Vaparshun: Improving toilet use in rural India

Pre Analysis Plan – July 2018

Study led by:
Indian Institute of Public Health Gandhinagar and London School of Hygiene and Tropical Medicine

Study funded by:
The International Initiaitve for Impact Evauation

1. Intervention

1.1. Theoretical framework

Vaparshun! used the Behaviour Centred Design (BCD) framework and theory of change (ToC) to design its intervention (Aunger & Curtis 2016). BCD has provided guidance to the design and delivery of successful behaviour change interventions in India for handwashing with soap (Biran et al. 2014), ORS use in Zambia (Greenland et al. 2017), food hygiene in Nepal (Gautam et al. 2017), infant feeding behaviour in Indonesia (White, Schmidt, et al. 2016), sanitation promotion in Tanzania, post-operative exercise Ireland (Doyle 2015) and has also been applied to the marketing of sanitation and hygiene products (clients include Kimberly Clarke, GoJo and Unilever).

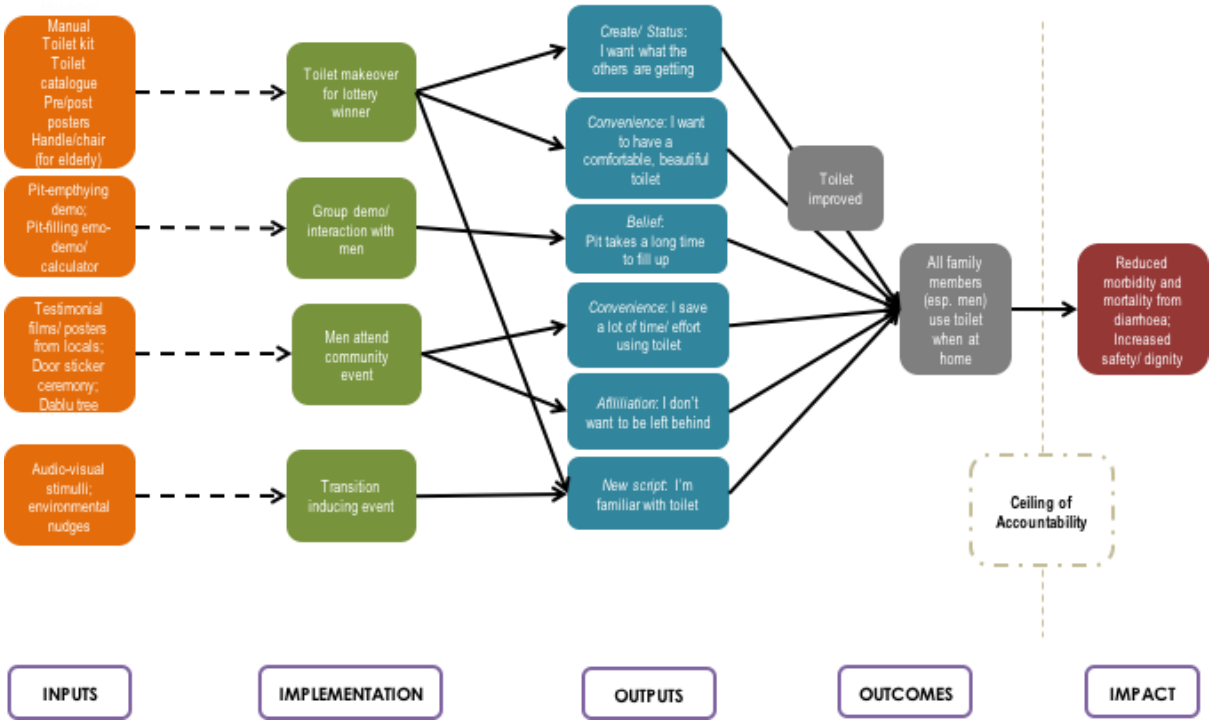
BCD is a logical and comprehensive approach to designing and evaluating behaviour change programmes (Aunger & Curtis 2016). This model, derived from reinforcement learning theory (Sutton & Barto 2017), develops a fundamental taxonomy of needs based in evolutionary biology, shows how the disruption of 'behaviour settings' is key, and sets out the steps involved in programming for behaviour change. In addition, it provides means of identifying the levers to change behaviour, provides guidance for intervention and tool design process for creating, delivering and measuring behaviour change programmes (Aunger & Curtis 2015). Thus, a BCD ToC indicates how an intervention aims to change the environment of the target population, how exposure to this environmental change influences the psychology of those in the target population, and how this prompts them to change their behaviour (which, in turn, impacts health and well-being). The intervention has to initiate this cascade of changes by providing activities that are surprising, cause revaluation of the target behaviour and affect the performance of the behaviour in its setting.

Vaparshun! intervention aims to inspire the target audience to revalue their toilets by recognizing that they provide benefits associated with the motives of hoard, create, convenience (comfort) and affiliation, and provide a reward pathway for transitioning to a new toilet use routine (please see Figure 1).

1.2. Intervention summary

Vaparshun! is using the BCD framework and ToC (Please see Figure 1) to design its intervention.

Figure 1: Vaparshun! Theory of Change



Our intervention consists of four different streams of activity, each of which has its own logic in the ToC (see Figure 1). The outcome is that family members and men improve and use their contractor-built toilets.

1. Create Motive:

Toilet Makeover: Conduct a lottery and perform makeovers of select government built toilets in the village, with the involvement of the community. Demonstrate improvements in comfort (light, space, ventilation, latrine chair/handle) and aesthetics (stencil painting of door and walls).

Challenge/Opportunity: Many of the 'contractor' toilets are built with low engagement from family members and are uncomfortable to use. People are left with toilets they are not proud of or engaged with.

Insight (from 'makeover' trial): If families invest in creating an attractive toilet they are be more likely adopt and use them.

Inputs: Materials for the physical and aesthetic improvement of a toilet, manuals for conducting the community event.

Outputs: Greater engagement with, and pride in, the toilet after makeover; others in the village inspired to conduct their own makeover. Those who use the upgraded toilet find it a more comfortable experience than they had expected causing reinforcement learning.

2. Hoard motive:

Pit Emptying Demo and Pit Filling Estimation Demo" A community event-based experience of the 'real' aspects related to pit filling/emptying designed to graphically overcome their perceptual barriers (e.g., squeezing a watermelon to show how little material there is in faeces).

Challenge/Opportunity: People over-estimate the speed at which a pit fills and are uncertain about the emptying process. Therefore they hoard the 'limited' pit space by using the toilet only partially.

Insight from FR: There are gaps between perception and reality which can be addressed. For example; water doesn't stay in the pit but seeps into the soil, faeces are composed mainly of

water, decomposition reduces volume, compost doesn't smell and twin pits can be used interchangeably forever.
Inputs: Scripts for 'emo-demos' (emotional demonstrations).
Outputs: Participants are less anxious about pit filling and emptying.

3. Affiliation/convenience motives:
Community Motivational Events: Small and large community events such as street plays, films, posters and pledging activities to bring alive convenience/comfort motives by amplifying problems associated with OD and rewards of using toilets; use of affiliation through testimonials films, posters, village maps, etc.

Challenge/Opportunity: Even if the barriers around pits and toilet comfort are addressed, it may still not be enough to motivate men with entrenched habits of OD to start using toilets.
Insights: Convenience/comfort can be a powerful drive for toilet usage. Those who use toilets in the village (women, children and elderly) find it is much more convenient and therefore do not return to OD. However, men who are non-users may not have experienced this and need to be convinced. Affiliation can be another strong drive for toilet usage. It is possible to exploit the emerging norms of toilet use and encourage men 'not to be left behind'.
Inputs: Scripts, props, invitations, loud hailers, audio-visual equipment, etc.
Outputs: Men use toilets because they 'get' how convenient they are, and so as not to be 'left behind'.

4. Transition Nudge
This would include incentives for usage and/or environmental or audio-visual nudges to support initialization of toilet use and habit formation. Every household can participate in the game. Those who have used the toilet 100% enter into a lottery to win. There is no cash incentive. The gifts would be either for adults (e.g. mobile phone) or children (e.g. bicycle). Every household can participate in the game. Those who have used the toilet 100% enter into a lottery to win. There is no cash. The gifts would be either for adults (e.g. mobile phone) or children (e.g. bicycle).

Challenge/Opportunity: Those who use toilets for a specific period tend to stick with the habit; however, some people, especially men, do not try out the toilet or find the first experience unpleasant.
Insight: Reward the use of toilets for a specific period so new habits can form.
Inputs: Stimuli and nudges.
Outputs: The entire family, especially men, form the habit of using a toilet.

In addition, village authorities will also be recruited to support delivery of the intervention.

2. Evaluation Questions and Hypotheses

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

Outcome Evaluation

1. How far can an innovative theory-based, scalable intervention improve toilet use behaviour of all family members amongst households with government/contractor-built toilets in areas of high coverage in rural Gujarat? (Primary outcome)
2. How much can the intervention affect toilet use in men?

3. How comparable are the study methods used (survey, sticker diaries and structured observations)?

Process Evaluation

4. Which components worked or did not work as expected according to the Theory of Change (ToC), and where did unexpected consequences arise? (Process evaluation)

Based on the formative research, we found that HHs with a government built toilet were less likely to use their toilets regularly. However, HHs with self-built toilet had better quality construction and features. They were also more likely to use it regularly. Therefore, we aim to focus on HHs with contractor built/government built toilets.

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

The proposed intervention aims to inspire the target audience to revalue their toilets by recognizing that they provide benefits associated with the motives of hoard, create, convenience (comfort) and affiliation, and provide a reward pathway for transitioning to a new toilet use routine.

Hypothesis: Vaparshun! uses the Behaviour Centred Design (BCD) framework and theory of change (ToC) to design its intervention (Aunger & Curtis 2016). Vaparshun's theory of change consists of four different streams of activity, each of which has its own logic in the ToC (see Figure). The outcome is that family members and men will improve and use their contractor-built toilets, as measured by follow up evaluation 2 months post-intervention delivery. The intervention components of toilet makeover demonstration, pit emptying demonstration, community motivational events to create new social norms and transition nudges aiming to change the environment of the target population by inspiring them to revalue their toilets by recognizing that they provide benefits associated with the motives of hoard, create, convenience (comfort) and affiliation, and provide a reward pathway for transitioning to a new toilet use routine, which we hope households will continue to practice.

Our assumption is that exposure to this environmental change will influence the psychology of those in the target population (all members in a household, especially men) to value their toilets, to modify their government built toilets (by making changes to the infrastructure, making toilets beautiful by painting the walls and installing features like handle, ventilation, light, toilet chair for disabled or old people etc. that enhance the user experience). This will prompt them to change their behaviour from open defecation to using their contractor built toilets (which, in turn, may impact health and well-being in the long term). The intervention will initiate a cascade of changes by providing activities that are surprising, cause revaluation of the target behaviour and affect the performance of the behaviour in its setting.

As the intervention will be delivered at cluster level a cluster randomised trial is the most suitable study design. The intervention to be studied will be delivered to and affect households with contractor built toilets, rather than individuals. Since people within a cluster are more likely to be similar, the outcome for each participant cannot be assumed to be independent of that for any other participant. The CRT will be an assessment of a complex intervention (addressing the complex determinants of low toilet use through activities delivered at cluster level), with the analyses of endpoints measuring multiple behaviours. The intervention will not measure health outcomes.

Primary hypothesis:

- An innovative theory-based, behavioural intervention can improve toilet use amongst households with government/ contractor-built toilets in high coverage areas of rural Gujarat. Toilet use for the primary outcome is defined as the proportion of households where all members use the toilet (the last time they defecate), measured through self-report or as reported by the questionnaire respondent on behalf of other members. We will use the 3ie standardised questionnaire to measure this outcome.

Secondary hypothesis:

- The intervention increases toilet use among household members as observed through structured observation and the newly developed sticker diary methodology.
- The intervention increases toilet use among men as observed through structured observation and the newly developed sticker diary methodology.
- The exposure to intervention (toilet makeover, emo-demo's and community events targeted at men) will exposure to our intervention will lead people to valuing their toilets and adoption of improved practices and use (less anxiety around pit filling (hoard), recognizing that they provide benefits associated with the motives create, convenience (comfort) and affiliation (pathways to change). Measured through questionnaire survey, structured observations and newly developed sticker diary methodology.

3. Sampling

3.1. Sampling frame

Sampling frame for the Census and Baseline will be the 2011 Census. The target enrolment for each of the clusters will be (45 control and 45 intervention) will be 30 HHs per cluster. Only one individual per HHs will be sampled for recruitment. Eligibility criteria for participation in the survey included: resides in the home is above 18 years of age per confirmed date of birth. Written informed consent will be obtained from all participants in the survey and will be offered in the local language Gujarati. The survey will be interviewer-administered in Gujarati using a tablet in which the interviewer will directly enter responses into a tablet.

The CRT will involve three blocks (taluks) of Bhavnagar district in Gujarat. Baseline data will be collected from randomly selected eligible households in identified clusters of Bhavnagar, Gujarat. Bhavnagar has relatively high rates of toilet coverage but also high rates of non-use of toilets (particularly by men, and particularly with respect to contractor-built toilets). Bhavnagar is typical of rural India in many respects with high levels of agricultural production alongside the rapid growth of industry (for example onion processing, ship-breaking and diamond polishing). It will happen in the context of existing Government efforts to improve sanitation coverage. The eligible population for the study is households that have functional latrines (defined by having a pit, pan, and pipe connecting the two).

Information will be collected on sanitation coverage and health indicators etc. through Census. As of now, all districts in Gujarat have been declared ODF. However, as per our formative research and discussion with partners working in the field, not all toilets built by the government support are being used and in some villages even if toilets are sanctioned by the government but the construction is pending, villages were declared ODF.

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

Government built/ contractor built functional latrine (defined by having a pit, pan, and pipe connecting the two).

3.1.2. What are the main characteristics of your population?

The CRT will involve three blocks (taluks) of Bhavnagar district in Gujarat. In 2011, Bhavnagar had population of 2,880,365 of which male and female were 1,490,201 and 1,390,164 respectively. Average literacy rate of Bhavnagar in 2011 were 75% compared to 66% of 2001. Male and female literacy were 84% and 66% respectively. Total literate in Bhavnagar District were 1,887,255 of which male and female were 1,087,371 and 799,884 respectively. In 2011 census, child sex ratio is 891 girls per 1000 boys. More than 91 % of the population is Hindu, followed by 7% Muslims and remaining population includes Jain, Buddhist, Christian and Sikhs. Bhavnagar has relatively high rates of toilet coverage but also high rates of non-use of toilets (particularly by men, and particularly with respect to contractor-built toilets). Bhavnagar is typical of rural India in many respects with high levels of agricultural production alongside the rapid growth of industry (for example onion processing, ship-breaking and diamond polishing).

3.1.3. What is the expected sample size?

30 HHs per cluster in 45 treatment and 45 control clusters. Clusters are defined as villages. The study uses the Census definition of household i.e. a 'household' is a group of persons related or unrelated or a mix of both, who normally live together and take their meals from a common kitchen, unless the exigencies of work prevent any of them from doing so. However, if a group of unrelated persons live in a Census house but do not take their meals from the common kitchen, then they are not constituent of a common household.

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

Sample will be HHs with a government built/ contractor built latrine.

Clusters in this study are defined as villages (with up to 200-300 HHs) with high toilet coverage (>75%). In larger villages (>300 HH), consisting of several hamlets (smaller settlements, usually a sub-division of a village) which are spread out, only one hamlet with high contractor built toilet coverage will be randomly selected for intervention and evaluation. All households in the clusters are eligible for participation. Clusters with high toilet coverage i.e. >75%. Households' within these clusters may include HHs with (government built and self-built) or without toilets. The study uses the Census definition of household i.e. a 'household' is a group of persons related or unrelated or a mix of both, who normally live together and take their meals from a common kitchen, unless the exigencies of work prevent any of them from doing so. However, if a group of unrelated persons live in a Census house but do not take their meals from the common kitchen, then they are not constituent of a common household.

Households will be recruited based on study selection criteria which includes a shared kitchen, having a government built/ contractor built, functional latrine. A functional latrine includes having 1) a pan that is not broken, and 2) a functional connection to a pit (single or twin pits) that exists. This is a subset of the total population living in the area, as the latter also includes households that built toilets on their own initiative without external funding/support and households that do not have any toilet access.

3.1.5. Please describe the anticipated subgroups, which will be studied, if relevant. For quantitative sub-group analysis, please explain how you are powered to do so. If you intend to conduct qualitative sub-group analysis, please clarify how you will do this.

Latrine use among men will be studied in sub-group analysis. This is because during formative phase this group was found to be most reluctant to use a toilet. The power of this analysis depends on the proportion of households with male inhabitants, and the proportion of households where some males do not use the latrine. This is difficult to anticipate at this stage. Power may be only slightly lower than the main analysis as most households will have male members, and men among the household members are most likely not to use the latrine.

Note: Since behaviour change interventions require village-level clustering to prevent spillovers, studies will likely not be adequately powered to conduct subgroup 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.

3.2. Statistical power

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

For our sample size calculation we assume that 65% of households with a government supported latrine will be using this latrine consistently. This figure is based on our formative research that found that about 44% of households have members who go for open defecation. We expect full-use households to increase to 75% after the intervention, which is an effect size of public health interest. Using a sample size formula for the comparison of two proportions results in 349 households per arm to detect this difference with 80% power and an alpha of 0.05. Assuming an ICC of 0.1 results in relatively large design effects, which means that sampling many households in a single village will be inefficient. We are choosing 30 households per cluster as enrolling more than that only marginally reduces the number of required clusters. As a result, we will enrol 45 villages per arm, and 30 households per village at a design effect of 3.9.

3.2.1.1. What are your assumptions about your alpha level?

0.05

3.2.1.2. What are your assumptions about your statistical power?

80%

3.2.1.3. What are your assumptions about variability in your effect size? The effect size will be relative to the variability in the population and sample. Practically, this is a justification of your chosen intra-cluster correlation coefficient and standard deviation. You may consider presenting references to previous literature (including rice's work) in support of this point.

Our outcomes are binary hence there is no requirement of specifying a standard deviation. We chose the ICC based on our data on reported latrine use from the Orissa trial. The ICC in that trial was 0.106. These data are unpublished to date but we are happy to share them if needed.

3.2.1.4. How many clusters will you have?

There are 90 clusters for outcome evaluation i.e. 45 clusters in intervention arm and 45 clusters in control arm.

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

We will have 30 HHs per cluster in each intervention and control arms (a total of 2700 HHs).

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

The sample size is sensitive to changes in the ICC. A lower effect size would also lead to a larger sample size but we agree a lower effect size would be of little public health relevance and is hence not accounted for.

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

NA.

3.3. Assignment to treatment

3.3.1. How will individuals be assigned to treatment and control conditions? Please list the characteristics and justification on which you will match the clusters?

We currently favour stratified randomisation with strata chosen based on variables deemed predictive of the outcome, or identified as such in the baseline survey. Most likely we will randomise within 5-10 strata of village level toilet coverage (depending on the distribution of this indicator). We may add substrata of a socio-economic / socio-demographic summary indicator and randomise within. In addition we may reject randomisations where a relevant number of intervention villages is within 3km of a control village. Pair matching remains an option but we do not currently see a need for it. The final decision will be made based on the census and baseline data. Matching / stratifying variables may include for example, population density, toilet coverage (government/contractor built toilets), level of education, and the number of different sub-castes/ communities (size of schedule caste community vs to other backward classes vs general category).

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

A detailed process evaluation, following the theory of change will be conducted. Activity logs will be checked as well.

4. Data Collection

4.1. Primary data collection instruments

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

Quantitative methods (Outcome evaluation)

- Sticker Diaries (30 HHs per cluster in 90 clusters)
- Structured Observations (200 HHs per intervention arm i.e. 400 HHs in total)
- Survey questionnaire (30 HHs per cluster in 90 clusters)

Qualitative methods (Process evaluation)

- Field observations

- Semi structured interviews
- Focus group discussions

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

Instrument	**Respondent**	**No.**
Survey	Head of the household or an elder member of the household. Caretakers of children under 7 years would be interviewed and each present member of the HH will be asked questions about his/her defecation behaviour.	Baseline: 10 HHs per cluster in intervention and control arms Outcome:30 HHs per cluster in intervention and control arms
Sticker diaries	One member only per HH (men/ women)	30 HHs per cluster in intervention and control arms
Structured observations	Observations only	400 observations (200 per study arm) Structured observations will be conducted on all members present in the household.
In-depth interviews	Implementation team, village leaders and representatives of the target population (men/women)	20 interviews
Focus Group Discussions	Representatives of intervention recipients	12 focus group discussions
Field observations	5 per intervention component	20 field observations of intervention events. This involves observation of how intervention is being delivered, response and engagement of participants and their reaction.

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

Instrument (Outcome Evaluation)	**Indicator**
Survey	Primary outcome and sub group analysis
Sticker diaries	Primary outcome and sub group analysis
Structured observations	Primary and sub group analysis
Instrument (Process Evaluation)	**Information to be obtained.**
In-depth interviews	Recruitment strategies, ffidelity, dose, pathways to change
Focus Group Discussions	Pathways to change, Reception- participant engagement and acceptability and participant response
Field observations	Recruitment strategies, ffidelity, dose

We will use structured observations and sticker diaries to assess how comparable are the study methods used (survey, sticker diaries and structured observations)

Outcome	Indicator	Definition	Measurement
Households with contractor-built toilets in intervention clusters report toilet use by all family members during the last time they defecated and in the last 24 hours	Post intervention reported use of HH toilets by members of household Number of members in a household that report toilet use (during all times in last 24 hours and the last time they defecated) compared with total number of members in a household.	Members in a household that report toilet use during last time they defecated. This will be self-reported using a standardised questionnaire (in a household roster for each household member individually in households that own government/ contractor built toilets) for all members in a household. In case members are not present, other family members or the primary respondent will be asked about where the person defecated last time. Mothers will be asked about the defecation behaviour of younger children. Information about all members in a household will be obtained.	Assessed 2 months after intervention delivery. 3ie prescribed Survey questionnaire (30 HHs per cluster in 90 clusters)

Table 1: Process evaluation dimensions

Component	Questions	Intervention	Method	Data to be collected	Phase	# Clusters	Data Analysis
Objective 1: **To understand the context and participant recruitment process:**							
Context Elements of physical (location, staff skills, resources), social (culture, caste) and political environment (existing programs, elections) that may directly/indirectly affect the intervention delivery and assess generalisability	What contextual factors (coverage of toilets,existing programs, availability of men to participate in intervention) in Bhavnagar enabled or impeded the implementation of intervention in the setting? How did it affect the delivery of intervention? What was done to address those factors (mid-course corrections)?	All four components	Review of project reports Interviews with implementers	Key features of clusters, ongoing activities in those clusters, data on delivery and receipt in those clusters. Variables include influence of external programmes, secular trends in related behaviours etc.	Through intervention lifecycle	4 clusters and overall experience from select clusters	Description of the context External and internal influence or contanimation if any.
Recruitment Enrolment or mobilisatio of participants into the intervention activities.	How were participants recruited for each intervention component? Did it affect the reach? Which sub group of individuals were more or less likely to be recruited? Why? Was the	All four components	Semi structured interviews Implementation reports Field observations Routine data	Steps taken to recruit participants, challenges faced if any and how were they addressed Patterns of reach	Inception phase of the intervention and 2 week post intervention delivery phase	2 intervention clusters	Description of activity areas, participnt selection, recruitment and mobilization strategy

Component	Questions	Intervention	Method	Data to be collected	Phase	# Clusters	Data Analysis
	recruitment process consistently appied across all clusters?						
Objective 2: **To understand the factors that affected implementation of the intervention:**							
Fidelity Adhering to protocol of intervention delivery and competency to deliver the intervention by implementers	How (structure, sequence and content of 4 intervention activitie- toilet makeover, emo-demo, community events, transition nudges) and what was the quality of intervention delivered as compared to intended plan? Did it depart from what was originally intended? If yes, how and what explains it?	All four components	Routine data from implementing partner Observations Semi structured interviews Implementation reports Field observations	Nuber and type of interventions delivered Checklist and creative partner's (Upward Spiral) perspective on content delivery Success and challenges faced by implementers Participant perspectives on the content and quality of intervention activities Any deviations from planned activities? Reasons?	Beginning and midway	All 90 clusters 2 clusters	Actual no of activities delivered (extent) over the planned number of activities (fidelity of implementaiton) Conent, timings and locaitons of interveniton delivery Methods of delivery and explanation provided
Dose *Quanity/number of activities/events delivered*	*Dose delivered:* What was delivered to the participants and what prortion of the intended intervention was	All components	Routine data (activity logs) from implementing partner Survey	Recall and recognition of intervention components delivered and messages accompanying those creative concepts.	2 months post intervention	All 90 clusters 2 intervention clusters	Proportion of participants that correclty recall key messges and activities delivered under Vaparshun!

Component	Questions	Intervention	Method	Data to be collected	Phase	# Clusters	Data Analysis
	actually delivered to the intended audience? *Dose received:* what proportion of creative material/messages did partiicpants receive?		Focus Group Discusison				What percentage or creative materials and messages was used by participants?
Reach *The extent to which the intended audience particpates in the intervention.*	To what extent does the intervention contact target population? Which sub-groups (men, women, young people, older people) are exposed to/participate in the intervention events? What explains the pattern of reach?	All four components	Household survey to understand exposure to the intervention components	Proportion of sample reporting participation in each intervention activity in intervention arm Sub-group of participants that attend each intewrvention component	2 months post intervention delivery	All 90 clusters (30HHs per cluster)	Number of events/activities over the number receiving the interventions Barriers and facilitators
Objecitve 3: **To understand the hypothesised pathways to change:**							
Participant engagement and response *The extent to which the target population engages with the*	Did the intervention meet the information needs of the target population? Do they understand and retain the key messages related to pit fillling	All four components *Toilet makeover & community events targeted at men*	Semi structured interviews Implementation reports Field observations Focus group discussions	Comprehension of messages and response to the intervention components Verifications questions will provide more detail about the event	Through intervention lifecycle Midway and 2 weeks post	4 clusters	Proportion of sample able to recall messages and recognize intervention concepts

Component	Questions	Intervention	Method	Data to be collected	Phase	# Clusters	Data Analysis
intervention events/activities.	esimation, conveneint and comfort of using tilets and toilet makeover? Did the implementers accept the activities delivered?			attended/ exposure to intervention component. Participant and implementer perspectives on messages and activities delivered as part of the intervention	intervention delivery		Preferred intervention components
Mediators *Intermediate processes that explain the change in outcome.*	Do behavioural determinants (affiliaiton, convenience and comfort) change as a result of the intervention delivery according to Vaparshun's theory of change?	*Toilet makeover & emo- demo, community events targeted at men*	Survey Focus group discussions (recipients and non-recipients) In-depth interviews	Indicators related to hypothesized behavioural determinants - understanding the motives associated with intervention components.	2 weeks post intervention delivery	All 90 clusters (30HHs per cluster) 4 clusters	Receptiveness to the intervention Particionants' response on norms and motivators i.e. desire for imprived toilets, belief that toilet use saves time and is conveneint and knowledge about time it takes for a toilet pit to fill.

The Process evaluation will employ a combination of data sources analysed according to the categories in the table and will be published. Process evaluation data will be analysed in two stages as done in other trials (Oakley et. all, 2006), (Elford J et al, 2002). In the first stage, process data will be analysed separately from the outcome data to minimize bias in interpretation of results. Descriptive statistics for implementation of intervention components such as number of sessions delivered, number of events held, and number of participants will be used to characterize the sample and to analyse the process measures.

In the second stage, we will conduct analysis to understand the relation between study outcomes and variation between the quality and extent of implementation of the intervention (fidelity, reach). This will also be used to understand the process that might mediate the observed relation between intervention components and outcomes (pathways to change) and to understand if and why toilet use among men differs in the intervention arms.

4.1.4. How will each instrument be developed?

Survey: We will use measurement questions suggested by 3ie and add them to the instrument developed during the formative research phase. The questionnaire will be field tested before the survey.

In-depth interviews: A questionnaire will be developed to assess the key components of the pathways of change**.**

Structured observations: These data will be collected through structured observation by a team of female enumerators. A structured observation checklist will be developed and enumerators will be trained on conducting structured observations. Observations will take place when most householders are present and when the behaviours of interest are likely to be seen. Structured observation requires a trained enumerator to visit a household around dawn as daily activities begin. The fieldworker remains at the household for 2 hours recording defection practices.

Sticker diaries: LSHTM has recently used sticker diaries to evaluate a school-based handwashing programme in the Indian state of Bihar. The study confirmed that over-reporting of desirable behaviour is much reduced, although not eliminated. We will develop stickers of all the different tasks performed by people in target population. One respondent from each household will be asked to create a "diary" of daily tasks under the guidance of the enumerator. The diary sheet is filled using stickers illustrating different tasks. To mask the tasks of interest (here latrine use), respondents **will be** provided with a comprehensive list of stickers reflecting activities that they may have undertaken the previous day, covering a wide range of activities including the daily regimen and personal care of the respondent. Participants will be able to choose different stickers for defecation (open defecation, latrine), providing a secondary rapid indicator of toilet use behaviour.

FGD Guide: FGD guide will be developed for discussions with intervention recipients, field workers implementing the intervention and also other key informants.

Field observations checklist: A structured reporting form will be developed to record details about the setting, fidelity according to criteria related to adherence to the protocol, the competence of delivery and participants' reactions to the event.

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

We will use android based smart phones/ tablets for data collection.

We will first use less obtrusive methods (sticker diary and structured observations) before the questionnaire to avoid the objective of the survey becoming clear to the study participants through direct questioning.

For local adaptation of the instruments translation and back-translation and checking of cultural and functional equivalence will be performed, again using inputs from implementing partners and the community. For adaption of the tools, they will be pilot tested and further modifications will be made to establish local norms. All field enumerators will be adequately trained and monitored with interrater reliability and accuracy testing performed.

Research will be carefully framed to assure anonymity of responses and neutrally worded questions and to reduce social desirability bias.

Staff will be blinded as to intervention status to reduce bias. Staff will be adequately trained on interview techniques/ filling questionnaire survey. Staff will be observed for short periods, with observations being conducted before surveying.

During formative research, we tested the survey instrument for measuring latrine use.

4.2. Secondary data sources

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

NA

5. Analysis

5.1. Outcome Variables

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

Outcome	Description	Hypothesis	Level
Households with contractor-built toilets in intervention clusters report toilet use by all family members during the last time they defecated and in the last 24 hours	Proportion of households with contractor-built toilets in intervention clusters compared to proportion of households with contractor-built toilets in control clusters that report toilet use by all family members in a household	The proposed intervention (through 4 intervention components) aims to inspire the target audience to revalue their toilets by recognizing that they provide benefits associated with the motives of hoard, create, convenience (comfort) and affiliation, and provide a reward pathway for transitioning to a new toilet use routine.	Household level
Latrine use among men *(sub group analysis)*	Proportion of households with contractor-built toilets in intervention clusters compared to proportion of households with contractor-built toilets in control clusters that report toilet use by men	Men are inspired to revalue their toilets by recognizing that they provide benefits associated with the motives of hoard, create, convenience (comfort) and affiliation, and provide a reward pathway for transitioning to a new toilet use routine.	Household level

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

Response: We do not intend to adjust primary outcomes for covariates.

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

We do not plan to do so.

5.2. Qualitative Analysis

What questions will be analysed using qualitative methods? Please also describe the qualitative methods that will be used (e.g. content analysis with criteria for codification).

An interview guide will be prepared to facilitate all interviews and discussions. These discussions and interviews will be voice recorded and transcribed verbatim, then analysed thematically following the six-

step method of Braun and Clarke which includes familiarisation with data, generating initial codes, searching for themes, reviewing themes, defining and naming themes and writing the report.

Objectives and research questions

Objective 1: To understand how context and participant recruitment process affects change.
Research questions:

1.1 What were the key contextual factors at time point of intervention (other programmes, events, socio-political, demographic, cultural factors) that might have influenced implementation and/or outcomes?

1.2 How were participants recruited?

Objective 2: To understand the implementation and delivery of Vaparshun.
Research questions:

1.1. Was Vaparshun intervention delivered as intended (fidelity- quality and extent)

1.2. What was the quantity of intervention delivered (dose delivered and dose received)?

1.3. Did the target audience come into contact with the intervention and how (reach)?

Objective 3: To understand participant engagement, response and hypothesised pathways to change.
Research questions:

3.1. Does exposure to the intervention components affect behavioural motives (i.e. enhance status, affiliation and convenience) among men in favour of toilet use (outcome of interest)?

3.2. Do these motives mediate any observed relation between intervention (toilet makeover, emo-demo- community events) and outcome (i.e. improved toilet use)?

3.3. What were the unexpected consequences?

Reference: Clarke, V. & Braun, V. (2013) Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. The Psychologist, 26(2), 120-123

5.3. Quantitative Analysis

5.3.1. Balance Checks

5.3.1.1. How will you check balance between treatment and control groups? Please specify the statistical test used to check for balance, as this is the main point of a pre-analysis plan. Additionally, please clarify why the same households are not being sampled twice; attrition could also be in the form of seasonal migration at the village level.

We will compare main socio-economic and demographic variables across arms without using statistical tests, as this is part of a randomised procedure. We consider it unnecessary to conduct statistical tests to check for balance in randomised controlled trials (see for example the CONSORT statement: “significance testing of baseline differences in randomized controlled trials (RCTs) should not be performed, because it is superfluous and can mislead investigators and their readers” *Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. 2010;2010(63):e1–37.).*

Households are surveyed twice: before the intervention in the form of a census. After the intervention to measure the outcomes. As discussed we do not measure the outcomes at baseline in the households surveyed after the intervention. Please see the flow diagram of the study attached with this submission. We avoid measuring the outcomes twice in the same households (e.g. at baseline and follow up) as this risks reactivity.

We will decide on which randomisation method to use after receiving the census data. Restricted randomisation will assure balance on the variables we use for stratification. If there is no balance in one randomisation round, then we will re-randomise. However, whether or not there is balance will not be decided based on significance tests, but based on pre-set limits of what is deemed acceptable imbalance for each variable included in the restricted randomisation. After the study is completed, it will be judged

based on whether the difference between groups is deemed serious enough. This is similar to assessing confounding for which also no significance tests exists.

If there is no balance in one randomisation round, then we will re-randomise. This will be decided based on pre-set limits of what is deemed acceptable imbalance for each variable included in the restricted randomisation.

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

Main socio-economic and demographic variables. Please clarify what is meant by "specification" in this context. Do you mean model equation?

Stratified design:

$$P(Y_i|T_v, S) = a \cdot S + b \cdot T_v$$

Where **S** is a matrix of indicator variables for all strata used in the randomization and *a* is a vector of coefficients for stratum-specific fixed effects. T denotes treatment, b treatment effect, *i* and v are indices for household and village.

Matched design:

$$P(Y_i|T_v, M) = a \cdot M + b \cdot T_v$$

Where **M** is a matrix of indicator variables for all strata used in the randomization and *a* is a vector of coefficients for stratum-specific fixed effects. T denotes treatment, b treatment effect, *i* and v are indices for household and village.

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? If your treatment and control groups are imbalanced at baseline, the treatment is not the only difference between them, which could confound your results.

Imbalances are unlikely to affect the main analyses especially since we use some form of restricted randomisation. We may however include variables with major imbalances in secondary analyses. For the primary analysis we do not wish to adjust the effect for baseline imbalances as this would go counter the idea of the randomised design and is not commonly done in randomised controlled trials in public health. We cannot see any circumstances under which we would consider specifying the possibility to adjust the main analysis (primary outcome) in the protocol. There is however no problem with doing such additional analysis as a sensitivity analysis. We will use multivariable regression analysis for these purposes.

The unadjusted primary endpoint analysis is what counts and is what will be emphasised in the paper to be written. If sensitivity analyses do not confirm the primary endpoint analysis it simply means that we are less confident in the results especially if other trials when combined in a systematic review should show results different from the primary endpoint result.

5.3.2. **Contamination:** How will you detect and manage any potential differential contamination between treatment and control groups?

Response: A minimum 3 km distance will be ensured between the boundaries of intervention and control villages. This will be achieved in a first step by randomising whole panchayats, not villages within panchayats, whilst only choosing one village per panchayat for the study. In cases where an intervention village is still less than 3km away from a control village (even though in a different panchayat), we will randomly select a new panchayat.

5.3.3. Attrition

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

We do not expect attrition as we do not sample the same households twice. Households included at baseline will be excluded at follow up.

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

NA

5.3.3.3. How does expected attrition change your power calculations?

NA

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?

NA

5.3.4. Missing Data

How will you deal with incomplete or missing data?

We will explore missingness for imbalances across arms. We may resort to imputation methods if missingness turns out a major issue.

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?

Primarily as intention to treat. We will calculate prevalence differences using GLM with binomial distribution and identity link.

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

Stratified design:

$$\boldsymbol{P(Y_i|T_v, S) = a \cdot S + b \cdot T_v}$$

Where **S** is a matrix of indicator variables for all strata used in the randomization and *a* is a vector of coefficients for stratum-specific fixed effects. T denotes treatment, b treatment effect, *i* and v are indices for household and village.

Matched design:

$$\boldsymbol{P(Y_i|T_v, M) = a \cdot M + b \cdot T_v}$$

Where **M** is a matrix of indicator variables for all strata used in the randomization and *a* is a vector of coefficients for stratum-specific fixed effects. T denotes treatment, b treatment effect, *i* and v are indices for household and village.

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?

We will attempt IV regression wile accounting for the limitations of this method in cluster randomised trials where observations within a cluster are not necessarily independent.
The intervention design and piloting is currently being finalised. Once we have finalised the design, we will decide which components are “essential” and for which exposure should be as high as possible. Most likely we will use exposure to cluster level activities to define whether a household was exposed or not. Household level activities that only target selected households will not be used for such purposes.

After discussion with Richard Hayes (LSHTM) we have decided to delete CACE analysis from this study on the basis of assumptions that are unmet in cluster randomised trials.

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

After discussion with Richard Hayes (LSHTM) we have decided to delete CACE analysis from this study on the basis of assumptions that are unmet in cluster randomised trials.

5.4. Heterogeneous Effects

Note: Since behaviour change interventions require village-level clustering to prevent spillovers, studies will likely not be adequately powered to conduct subgroup 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?

Men were found to be an important stakeholder group for the intervention based on our formative research and review of background literature. We may conduct subgroup analyses by gender.

5.4.2. What is the broad theory of action that leads you to anticipate these effects?

Please provide a more detailed explanation here.

Men are primarily responsible for building toilets in homes and often ten to be the ones defecating in the open.

Vaparshun’s theory of change, outlines the steps and hypothesised mechanisms of change towards improving toilet use among all members in a household.

The intervention aims to increase toilet use among all members in a household (especially men) by delivering a cascade of activities, at the cluster level, that will help people understand (functionality, benefits and features) and value their toilets. We hypothesise that exposure to our intervention will lead people to find it convenient and comfortable to use their government/contractor built toilets and will make them usable by carrying out suitable repair and/or modifications (i.e. toilet makeover) of the structure (such as painting walls, creating ventilation, installing tap/water station, handles, toilet chairs for differently abled/ old people). The intervention will deliver components (pit emptying/ filling emotional-demonstration, transition nudges and community events) that we anticipate will make people feel less anxious about pit filling and emptying, which are likely to hamper their motivation to use a toilet, and will reduce the tendency to ‘hoard’ (i.e. save it for later) pit space (due to fear of pit filling up quickly and the anxiety of emptying it in absence of available services in the area). We hypothesise that this will make the experience of toilet use comfortable and desirable and will lead to changes in behaviours, such that toilet use becomes ‘normal’ for all members in a household.

Vaparshun’s hypothetical ToC was developed and pre-tested in the formative phase of the study as described in previous section 2 of this document. Vaparshun’s process evaluation is aligned to its theory of change. The process evaluation approach is based on components suggested by Linan and Steckler, 2002 in their process evaluation framework and is adapted from similar studies (Greenland et al. 2017)(Roma et al. 2014)(Boisson et al. 2014)(Bonell et al. 2006) (Grant et al. 2013).

5.5. Standard Error Adjustments

5.5.1. How will you address clustering in your data?

GEE and robust standard errors.

5.5.2. How will you address false positives from multiple hypothesis testing?

We will not adjust for 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.)

NA

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.

We do not plan such procedures.

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.

Census and Baseline data collection questionnaires submitted and approved.