confidence (2 items). Items were re-coded as needed such that valence would be consistent. To create scores, we simply summed each item in the domain and divided by the number of items to arrive at scores that could range from 1 (optimal score, highest level of agreement) to 4 (least optimal score).

Consistent with our model approach, we fit a linear DID (difference in difference) model with robust standard errors and adjusting for age, sex, household size, SES, and education of male and female heads of household. Only 729 (58%) respondents had complete behavioral determinants and covariates.

Scores were similar at baseline and endline for all behavioral determinants. Most determinants (ability, motivation, physical opportunity, risk perceptions, and social norms) had scores in the 1.3 to 1.7 range indicating general agreement with the questions posed. Self-regulation scores ranged from 2.3 to 2.4, indicating responses were between mildly agree and mildly disagree (See Table 18). We found no increase in reported scores at endline as hypothesized (See Table 19).

We ran GEE models to test the association of each individual behavioral determinant on latrine use, adjusting for respondent age, sex, household size, and male and female head of household education attainment. Physical opportunity, ability, and social norms scores were all significant predictors of latrine use as hypothesized (favorable scores were associated with latrine use) (see Table 20). Scores for motivation, risk perception, and self-regulation were not.

Table 18:
Behavioral Determinants Scores at Baseline and Endline (n=729)

Determinant	Baseline				Endline			
	Intervention		Control		Intervention		Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Self Regulation	2.3	0.3	2.4	0.3	2.4	0.3	2.4	0.3
Physical Opportunity	1.7	1.0	1.7	1.0	1.5	0.9	1.5	0.9
Ability	1.7	0.5	1.7	0.5	1.6	0.5	1.6	0.5
Social Norms	1.6	0.5	1.7	0.5	1.5	0.4	1.5	0.4
Motivation	1.4	0.5	1.4	0.5	1.3	0.4	1.3	0.4
Risk Perception	1.7	0.7	1.9	0.8	1.7	0.7	1.7	0.8

Table 19:
Effect of Intervention on Behavioral Determinants Scores*

Difference in Difference	Effect Size	SE	CI	p value
Physical Opportunity	0.99	0.083	0.84-1.16	0.888
Ability	1.28	0.048	0.93-1.13	0.600
Social Norms	1.02	0.042	0.94-1.11	0.620
Motivation	1.00	0.042	0.92-1.09	0.969
Risk Perception	1.15	0.078	0.98-1.34	0.079
Self Regulation	1.01	0.035	0.94-1.08	0.788

*Models adjusted for age, sex, household size, SES, sex, education of male and female household heads.

Table 20:
Association of Behavioral Determinants on Latrine Use at Endline*

Difference in Difference	OR	SE	CI			p value
Physical Opportunity (N=899)	0.46	0.03	0.40	-	0.54	0.000
Ability (n=899)	0.70	0.10	0.53	-	0.92	0.010
Social Norms (n=899)	0.32	0.06	0.21	-	0.47	0.000
Motivation (n=899)	0.84	0.17	0.56	-	1.25	0.384
Risk Perception (n=899)	0.84	0.12	0.64	-	1.11	0.228
Self Regulation (n=901)	1.37	0.26	0.94	-	2.00	0.099

*Models for each behavioral determinant run independently and adjusted for age, sex, household size, SES, male and female head of household education.

4.2.4. Heterogeneity of Impacts

Additionally, while we were not powered to carry out sub-group analyses, we also present models stratified by gender.

Our intention to treat (ITT) analysis revealed an increase in reported latrine use among females age 5 and older of 6.6% (95% CI 2.2%-11.0%, p=0.003) in the intervention group at endline after accounting for the increase in latrine use observed in the control group. (See Table 21).

Our intention to treat (ITT) analysis revealed an increase in reported latrine use among males age 5 and older of 6.1% (95% CI 1.3%-10.8%, p=0.011) in the intervention group at endline after accounting for the increase in latrine use observed in the control group. (See Table 22).

Table 21: Effect of Intervention on Latrine Use among Females
(adjusted for clustering, intention-to-treat analysis)

	Effect Size	SE	95% CI%			*P*
Difference in Difference	0.06	0.02	0.02	-	0.11	0.00
Intervention Arm	-0.03	0.03	-0.09	-	0.04	0.42
Baseline Latrine Use	0.13	0.01	0.10	-	0.16	0.00
Number of Household Members	-0.01	0.00	-0.01	-	0.00	0.03
SES						
Quintile 1	*Ref*					
Quintile 2	0.09	0.03	0.04	-	0.14	0.00
Quintile 3	0.17	0.02	0.12	-	0.21	0.00
Quintile 4	0.24	0.02	0.20	-	0.29	0.00
Quintile 5	0.32	0.02	0.27	-	0.37	0.00
Education of Male HH Head	0.01	0.00	0.01	-	0.01	0.00
Education of Female HH Head	0.00	0.00	0.00	-	0.01	0.03
Age Category						
5-12	0.02	0.02	-0.01	-	0.05	0.16
13-19	0.06	0.01	0.03	-	0.09	0.00
20-29	0.06	0.01	0.04	-	0.08	0.00
30-39	*Ref*					
40-49	-0.01	0.01	-0.04	-	0.02	0.57
50-59	-0.05	0.02	-0.08	-	-0.02	0.00
60+	-0.08	0.02	-0.11	-	-0.05	0.00
Intercept	0.45	0.04	0.37	-	0.53	0.00

Table 22. Effect of Intervention on Latrine Use among Males
(adjusted for clustering, intention-to-treat analysis)

	Effect Size	SE	95% CI%			*P*
Difference in Difference	0.06	0.02	0.01	-	0.11	0.01
Intervention Arm	0.00	0.04	-0.07	-	0.07	0.92
Baseline Latrine Use	0.14	0.01	0.12	-	0.17	0.00
Number of Household Members	-0.01	0.00	-0.01	-	0.00	0.04
SES						
Quintile 1	*Ref*					
Quintile 2	0.08	0.03	0.03	-	0.13	0.00
Quintile 3	0.14	0.02	0.09	-	0.19	0.00
Quintile 4	0.24	0.02	0.20	-	0.27	0.00
Quintile 5	0.33	0.03	0.28	-	0.38	0.00
Education of Male HH Head	0.01	0.00	0.01	-	0.01	0.00
Education of Female HH Head	0.00	0.00	0.00	-	0.01	0.00
Age Category						
5-12	0.05	0.02	0.02	-	0.09	0.00
13-19	0.02	0.02	-0.02	-	0.05	0.37
20-29	0.03	0.02	0.00	-	0.06	0.07
30-39	*Ref*					
40-49	-0.05	0.02	-0.08	-	-0.02	0.00
50-59	-0.05	0.02	-0.08	-	-0.01	0.01
60+	-0.10	0.02	-0.13	-	-0.07	0.00
Intercept	0.35	0.04	0.27	-	0.43	0.00

5. Cost analysis

5.1. Cost of 'Sundara Grama' Intervention

The *Sundara Grama* intervention cost $42,065.29 to roll out across 36 intervention villages (33 for the trial, 3 for the sub-study) over a period of 8 months (1 month for training, 1.5 months for software activities, and another 5.5 months for latrine assessments and repairs) See Table 23 for a summary of intervention costs and Appendices 7-9 for detailed description of program administration, staff training, and implementation costs).

Table 23. Summary of Intervention Costs

Item	Budgetary Total (Rs.)	Budgetary Total (USD)	Unit	Unit cost (Rs.)	Description
Program administration	₹ 1,017,800.00	$14,540.00	-	-	20 RWI staff hired for 3 months, office rental for 3 months, office stationary for 3 months, office organizational for 3 months
Staff training	₹ 131,303.20	$1,875.76	-	-	20 RWI staff trained on all intervention activities during the course of 12 training days which included both in-house and field practice. All staff provided with print outs of intervention materials (manual, facilitator guides, logsheets) and organizational folder to safely keep materials.
Staff transportation	₹ 141,960.00	$2,028.00	-	-	20 RWI staff provided with transportation stipend to cover cost of travel to intervention villages during implementation.
Palla performances	₹ 422,233.36	$6,031.91	43 palla performances	₹ 9,819.38	43 palla performances implemented at a cost of 9803.26Rs/performance
Transect walks	₹ 7,630.00	$109.00	36 transect walks	₹ 211.94	36 transect walks implemented at a cost of 211.94Rs/walk for activity materials
Community meetings	₹ 1,146.68	$16.38	36 community meetings	₹ 31.85	36 community meetings implemented at a cost of 10.83Rs/meeting for activity materials
Mother's group meetings	₹ 232,754.00	$3,325.06	36 mother's group meetings	₹ 6,465.39	36 mother's group meetings implemented at a cost of 6382.33Rs/meeting for activity materials
Household Visits	₹ 159,359.00	$2,276.56	2189 posters	₹ 72.80	2189 household posters
Wall paintings	₹ 360,000.00	$5,142.86	36 wall paintings	₹ 10,000.00	36 wall paintings implemented at a cost of 10000Rs/painting
Latrine assessments	₹ 48,815.00	$697.36	815 latrines assessed	₹ 59.90	815 latrines assessed at a cost of 59.90Rs/latrine
Latrine repairs	₹ 421,569.00	$6,022.41	457 latrines repaired	₹ 922.47	457 latrines repaired at a cost of 922.47Rs/latrine
TOTAL	**₹ 2,944,570.24**	**$42,065.29**			

The intervention specifically targeted households that had a latrine before implementation (i.e. when asked at baseline survey). The 36 intervention villages consisted of 2173 households with at least 1 latrine. The cost of the intervention per *targeted* household was $19.36 (see Table 24). We also calculated the cost of the intervention per household *reached*. The 33 trial intervention villages included 2828 households. Among these households, 2520 (89%) reported attending at least 1 of the 6 software intervention activities. As such, the cost of the intervention per household *reached* was $16.69. This cost is slightly inflated as the cost of the intervention includes all 36 intervention villages but quantitative data on activity attendance was only collected in the 33 trial intervention villages (3 intervention villages only had qualitative data collection). In addition, the exposure to the intervention measure only accounts for the software activities and does not include the roughly 457 households that also received latrine repairs. With an 89% exposure to software activities, however, it is very likely these households are already accounted for in the exposure denominator.

Table 24. Cost of Intervention per Household Type

Total cost of intervention = 42,065.29 USD	**N**	**Cost per HH (USD)**
Intervention households	3205	13.12
Reached households (i.e. attended at least 1 activity)	2520	16.69
Targeted households (i.e. with latrine)	2173	19.36

5.2. Cost-Effectiveness of 'Sundara Grama' Intervention

We conducted a cost-effectiveness analysis of the *Sundara Grama* intervention. The analysis was restricted to only those households that had a latrine and provided self-reported latrine use behavior at both baseline and endline (N = 3723 HHs). First, we examined how different exposure levels to the intervention impacted household latrine use. During endline data collection, intervention households were read a description of a given software activity and then asked "Did you or other members in your household attend X?" The total number of activities attended was then tallied to determine an 'exposure level.' Households were categorized as having null/low exposure to the intervention if the household attended 0 or 1 activity (427 HHs, 23.7%), medium exposure if 2 or 3 activities (855, 47.4%), and high exposure if 4 to 6 activities (521, 28.9%); control households were automatically categorized as null exposure. The measure of effect was based on whether or not a household reported improved latrine use – that is, the household reported more members using the latrine at endline compared to baseline. In total, 894 intervention households improved in their latrine use and 770 control households improved in their latrine use. We then calculated the inter-quartile unadjusted odds ratios. Households with high exposure to the intervention had 1.71 greater odds (95% CI 1.32, 2.21) of improving their latrine use compared to households with null/low exposure and 1.28 greater odds (95% CI 1.03, 1.60) compared to households with medium exposure (see Table 25). As expected, we found even higher odds of improving household latrine use when comparing intervention households with high or medium levels of exposure to *control households* with null exposure to the intervention (high exposure vs. null control = OR 1.86; 95% CI 1.54, 2.28; medium exposure vs. null control = OR 1.46; 95% CI 1.24, 1.72).

The incremental cost-effectiveness ratio (ICER) of the *Sundara Grama* intervention was calculated based on a total implementation cost of $42,065.29 and dividing by an effect on 894 intervention households (i.e. improving latrine use). The *Sundara Grama* intervention has an ICER of 47.05 USD/household – that is, a cost of 47.05 USD to improve latrine use of one household that experiences some level of exposure to the intervention.

Table 25. Impact of Intervention Exposure on the Odds of Improved Household Latrine Use				
	Intervention Exposure	**Odds ratio**	**95% CI**	**p-value**
Intervention HHs only	medium vs. null/low	1.331	1.05, 1.68	0.0083
	high vs. medium	1.282	1.03, 1.60	0.013
	high vs. null/low	1.706	1.32, 2.21	<0.0001
Control HHs (null) vs. Intervention HHs of certain exposure level	null/low vs. null (control)	1.099	0.89, 1.36	0.1922
	medium vs. null (control) vs. medium	1.462	1.24, 1.72	<0.0001
	high vs. null (control)	1.875	1.54, 2.28	<0.0001

6. Discussion

6.1. Discussion Overview

We found a reported increase in latrine use among individuals age 5 and older of 6.4% (95% CI 2.0%-10.7%) in the intervention group at endline after accounting for the increase in latrine use observed in the control group. An increase in control communities is not surprising. Given the current push by the government to declare India open defecation free by 20 October 2019, a reported increase in use across the country is expected. Even though only 3% of respondents from control villages reported that activities took place *in their villages* in the preceding year, it is highly likely that other campaigns that were not 'in the village' influenced use across both arms, including radio, television, and newspaper campaigns as well as campaigns that targeted urban areas that rural residents visited. We expected changes to occur in both control and intervention villages and our study was designed to detect a difference in intervention arm villages despite increases in latrine use reported in both intervention and control villages. While not a large effect size, our findings demonstrate that the intervention did influence reported use. The intervention resulted in a greater increase in use than what would have occurred if the intervention did not take place. Villages in rural Odisha have been recipients of sanitation interventions for decades, with mixed results, and that this trial found the intervention evaluated to have an impact is promising.

A stratified analysis revealed that the increases in latrine use were comparable among females (6.6%, 95% CI 2.2%-11.0%, p=0.003) and males (6.1%, 95% CI 1.3%-10.8%, p=0.011). Our intervention aimed to increase latrine use by all and deliberately included messaging and activities that targeted both men and women. For example, in our formative research, we learned that not all household members, including both men and women, knew how to use their latrines or that defecating in the open could pollute the environment and harm health. In the palla, we deliberately included skits that had demonstrations of how to use the latrine and that humorously depicted how flies could be deadly transmitters of fecal pathogens when feces were in the environment. The palla also included messages specifically targeting men, encouraging them to be role models for other men in the community by using the latrine, and for women, acknowledging that open defecation is often a time to connect with other women but pointing out that finding other places and reasons to walk and talk would be less dirty.

In the spirit of increasing latrine use for all, our intervention also aimed to increase the safe disposal of child feces. We found a reported increase of safe disposal of child feces of 20.4% (95% CI 11.7%-29.2%, p<.001) in the intervention group at endline after accounting for the increase in safe child feces disposal observed in the control group. To our knowledge, this is the first evaluation of an intervention specifically designed to increase feces disposal. Many mothers indicated that they did not realize the importance of safely disposing child feces. We believe that simply enabling caregivers to understand the importance of disposing feces in the latrine contributed considerably to reported changes in behavior.

Despite impact, our robust, mixed-methods, process evaluation revealed that specific intervention components, including recruitment, reach, fidelity and satisfaction of each of the community activities, could be improved to potentially achieve greater impact. At endline we learned that 66% of households said they had one member attend the Palla and 80% reported without prompting that the palla took place in their community, indicating reach was strong, but could also be improved. Fidelity scores for this activity were also strong. Most importantly, participants

discussed enjoying the Palla performance, remembering several key stories that communicated health messages and the importance of a clean village, and wanting to see the Palla again.

People also recalled the Transect Walk, either because it illuminated their understanding of the condition of their village or because it caused them to feel shame and disgust at the filth. The qualitative substudy helped to reveal that this activity also has the potential to be harmful. When done early in the morning, people may be caught in the act of defecating, an experience which could be particularly detrimental to women (Routray et al., 2015, Caruso et al., 2017, Sahoo et al., 2015). Holding the walk at a different time could eliminate the risk of causing harm. Additionally, because the Transect Walk is inexpensive and does not have to be intrusive, it could be rolled out more frequently to allow for more frequent assessment of and reflection about, the village condition.

The community meetings and household visits were not delivered as originally designed because of time constraints. Further, not all of the behavioral techniques planned for the community meetings were delivered as planned. Still, some community members reported changes in their village, like cleaning parts for alternative uses, that resulted from the action planning in the meetings. Participants discussed how they did not know they could come together in such a way. Members of other villages commented that tensions are so high in their villages that it would not be likely for any collaboration. These mixed findings illuminate the importance of context. The success of an intervention strategy that emphasizes key motivators at the community level, like status and cleanliness and beauty, may not be possible in villages that do not have a sense of collective identity and pride. In villages that may be more receptive, additional intervention components that serve to strengthen collective efficacy may be useful (Delea et al., 2018).

We learned that the wall painting, intended as a reminder of the community action plan and to aspire to be a 'positive deviant' household, was not seen by many households, particularly women who have little mobility outside the house. If adapting the intervention, we would consider eliminating the map to divert the time and resources to strengthening reach, the content of the community meeting and household visits, and re-introducing sex-segregated community meetings to each village to enable greater access and participation of women.

Our intervention did not include a heavy focus on pit emptying. From our formative research, we did not find a concern for pit emptying to be an impediment to latrine use among those we engaged, which is contrary to what has been found in other Indian states (Coffey et al., 2017). In the present study, only 5% (189) of the baseline population reported having had a full pit, and *of those* 86% practiced healthy behaviors: 74% (138) emptied their pits and 12% (22) switched to or built a new pit. Only 13% (25) restricted members or stopped using altogether (see Appendix 11). We do acknowledge that pit emptying is an important consideration for the sustainability of latrine use, and do not dey that pit emptying may be a concern in the coming years as households continue to use their latrines. To this end, the palla performances did include some messages about pit emptying.

In regard to safe child feces disposal, the specific contribution of the hardware, specifically the potties, needs further investigation. From the qualitative research, we know that the potties were not universally used or liked, whether because they were not appropriate for all ages of children in the household or they broke and caused more challenges for mothers. While we managed to reach nearly 96% of eligible *households* through the Mother's Group activity, through our qualitative work we realized that we were not reaching *mothers of small children* directly. Many households sent other family members to the meetings and mothers reported not being allowed to attend. It is conceivable that the intervention could have greater impact if strategies were

employed to ensure greater participation of mothers or to modify the activity to have specific messages for grandmothers/mothers-in-law given they were common proxy attendees.

We found no difference in behavioral determinant scores among intervention and control villages at endline compared to baseline. It is possible that our intervention did not improve these behavioral determinants as expected. It is also possible that the measures we created to assess change could have been improved or that we did not have a sample large enough to detect a difference.

There was no difference in latrine construction at endline among non-latrine owning households in intervention villages compared to controls. This finding is not surprising as our intervention was not designed to motivate an increase latrine construction. Our finding that latrine coverage increased at a similar rate across study arms suggests that background sanitation efforts that we were not aware of may have been underway across the study area. Construction efforts can contribute to increases in latrine use. In an ad hoc analysis, Garn et al found a 5.8% increase in use for every 10% of latrine coverage (2017). The Research Institute for Compassionate Economics found that while a reduction in open defecation was observed between 2014 and 2018 in Bihar, Madhya Pradesh, Rajasthan, and Uttar Pradesh among rural households, the proportion of latrine owners practicing open defecation during that time period remained the same, suggesting that reductions in open defecation were likely driven by latrine coverage increases alone (Gupta, 2019). Our analysis only includes households that had latrines at both baseline and endline, thus the increase in latrine coverage in both arms does not explain the increases in use observed in our sample. Further analysis of our data is needed to understand what proportion of members in households that constructed latrines since baseline are using them.

6.2. Policy and Program Relevance: Evidence Uptake and Use

We engaged three key types of stakeholders: researchers (as well as donors), practitioners, and community members. Below we describe the strategies used to engage these stakeholders and the resultant impacts.

6.2.1. Research Stakeholders (researchers, donors and implementing partners)

We engaged researchers in the sanitation and public health fields by participating in a number of workshops and conferences. We presented at the Social Norms Workshop hosted by PENN Song, 3ie Delhi Evidence Week, Tech4Dev conference hosted by UNESCO, and presented multiple times at the 'Water and Health' conference hosted by University of Chapel Hill.

By engaging and presenting at these workshops we were able to shed light on the dearth of rigorous trials that examine changes in water, sanitation and hygiene (WASH) behaviors with *behavior* as the primary outcome. We emphasized the importance of investing in behavior change intervention design and evaluation and how the 3ie Thematic Window 14 is a much needed start. We believe such engagement helped continue the momentum for these types of behavioral impact evaluations and communicate to both researchers and donors alike the need for such investments and continued learning.

Our presentations also highlighted the current "black box" of intervention design and how the public health field needs to do a better job showcasing *how* interventions are designed, and what evidence and theory they are based upon (if used). We were able to offer a tangible example by showcasing our team's process of applying behavioral theory and conducting formative research to then design the *Sundara Grama* intervention in Phase 1 of this grant. We also shared our

intervention design process at the Social Norms Workshop and 3ie Delhi Evidence week, which included a diverse audience of researchers, donors and practitioners – many of whom approached us after the presentation to continue the dialog, demonstrating the interest varied stakeholders have in learning more about intervention development.

Finally, we will continue to engage research stakeholders through future publications and conferences. In this way we will continue to shape the sanitation and public health fields by offering lessons learned and future directions. For example, we plan to showcase both the impact evaluation findings as well as our rigorous process evaluation findings, emphasizing the importance of such data to work in concert to better inform policy and practice.

6.2.2. Community Stakeholders

The community results sharing meetings were a successful stakeholder engagement activity for community members. As described in section 3.3 Ethics, in each of the 33 trial intervention villages we held a community meeting where Emory research assistants (RAs) shared the overall impact results as well as the changes in latrine use and safe disposal of child feces behaviors specific to that particular village. A banner was used to share the findings and acted as an effective visual aid (see Appendix 10). The RAs walked through each of the numbers displayed on the banner and checked with community audience members that the information was clear. The facilitated discussion at the end of the meeting was also an effective way to create space for reflection and elicit feedback from community members on both their village-specific results and the intervention activities. Some key take-aways from the feedback discussions are noted below:

(1) Transect walks are an effective technique for causing re-evaluation of the environment and as such, reflections of shame and disgust. At many of the community meetings, villagers discussed the impact of the transect walk. While some villagers did not like the activity and felt it was disrespectful, most villagers thought the activity was effective and 'motivating.'
(2) Effective sanitation behavior change programs must involve continued engagement of communities. At some of the community meetings, villagers discussed how the repeated visits by mobilizers, as well as the survey team, caused them to continue to reflect on the open defecation in their village and feel a sense of shame. Villagers discussed that the repeated visits were motivating their fellow community members to change their behavior.
(3) Overall, the community meetings highlighted a mixed response as to whether or not the *Sundara Grama* intervention was effective at changing latrine use behavior.

While we are not able to follow-up, we believe the meetings will likely impact community members in the future. Local stakeholders such as ward members now have an understanding of their village's latrine use compared to neighboring villages, and as such may continue efforts towards improving their village's sanitation. Moreover, the reflection discussions provided important insights on the intervention activities and elucidated recommendations for future iterations of the intervention should it be implemented again. In many ways we found the debrief notes on these discussions to be just as rich, if not more so, than the data from the post-endline qualitative research.

6.2.3. NGO/Practitioner Stakeholders

We heavily engaged with the Orissa-based NGO Gram Vikas throughout our impact evaluation. Gram Vikas just entered their fifth decade of community development work and plan to develop various WASH-related behavior change interventions moving forward. As such, after Phase 1, we reached out to Gram Vikas and presented to Director Liby Johnson and other Gram Vikas staff

on our process for applying behavioral theory and formative research to intervention design and how we used this process to develop the *Sundara Grama* intervention. We then collaborated with Gram Vikas and 3ie team members on a Gates funded MEDS Learning Exchange. The goal of the learning exchange was to explore what worked and what did not work when it came to latrine use behavior change interventions in rural Odisha – comparing and examining the Emory *Sundara Grama* intervention and Gram Vikas *MANTRA*. This was another successful engagement with Gram Vikas as we continued to share lessons learned from our different experiences in intervention design.

We are now working with Gram Vikas on a future study that will use a similar intervention design process as the one we used for *Sundara Grama.* We will design a behavior change program that promotes safe child feces management (CFM) practices among households with small children in MANTRA villages. This study collaboration developed in part from our continued and successful stakeholder engagement activities with Gram Vikas over the past two years. Moreover, findings from our impact evaluation, especially findings on the mother's group meeting activity which focused on safe CFM practices, will be used to inform aspects of this future study with Gram Vikas.

6.3. Challenges and Lessons

Identifying villages for evaluation: While we invested our own resources in terms of time, capacity, and funds to identify appropriate villages with a rapid assessment before funding was awarded in June/July 2017, we needed to invest additional time from October-December to map all potential villages to ensure they fit eligibility criteria. Specifically, we needed to make sure that coverage and size were adequate to ensure that we were adequately powered. This consumed both time and resources. We will take this as a lessons learned and plan for a mapping phase in future studies.

Limited time frame for evaluation and intervention activities: The limited time frame for completing all activities posed challenges. First, it would have been ideal to collect baseline and endline data exactly one year apart. Mapping delayed baseline, so the window between data collection events was narrowed. In addition, due to the limited time frame, study households were engaged several times in a short period of time, especially intervention households, which could have led to participant fatigue. In fact, during recruitment of post-endline qualitative participants, some households expressed frustration and anger at being engaged again on the topic of sanitation. Intervention households in particular experienced multiple activities: a baseline survey, measurement team baseline survey, intervention activities, possibly qualitative activities, endline survey, measurement team endline survey, and possibly post-endline qualitative activities. A household could have been approached seven or more times in the course of 1.5 years about latrine use. We could not resolve this challenge and simply take it as an important reflection to consider in future studies. Ultimately, all the engagement activities were important, but a wider window of time for activities may have enabled greater success, particularly with post-endline qualitative activities.

Resource constraints for implementation: As described previously, we were not able to roll-out the household visit as originally intended due to resource constraints. Specifically, training and roll-out of the other intervention activities took longer than expected. As such there was limited time for the household visit activity and we could not employ RWI staff longer due to the 20 USD cap on the intervention cost. To resolve this challenge, we revised the household visit so that it consisted of only 3 key activities and took approximately 10 minutes to deliver (a very brief visit).

The lesson learned is that there is always a fine balance between the time and financial constraints of a program and the ideal program activities. When having to revise behavior change program activities, we needed to carefully consider which behavior change techniques should remain and which could be cut.

Enabling environment for latrine use behavior (i.e. functional latrines): We felt it was important to include latrine repairs as part of the intervention since we knew from past research that many government-built latrines are often in disrepair and not functional. In addition, based on behavioral theory, it is imperative that the enabling environment is in place when trying to change a behavior. Otherwise, the environment poses another barrier to the behavior. As such, we spent considerable time and resources on the latrine repairs. While this activity was successful, it pulled time and funds from the behavior change activities. However, it was also a policy-relevant component. As India continues to increase sanitation coverage and use, there will be a need to continue investment in sustaining both facilities and behavior to ensure that any progress towards an open defecation free India are maintained.

Seasonal variation: It is important to consider how seasonal variation can influence study activities. While we tried to avoid having any research activities, including data collection or intervention delivery, take place during extremely hot summer or monsoon seasons, the mapping exercise that we undertook at the outset of the study to identify eligible villages took longer than expected and pushed our timeline back. Intense heat and rains can both interrupt or slow study activities. Most importantly, data collection at baseline and endline should occur in similar seasons, if not exactly a year apart, so that variation in climate is not a factor influencing the outcome of interest.

Information asymmetry: Residents of the intervention villages likely received asymmetric information because of their ability and/or willingness to attend the intervention activities. This is true even for the mother's group meetings, which was the most well-attended. We know from the qualitative research as part of the process evaluation that mothers-in-law or other family members sometimes attended on behalf of mothers. We opened the meeting to caregivers to enable our reach to all households with children under age five. However, the qualitative research revealed that the mothers who had someone attend on their behalf may or may not have received any of the information or messaging about the importance of safe child feces disposal. In other words, the person who attended did not always tell mothers about what was learned at the meetings.

Hardware distribution strategies: It is important to be mindful of community members' expectations when distributing hardware to community members and to anticipate and plan for any potential challenges. As part of our intervention we distributed potties and scoops to facilitate child feces disposal. We planned for mothers and caregivers *who had attended the meeting that instructed those present on safe use and care* to receive the hardware after the meeting. In some instances, people heard about the hardware in advance and just wanted to come to take the items and leave or men in the community came just to collect hardware. Also, in some villages, anganwadi workers wanted some hardware for themselves. We had not anticipated these challenges and had to devise plans to ameliorate tensions with community members, if they resulted.

7. Conclusions and Recommendations

Our theory-based intervention increased latrine use and safe disposal of child feces in intervention villages compared to controls. Our process evaluations demonstrated that fidelity was strong, but can be improved. Efforts to reach participants, particularly women, can also be strengthened, potentially further increasing the impact of the intervention on both behaviors.

Moving forward, we offer the following recommendations for key stakeholders in the realm of sanitation and behavior change:

- Policymakers: We recommend that policymakers recognize that behavior change takes time and is 'a moving train.' That is, behavior change programs should seek to catalyze the *initial* adoption of a behavior but also the long-term *maintenance* of the behavior as well. We believe sanitation, and latrine use in particular, should be viewed as a long-term, continuous investment. As such, in addition to further investment to make sure those who have yet to change behavior and use the latrine for all defecation purposes, the government should invest in strategies to maintain the progress made thus far. Specifically, to sustain latrine use and make sure latrine users do not revert to open defecation, continued programming is needed to make sure that users still remain motivated and convinced by behavior change messages and that latrines themselves are useable. Latrines will always need to be repaired—whether due to expected wear or from unexpected events like cyclones—and people will need resources and support to fix them. Further, if used as expected, latrines will need to be emptied. In Puri, Odisha, single pit latrines were the most common. People in Puri will need support, whether guidance or resources, to empty latrines in the years to come if they are still to be used.

  These recommendation comes from our findings, which highlight the variability households and individuals experience in their history of practicing latrine use, namely how some consistently practice latrine use and others do not. As such, it is not enough to view behavior change programs as a "one-time" need. Our extensive investment in repairing latrines further supports this point. Latrine construction has been considered a one-time investment by the government. No mechanisms exist to support households to fix their latrines if they break, or to even get them to working order if built poorly.

  Program managers: We recommend that program managers are trained to understand the behavior change techniques being employed in each intervention activity. This recommendation comes from our experience training RWI staff and the findings from the process evaluation data, which show key behavior change techniques were not effectively employed. It is possible that RWI staff did not have a strong understanding of how the activities should operate. NGO program staff often have past work experience implementing behavior change programs. However, many behavior change programs focus solely on knowledge dissemination and general awareness campaigns. As such, it is imperative to adequately train program managers when behavior change interventions employ a more diverse and complex set of behavior change techniques so the program staff truly understand the goal of the activity and how it should operate. This will ensure behavior change interventions are implemented with fidelity.

  We also recommend that program managers, particularly those who are delivering interventions at a village or community level, invest time in understanding village dynamics that may influence program delivery. During the formative research that preceded this trial, we found that one villages had unexplained 'tensions' that prevented some members from

attending, and that another village prevented women from attending public events, like the palla. We had community mobilizers make an initial visit to meet stakeholders and identify any potential community dynamics that may influence the intervention delivery and who may be able to participate in activities. Despite this effort, more time meeting with stakeholders would have been beneficial. When we shared results back to community members, we learned that some villages had members who did not benefit from activities because they did not know they occurred or were not permitted to attend.

- Program participants: We recommend that program participants are actively engaged in the piloting, process evaluation and results sharing of behavior change programs as they are often the true experts in identifying why a program activity is successful or is unsuccessful. The formative research phase that preceded this trial, the qualitative research in the sub-study villages, and the post-endline qualitative research all enabled invaluable opportunities to collect insights from participants who provided critical reflections that improved our understanding of the program and how we would change future iterations of the program if given the opportunity.

- Researchers: We recommend that researchers conduct qualitative sub studies and a rigorous process evaluation in order to have a rich understanding of how a behavior change intervention was actually implemented (i.e. fidelity of the treatment) and to better understand the impact evaluation results. This type of data also offers rich findings on what aspects of the intervention were successful and what aspects were not successful/failed, which offers more fruitful findings for future researchers and practitioners alike.

- Donors: We recommend that donors invest more time and funding towards the intervention design process. The WASH field still does not have a good grasp on what behavior change techniques are effective at changing which WASH behaviors and which communication channels are best for delivering those techniques (i.e. community-level activities, group discussion, household visits, etc.). As such, providing adequate time and funding will better ensure that effective behavior change interventions are being designed and are subsequently worth evaluating through rigorous impact evaluations.

Appendices

Appendix 1: Pre-Analysis Plan

TW14 Pre-analysis plan: [TW14.1006 *Implementing and evaluating low-cost interventions to improve latrine use among rural households in Odisha, India*]

1. Intervention

1.1. Theoretical framework

Please describe the underlying behavioural theory - which will be used to guide your strategy for eliciting behavior change through your intervention.

Note: The behavior change theory provides the rationale for the hypotheses to be tested along the causal chain, in addition to the intervention and related activities. We assume that the concepts within the theory will become more context-specific during the course of the evaluation. In this way, describing the underlying theory in the pre-analysis plan will facilitate the interpretation of findings, which might vary from what was originally expected. Documenting changes or additions will better inform the underlying theory.

Based on our formative research and a thorough literature review, latrine use in rural Puri District of Odisha State is influenced by: Risk perceptions, ability to use a latrine, Social Norms, Motivations, Self-regulation of the behavior, and Physical Opportunity (i.e. the presence of a functional latrine in the household).

We expect that the various interventions we propose will impact these noted behavioral determinants and therefore encourage latrine use. Specifically, the intervention aims to make people aware of risks associated with open defecation and the health benefits of latrine use; instruct individuals as to how to use latrines and potties for safe child feces disposal; shift social norms around latrine use; motivate latrine use; encourage continued use; and provide facilities (latrine repairs) and hardware (potties, scoops) to enable use by all. The overall delivery of the intervention is informed by the Trans-theoretical model' (Prochaska, 2013) and is described in more depth in section 1.2.

A full depiction of our theory of change is in Appendix A at the end of this document, and a simplified theory of change is in Appendix B (Images could not be places within this template accurately, hence are at end).

1.2. Intervention summary

Please summarise your intervention.

a) In treatment villages, a local NGO, Rural Welfare Institute (RWI), will deliver an intervention package that includes activities at the community and household levels, with

additional activities and hardware for mothers of children under five. For households with non-functional latrines, RWI will hire masons to provide the appropriate repairs.

Community Level Activities:

At the community level, there are four activities.

Palla: The first activity is a traditional performance known as a '*palla*', which includes a series of songs and skits that aim to increase practical knowledge of latrine use, pit emptying, and safe feces disposal and perceived benefit of latrines, and decrease preference for OD. The palla, which will be performed by local troops hired, managed and trained by RWI, introduces and repeats a slogan in Oriya: 'our clean, healthy, beautiful community'. The palla targets the 'status' motive, encouraging villagers to think about their households and their village compared to others, to consider what others may see when they visit.

Transect Walk: The second activity is a 'transect walk', which aims to make community members aware of the true state of their village. Community mobilizers from RWI make a surprise visit to walk around and observe the village with members. They use holi powder to mark feces, which enables those not on the walk to observe its effects.

Community meetings: The third activity is a series of community meetings organized and led by RWI, one for women and one for men. During these meetings, participants come up with a vision for their community (like a goal) and decide on an action plan to achieve the vision. At this time, the participants are also asked to identify 'positive deviant' households that always use their latrines. These households are later given a banner by RWI to place in front of their house to publically recognize and praise their contribution to making the community cleaner, healthier and more beautiful.

Wall Painting: The fourth activity is a wall painting of a map of the village that identifies the positive deviant households. The painting aims to remind the village of their goal and motivate other households to use their latrines. RWI will hire and supervise local artisans to paint these murals.

Household Level Activities

Household visits: At the household level, an RWI mobilizer will make *household visits* to all households with latrines, providing targeted information, repeating key messages, and encouraging commitment to the community vision of latrine use by all members of the household at all times. These visits personally reinforce community messages and secure household pledges; they are important for those members who may or may not have been able to participate in other community-level activities.

Latrine Repairs: Households that need latrine repairs will receive repairs from masons hired by RWI so that a key barrier to use—unavailability of a functional latrine— is removed. We will include latrine observation measures in our census survey of eligible households so we can determine if repairs are needed and if so, what the exact

structural problems are for those non-functional latrines. In our experience from previous research in the area, many households had unfinished latrines. Organizations came in to the villages to build latrines and once completed enough, took a picture to send in and receive their payment. These pictures could easily hide unfinished aspects, like missing pipes. Households were then left with no recourse and no funds to make their latrines functional. We will develop a 'latrine repair eligibility criteria' as there may be some latrines beyond repair or that would cost too much too repair. From formative research, that included a census of 7 villages (N=1405 HHs), we expect the following to be common structural problems: missing slabs, missing pipes connecting pans and pits, and unconnected pans and pits (though pipe present). These repairs are well-within the capacity of masons. There may be other repairs needed that will be discovered via observation during the baseline census. These repairs will be considered based on mason capacity, cost, and time needed. After the baseline census of households with latrines, we can apply these criteria to identify the final list of eligible households for the study sample. Local masons will make repairs. By brining in masons to complete these latrines or make minor repairs to broken structures, we are simply providing households with the latrine they thought they were going to receive. We are not considering training masons at this stage. However, we will make it known to all villagers when the mason will be coming. We will provide very specific repairs, but villagers will be able to hire the masons when they are there to do additional work. They will also have their contact information should they need them in the future.

Mother Group Meetings

Finally, RWI will hold a targeted meeting open to all mothers/caregivers of children under age five to provide them with action knowledge and hardware (scoops, potties) for safe child feces disposal. Costs are not high: scoops cost 39 rupees each and potties cost 300 rupees each, though there may be even cheaper models on the market. Members of RWI will provide the necessary instruction for how to use this hardware and why it is important. We believe this aspect of the intervention is sustainable and scalable because we believe Anganwadi workers or members of women's groups could be trained to lead such sessions in the future if brought to scale. Additionally, once women start using potties and scoops, they could also train one another. In terms of hardware, there is an initial investment, but women could re-use the potties if they have another child or give them to another family if their children outgrow them. If the potties do break, the women can use the scoops. They can have the child squat over the scoop, instead of the potty, and simply use the scoop to dispose of the feces. If the scoop breaks, they can improvise by using other materials, like plastic or metal.

Overall process of intervention delivery

The 'processes of change' from the trans-theoretical model informed the order of intervention activity delivery (Prochaska, 2013). Prochaska et al note various stages of change, from pre-contemplation (has not considered change), to contemplation (intend to take action in future), preparation (some steps taken in appropriate direction of

change), action, and maintenance (sustained behavior). Based on our formative research, we recognize that when the intervention team first enters the villages, most dwellers will not have participated in any pro-latrine use behavior change activities and may not be thinking at all about latrine use. They may be in the 'pre-contemplation' phase. As such, the 'processes of change', also outlined by Prochaska et al, are useful strategies for moving people from moving through these stages of change towards healthy behavior.

Specifically, the processes are as follows:

- 'Consciousness raising': This involves introducing people to new ideas and facts that support the healthy behavior change.
    - The palla performance, as a community wide event, introduces (and for some, perhaps) re-introduces facts and ideas about latrine use and safe child feces disposal. The 'facts' and 'ideas' of focus are those identified in the formative research that pre-ceded this trial.
- 'Dramatic Relief': This entails an experience of negative emotions that may be associated with the unhealthy behavior, which is open defecation in this case.
    - Following the palla, RWI holds a surprise transect walk to observe the village. In so doing, RWI also walks through open defecation sites and deposits holi powder on feces to make them stand out in the environment. This activity is intended to make people aware of the filth of the community and the potential risks to health that such filth could cause.
- 'Self-reevaluation': This is the realization that the behavior is associated with identity.
- 'Environmental reevaluation': This is the realization that the behavior—whether good or bad-- has an impact on a person's physical and social environment.
- Self-liberation': This is the commitment to make the behavioral change.
    - The community meetings combine self-reevaluation, environmental reevaluation and self-liberation ideas. They start with 'consciousness raising', recognizing that all participants may not have been able to attend the palls. They then discuss how the behavior impacts all, and specifically the identity of the community as a whole and how it is viewed by outsiders (self-reevaluation). It takes time to reflect on the transect walk and to discuss with all what open defecation does to their shared physical environment, and it also acknowledges and celebrates positive deviants in the community who use their latrine all the time, establishing a person or people to be considered socially positive contributors (environmental re-evaluation). The meetings then have participants identify action steps they can take and ends with a commitment to these actions (self-liberation).
- 'Helping relationships': This involves the use of social support to enable the behavior change.
- 'Counter conditioning': This involves the uptake of the alternative healthier behavior.
- 'Stimulus control': This is the addition of cues to engage the behavior.

    - While the palla performance and the community meetings do touch on 'helping relationships', 'counter conditioning' and 'stimulus control', the visits made to each household focus on them. The RWI mobilizer holds a meeting with each family as a group, focusing on promoting commitment of use among all household members (helping relationships), seeks commitment to establish the new latrine use behavior (self-liberation and counter-conditioning), and provides a banner to help people remember to use the latrine (stimulus control).
- 'Reinforcement management': This entails providing rewards or the positive behavior change.
    - This typically may come a step earlier, but in this intervention, we engage this with a wall painting in the village. This paining is a map of the village. All households that have been identified as using the latrine all the time are painted in a special color. Those that later become recognized as having all members use the latrine can then have the mural changed to reflect this as a reward.
- 'Social liberation': This is the realization that the social norms are changing to promote the behavior.
    - We recognize that great change may not happen in a short period of time and social norms may not be shifted by the close of the intervention activities, but RWI engages this process of change at the end the intervention by having a closing ceremony that reviews all that has been done in the village, celebrating the engagement of households throughout all the activities, and by acknowledging that they are all taking positive steps to making their village cleaner and healthier.

**Note: Those engaged in the mothers group meeting will go through the same activities as other members of the community, ideally. The primary difference is that they will also participate in a group meeting that will provide them with information about safe feces disposal, hardware to facilitate this behavior, and skills for practicing it effectively. Several of the 'processes' of change noted above will be re-engaged in these specific meetings.

2. Evaluation Questions and Hypotheses

2.1. What are the main evaluation question(s) the study seeks to answer?

Main research question: Is latrine use among people who own a latrine in communities that received the intervention significantly different at endline than among people who own a latrine in communities that did not receive the intervention?
Secondary research question 1: Is latrine construction by endline among people who do not own a latrine in communities that received the intervention significantly different than among people who do not own a latrine in communities did not receive the intervention?

Secondary research question 2: Are behavioral determinant scores (i.e. scores for social norms, abilities, physical opportunity, risk perception, motivation, and self-regulation)

significantly different at end line among owners of latrines in intervention villages compared to owners of latrines in control villages?

Secondary research question 3: Are behavioral determinant scores (i.e. scores for social norms, abilities, physical opportunity, risk perception, motivation, and self-regulation) associated with latrine use?

2.2. What are the hypotheses to be tested throughout the causal chain?

Main research question:

H_1: Latrine use among latrine owning households in in communities that received the intervention will be significantly higher compared to latrine owning households in control communities.

Secondary research question 1:

H_2: Latrine construction among non-latrine owners in communities that received the intervention will be significantly higher compared to controls.

Secondary research question 2:

H_3: Latrine use behavioral determinant scores are significantly higher at endline among latrine owners in intervention villages compared to latrine owners in control villages.

H_4: Child feces behavioral determinant scores are significantly higher at endline among latrine owners with children under age 5 in intervention villages compared to latrine owners in control villages.

Secondary research question 3:

H_5: High latrine use behavioral determinant scores are significantly associated with latrine use.

H_6: High child feces behavioral determinant scores are significantly associated with safe child fees disposal.

3. Sampling

3.1. Sampling frame

The eligible population for the study is households that have latrines (defined as having a pit, pan, and pipe connecting the two).

3.1.1. Please list any additional inclusion and/or exclusion criteria for the eligible population.

We will include all households that own latrines, not just households that own functional latrines. As part of the intervention, we will make latrines functional by making low cost repairs (pipe connections, etc.).

Research (i.e. survey administration, qualitative interviews and discussions) will only be carried out with women and men over age 18. We will specifically target women in the household to respond to questions because we feel most confident that they would be able to most accurately respond about the behavior of others in the household, especially children and other dependents.

3.1.2. What are the main characteristics of your population?

The current research will take place in the rural district of Puri, Odisha State, India. Some of the communities will potentially have been engaged previously in a cluster-randomized trial (CRT) (Registration No. NCT01214785)(Clasen et al., 2012). In Puri district, 86% of households live in rural areas, 57% of households are living below the poverty line (BPL) and 19% of women are illiterate(MDWS, 2011). Approximately 15% of households in rural areas were estimated to have had sanitation in 2008, prior to the previous trial (MDWS, 2011).

According to the 2015-2016 Indian National Family Health Survey, 37% of rural households in Puri district have an improved sanitation facility – the same percentage at the national level (37% in rural India)—albeit somewhat higher than the overall state level (23% in Odisha) (NFHS-4, 2015-2016).

3.1.3. What is the expected sample size?

Main trial villages

66 communities will be engaged in the trial, 33 in the intervention arm and 33 in the control arm

We will aim for community sizes that range from 50-150 households each and that have 60% latrine coverage or higher. However, community sizes and coverage may be outside of those ranges as we go through a process of identifying suitable villages via a mapping exercise, which will provide the most up-to-date data.

We anticipate households to have an average of 4 people each. In these main trial villages, we will conduct a baseline and endline census. All households in each village will be asked if they have a latrine. If they do not, the survey ends; if they do, further questions will be asked, including questions about latrine use for all family members. Observations of latrine facilities will also take place.

We anticipate an average of 100 households per village with an average of 4 members each. Given that we will be targeting villages that we expect to have 60% latrine coverage or more, we expect to collect latrine use information from at least 3,960 households (66 villages*100 households * 0.6 coverage) with data on 15,840 individuals (3,960 latrine owning households * 4 persons per household).

We will also randomly select a subset of 20 households per trial village to answer questions on determinants to latrine use behaviors. We are aiming for 10 women and 10 men. With 66 villages and 20 individuals per household, we aim to engage a total of 1320 individuals (660 women and 660 men).

Further, all caregivers of children under age 5 who own a latrine will be asked about child feces disposal knowledge and practices. From our previous work, we anticipate that there will be approximately 10 households per village (average size 100) that will have at least 1 child under age 5. As such, we anticipate engaging 660 caregivers of children under age 5 across the 66 trial villages.

Finally, after endline, we aim to conduct follow-up qualitative research to get more detailed feedback about how the intervention may or may not have influenced their reported behaviors. Specifically, we will carry out in-depth interviews (IDIs) with individuals in the intervention communities that initiated latrine use by endline compared to baseline, with individuals who did not change behavior at all, and with individuals from households that did not have a latrine at baseline but had one by endline. We will also carry out focus group discussions (FGDs) with men and women in intervention communities that saw the greatest overall change in latrine use and the least overall use to determine what factors contributed to change and if and how the intervention components influenced that change. The endline data will enable us to identify who to target and how many individuals. As such, sample size numbers are not final at this stage. We anticipate approximately 30 IDIs and 8 FGDs with approximately 6-10 participants in each. Further information about activities is provided in section 4.1.1.

Subset of Qualitative Villages

We will engage 6 additional non-trial villages for qualitative research. Three of these villages will receive the intervention and 6 will not. The establishment of these village will allow us to ask questions about the intervention and latrine use more broadly in the weeks and months following intervention delivery without disrupting the actual trial. We anticipate approximately 30 IDIs and 8 FGDs with approximately 6-10 participants in each. Further information about activities is provided in section 4.1.1.

3.1.4. Is there any reason to believe that the sample differs from the population? If so, how does it differ?

To be eligible for inclusion in the study, we are seeking villages with 50-150 households (mean 100 households) and a minimum of 60% latrine coverage. These criteria reflect the both the study needs (coverage and household size influence sample size estimates) and the limits of our funding (larger villages would require more inputs and thus more costs).

As such, it is reasonable to believe that our sample does not reflect all villages in Puri or even in the three blocks we will be working specifically. Our pre-trial rapid assessment identified villages that were smaller, larger, and that had less that 60% latrine coverage.

3.1.5. Please describe the anticipated subgroups, if relevant, which will be studied.

Note: Since behavior change interventions require village-level clustering to prevent spillovers, studies will likely not be adequately powered to conduct sub-group analysis, and subgroup analysis is not expected. Proposals to do subgroup analysis should be accompanied by an explanation of how studies will be able to detect differences between subgroups.

We do not anticipate conducting any sub-group analyses.

3.2. Statistical power

3.2.1. What is the effect size that you will be able to detect?

We are powered to detect an absolute increase of **10%** in latrine use among intervention group.

We used a Monte Carlo simulation approach to determine the sample size required for the proposed cluster-randomized controlled trial of latrine use (Arnold et al 2011). The simulation parameters were estimated from previously collected data on latrine use among participants in the Odisha sanitation trial (Clasen et al 2014). Because the proposed intervention is targeted to households with existing latrines, only households with a latrine at baseline of the Odisha sanitation trial were used to generate simulation parameters. A rapid assessment of villages in the study area was conducted to determine the number of households and latrine coverage rates. As the design of the proposed study includes both randomization at the village level and repeated measures (baseline and follow-up) among study participants, the data-generating model used for these simulations was a 3-level mixed effects model with cluster (village) level and person-level random effects. The parameters used in our sample size simulations were as follows:

1. Average number of households in eligible villages = **98**
2. Coefficient of variation in village size = **0.35** (eligible villages ranged in size from 50-150 households)
3. Latrine coverage among eligible villages = **75%** (minimum of 60% coverage required for eligibility)
4. Average number of households per village with latrines = 98*0.75 = **73**
5. Average number of persons per household = **4**

6. Average number of eligible persons per village=73*4 = **292**
7. Proportion of persons who reported using latrine the previous day = 0.47 (**0.45** used in simulation)
8. Village--level standard deviation = **1.43**
9. Person--level standard deviation = **3.88**
10. Village--level ICC = **0.10**
11. Correlation of individual use over time (rainy, winter, and summer seasons) = lowest was **.60** (most conservative)

In addition, the following parameters were specified by the study team

1. Alpha = **0.05**
2. Beta = **0.80**
3. Effect size = Absolute increase of **10%** in latrine use among intervention group
4. Minimum Latrine coverage = **60%**
5. Attrition = **10%** loss to follow-up between baseline and follow-up.

Latrine use is a binary outcome for each individual. It will be aggregated to a percentage for the village. This question is in use by all 3ie grant recipients. The question is asked for each household member is:

> *The last time [NAME] defecated, did [NAME] defecate in the open or use the latrine?*
>
> 01= Open; 02= Latrine; 03= Somewhere else (potty, nappy, etc.)
> *if the respondent chooses 3, the follow-up questions enable classification of the feces disposal as safely disposed of in the latrine or not.

3.2.1.1. What are your assumptions about your alpha level?

The alpha level is 0.05.

We assume village size between 50-150 households, an average of 5 persons per household, minimum village level latrine coverage of 60%, 10% loss to follow-up between baseline and endline, village level ICC of 0.1, village level standard deviation of 1.43, person-level standard deviation of 3.88, and correlation of individual use over time of 0.6.

3.2.1.2. What are your assumptions about your statistical power?

The beta is 0.80.

3.2.1.3. What are your assumptions about variability in your effect size?

Based on our analysis (mixed-effects modelling) of latrine use data from a previous sanitation trial in Odisha state, we assume a village-level standard deviation of 1.43 and a person-level standard deviation of 3.88. These parameters were used in empirical sample size calculations via Monte Carlo simulation.

3.2.1.4. How many clusters will you have?

We will have 66 clusters (villages); 33 will receive the intervention, 33 will act as controls.

3.2.1.5. How many people will you have in each cluster?

We anticipate collecting latrine use data on all members of households that own latrines. We will aim for clusters sizes that will range from 50-150 households, have at least 60% latrine coverage and expect that households will have an average of 4 people each. As such:

On average, we anticipate villages will have an average of 97 households with 75% coverage and 4 people per household. Thus, we anticipate collecting latrine use data on an average of 292 people per cluster. [97 * 0.75 * 4 = 292] or 9,636 people per arm.

At minimum, in a village of 50 households with 60% coverage and 4 people per household, we anticipate collecting latrine use data on 120 people per cluster.
[50 * 0.6 * 4 = 120].

At maximum, in a village of 150 households with 100% coverage and 4 people per household, we anticipate collecting latrine use data on 600 people per cluster.
[150 * 1 *4 = 600].

3.2.1.6. How sensitive is your effect size to changes in your parameters?

Our minimum detectable effect size (MDES) for latrine use is fairly robust to potential changes in the parameters of our sample size calculations. For example, with an increase in the village-level ICC from 0.10 to 0.15 we are powered for a 12.1% increase in latrine use. Similarly, a reduction in the within-person (pre-post) correlation in latrine use from 0.6 to 0.5 would result in our MDES increasing from 10% to 11%. Nevertheless, we are confident in the accuracy of the parameters used in our sample size calculations as they were calculated from recent longitudinal latrine use data collected in Odisha state using a similar methodology to that used in the current study.

3.2.2. If you plan to include covariates in your analysis, what share of variance do you expect to predict with your co-variates?
Note: It is not required that you include covariates

Covariates we aim to include are noted in response to 5.1.2. We did not calculate our sample size based on assumptions about shared variance from covariates. Ours is a more conservative approach.

3.3. Assignment to treatment

3.3.1. How will individuals be assigned to treatment and control conditions?
We will assign entire villages to treatment and control conditions, not individuals. We will engage 66 villages in the main trial, 33 of which will receive the intervention and 33 that will serve as controls. We will engage another 6 villages in qualitative work alone so that we may learn about experiences and perceptions of the intervention closer to the time of implementation and not post endline. Three of these 6 villages will receive the intervention; three will not and will enable us to ask other latrine use related questions.

Village selection: Trial Villages

In the summer of 2017, our research team carried out a rapid assessment of villages in three block in Puri: Delang, Pipili, and Nimapada in order to generate approximate data on village sizes (number of households) and latrine coverage. We focus on these blocks due to the ability of our team and our partner to access them. We carried out this exercise as we know village sizes and coverage are always changing and that visiting and talking with village leaders would give the most up-to-date sense of current village status. In total, 282 villages were visited. We use this list and the data generated as our sampling frame. From this list, we identified all potentially eligible villages that are suitable for inclusion (between 50 and 150 households per village, at least 60% latrine coverage, and not declared open-defecation free). Prior to baseline we will visit the villages to map the villages for future visits and, during this process, also verify the total number of households in the villages and the number that have latrines to confirm latrine coverage. At this time, we will also be able to assess the location of villages and determine their proximity to one another. If villages are immediately adjacent, we will only select one so as to minimize the likelihood of spillover. We will collect baseline data from those villages that we have verified to fit our criteria.

Following the baseline, we will randomize the 66 villages, 33 to an intervention group that will receive the intervention package and 33 to a control group that will not receive any intervention and will serve as comparisons. We will use stratified randomization to ensure balance on significant criteria. While there are many potential criteria (program under which latrines were provided, proportion

constructed with household funds versus government subsidy, etc.), we will prioritize village size and latrine coverage. We also will ensure that villages are not immediately adjacent to minimize the likelihood of spillover or might otherwise be influenced by contamination.

Village selection: Qualitative Villages

We will engage 6 additional villages in qualitative activities only. These villages will be selected from the same pool as the 66 trial villages using the same processes noted. We will not collect baseline data on these villages, but will collect basic information from community stakeholders to understand under what program (TSC, NBA, SBM) the village received their latrines, whether a large portion was constructed at the cost of the householder rather than by subsidy, village size, and latrine coverage. As with the 66 villages engaged in the trial, we will randomize the 6 villages along these criteria, assigning 3 to the intervention and 3 to serve as controls.

The implementers will not be made aware which villages are in the formal trial (with endline and baseline data collection on latrine use, etc.) and which will be engaged in qualitative work. We want to be sure that all villages receiving the intervention do so as consistently as possible.

3.3.2. How will you check that individuals in the treatment condition received treatment as anticipated?

We will monitor but not participate in the intervention delivery. Guided by the approach noted by Saunders et al (2005), we will use quantitative and qualitative tools to carry out a process evaluation of the intervention, which will include observations of the activities to understand recruitment, reach, dose delivered, fidelity, and dose received. To understand *satisfaction* with the intervention, we will carry out qualitative research with household members who received the intervention after endline and with a subset of households shortly after intervention delivery from a distinct set of villages not engaged in the impact evaluation. We will also conduct interviews with the partner staff to understand contextual factors that may have influenced delivery, challenges associated with delivery, and recommendations for improvements.

4. Data Collection

4.1. Primary data collection instruments

4.1.1. What data collections instruments will you employ for quantitative and qualitative analysis?

The table below provides details regarding the proposed instruments (Q 4.1.1), and the target interviewees (Q 4.1.2).

Activities during the 66 village impact evaluation (CRT)			
Activity	**Participant type**	**Sample Size**	**Purpose**
Census *To be carried out at baseline and endline.*	Representative from each village household.	~6600-7260 HHs per round; 13,200-14,520 inclusive of baseline and endline (assuming average community size is 100-110 households across 66 villages)	To determine latrine coverage rates; To identify households with latrines for additional questions on use, latrine history/funding. Non latrine owning households will provide basic demographic information. Latrine owning households will provide expanded information on demographics, latrine use, etc. See following table for more information.
Latrine Observations *To be carried out at baseline and endline.*	To be carried out after administration of census at all households with latrines	~ 5808 [Assuming 110 households per village with 80% latrine coverage]	Aim is to understand if latrines are functional, if repairs are needed, what repairs are needed, and if latrine meets criteria for repair. To be carried out at baseline and endline.
Child Feaces Behavioral Determinants Supplement *To be carried out at baseline and endline.*	Caregivers from all latrine-owning households with children under age 5 per village	660 per round; 1,320 inclusive of baseline and endline [assuming average of 10 households per village with latrines *and* children under age 5: 66 villages * 10 households]	To identify drivers of child feces handling behavior in latrine owning households that have children under age 5.
Latrine Use Determinants Supplement *To be carried out at baseline and endline.*	Representative from 20 randomly selected latrine-owning households per village	1320 per round [66 villages * 20 households]	To identify drivers of latrine use among a subset of adults in latrine owning households. The same respondents answering at baseline will be followed-up with at endline.
Total for Impact evaluation/CRT related activities in 66 villages, at both baseline and endline		18,480 [Max, for survey activities]	

		11,616 for observation activities	
Qualitative Activities post trial endline			
Activity	**Participant type**	**Sample Size**	**Purpose**
In Depth Interview	Adults (men and women) with latrines	10	To identify perceptions of various intervention components and if/how the intervention may have influenced personal behavior
In Depth Interview	Adults (men and women) without latrines	10	To identify perceptions of various intervention components and if/how the intervention may have influenced perceptions of latrine use and interest in building a household latrine
In Depth Interview	Adults (men and women) with latrines	10	To identify emerging barriers and drivers to latrine use and not-use (in villages not receiving interventions)
Focus Group Discussion	Adults (men and women) who became latrine users in intervention villages	40 Approx. 4 FGDs (sex specific) with 6-10 participants each	To identify perceptions of various intervention components and if/how the intervention may have influenced personal behavior; to discuss if and how the intervention influence social norms
Focus Group Discussion	Adults (men and women) who continued to not use latrines in intervention villages	40 Approx. 4 FGDs (sex specific) with 6-10 participants each	To identify perceptions of various intervention components and why the intervention did not influence behavior
In Depth Interview	NGO Partner Staff (men and women)	10	To gin perceptions of the intervention from the implementer, how it can be improved, challenges faced, etc.
Total Qualitative Activities post trial endline		120	
Qualitative Activities in a subset of 6 villages not engaged in the trial (3 receiving intervention, 3 not)			
Activity	**Participant type**	**Sample Size**	**Purpose**

In Depth Interview	Adults (men and women) with latrines	10	To identify perceptions of various intervention components and if/how the intervention may have influenced personal behavior
In Depth Interview	Adults (men and women) without latrines	10	To identify perceptions of various intervention components and if/how the intervention may have influenced perceptions of latrine use and interest in building a household latrine
In Depth Interview	Adults (men and women) with latrines	10	To identify emerging barriers and drivers to latrine use and not-use (in villages not receiving interventions)
Focus Group Discussion	Adults (men and women) who became latrine users in intervention villages	40 Approx. 4 FGDs (sex specific) with 6-10 participants each	To identify perceptions of various intervention components and if/how the intervention may have influenced personal behavior; to discuss if and how the intervention influence social norms
Focus Group Discussion	Adults (men and women) who continued to not use latrines in intervention villages	40 Approx. 4 FGDs (sex specific) with 6-10 participants each	To identify perceptions of various intervention components and why the intervention did not influence behavior
Total for qualitative activities in subset of 6 villages		110	

4.1.2. What is the hypothesised list of interviewees/targets (i.e. types of actors or stakeholders who will be interviewed, anticipated interview formats and expected number of respondents)? You may wish to present this information in a table.

See table above.

4.1.3. What (groups of) indicators will each instrument cover?

The table below describes the indicators each instrument will cover (Q 4.1.3) as well as how each tool was or will e developed (Q4.1.4).

4.1.4. How will each instrument be developed?

The table above describes how each instrument was or will be developed.

4.1.5. Please comment on the validity and reliability of each instrument, including any anticipated validation checks.

For Latrine Use*:* Researchers, and 3ie and rice affiliates determined the key indicators for latrine use collectively. Observations of the latrines will be carried out as a validation check. If latrines are deemed to be non-functional (i.e. missing pipe connections, etc.) or to not have evidence of any recent use (i.e. used for storage, no water vessels/shoes/ worn path to facility, etc.), reports of use will be called into question.

For Behavioral Determinants: Indicators for the various behavioral determinants under investigation (see table above and survey document) were created based on a review of relevant literature and with support from relevant technical documents (See: Mosler, H.-J., & Contzen, N. (2016). Systematic behavior change in water, sanitation and hygiene. A practical guide using the RANAS approach. Version 1.1. Dübendorf, Switzerland: Eawag.; Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press). Construct validity for each of these determinants will be assessed using confirmatory factor analysis (CFA). Reliability will be assessed using chronbach's alpha.

For Socio-economic status: We will use principle components analysis (PCA) to assess construct validity of our SES index. Reliability will be assessed using chronbach's alpha.

4.2. Secondary data sources

Please describe the anticipated secondary sources of data, if any, which will be used for this study.

Not applicable.

5. Analysis

5.1. Outcome Variables

5.1.1. Your primary outcome is latrine use. Please describe the primary and any secondary outcome variables of interest using the following table:

Outcome	Description	Hypothesis	Level
"Outcome 1"	*(brief description of outcome with*	*(E.g. related to Hypothesis 1; related*	*(E.g. individual,*

(e.g. latrine use)	*associated indicators and constituent parts)*	*to H1 and H2, etc.)*	*household, etc.)*
Latrine Use	As defined a priori in the guidelines: "For every household member, as part of a household roster (where household is defined as living under this roof): "The last time [NAME] defecated, did [NAME] defecate in the open or use the latrine?" Options: Open, Latrine, Somewhere else	Related to H_1 and H_5	Individual
Latrine Coverage	Assessment of latrine coverage in villages. Noted by self-report and confirmed by observation.	Related to H_2	Village/ Community
Latrine Use Determinants, specifically Risk Perceptions, Ability, Social Norms, Motivation, Physical Opportunity, and Self-Regulation	For each determinant, there are several indicators that define the associated constructs.	Related to H_3 and H_5	Individual
Latrine Use Determinants, specifically Risk Perceptions, Ability, Social Norms, Motivation, Physical Opportunity, and Self-Regulation	For each determinant, there are several indicators that define the associated constructs.	Related to H_4 and H_6	Individual

5.1.2. If you plan on including covariates in your analysis, please provide a list of covariates that may be included.

Baseline latrine use, age, sex, educational attainment, household size, household socio-economic status

We will run both unadjusted and adjusted models and compare models as a sensitivity analysis.

5.1.3. If you plan to aggregate multiple variables into an index, which variables will you aggregate and how?

1. Socioeconomic status: We will ask respondents if they own various household items and use that information to create an asset index using Principle Components Analysis (PCA), resulting in a score for each household.

2. Latrine Use Determinants: For each of the latrine use determinants (Risk Perceptions, Ability, Social Norms, Motivation, Physical Opportunity, and Self-Regulation), we have a series of questions. (See Survey; determinants to which the questions correspond are noted on the left-side column). The number of questions per determinant ranges from 3-20 per. We will carry out confirmatory factor analysis to confirm the factor structure for each determinant (eliminating questions as appropriate), check reliability (cronbach's alpha), and then determine scores based on final questions per determinant and the responses provided.

5.2. Qualitative Analysis

Which methods will be used to analyse qualitative data (e.g. content analysis with criteria for codification)?

Transcripts from qualitative activities will be analysed using thematic content analysis.

Members of the research team will begin analysis by reading through transcripts and writing memos about the issues discussed in the location where they collected the data. The memos will inform the creation of a preliminary codebook. The preliminary codebook will be shared among members of the research team and refined. Researchers will then use the final codebook to apply codes to the data collected. This may be done in duplicate to compare coding strategies and make certain that coding is consistent across all researchers.

Once coding is complete, researchers will write thematic memos. Some memo topics will be pre-determined (deductive). Other memos will be created that are not anticipated based on what is learned from the data collected (inductive).

5.3. Quantitative Analysis

5.3.1. Balance Checks

5.3.1.1. How will you check balance between treatment and control groups?

We will use the baseline data to calculate descriptive statistics and compute the standardized difference between arms in order to compare groups on all outcome measures as well as individual and village-level sociodemographic characteristics. The standardized difference is a metric that expresses the difference between groups in standard deviation units.

For continuous covariates, the standardized difference (*d*) is calculated as

$$d = \frac{(\bar{x}_{intervention} - \bar{x}_{control})}{\sqrt{\frac{s^2{}_{intervention} + s^2{}_{control}}{2}}}$$

where $\bar{x}_{intervention}$ and $\bar{x}_{control}$ denote the sample mean of the covariate and $s^2{}_{intervention}$ and $s^2{}_{control}$ denote the sample variance of the covariate in the intervention and control groups.

For dichotomous covariates, the standardized difference (*d*) is calculated as

$$d = \frac{(\hat{p}_{intervention} - \hat{p}_{control})}{\sqrt{\frac{\hat{p}_{intervention}(1 - \hat{p}_{intervention}) + \hat{p}_{control}(1 - \hat{p}_{control})}{2}}}$$

where $\hat{p}_{intervention}$ and $\hat{p}_{control}$ denote the prevalence in the intervention and control groups.

If the standardized difference indicates an important imbalance between groups at baseline, we will perform sensitivity analysis by adjusting for that variable in statistical models and comparing effect estimates to the unadjusted models.

In accordance with CONSORT guidelines, we will not perform significance testing of between-group differences in baseline characteristics or outcomes to assess balance.

5.3.1.2. What is the specification that you will run and what variables will you include?

Not applicable. The computation for standardized difference is presented above

5.3.1.3. If there is an imbalance (between treatment and control groups) in one or more baseline covariates, how do you plan to address this?

Statistical models will be adjusted for covariates that are observed to be severely imbalanced between groups at baseline.

5.3.2. Contamination

How will you detect and manage any potential differential contamination between treatment and control groups?

Given the nature of the intervention (community performances, meetings, household visits, latrine repairs) and our intention to have intervention and control arms geographically separated, we do not expect contamination. However, at endline, we will ask a series of questions to determine exposure to the intervention in both the intervention and control villages.

5.3.3. Attrition

5.3.3.1. What is your anticipated attrition rate and what evidence is this prediction based on?

The expected attrition rate is 10%. This estimate is based on our previous experience conducting sanitation trials in the region.

5.3.3.2. What can you do anything to prevent or remedy sample attrition?

For our primary outcome, latrine use, we are aiming to get latrine use data from all households in the villages that have latrines. Given that we are sampling all eligible households that have latrines, we are not able to do anything to increase the sample size to remedy attrition. To prevent attrition, we can simply train our field team to carry out the data collection in a respectful and confidential manner so as to encourage participation again at endline.

5.3.3.3. How does expected attrition change your power calculations?

Our power calculations incorporate the estimated 10% attrition rate.

5.3.3.4. How will you check balance between attritors and non-attritors? What is the specification that you will run and what variables will you include in these balancing checks?

We will compare attritors and non-attritors on baseline latrine use and coverage as well as sociodemographic characteristics (age, gender, SES, household size, and caste).

5.3.4. Missing Data

How will you deal with incomplete or missing data?

Given our study design, we anticipate that the primary cause of missing data will be loss to follow-up. As GEE is robust to missing data under the MCAR mechanism, we will conduct Little's test to evaluate whether the MCAR assumption is supported. If the MCAR assumption is not supported, we will assume a MAR mechanism and conduct sensitivity analyses by conducting weighted (inverse-probability weighting) GEE and comparing model results.

5.3.5. Treatment Effects

Note: Many studies may have awareness campaigns where one may not be able to know whether a household participated or heard the message or not. In these cases, it may not be possible to estimate a Treatment on the Treated (TOT) effect. We therefore do not expect that all studies will provide estimates of TOT.

5.3.5.1. Intent to Treat

5.3.5.1.1. How will you estimate the (causal) effect of the offer of the treatment?

We will conduct an intent to treat analysis of differences in the specified outcomes between the treatment and control groups following delivery of the intervention. We will employ generalized estimating equations (GEE) with robust standard errors to account for village-level clustering in the outcome. Models will be adjusted for baseline latrine use and pre-specified confounders such as age and sex. We will also report a "difference-in-difference" between the study arms, though with balanced study arms, this is not expected to be different from the main comparison.

5.3.5.1.2. What is the specification that you will run and what controls will you include in your specification?

We will use GEE with robust standard errors to estimate a marginal (population average) model with the general form

$$g\left(E\left[Y_{ij}|x_{ij}\right]\right) = \boldsymbol{x'_{ij}\beta}_{ij}$$

where g(.) is the link function**,** Y_{ij} is the outcome of interest for the j_{th} observation in the i_{th} cluster, $\boldsymbol{x'_{ij}}$ is a vector of covariates**, and** $\boldsymbol{\beta}_{ij}$ is a vector of regression coefficients. We will specify an exchangeable correlation matrix as the most plausible and parsimonious choice of working correlation structure, noting that GEE with robust estimation yields valid estimates of model coefficients and standard errors when the correlation structure is misspecified. We will use a log-binomial link function, which will yield the prevalence ratio of post-intervention latrine use in persons receiving the intervention relative to controls, adjusting for baseline latrine use and control variables. Specifically, we will estimate the model

$$\begin{aligned} &log(probability\ of\ latrine\ use) \\ &\quad = \beta_0 + \beta_1(\text{post} - \text{intervention latrine use}) + \beta_2(\text{baseline latrine use}) \\ &\quad + \beta_3(\text{age}) + \beta_4(\text{sex}) + \beta_5(\text{educational attainment}) + \beta_6(\text{household size}) \\ &\quad + \beta_7(\text{household SES}) \end{aligned}$$

5.3.5.2. Treatment on the Treated

5.3.5.2.1. How will you estimate the (causal) effect of the receipt of the treatment?

It will not be possible for us to estimate a Treatment on the Treated effect. Our intervention involves various events, including community level performances to raise awareness, community meetings and demonstrations, and household level visits and latrine repairs. While tracking household level engagement will be possible, it will not be feasible to accurately track anticipation in the community-wide events and understand if messages were received.

5.3.5.2.2. What is the specification that you will run and what controls will you include in your specification?

Not applicable.

5.4. Heterogeneous Effects

Note: Since behavior change interventions require village-level clustering to prevent spillovers, studies will likely not be adequately powered to conduct sub-group analysis, and subgroup analysis is not expected. Proposals to do subgroup analysis should be accompanied by an explanation of how studies will be able to detect differences between subgroups.

5.4.1. Which groups do you anticipate will display heterogeneous effects?
We are not planning sub-group analyses.

5.4.2. What is the broad theory of action that leads you to anticipate these effects?
Not applicable.

5.5. Standard Error Adjustments

5.5.1. How will you address clustering in your data?
We will employ generalized estimating equations (GEE) with robust standard errors to account for village-level clustering in the outcome.

5.5.2. How will you address false positives from multiple hypothesis testing?
As described in this analysis plan, we are fitting a small number of pre-specified models. The number of planned analyses is not sufficient to warrant concerns about multiple testing.

5.5.2.1. If you plan to adjust your standard errors, what adjustment procedure will you use? (e.g., Family Wise Error Rate, False Discovery Rates, etc.)
Not applicable.

5.5.2.2. How will you deal with outcomes with limited variation? For instance, one option could be to decide in advance that outcomes that vary below a certain threshold will be omitted from the analysis.

Based on our previous experience, we do not anticipate there to be limited variation in outcomes. If we have outcomes with limited variation, we will test those outcomes. If we have covariates for which over 95% carry the same value, we will drop those.

List of optional attachments

Script (Optional)

You may wish to upload an analysis script with clear comments. This optional step is helpful in order to create a process that is completely transparent and increase the likelihood that your analysis can be replicated. We recommend that you run the code on a simulated dataset in order to check that it will run without errors.

Data Collection Tools (Optional)

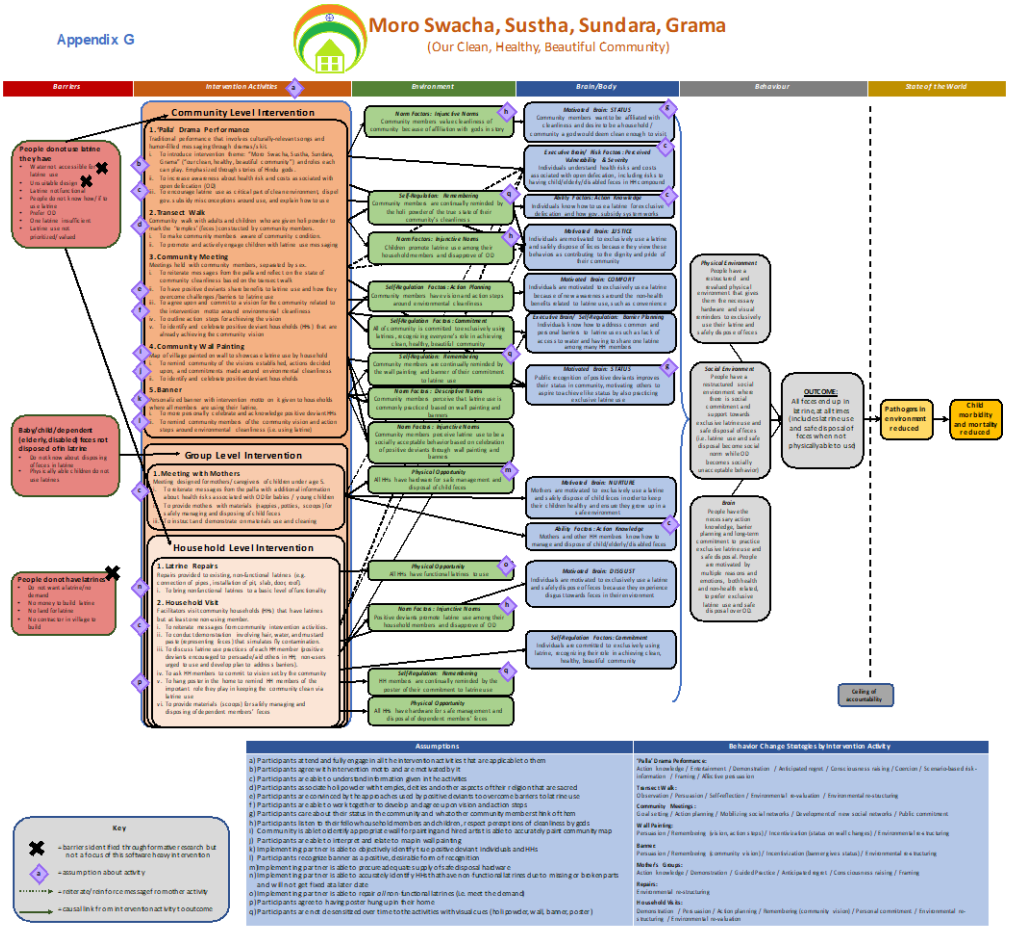
You may wish to attach any qualitative or quantitative data collection tools, if available.

PAP References

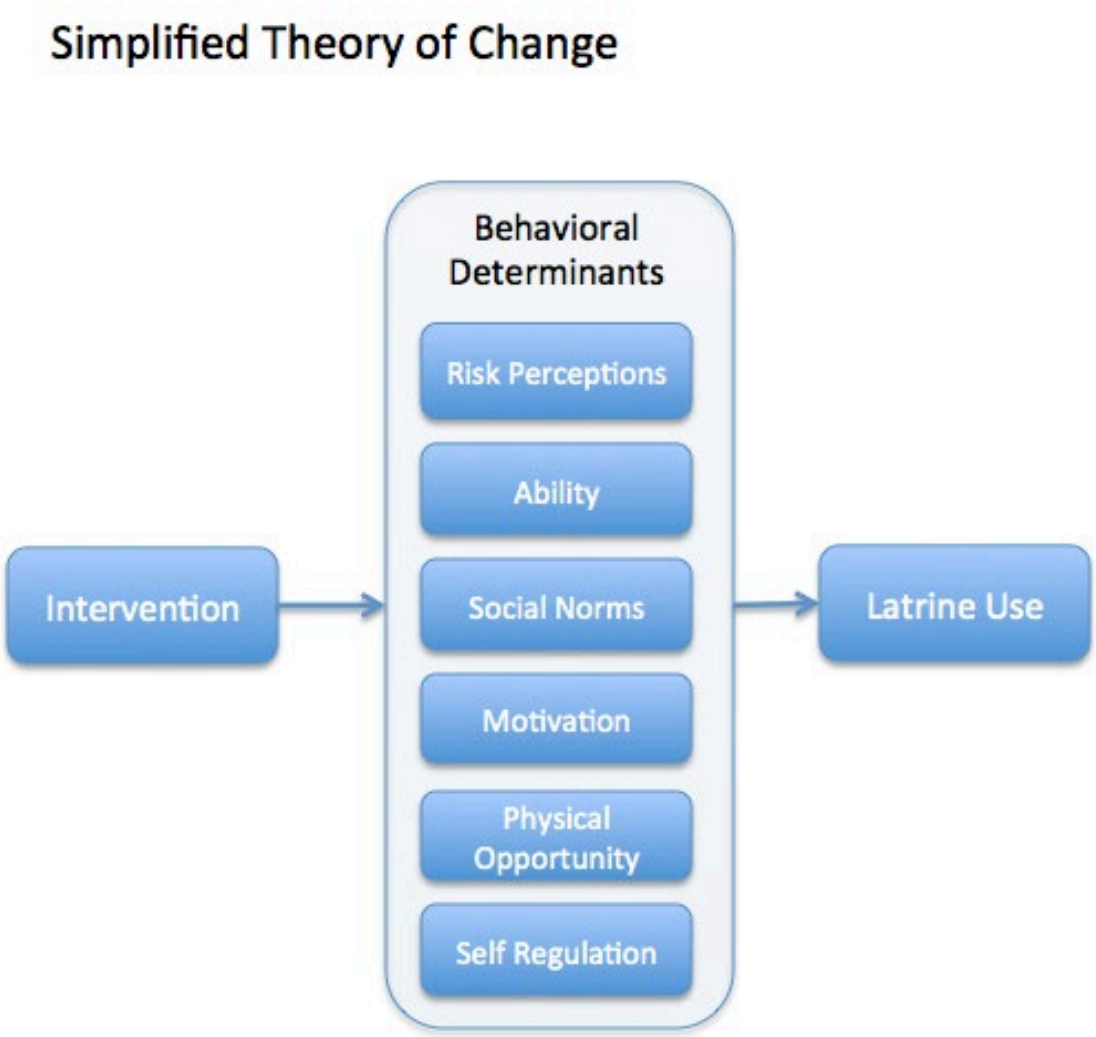
Prochaska, J. O. (2013). Transtheoretical model of behavior change. In *Encyclopedia of behavioral medicine* (pp. 1997-2000). Springer New York.

PAP APPENDICES

PAP Appendix A.



PAP Appendix B.



Appendix 2: Fidelity/Dose Scoring for the Palla

Fidelity/Dose Scoring for the Palla		
Metric	**Scoring**	**Possible Points**
Fidelity		
Stakeholders in attendance	0=no stakeholders 1= at least 1 stakeholder	0-1
Length of Palla	0 = <1hr long 1 = if >1hr long	0-1
Pre-Palla Activities and Preparations Carried Out as Expected	0= if not all expected pre-palla activities carried out (e.g. music to attract people, Welcome by stakeholder, etc.) 1= if all expected pre-palla activities carried out	0-1
Palla Skits/Messages Delivered in Correct Order	0 = not in correct order 0.5 = somewhat in order 1 = correct order	0-1
Dose		
Performance Score A. Intro opening B. Lakshmi story C. Uncle Nidhi Story D. Runi's Story E. Pit Emptying Story F. Sarpanch & Ward member Skit G. Child feces Story H. OD and latrine use messages I. Benefit of latrine use J. Closing	For A-J: 0 = > 50% of components included 0.5pt = 50 to 99% of components included 1 = ALL components included	0-10
Total Possible Points		**14**

Appendix 3: Fidelity/Dose Scoring for the Transect Walk

Fidelity/Dose Scoring for the Transect Walk		
Metric	**Scoring**	**Possible Points**
Fidelity		
Stakeholders in attendance	0=no stakeholders 1= at least 1 stakeholder	0-1
Participant Engagement	0= Engagement in < 50% facets 1= Engagement in 50-99% facets 2= Engagement in all facets	0-2
Activity Start Time	0= If TW started at time other than 5-8am. 1= If TW started between 5-8am.	0-1
Preparations/ Materials	0= No preparations 1= Have all of following: colored powdered (not red or orange), soap, noisemakers	0-1
TW Components Delivered in Correct Order	0 = not in correct order 0.5 = somewhat in order 1 = correct order	0-1
Dose		
Walk Score A. Opening Procedures B. Middle Procedures C. Powder Distribution for All D. Message Delivery E. Closing Procedures	<u>For A-E:</u> 0 = > 50% of components included 0.5pt = 50 to 99% of components included 1 = ALL components included	0-5
Total Possible Points		**11**

Appendix 4: Fidelity/Dose Scoring for the Community Meeting

Fidelity/Dose Scoring for the Community Meeting		
Metric	**Scoring**	**Possible Points**
Fidelity		
Stakeholders in attendance	0=no stakeholders 1= at least 1 stakeholder	0-1
Participant Engagement	0= Engagement in < 50% facets 1= Engagement in 50-99% facets 2= Engagement in all facets	0-2
Sex-Segregated Meetings	0 = Meetings not sex segregated 1 = Separate men's / women's meetings	0-1
Length of Community Meeting	0 = <1hr long 1 = if >1hr long	0-1
Preparations/ Materials	0= No preparations 1= Banner hung and Meeting held in private place	0-1
Meeting Messages Delivered in Correct Order	0 = not in correct order 0.5 = somewhat in order 1 = correct order	0-1
Dose		
Community Meeting Score A. Opening Procedures B. Purpose of meeting stated C. Transect walk reflection D. Motto stated E. Sanitation/OD statistics shared F. Dignity and pride discussed G. Health messages shared H. Role modeling discussed F. Recognition of challenges G. Positive deviants Identified H. Model village discussed J. Action steps discussed K. Commitment and closing	For A-J: 0 = > 50% of components included 0.5pt = 50 to 99% of components included 1 = ALL components included	0-13
Total Possible Points		**20**

Appendix 5: Fidelity/Dose Scoring for the Mother's Group

Fidelity/Dose Scoring for the Mother's Group Meeting		
Metric	**Scoring**	**Possible Points**
Fidelity		
Stakeholders in attendance	0=no stakeholders 1= at least 1 (ASHA/Anganwadi)	0-1
Participant Engagement	0= Engagement in < 50% facets 1= Engagement in 50-99% facets 2= Engagement in all facets	0-2
Mother's Group Meeting Length	0 = <1hr long 1 = if >1hr long	0-1
Preparations/ Materials	0= No preparations 1= Anganwadi consulted for organizing and banner held	0-1
MG Components Delivered in Correct Order	0 = not in correct order 0.5 = somewhat in order 1 = correct order	0-1
Hardware Distribution	0 = Not all got potty/scoop 1 = All got potty/scoop	0-1
Dose		
Mother's Group Score A. Opening B. Discussion of current practices C. Health messages discussed D. F-diagram Discussed E. Infant feces discussed F. Potty discussed G. Scoop discussed H. Training I. Closing	For A-E: 0 = > 50% of components included 0.5pt = 50 to 99% of components included 1 = ALL components included	0-9
Total Possible Points		**16**

Appendix 6: Estimated and Actual Parameters Informing Sample Size Calculations

Estimated and Actual Parameters that Informed Sample Size Calculations		
Parameters	**Estimated**	**Actual**
Average number of households in eligible villages	98	88
Latrine coverage among eligible villages	75%	68%
Average number of households per village with latrines	73	60
Average number of persons per household (Excluding children < 5)	4	4.3
Average number of eligible persons per village	292	256
Coefficient of variation in village size	0.35	0.37
Total number of persons in households with latrines	17344 (accounting for 10% non-response)	16880 (Control=8654, Intervention=8226)
Proportion of persons who reported using latrine the previous day	0.47	0.60
Estimated Village-level ICC	0.10	0.103
Correlation of individual use over time	.60	.60
Alpha	0.05	0.05
Beta	Beta = 0.80	Beta = 0.80
Effect Size	**Absolute increase of 10% in latrine use**	**Absolute increase of 10% in latrine use**
Loss to follow-up between baseline and follow-up.	10%	10%

Appendix 7: Program Administration Costs

Program Administration Costs					
Item	Budgetary Total (Rs.)	Budgetary Total (USD)	Unit	Unit cost (Rs.)	Description
Community mobilizer staff	₹ 719,320.00	$10,276.00	48 months of community mobilizer pay	₹ 14,985.83	16 RWI staff members were hired as community mobilizers for 3 months at a cost of 14985.83Rs/month
Supervisor staff	₹ 239,610.00	$3,423.00	12 months of supervisor pay	₹ 19,967.50	4 RWI staff members were hired as supervisors for 3 months at cost of 19967.5Rs/month
Office rent	₹ 22,890.00	$327.00	3 months	₹ 7,630.00	Office rent for 3 months at 7630Rs/month
Office stationary	₹ 3,290.00	$47.00	3 months	₹ 1,096.67	Stationary for general office needs for 3 months at 1096.67Rs/month
Office organizational	₹ 32,690.00	$467.00	3 months	₹ 10,896.67	?
TOTAL	**₹ 1,017,800.00**	**$14,540.00**			

Appendix 8: Intervention Staff Training Costs

Intervention Staff training costs					
Item	**Budgetary Total (Rs.)**	**Budgetary Total (USD)**	**Unit**	**Unit cost (Rs.)**	**Description**
Lunch	₹ 50,858.50	$726.55	408 meals	₹ 124.65	Lunch provided to RWI staff and Emory team for 12 days of intervention training (5 days of in-house training, 5 days of in-house & field training, and 2 days of field training). Total training participants included 23 RWI staff (16 community mobilizers, 4 supervisors, 3 administrators) and 12 Emory team members. 34 training participants * 12 days = 408 meals at 124.65Rs/meal
Lunch materials	₹ 1,980.30	$28.29	-	-	Purchase of paper plates, spoons, water, soap for handwashing and other items needed for serving and administering lunch during the 12 training days
Tea and biscuits	₹ 6,946.10	$99.23	34 participants	₹ 204.30	Tea and biscuits provided to RWI staff and Emory team throughout the 12 days of intervention training at a price of 204.30Rs/participant
Transportation stipend for RWI staff to attend training	₹ 42,191.10	$602.73	20 participants	₹ 2,109.56	Transportation stipend given to each RWI staff member (16 community mobilizers, 4 supervisors) to cover cost of attending the 12 days of intervention training at RWI office. Stipend ranged from 120 to 200 Rs per day per person depending on how far away the staff member lived for an average total transportation stipend of 2109.56Rs/staff member
Transportation for field training	₹ 13,699.70	$195.71	-	-	Cost to transport RWI staff and Emory team to 'pilot villages' to practice intervention activities during the 7 field training days.
Training hall	₹ 5,351.50	$76.45	10 days	₹ 535.15	Fee for using RWI's training hall at a cost of 535.15Rs/day for a total of 10 in-house training days
Notebooks and pens	₹ 2,072.00	$29.60	20 participants	₹ 103.60	Purchase of notebooks and pens for the 20 RWI staff (16 community mobilizers and 4 supervisors) to use for note-taking during intervention training at a cost of 103.60Rs/participant
Printing and photo copying of training materials	₹ 2,472.40	$35.32	-	-	Cost to print training materials and make photo copies for the 20 RWI staff. Training materials for each participant included 1 intervention manual, 5 facilitator guides, 1 palla script, 3 logsheets, and 12 daily training evaluation forms.
Materials for training demonstrations	₹ 1,450.40	$20.72	-	-	Cost of materials needed to practice intervention activities during training such as colored powder, soap, chocolate (transect walk activity), potty and scoop (mother's group meeting), and drinking glasses and mustard paste (household visit activity).
Practice/pilot palla performances	₹ 4,281.20	$61.16	2 palla groups	₹ 2,140.60	Payment to 2 palla performance groups for conducting pallas during the field training days at a cost of 2140.60Rs/palla group
TOTAL	**₹ 131,303.20**	**$1,875.76**			

Appendix 9: Intervention Implementation Costs

Intervention Implementation Costs					
Item	**Budgetary Total in INR**	**Budgetary Total in USD**	**Units**	**Unit cost**	**Description**
Transportation					
Transportation stipend for community mobilizers	₹ 106,190.00	$1,517.00	16 community mobilizers	₹ 6,636.88	16 RWI community mobilizers were given a daily transportation stipend to cover their travel costs to intervention villages during implementation. Cost of 6636.88 Rs/person for 1.5 months of implementation
Transportation stipend for supervisors	₹ 35,770.00	$511.00	4 supervisors	₹ 8,942.50	4 RWI supervisors were given a daily transportation stipend to cover their travel costs to intervention villages during implementation. Cost of 8942.50Rs/person for 1.5 months of implementation
Palla					
Hired palla performers	₹ 281,050.00	$4,015.00	43 palla performances	₹ 6,536.05	2 palla performance troupes consisting of 5 to 6 members each conducted 43 pallas across the intervention villages at a cost of 6536.05Rs/palla performance
Transportation stipend for palla performers	₹ 93,660.00	$1,338.00	43 palla performances	₹ 2,178.14	2 palla performance groups were given a stipend to cover their travel costs to intervention villages. Cost of 2178.14Rs/palla performance
Materials for performance	₹ 46,830.00	$669.00	-	-	Cost of costumes, sound system and other materials needed for the palla performances
Banner	₹ 693.36	$9.91	8 banners	₹ 86.67	A banner with the intervention motto and logo was hung up before the palla performance started as a visual signal to help gather community members to the performance location. Eight banners were printed at a cost of 86.67Rs/banner. (The banners were re-used across the 43 palla performances)
Transect Walks					
Colored powder	₹ 7,630.00	$109.00	36 transect walks	₹ 211.94	Cost of colored powder which was distributed to participants during transect walk activity in order to mark feces seen during the walk. Cost of 211.94Rs/intervention village

Intervention Implementation Costs *(continued)*					
Item	**Budgetary Total in INR**	**Budgetary Total in USD**	**Units**	**Unit cost**	**Description**
Community Meetings					
Banner	₹ 346.68	$4.95	4 banners	₹ 86.67	A banner with the intervention motto and logo was hung up in the location of the community meeting as a visual signal to help gather community members to the location. Four banners were printed at a cost of 86.67Rs/banner. (One banner was given to each RWI supervisor team and re-used across their community meetings)
Paper and pens for action steps	₹ 800.00	$11.43	-	-	Colored paper and pens were purchased for the 'action steps' activity during the community meeting.
Mother's Group Meeting					
Potties	₹ 182,500.00	$2,607.14	730 potties	₹ 250.00	730 potties were purchased for distribution during the mother's group meeting across 36 intervention villages (~20 potties/village) at a cost of 250Rs/potty
Scoops	₹ 21,878.10	$312.54	730 scoops	₹ 29.97	730 scoops were purchased for distribution during the mother's group meeting across 36 intervention villages (~20 potties/village) at a cost of 29.97Rs/potty
Banner	₹ 4,680.00	$66.86	8 banners	₹ 585.00	A banner with instructional drawings on how to safely dispose the feces of infants/babies, toddlers and young children was hung up during the mother's group meeting. Eight banners were printed at a cost of 585Rs/banner. (Two banners were provided to each RWI supervisor team to re-use across their mother's group meetings).
Transportation costs of hardware	₹ 2,600.00	$37.14	-	-	In some cases a transportation service was hired (i.e. auto driver) to help transport the potties and scoops hardware to the mother's group meeting location.
Hired artist for banner drawings and design	₹ 21,095.90	$301.37	-	-	A local artisan was hired to create the instructional drawings used for the mother's group banner. The artist charged around 900Rs/drawing
Household Visits					
Posters	₹ 159,359.00	$2,276.56	2189 posters	₹ 72.80	Households with a latrine were visited by the RWI community mobilizers and given a poster as part of the household visit activity. The poster acted as a reminder on latrine use or to celebrate the household for already using their latrine. 2189 posters were printed at a cost of 72.80Rs/poster.

Intervention Implementation Costs *(continued)*					
Item	**Budgetary Total in INR**	**Budgetary Total in USD**	**Units**	**Unit cost**	**Description**
Community Wall Paintings					
Hired artisan team	₹ 360,000.00	$5,142.86	36 wall paintings	₹ 10,000.00	Two artisan teams with 4 to 5 members each were hired to complete a wall painting in each of the 36 intervention villages. Artisans were responsible for their own transportation, painting materials and anything else needed to complete the wall paintings. Cost of 10000Rs/wall painting.
Latrine Assessment and Repairs					
Latrine assessments	₹ 48,815.00	$697.36	815 latrines assessed	₹ 59.90	Two contracting groups were engaged for the latrine assessment and repair activity. An assessor from each group was paid to go to each selected household and complete a full assessment of the household's latrine. About 815 households were selected for latrine assessment across the 36 intervention villages. The two assessors completed their work over the course of 3 months at a cost of 59.90Rs/latrine assessed.
Latrine repairs	₹ 421,569.00	$6,022.41	457 latrines repaired	₹ 922.47	The two contracting groups completed latrine repairs in about 457 households across the 36 intervention villages. Households that planned to destroy their latrine, required an entirely new superstructure or pit, or used their latrine as storage and did not remove the storage items during the assessment were not eligible for repairs. The types of repairs included about 233 slab repairs, 153 I-pipe repairs, 43 pan repairs, 363 door repairs, 130 pit repairs, and 117 floor repairs. The average cost of latrine repairs per household, including material and labor, was 922.47Rs.
TOTAL	**₹ 1,795,467.04**	**$25,649.53**			

Appendix 10: Village Banner for Post-Endline Results Sharing

Moro Swaccha, Sustha, Sundara Grama!

Latrine Use In _____

_____

___ households, ___ individuals

___ of households have latrine

_____ complete
_____ in construction
_____ abandoned

_____ single-pit pour flush
_____ flush to tank

STUDY MEASURES	_____ Surveys (before / after)		% Change	All Study Villages Surveys (before / after)		% Change
Households that have a latrine	____	____	____	____	____	____
Individuals that use a latrine	____	____	____	____	____	____
Adult Men	____	____	____	____	____	____
Adult Women	____	____	____	____	____	____
Young Adults & Children (5-18)	____	____	____	____	____	____
Households w/ everyone using latrine	____	____	____	____	____	____
Households (with child <5) that dispose of child feces into latrine	____	____	____	____	____	____

ଗ୍ରାମରେ ପାଇଖାନା ବ୍ୟବହାର

_____

..... ଘର, ... ଲୋକସଂଖ୍ୟା

..... ଟି ଘର ପାଇଖାନା ଅଛି

..... ସଂପୂର୍ଣ୍ଣ
..... ତିଆରି ଚାଲିଛି
..... ପରିତ୍ୟକ୍ତ

..... ଗୋଡିକିଆ ନଳ ଥିବା ପାଇଖାନା ଚାଳି
..... ବଡ଼ ଚାଳି ଥିବା ପାଇଖାନା

ଗବେଷଣା ଅନୁଧ୍ୟାନ	_____ ସର୍ଭେ (ପୂର୍ବ ଅବସ୍ଥା / ପର ଅବସ୍ଥା)		% ପରିବର୍ତ୍ତନ	ସମସ୍ତ ଅଧ୍ୟୟନ ଗ୍ରାମ ସର୍ଭେ (ପୂର୍ବ ଅବସ୍ଥା / ପର ଅବସ୍ଥା)		% ପରିବର୍ତ୍ତନ
ପାଇଖାନା ଥିବା ଘର	___	___	___	___	___	___
ପାଇଖାନା ବ୍ୟବହାର କରୁଥିବା ଲୋକ ସଂଖ୍ୟା	___	___	___	___	___	___
ପୁରୁଷ	___	___	___	___	___	___
ସ୍ତ୍ରୀ	___	___	___	___	___	___
ଅନ୍ୟ ଏବଂ ଯୁଆ (୫-୧୮)	___	___	___	___	___	___
ସମସ୍ତ ପରିବାର ଲୋକେ ପାଇଖାନା ବ୍ୟବହାର କରୁଥିବା ଘର	___	___	___	___	___	___
୫ ବର୍ଷରୁ ଛୋଟ ପିଲାଥିବା ଘର, ଯିଏ ଯୁଆର ଆଡ଼ାକୁ ପାଇଖାନାରେ ନିଶ୍ଚାସନ କରନ୍ତି	___	___	___	___	___	___

Appendix 11: Water and Sanitation Characteristics of the Full Study Population

Water and sanitation-related characteristics of the full study population			
	Intervention	**Control**	**Total**
Latrine coverage, per village [Mean (SD)]	63.7 (23.2)	67.3 (23.7)	
Households sharing latrines, per village [Mean (SD)]	3.1 (1.7)	4.1 (2.6)	
Households that own more than one latrine			
Households (proportion)	147 (5.2)	151 (5.0)	298
Type of latrine (n = 4287 latrines)			
Flush to septic system	336	299	635
Flush to pit (latrine)	1507	1625	3132
Flush to somewhere else	3	4	7
Flush to unknown place	1	1	2
VIP	10	11	21
Pit latrine with slab	62	53	115
Pit latrine without slab	20	17	37
Latrine in construction	127	155	282
Other	26	30	56
Number of pits (n=4308 latrines)			
One pit	1451	1543	2994
Two pits	154	187	341
Septic tank	358	320	678
No pit or tank	128	150	278
Don't know	10	7	17
Latrine history and status			
Received money/materials for latrine (n = 3971)			
Yes	1480	1588	3068
No	447	456	903
Source of funding for latrine (n = 4307)			
TSC	4	5	9
TSC / SBM	1	0	1
TSC / self-financed	12	5	17
NBA	4	5	9
NBA / SBM	1	0	1
NBA / SBM / self-financed	1	0	1
NBA / self-financed	8	2	10
SBM	939	951	1890
SBM / self-financed	431	481	912
SBM / other	1	0	1
Self-financed	537	561	1098
Self-financed / other	49	36	85
Other	55	105	160
Don't know	58	55	113

Water and sanitation-related characteristics of the study population *(continued)*

	Intervention	Control	Total
Status of latrine (n = 4310)			
Complete	1841	1908	3749
In construction / incomplete	220	262	482
Abandoned before complete	41	38	79
Latrine repairs needed (n = 4287)			
Yes	598	750	1348
No	1494	1445	2939
Latrine reportedly used for defecation (n = 4310)			
Yes	1598	1674	3272
No	504	534	1038
Latrine looks like it is used (n = 4287)			
Yes	1452	1525	2977
No	640	670	1310
Fecal sludge management			
Latrine pit ever filled (n = 4012)			
Yes	90	99	189
No	1873	1950	3823
Action taken when pit filled (n = 189)			
Emptied	67	60	127
Emptied / built new pit / everyone stopped using	1	0	1
Emptied / everyone stopped using latrine	7	4	11
Built a new pit	0	6	6
Built a new pit / switched to a second pit	0	1	1
Built a new pit / everyone stopped using latrine	0	1	1
Switched to using a second pit	0	14	14
Everyone stopped using latrine	11	11	22
Restricted use to a select few members	3	0	3
Don't know	1	2	3
How latrine pit was emptied (n=139)			
Hired someone to manually empty	28	18	46
Hired tanker to empty	18	21	39
Someone in family manually emptied	25	20	45
Other/Don’t Know	4	5	9

Water and sanitation-related characteristics of the study population *(continued)*

	Intervention	Control	Total
How sludge was disposed of after emptying			
Put in cultivation fields	18	14	32
Sludge buried in a hole	21	17	38
Sludge left in open	17	7	22
Taken by municipality or in private sewer	12	14	26
Other/Don't Know	9	12	21
Access to WASH			
Handwashing location present in / near latrine (n = 4287)*			
Yes	881	881	1762
No	1182	1286	2468
Water source within 30 feet of latrine (n = 4287)			
Yes	1439	1508	2947
No	653	687	1340
Main source of drinking water for household (n = 3971)			
Public tap	73	65	138
Private tap	144	165	309
Shallow tube well (Popular 6)	1222	1347	2569
Deep borehole with hand pump (India Mark II / III)	407	439	846
Protected well	1	4	5
Unprotected well	41	3	44
Tanker truck	0	2	2
Surface water	4	3	7
Bottled water	1	0	1
Other	34	16	50
Location of drinking water source (n = 3971)			
In own dwelling	760	754	1514
In own yard / plot / compound	566	634	1200
Outside the compound	601	656	1257
Access to enclosed bathing area (n = 3971)			
Yes	618	553	1171
No	1309	1491	2800

*Question incorrectly coded in survey - 57 incorrectly coded responses not included.

Appendix 12: Demographic Characteristics of Participants in FGDs Concerning Perceptions and Spillover of the Sundara Grama Intervention

Focus Group Discussions (n=152)	n	%
Sex		
Female	72	47.4%
Male	80	52.6%
Intervention Status		
Intervention	65	42.8%
Control	87	57.2%
Caste		
Scheduled Caste	13	8.6%
Scheduled Tribe	3	2.0%
General Caste	99	65.1%
Other Backward Castes	37	24.3%
Latrine Ownership		
Yes, functional	101	66.4%
Yes, non-functional	9	6.0%
No	42	27.6%
Frequency of Latrine Use for Defecation, among latrine owners		
Always	96	87.2%
Sometimes	10	9.0%
Never	4	3.6%

Appendix 13: Demographic Characteristics of Participants in IDIs and FGDs Concerning Perceptions and Spillover of the Mother's Group Activity from the Sundara Grama Intervention

In Depth Interviews (n=24) and Focus Group Discussions (n=37)	n	%
Female	61	100%
Intervention Status		
Intervention (IDIs)	24	39.3%
Control (FGDs)	37	60.7%
Caste		
Scheduled Caste	6	9.8%
Scheduled Tribe	0	0.0%
General Caste	35	57.4%
Other Backward Castes	20	32.8%
Latrine Ownership		
Yes, functional	36	59.0%
No	20	32.8%
No Response	5	8.2%
Frequency of Latrine Use for Defecation (IDIs Only)		
Always	10	41.7%
Sometimes	1	4.2%
Never	10	41.7%
No Response	3	12.5%

References

ARNOLD, B. F., HOGAN, D. R., COLFORD, J. M. & HUBBARD, A. E. 2011. Simulation methods to estimate design power: an overview for applied research. *BMC medical research methodology,* 11**,** 94.

AUNGER, R. & CURTIS, V. 2016. Behaviour Centred Design: Towards an applied science of behaviour change. *Health psychology review,* 10**,** 425-446.

BANDA, K., SARKAR, R., GOPAL, S., GOVINDARAJAN, J., HARIJAN, B. B., JEYAKUMAR, M. B., MITTA, P., SADANALA, M. E., SELWYN, T. & SURESH, C. R. 2007. Water handling, sanitation and defecation practices in rural southern India: a knowledge, attitudes and practices study. *Transactions of the royal society of tropical medicine and hygiene,* 101**,** 1124-1130.

BARNARD, S., ROUTRAY, P., MAJORIN, F., PELETZ, R., BOISSON, S., SINHA, A. & CLASEN, T. 2013. Impact of Indian Total Sanitation Campaign on Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years Following Programme Implementation. *PloS one,* 8**,** e71438.

BARTHOLOMEW, L. K., PARCEL, G. S., KOK, G., GOTTLIEB, N. H. & FERNANDEZ, M. E. 2011. *Planning health promotion programs: An intervention mapping approach*, Jossey-Bass.

CARUSO, B. A., CLASEN, T., YOUNT, K. M., COOPER, H. L., HADLEY, C. & HAARDÖRFER, R. 2017. Assessing Women's Negative Sanitation Experiences and Concerns: The Development of a Novel Sanitation Insecurity Measure. *International journal of environmental research and public health,* 14**,** 755.

CLASEN, T., BOISSON, S., ROUTRAY, P., CUMMING, O., JENKINS, M., ENSINK, J. H., BELL, M., FREEMAN, M. C., PEPPIN, S. & SCHMIDT, W.-P. 2012. The effect of improved rural sanitation on diarrhoea and helminth infection: design of a cluster-randomized trial in Orissa, India. *Emerging themes in epidemiology,* 9**,** 7.

CLASEN, T., BOISSON, S., ROUTRAY, P., TORONDEL, B., BELL, M., CUMMING, O., ENSINK, J., FREEMAN, M., JENKINS, M. & ODAGIRI, M. 2014. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health,* 2**,** e645-e653.

COFFEY, D., GUPTA, A., HATHI, P., SPEARS, D., SRIVASTAV, N. & VYAS, S. 2015. The puzzle of widespread open defecation in rural India: Evidence from new qualitative and quantitative data. *Working Paper*.

COFFEY, D., GUPTA, A., HATHI, P., SPEARS, D., SRIVASTAV, N. & VYAS, S. 2017. Understanding open defecation in rural India: Untouchability, pollution, and latrine pits. *Economic and Political Weekly,* 52**,** 59-66.

DE SILVA, M. J., BREUER, E., LEE, L., ASHER, L., CHOWDHARY, N., LUND, C. & PATEL, V. 2014. Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials,* 15**,** 267.

DELEA, M., SCLAR, G., WORETA, M., HAARDÖRFER, R., NAGEL, C., CARUSO, B., DREIBELBIS, R., GOBEZAYEHU, A., CLASEN, T. & FREEMAN, M. 2018. Collective efficacy: Development and validation of a measurement scale for use in public health and development programmes. *International journal of environmental research and public health,* 15**,** 2139.

FREEMAN, M. C., GARN, J. V., SCLAR, G. D., BOISSON, S., MEDLICOTT, K., ALEXANDER, K. T., PENAKALAPATI, G., ANDERSON, D., MAHTANI, A. G. & GRIMES, J. E. 2017. The impact of sanitation on infectious disease and nutritional status: A systematic review and meta-analysis. *International journal of hygiene and environmental health,* 220**,** 928-949.

FREEMAN, M. C., STRUNZ, E., UTZINGER, J. & ADDISS, D. G. 2016. Interventions to improve water, sanitation, and hygiene for preventing soil-transmitted helminth infection. *Cochrane Database Syst Rev***,** 1-12.

GARN, J. V., SCLAR, G. D., FREEMAN, M. C., PENAKALAPATI, G., ALEXANDER, K. T., BROOKS, P., REHFUESS, E. A., BOISSON, S., MEDLICOTT, K. O. & CLASEN, T. F. 2017. The impact of sanitation interventions on latrine coverage and latrine use: A systematic review and meta-analysis. *International journal of hygiene and environmental health,* 220**,** 329-340.

GUPTA, A., KHALID, N., DESPHANDE, D., HATHI, P., KAPUR, A., SRIVASTAV, N., VYAS, S., SPEARS, D. & COFFEY, D. 2019. Changes in Open Defecation in Rural North India: 2014-2018.

HIRVE, S., LELE, P., SUNDARAM, N., CHAVAN, U., WEISS, M., STEINMANN, P. & JUVEKAR, S. 2015. Psychosocial stress associated with sanitation practices: experiences of women in a rural community in India. *Journal of Water, Sanitation and Hygiene for Development,* 5**,** 115-126.

HUBBARD, A. E., AHERN, J., FLEISCHER, N. L., VAN DER LAAN, M., SATARIANO, S. A., JEWELL, N., BRUCKNER, T. & SATARIANO, W. A. 2010. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology***,** 467-474.

HULLAND, K. R., CHASE, R. P., CARUSO, B. A., SWAIN, R., BISWAL, B., SAHOO, K. C., PANIGRAHI, P. & DREIBELBIS, R. 2015. Sanitation, Stress, and Life Stage: A Systematic Data Collection Study among Women in Odisha, India. *PloS one,* 10**,** e0141883.

INTERNATIONAL INSTITUTE FOR POPULATION SCIENCES (IIPS) AND ICF 2017a. National Family Health Survey (NFHS-4), 2015-16: District Fact Sheet Puri Odisha. Mumbai: IIPS.

INTERNATIONAL INSTITUTE FOR POPULATION SCIENCES (IIPS) AND ICF 2017b. National Family Health Survey (NFHS-4), 2015-16: Odisha State Report
. Mumbai: IIPS.

MDWS. 2011. *NGP* [Online]. Ministry of Drinking Water and Sanitation, Government of India. Available: http://www.mdws.gov.in/TSC [Accessed 25 November 2013].

MICHIE, S., VAN STRALEN, M. M. & WEST, R. 2011. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science,* 6**,** 42.

MINISTRY OF DRINKING WATER AND SANITATION. *Swachh Bharat Mission- Gramin Dashboard* [Online]. Available: http://sbm.gov.in/sbmdashboard/Default.aspx [Accessed 3 December 2018].

MOSLER, H.-J. 2012. A systematic approach to behavior change interventions for the water and sanitation sector in developing countries: a conceptual model, a review, and a guideline. *International journal of environmental health research,* 22**,** 431-449.

O'REILLY, K. & LOUIS, E. 2014. The toilet tripod: Understanding successful sanitation in rural India. *Health & PLace,* 29**,** 43-51.
ROUTRAY, P., SCHMIDT, W.-P., BOISSON, S., CLASEN, T. & JENKINS, M. W. 2015. Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study. *BMC public health,* 15**,** 880.
SAHOO, K. C., HULLAND, K. R., CARUSO, B. A., SWAIN, R., FREEMAN, M. C., PANIGRAHI, P. & DREIBELBIS, R. 2015. Sanitation-related psychosocial stress: a grounded theory study of women across the life-course in Odisha, India. *Social science & medicine*.
SAUNDERS, R. P., EVANS, M. H. & JOSHI, P. 2005. Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide. *Health promotion practice,* 6**,** 134-147.
SCLAR, G., PENAKALAPATI, G., CARUSO, B. A., REHFUESS, E. A., GARN, J. V., ALEXANDER, K. T., FREEMAN, M. C., BOISSON, S., MEDLICOTT, K. & CLASEN, T. F. Exploring the Relationship Between Sanitation and Well-being: A Systematic Review and Qualitative Synthesis.
SCLAR, G. D., PENAKALAPATI, G., CARUSO, B. A., REHFUESS, E. A., GARN, J. V., ALEXANDER, K. T., FREEMAN, M. C., BOISSON, S., MEDLICOTT, K. & CLASEN, T. 2018. Exploring the relationship between sanitation and mental and social well-being: A systematic review and qualitative synthesis. *Soc Sci Med,* 217**,** 121-134.
SNOWDON, W., SCHULTZ, J. & SWINBURN, B. 2008. Problem and solution trees: a practical approach for identifying potential interventions to improve population nutrition. *Health promotion international,* 23**,** 345-353.
STARR, L. & FORNOFF, M. 2016. Theory of Change: Facilitators's Guide. *In:* USAID (ed.). Washington DC.
TORONDEL, B. 2015. Coverage, Use and Faecal Sludge Management of Latrines in Orissa, India: a Cross-Sectional Study. *Water and Health Conference: Where Science Meets Policy.* University of North Carolina, Chapel Hill.