Improving demand for quality water: salience, water quality judgement and willingness to pay for water testing

A Study of the Jal Jeevan Mission in Gujarat, India.

PRE-ANALYSIS PLAN 2023

JUNE 2023

PRIMARY INVESTIGATORS

Dr Sharon Barnhardt, CSBC

Dr Pavan Mamidi, CSBC

Dr Shagat Mukherjee, CSBC

CONTRIBUTORS

A.R. Selva Swetha, CSBC

This project is officially documented here, on the AEA RCT Registry.

1. Study Context

The Jal Jeevan Mission is a flagship scheme of the government. It seeks to provide functional household tap connections for all rural households by the year 2024, as well as promote the holistic management of local water resources. The core objectives are Safe drinking water, Women empowerment, and Core Water Management of resources.

2. Motivation

This study examines behavioural levers and pathways for improving user demand for safe drinking water and hence demand for the national scheme. Specifically, we hope to understand the effect of different behavioural types of framing of water quality information on the value perception of water quality.

We look to add to the existing body of research on the links between water quality judgement and knowledge of ill effects of water quality and salience of water quality in day-to-day decision making. Literature suggests that information is insufficient to change behaviour (information may be important because of psychological factors, such as increasing the salience of water contamination, but the mechanism isn't through a Bayesian updating). Information campaigns alone have not induced behaviour change (Dupas, 2009; Dupas, 2014). In contrast, psychologically targeted elements added to information provision interventions improve self-efficacy and the salience of the targeted behaviour, and may achieve measurable improvements in outcomes (John and Orkin, 2022; Haushofer et al., 2019; Jalan and Somanathan, 2008). Thus, combining multiple behavioural change elements like disgust, salience, self-efficacy, etc., can be a valuable approach to achieving substantial behavioural change.

3. Research Design

3.1 Research Questions

1. Does increasing knowledge and salience of invisible and imperceptible contaminants in water increase demand for scientific testing of drinking water quality?
2. Does demonstrating evocative water quality tests increase the demand for scientific testing of drinking water quality?

3.2 Treatment Arms

Each individual was assigned to two video treatments.

1. Treatment 1: Informational video to prompt water quality testing

a. T1a: Negative frame; Planting a seed of doubt regarding water quality (by increasing knowledge and salience of invisible contaminants).
b. T1b: Positive framing of the importance of quality.

2. Treatment 2: Demonstration and visualisation of water quality through evocative scientific testing.

Tables 3.1 and 3.2 contain a detailed description of the treatment and control videos for treatments 1 and 2 respectively. All videos featured live actors, and were in Gujarati.

Table 3.1: Description of the Treatment 1 arms

Experiment Arm	Description	Visual Elements	Call to Action	Tonality	Length
Control (C1)	How to withdraw money in an ATM? A woman exiting an ATM runs into her friend, and describes to her the process of withdrawing money.	An animation of an ATM machine is also shown, with the woman's voiceover.	Nil	Info for social good	2 mins, 33 seconds
Negative Frame (T1a)	How safe is your water?	* Visually zooming into water to show that sensorial perceptions can be misleading. * Kidney stone passing through the excretory system. * Growth of bacteria in the stomach.	Demand for clean drinking water; test your water to know your water.	Disgust and pain	2 mins, 46 seconds
Positive Frame (T1b)	How safe is your water? A conversation between an urban woman and her uncle in the village about the importance of clean water,	Why clean drinking water is important to health.	Demand for clean drinking water; test your water to know your water.	Aspiration and dignity	2 mins, 11 seconds

Table 3.2: Description of the Treatment 2 arms

Experiment Arm	Description	Visual Elements	Call to Action	Tonality	Length
Control (C2)	How to use a QR code to make payments. A friendly woman in her 20s is seen describing the process.	Step by step walkthrough, with the screen display of an UPI payment app. The video of the trainer overlaid on the app screen.	Nil	Information for social good	3 mins, 8 seconds
Evocative scientific testing of water quality (T2)	Live demonstration of three water quality tests 1. TDS 2. Bacterial test - H2S, 3. Residual Chlorine Test	Evocative change in colour of water for tests 2 & 3, in case of poor quality water, to pink and black respectively.	Demand for testing	Popular Science	2 mins, 58 seconds

3.3 Flow of the Experiment

CSBC, NYAS and Tata Trusts conducted the experiments in a lab-in-the-field setting in Gujarat.

Part one of the lab activities happens in a one-to-one booth setting, where the enumerator is the same gender as the participant. Here, we randomise individual participants screened for eligibility under the study to the experimental conditions. We show them the allocated treatment videos on a mobile phone (with earphones). Immediately following the video interventions, we measure two outcomes:

1. Water Taste-or-Test Game

   For a pre-randomised sample of water placed before them in a glass, the respondent plays three rounds of an incentive-compatible game. The respondent receives an endowment of Rs. 100 at the beginning of the game. For each round of the game, they are asked whether they would like to view a water quality report (of three types of scientific tests of water quality that were already performed on the water sample - TDS, Chlorine, H2S Bacteriological test) by paying a small fee, or would drink the water without viewing the results and receive an additional Rs. 100. The fee for testing is Rs.

30, Rs. 20, and Rs. 10, respectively, for the three game rounds. Following the three rounds, the respondent rolls a die to determine for which of the three rounds they will receive the payout.

The three samples of water for each lab activity are created the previous day, using a standardised formula, and are safe to consume. Mineral water formed the base for the all the samples. The three samples were: (1) plain mineral water, (2) cloudy-looking water, (3) chlorinated water.

The game is conducted in a closed one-to-one lab booth setting, to maintain privacy and eliminate learning effects/influence from other participants' playing of the game.

While the respondent is informed of their game earnings at the end of the game, they are not paid out until the completion of the post-game questionnaire to mitigate any effect of respondents' sharing their earnings amongst each other on the respondents' survey responses.

2. Post-game Questionnaire

   Immediately following the game, the respondent was administered a 20-minute survey through a different enumerator.

   This survey is conducted in an open setting in the lab outside of the booths. The enumerator was assigned based on availability and was not one-on-one gender-matched with the respondent.

Part three of data collection took place at the respondent's house:

3. Household Water Testing

We attempted to visited all respondents at their houses on the day immediately after the lab. If they were not available at home, up to two additional revisits were done, each on a different day. Those who remained unsurveyed even after a total of three visits, were then contacted and surveyed on the phone. (If phone respondents expressed interest in purchasing household testing, the testing was done through another (non-respondent) household member at the household.)

At the household, enumerators followed a standard script, describing the bundle of three tests being offered:

- TDS (Total Dissolved Solids, instantly measured using a TDS metre)
- H2S Bacteriological Test (sample taken in a vial, incubated for 48 hours, and result reported through WhatsApp/phone call/SMS)

- Residual Chlorine Test (result instantly observable through colour change)

While the price of the bundle was Rs. 100, the respondent was eligible for a discount, according to the discount coupon they would have received at the lab.

The enumerator scripts, testing procedures, and a few qualitative and quantitative questions to the respondent about reasons for choosing to avail / not avail the test were programmed into the SurveyCTO mobile application. Time-stamped pictures of the three test tubes/vial and the test report delivered to the respondent were collected on the application, for quality check purposes.

The household water testing enumerator was not gender-matched with the respondent. S/he was also blind to the respondent's game earnings through the lab activity and the two treatment assignments.

3.4 Treatment Assignment

Treatments T1 and T2 were assigned at the individual level, cross-randomised against each other, stratified by gender.

There are two types of control groups: (1) for each video treatment is a control group which is shown the control video, (2) a *pure* control group, which is not administered the game and is shown the control videos for both treatment. Thus, there are seven experimental groups.

Table 3.4 Experimental Groups

Group	Video 1	Video 2	Game Measurement	Post-Game Measurement
1	Negative Frame	T2 Control	Yes	Yes
2	T1 Control	T2 Control	Yes	Yes
3 (Pure Control)	T1 Control	T2 Control	No	Yes
4	Positive Frame	Testing	Yes	Yes
5	Negative Frame	Testing	Yes	Yes
6	T1 Control	Testing	Yes	Yes
7	Positive	T2 Control	Yes	Yes

In implementation, serially ordered unique IDs by gender, for each village, were pre-randomised (computer-randomised) into one of the seven cross-randomised T1x T2 groups.

Other (non-treatment) random assignments
In addition to the two treatments, there were two further randomised manipulations to be used as control variables for the two respective points of outcome measurement: lab and the household.

- water_sample: individuals were randomly assigned to one of three types of water samples for the water taste-or-test game.
- discount_voucher: individuals were randomly assigned to one of three levels of discount vouchers for the household water testing

Table 3: Description of the Experimental Flow in the Lab and at the Household.

<table>
<tr><th>Step 1</th><th>Step 2</th><th>Step 3</th><th>Step 4</th><th>Step 6</th><th>Step 7</th><th>Step 8</th><th>Step 9</th></tr>
<tr><td>Subject Comes to the Lab:

Screening & Consent for the Study</td><td>Video 1</td><td>Video 2</td><td>Game:

Water Taste or Test

(Token with the payout amount)</td><td>Post game Questionnaire</td><td>Discount Voucher for HH water testing
+
Receive Game Payment</td><td>Offer of HH Testing of Water

Residual Chlorine & TDS</td><td>H2S Test Results Reporting</td></tr>
<tr><td>3-4 mins</td><td>3 mins</td><td>4 mins</td><td>8-10 mins</td><td>16-25 mins</td><td>2-3 mins</td><td rowspan="2">8-15 mins</td><td rowspan="2">3-6 mins</td></tr>
<tr><td colspan="4">Part 1: In Booth: 23-25 mins [Enumerator 1][1]</td><td colspan="2">Outside the Booth {Enumerator 2}</td></tr>
<tr><td colspan="6">Day 1</td><td>Day 2</td><td>Day 3
(Up to three revisits on days 4-7)</td></tr>
<tr><td colspan="6">At the Lab-in-the-field</td><td>At respondent’s house</td><td>By SMS/Phone /WhatsApp</td></tr>
</table>

[1] Within the lab, the enumerator’s gender matched that of the participant’s. For the post-game survey the gender assignment was based on convenience/availability of enumerators. The survey enumerators’ gender ratio was *roughly* 1:2 :: M:F.

3. Sample Size and Sampling

We calculated that to detect a medium effect size of 0.3 in the proportion of people willing to pay for water quality testing, we would need 175 individuals per group, as described in Table 3. With at least 1225 individuals, the experiments would be powered to compare the T1 groups (dignity v. disgust v. control video 1) and T2 groups (testing video v. control video 2). We set our target sample for the field at 1300 individuals.

Assumptions for power calculations:

a. Effect size (Cohen's *h*): 0.3
b. Power: 80%,
c. Significance level: 5%
d. Treatment is at the individual level
e. Outcome: Proportion of people who choose to pay for testing
f. Difference of proportion power calculation for binomial distribution. Two-sample proportion test for equal samples (pwr.2p.test on R)

Table 3: Sample size for two treatments

Total N=1225 individuals		**T2: Water Testing Videos**		
		Control	**Treatment**	**T1 Totals**
T1 Framing Videos	**Control Video**	175 indiv	175 indiv	**350 indiv**
	Pure Control Group (No game measurement)	175 indiv		**175 indiv**
	Negative Framing	175 indiv	175 indiv	**350 indiv**
	Positive Framing	175 indiv	175 indiv	**350 indiv**
	T2 Totals	**525 + 175 indiv**	**525 indiv**	**1225 indiv**

At 1050 individuals, we are under-powered for determining the optimal level of discount for HH water testing, but powered to compare the T1 groups (dignity v. disgust v. control 1) and T2 groups (testing video v. control video 2).

(To be powered for comparing pricing in addition to T1 and T2, i.e., for **T1 X T2 X Pricing**, the above table will be multiplied by 3. We will need 1050x3 = **3150 individuals** in the sample.)

3.2 Village Sampling

- Villages where our implementation partner, Tata Trusts, works constitute our sample frame of villages for the study.
- Villages were stratified based on water quality testing results reported by Tata Trusts for the post-monsoon round of testing in November 2022. We grouped villages according to whether they were deviant from the acceptable range for the following parameters: TDS, Residual Chlorine and the H2S/bacteriological test.
- Within each group, villages were ranked based on descending order of the number of households with a tap connection.
- The overall sampling rank of a village was assigned by cycling through the testing groups in the following order: H2S, TDS and Residual Chlorine.
- An equal number of villages were sampled in Amreli and Bhavnagar districts. The temporal order of data collection followed the sampling rank within each district.

3.3 Study Population

The study population is **rural households with a household-level piped-water supply**. We covered 22 villages in Amreli and Bhavnagar districts in Gujarat, India.

Individuals who meet the following screening criteria are sampled in each of the study villages:

1. Is above 18.
2. Belongs to a HH with a functional household tap connection.
3. Is the male head of Household or the female head of household. Or, is a key decision maker for each gender.
4. Is normally resident in the village. (i.e., lives for at least *6* months in a year).
5. Only one member of the household is allowed to participate in the study.

3.3.1 Respondent Recruitment

Respondents were recruited on the day of the lab. In order to ensure representativeness at the village level, each village was divided into various zones based on the communities resident in them. Enumerators went door-to-door using either the Left Hand Rule or the Right Hand Rule, and in-field random sampling was followed. Participants were orally screened and issued tokens to produce when visiting the lab. Where a cluster of households resided in one dwelling, only one household was screened into the study. This was to mitigate (1) the spillover effect of the

lab interventions on the takeup of household testing of water, as well as (2) the potential impact (expected decrease) in likelihood of testing takeup when an immediate neighbour had taken up the test.

At the time of recruitment, respondents were not informed of the potential to earn money based on their responses to the game.

3.4 Data Collection

Enumerators from NYAS, a contracted data collection agency, were hired to administer the in-person survey on licensed software Survey CTO on an offline app on a mobile phone.

Only complete surveys having both lab and household measurement will be used for analysis, and no participants with partial surveys will be recontacted to resume the survey. While the enumerators were familiar with the experimental flow of the study, in order to mitigate potential bias, they were not made aware of the details of the treatment groups or outcomes.

4. Analysis

4.1 Outcomes

Table 4.1: Outcome Variables

S.No	Outcome Variable	Outcome Measures	Source	Measurement
	Knowledge Outcomes (Intermediate Outcome)			
1.	**Knowledge about kidney stone cause**	To what extent do you agree or disagree with the statement: “Prolonged consumption of hard water may cause kidney stone formation in adults.”	Post-game questionnaire	Variable Type: Ordinal Likert Scale.
2.	**Knowledge about kidney stone visible sign**	To what extent do you agree or disagree with the statement: Consider a healthy adult living in Gujarat.	Post-game questionnaire	Variable Type: Ordinal Likert Scale.

		“Being visibly healthy today is no indication that the person will not face issues like kidney stones if they drink bad water over a prolonged time.”		
	Attitudinal Outcomes (Intermediate Outcomes)			
3.	**Water testing - Usefulness**	How useful or not useful is it to test the water?	Post-game questionnaire	Variable Type: Ordinal Likert Scale.
4.	**Water testing - Intent**	If a quick and easy testing method were to be made available to test the water, would you be interested in testing your own water?	Post-game questionnaire	Variable Type: Ordinal Likert Scale.
5.	**Judgement of water quality through sensory perception**	To what extent do you agree or disagree with the statement: "The quality of the water my family drinks can be accurately judged through colour, taste and odour of the water"	Post-game questionnaire	Variable Type: Ordinal Likert Scale.
6.	**Judgement of safety of their household water for adults**	How safe do you think the water is to drink for ADULTS of your household?	Post-game questionnaire	Variable Type: Ordinal Likert Scale.
7.	**Judgement of safety of their household water for children**	How safe do you think the water is to drink for CHILDREN of your household? This outcome will be reported only for respondents in households with a	Post-game questionnaire	Variable Type: Ordinal Likert Scale.

		child less than 15 years.		
	Behaviour Outcomes (Primary Outcomes)			
8.	**Household testing takeup**	1 = Participant chose to take up testing 0 = Did not take up testing.	Household visit	Variable Type: Binary.
9.	**Bought game water tests - at any price**	1 = Participant chose to view test results before tasting water for any one of the game rounds. *At any price* (By sample type) 0 = Participant chose to drink the water without testing.	Post-intervention Taste or Test game	Variable Type: Binary.
10A.	**Bought game water tests - at Rs. 30**	1 = Participant chose to view test results before tasting water for round 1. (By sample type) 0 = Participant chose to drink the water without testing.	Post-intervention Taste or Test game	Variable Type: Binary.
10B.	**Bought game water tests - at Rs. 20**	1 = Participant chose to view test results before tasting water for round 2. 0 = Participant chose to drink the water without testing.	Post-intervention Taste or Test game	Variable Type: Binary.

4.3 Covariates

A. Demographic Covariates

Variable	Description	Measurement	Source
Household with children <=15 years	Do you have any children less than 15 living in your HH?	Variable Type: Binary	Post-game Survey
Household with children 0-5 years	● No of female children less than 5 years living in the household ● No of male children less than 5 years living in the household	Variable Type: Numeric.	Post-game Survey
Intra-household Status of Respondent	Self-reported status as (gender-level) head of household.	Variable Type: Binary	Pre-consent Screening
Female	Self-reported gender indicator for a female respondent.	Variable Type: Binary.	Pre-consent Screening
Age	Self-reported age of respondent.	Data Type: Numeric.	Pre-consent Screening
Marital Status of respondent	Self-reported marital status of respondent.	Variable Type: Binary	Post-game Survey
Home Ownership	Binary variable for respondent's household owning the home they live in.	Variable Type: Binary	Post-game Survey
Asset Ownership	An additive index to be constructed based on variables as determined by a Principal Component Analysis of the following variables. 1. Does your family own the house you live in?	Standardised Index	Post-game Survey

Variable	Description	Measurement	Source
	2. How many rooms does your house have? 3. How many two-wheelers does your household own? 4. How many four-wheelers does your household own? 5. How many computers/laptops does your household own? The index will be divided by the number of members (adults and children) in the household. The per capita asset index will be standardised to have its mean as zero and SD as 1.		
Monthly per capita household Income	Self-reported income per capita for the last month. This will be constructed by taking the mid-point of the categories reported for the income question below, and dividing by the number of total members (adults and children) in the household. What was approximately your total household income in the previous month? 1. Less than Rs. 5000 2. Rs. 5001 - 10,000 3. Rs. 10,001 - Rs., 15000 4. Rs. 15,001 - Rs.20,000 5. Rs., 20,001 - Rs., 25,000 6. Rs., 25,001 - Rs., 50,000 7. More than Rs. 50,000 Based on the frequency of the open-ended income class in the data, an approximation of the upper bound based on Gujarat rural income distributions will be used in the calculation of the midpoint.	Data Type: Ordinal categorical	Post-game Survey

Variable	Description	Measurement	Source
Highest education in household	Self-reported education category. Manipulation: Convert to 3 binary variables ● *prim*; 1 = Finished primary school; 0 = did not finish primary school ● *sec* 1= Finished secondary school; 0 = did not finish secondary school ● *college* 1= Finished college; 0 = did not finish college	Variable Type: Ordinal categorical	Post-game Survey
Respondent highest education	Self-reported education category of highest educated member in the household. Manipulation: Convert to 3 binary variables ● *prim*: 1 = Finished primary school; 0 = did not finish primary school ● *sec*: 1= Finished secondary school; 0 = did not finish secondary school ● *college* 1= Finished college; 0 = did not finish college	Variable Type: Ordinal categorical	Post-game Survey
Mobile phone ownership	Self-reported ownership of own mobile phone.	Variable Type: Binary	Post-game Survey
Smartphone ownership	Self-reported indicator for if respondent's own phone is a smart-phone.	Variable Type: Binary	Post-game Survey
District	District	Variable Type: Categorial	Lab dataset
Village	Village	Variable Type: Categorial	Lab dataset
Game water sample	Randomly assigned (stratified by gender) water sample type for the Taste of Test Game. Type 1: Cloudy sample Type 2: Chlorinated water sample	Variable Type: Categorial	Game

Variable	Description	Measurement	Source
	Type 3: Mineral water		
Test price at household	Randomly assigned (stratified by gender) level of discount for the household water testing bundle (TDS, Chlorine, and H2S Bacteriological test) Discount levels: 10%, 50%, 90% on a total price of Rs. 100 for the bundle of three tests. Manipulation: convert to a numeric variable denoting the test price levels of test price: Rs. 90, Rs. Rs. 50, Rs. 10	Variable Type: Numeric	Post-game survey discount randomization

B. Other Covariates

	Variable	Description	Measurement	Source
1.	**JJM water disruption**	Households will be classified into the categories based on the number of continuous days they last went without JJM water. 1. Less than 24 hours 2. 1 to 2 days 3. 3 to 4 days 4. 5 or more days	Variable Type: Ordinal categorical.	Post-game Survey
2.	**Game earnings**	Total earnings from the game in Rupees for each participant (except pure control)	Variable Type: Numeric	Lab game
3.	**Past exposure to testing**	Self-report of whether they have had their household water tested in the past.	Variable Type: Binary	Post-game Survey

4A.	**Risk preferences: risk aversion**	We will create three levels of a risk aversion variable based on the following two questions: R1. If you had a choice between gaining Rs. 500 for sure, and a lottery where we toss a fair coin and you have a gain of Rs. 0 with ½ chance [heads] and **a gain of Rs. 1000** with ½ chance [tails], you would: (1) Choose the sure gain of Rs. 500 (2) Choose the lottery (3) Be indifferent between the two options R2. If you had a choice between gaining Rs. 500 for sure, and a lottery where we toss a fair coin and you have a gain of Rs. 0 with ½ chance [heads] and **a loss of Rs. 1000** with ½ chance [tails], you would: (1) Choose the sure gain of Rs. 500 (2) Choose the lottery (3) Be indifferent between the two options Manipulation: Risk aversion=1 if R1==1 & R2==2, else the variable takes the value 0.	Variable Type: Binary	Post-game Survey

4B.	**Risk preferences: risk attitudes**	This variable takes on the following three categories based on R1 above. ● Risk averse if R1==1 ● Risk loving If R1==2 ● Risk neutral if R1==3	Variable Type: Categorical	Post-game Survey
5A.	**Time preferences: time impatient**	T3. If you had a choice between getting Rs. 2000 right now versus Rs. 4000 in six months' time, what would you choose? (1) Rs 2000 right now (2) Rs 4000 in six months (3) Be indifferent between the two options T4. If you had a choice between getting Rs. 2000 right now versus Rs. 4000 in six months' time, what would you choose? (1) Rs 2000 in a year from now (2) Rs 4000 in a year and six months from now (3) Be indifferent between the two options Manipulation: A respondent is coded as time impatient, if they always choose the smaller amount, closer in time: Time impatient=1 If T3==1 & T4==1	Variable Type: Binary	Post-game Survey

		Else, time impatient=0		
5B.	**Time preferences: present biased**	Manipulation: A respondent is coded as present biased, if they choose the smaller amount closer in time now, but is willing to be patient in the future. Time impatient=1 if T3=1 & T4=2) Else, time impatient=0		Post-game Survey
6A.	**Trust (general)**	Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?	Variable Type: Ordinal Categorical	Post-game Survey
6B.	**Trust (local)**	To what extent do you agree or disagree with the statement: Most people in your village/community can be trusted.	Variable Type: Ordinal Categorical	Post-game Survey
7.	**Household health Index**	Equally weighted, simple additive index (score 1 to 4) based on the following two questions: 1. How do you perceive your own health overall?	Standardis ed Index	Post-game Survey

		(4= very healthy, 1= very unhealthy) 2. How do you perceive the health of your children overall? (if the household has children) (4= very healthy, 1= very unhealthy)		
8.	**Diarrhoea Prevalence**	Equal-weighted index of the following variables: 1. In the past two weeks, have any of your children experienced diarrhoea (watery or loose stools)? Binary variable, with 1=Yes. 2. To what do you think diarrhoea, especially among children, is a problem in your neighbourhood? This is a categorical variable Binary variable, with 1=Yes.	Standardised Index	Post-game Survey
9.	**Days between lab and household takeup measurement**	The number of days elapsed between the date the lab was conducted and date the household testing takeup survey was completed for a given participant.	Variable type: numeric	Constructed variable using: household testing survey dataset and lab survey dataset.
10.	**Phone surveyed in household testing**	1=Respondent unable to be reached after three household visits, and hence phone-surveyed instead of in-person. 0= Surveyed in-person	Variable type: Binary	Household testing

3.2 Intervention Questions & Manipulation Checks

We will report treatment-wise descriptive statistics for *each* video, for the following intervention-related questions:

Table 3.2: Manipulation checks and intervention-related variables for each of the two videos

Variable	Question	Measurement	Source	Notes
Found the video interesting	Self-reported measures using Likert scale rating (1-5) from 1=Very Uninteresting to 5=Very Interesting. Manipulation: Binary variable for respondent answered 4 & 5.	Variable Type: Binary.	Post-intervention, pre-game.	
Found the video short	Self-reported measure of perception of length of video, using Likert scale rating (1-5) from 1=Too Long to 5=Too Short. Manipulation: Binary variable for respondent answered 4 & 5.	Variable Type: Binary	Post-intervention, pre-game.	
Reaction to video	Self-reported reaction immediately after watching the video: Which of the following words describe how you felt after watching this video? I will read a list of words. Please yes or no to each. Note: Read out all the options. Mark ALL that apply. 1. Fearful 2. Happy 3. Hopeful 4. Responsible 5. Angry 6. Worried 7. Healthy	Variable Type: Binary. For each reaction, reported as % who answered yes.	Post-intervention, pre-game.	We expect video T1a to register a different response to video T1b.

Reaction to video: Fear	A simple additive index of the binary variables created from the options to the question about their reaction to the video. The indices will take on values from 0 (not at all fearful/worried) to 3 (most fearful/worried). ***fear_anger_index:*** fearful + worried + angry	Variable Type: Ordinal categorical.	Post-game Survey.	We expect video T1a > video T1b.
Video Content Recall	Respondent correctly recalling the content of each video, from a single-select question (with prompted options). You were shown two videos earlier. What was the first video about?	Variable Type: Binary.	Post-game Survey.	
Learned something new from the video	Self-reported response to a 4-point Likert scale about whether they learned anything new in the video. Did you learn anything new from the video? 1= Definitely, yes 2= Somewhat, yes 3= Somewhat, no 4 = Definitely no Manipulation: Respondent answered 3 or 4.	Variable Type: Binary	Post-intervention, pre-game.	

4. Analysis

The research questions we seek to answer in analysis are the following. The models used are given in Table 4.1. The comparison groups refer to the groups listed in Table 3.4 (reproduced below)

Table 3.4 Experimental Groups

Group	Video 1	Video 2	Game Measurement	Post-Game Measurement
1	Negative Frame	T2 Control	Yes	Yes
2	T1 Control	T2 Control	Yes	Yes
3 (Pure Control)	T1 Control	T2 Control	No	Yes
4	Positive Frame	Testing	Yes	Yes
5	Negative Frame	Testing	Yes	Yes
6	T1 Control	Testing	Yes	Yes
7	Positive	T2 Control	Yes	Yes

Treatment 1

R1. Do the Treatment 1 videos increase the take-up of testing? (pooled)

A. In the game? (1,4,5,7 vs, 2,6)
B. At home? (1,4,5,7 vs, 2,3,6)

R2. Does the negatively framed video increase the take-up of testing more than the control video? (subsample) (1,2,3,5,6)

A. In the game? (1,5 vs. 2,6)
B. At home? (1,5 vs. 2,3,6)

R3. Does the positively framed video increase the take-up of testing more than the control video? (subsample) (2,3,4,6,7)

C. In the game? 4,7 vs. 2,6)
D. At home? (4,7 vs. 2,3,6)

R4. Does the negatively framed video increase the take-up of testing more than the positively framed video? (subsample)

E. In the game? (Group 1+5 vs. Group 4+7)
F. At home? (Group 1+5 vs. Group 4+7)

R5. Does the negatively framed video increase the take-up of testing (compared to control), among those subject to treatment 2? (subsample)

A. In the game? (Group 5 vs. Group 6)
B. At home? (Group 5 vs. Group 6)

R6. Does the positively framed video increase the take-up of testing (compared to control), among those subject to treatment 2? (subsample)

C. In the game? (Group 4 vs. Group 6)
D. At home? (Group 4 vs. Group 6)

R7. Does the negatively framed video increase the take-up of testing compared to positively framed video, among those subject to treatment 2? (subsample)

E. In the game? (Group 5 vs. Group 4)
F. At home? (Group 5 vs. Group 4)

Treatment 2

R8. Does the water testing video (Treatment 2) increase take up of test results? (Pooled, overall)

A. In the game? (Group 4+5+6 vs Group 1+2+7)
B. At home? (Group 4+5+6 vs Group 1+2+3+7)

R9. Is the overall effect different for those who saw the positively framed first video?

C. In the game? (Group 4 vs. Group 7)
D. At home? (Group 4 vs. Group 7) *(There is no pure control comparison possible.)*

R10. Is the overall effect different for those who saw the negatively framed first video?

E. In the game? (Group 5 vs. Group 1)
F. At home? (Group 5 vs. Group 1) *(There is no pure control comparison possible.)*

R11. What is the marginal effect of Treatment 2 on household testing take-up?
(T2 + T1 control) vs. Pure control. (Group 6 vs. Group 3)

Effect of Chlorine

R12. Does the smell of chlorine in the game water samples increase or decrease take up of testing? (1,4,5,7 vs, 2,3,6)

Price and Treatment

R13. What is the interaction of the price of the household test and video group on take up of household testing? (1,4,5,7 vs, 2,3,6)

4.1 Model Specifications

Model specifications for the primary outcomes are laid out in Table 4.1 below.

Table 4.1: Model Specifications

Model.	Outcome	Independent Variable of interest	Model Type	Experimental and Comparison Groups	Sample Size	Specific Covariates
R1A.	Game Testing Takeup	Treatment Video 1	Logistic regression	(T1 Positive + T1 Negative) v. T1 Game Control	350, 350, 175 N=850	• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R1B.	Household testing takeup	Treatment Video 1	Logistic regression	(T1 Positive + T1 Negative) v. T1 Pure Control	350, 350, 175 N=850	• HH water availability • Previous exposure to testing • Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village

R2A	Game Testing Takeup	T1 Postive Frame	Logistic regression	T1 Positive v. T1 Game Control	350, 175 (N=525)	• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R2B	Household testing takeup	T1 Postive Frame	Logistic regression	T1 Positive v. T1 Pure Control	350, 175 (N=525)	• HH water availability • Previous exposure to testing • Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village
R3A	Game Testing Takeup	T1 Negtative Frame	Logistic regression	T1 Negative v. T1 Game Control	350, 175 (N=525)	• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R3B	Household testing takeup	T1 Negative Frame	Logistic regression	T1 Negative v. T1 Pure Control	350, 175 (N=525)	• HH water availability • Previous exposure to testing • Game earnings

						• Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village
R4A	Game Testing Takeup	T1 Negative Frame	Logistic regression	T1 Negative v. T1 Positive Frame	350, 350 (N=700)	• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R4B	Household testing takeup	T1 Negative Frame	Logistic regression	T1 Negative v. T1 Positive Frame	350, 350 (N=700)	• HH water availability • Previous exposure to testing • Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village
R5A	Game Testing Takeup	Treatment Video 2	Logistic regression	T2 Experimental	525, 350	• Water sample • Trust • Risk preferences

				v. T2 Control	N=875	• Time preferences • Fixed effects: village • Enumerator
R5B	Household testing takeup	Treatment Video 2	Logistic regression	T2 Experimental v. T2 Pure Control	525, 175 N=700	• HH water availability • Previous exposure to testing • Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village
R5C	Game Testing Takeup	Treatment Video 2 | Positive Treatment 1	Logistic regression	T2 Experimental x Positive Video v. T2 Game Control x Positive Video	350, 350 N=700	• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R5D	Household testing takeup	Treatment Video 2 | Positive Treatment 1	Logistic regression	T2 Experimental x Positive		• HH water availability • Previous exposure to testing

				Video v. T2 Game Control x Positive Video		• Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village
R5E	Game Testing Takeup	Treatment Video 2 | Negative Treatment 1	Logistic regression	T2 Experimental x Negative Video v. T2 Game Control x Negative Video		• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R5F	Household testing takeup	Treatment Video 2 | Negative Treatment 1	Logistic regression	T2 Experimental x Negative Video v. T2 Pure Control x Negative Video		• HH water availability • Previous exposure to testing • Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price

						• Phone-surveyed • Fixed effects: village
R6	Game Testing Takeup	Chlorinated Sample	Logistic regression	Chlorinated sample v. Mineral sample	350, 700 N=1050	• Water sample • Trust • Risk preferences • Time preferences • Fixed effects: village • Enumerator
R7	Household testing takeup	Testing price x T1	Logistic regression	At each price: T1 Experimental v. T2 Pure Control	525, 175 N=700	• HH water availability • Previous exposure to testing • Game earnings • Days between lab and household takeup measurement • Trust • Risk preferences • Time preferences • Test price • Phone-surveyed • Fixed effects: village

In addition to the primary outcomes described above, we will also analyse the secondary and intermediate outcomes described in Table 4.1. Ordered Logit will be used for the ordered categorical outcome variables.

Models:

For each outcome measure, we will estimate two models, one controlling for demographic covariates and one without.

We will estimate models of the following forms:

(1) $Y_i = \beta_0 + \beta_1 * \text{treatment}_i + \varepsilon_i$

and

(2) $Y_i = \beta_0 + \beta_1 * \text{treatment}_i + \beta_2 * \text{demographics}_i + \beta_3 * \text{specific covariates}_i + \varepsilon_i$

and

(3) $Y_i = \beta_0 + \beta_1 * \text{treatment}_i + \beta_2 * \text{demographics}_i + \beta_3 * \text{specific covariates}_i + \beta_4 *$ village fixed effects $+ \varepsilon_i$

and

(4) $Y_i = \beta_0 + \beta_1 * \text{treatment}_i + \beta_2 * \text{demographics}_i + \beta_3 * \text{specific covariates}_i + \beta_4 *$ village fixed effects $+ \beta_5$ Treatment x specific covariates $+ \beta_6$ Treatment x demographics $+ \varepsilon_i$

- Y_i is the dependent variable for individual respondent i.
- treatment_i is a binary variable that takes on a value of 1 if observation i received the treatment (i.e., watched the video, in question) and 0 otherwise.
- β_0 is the intercept or constant term.
- β_1 is the coefficient for the treatment_video variable, which represents the effect of the treatment on the dependent variable.
- β_2 to β_4 are coefficients for the demographic, specific covariates, and other covariates, representing their respective effects on the dependent variable.

ε_i is the error term for observation i, which represents the deviation of the actual value of Y_i from the predicted value based on the regression equation.

4.2 Multiple Hypothesis Testing

We will use a multiple testing adjustment to mitigate the increased risk of falsely rejecting the null hypotheses introduced by testing the effect of the intervention on multiple outcomes. Specifically, we will use the Storey method to control the False Discovery Rate (FDR).[2] The total number of outcome variables will be considered as one family to be tested.

2 Storey, J. D. (2002). A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(3), 479-498.

4.3 Randomisation Balance Check

Treatment was at the individual level, and treatment status is the only difference between the various treatment and control group participants in the lab-in-the-field experiment. On average, all other characteristics of treatment and control group members, including demographics, should be balanced. Treatment effect estimates could be biased if there is an imbalance across experimental groups despite the randomisation process.

Model: X ~ treatment_assignment + error
X are the different pre-treatment covariates.

4.4 Attrition Analysis

There are two points of outcome measurement: day one in the lab, and at day two (or later) at the household. We expect attrition despite the total of three household visits attempted, each on different days, as well as attempts to schedule appointments by phone.

1. We will check for the balance of treatment assignment between attrited and non-attrited groups.

   Model: $Y_i = \beta_0 + \beta_1 * \text{treatment}_i + \varepsilon_i$

   Y is a binary variable, indicating the group attrited at the household measurement.

   For a given treatment, if one experimental group is found to have greater attrition than the other, we would then check for balance of pre-treatment covariates across groups. If the characteristics of the attrited sample were, in fact, different from the non-attritted sample, we would be misattributing the treatment effect to the videos, while it may be explained by the differences in pretreatment characteristics.

   Model: $Y_i = \beta_0 + \beta_1 * \text{pre-treatment variable}_i + \varepsilon_i$

   Y is a binary variable, indicating the group attrited at the household measurement.

2. Additionally, we will also check for the balance of phone vs in-person household surveys across experimental groups.

   Model: $Y_i = \beta_0 + \beta_1 * \text{treatment}_i + \varepsilon_i$

   Y is a binary variable, indicating the respondent was phone-surveyed at the household measurement.

For a given treatment, if one experimental group is found to have greater proportion of phone surveys than the other, we will alco check for balance of pre-treatment variables by phone vs. in-person household survey.

Model: $Y_i = \beta_0 + \beta_1 * \text{pre-treatment variable}_i + \varepsilon_i$

Y is a binary variable, indicating the respondent was phone-surveyed at the household measurement.

References

Ahuja, A., Kremer, M., and Zwane, A.P. (2010). Providing Safe Water: Evidence from Randomized Evaluations. Annual Review of Resource Economics 2010 2:1, 237-256

Dupas, Pascaline (2009). “What matters (and what does not) in households' decision to invest in malaria prevention?” American Economic Review: Papers and Proceedings, 99(2): 224-30.
Haushofer, Johannes and John, Anett and Orkin, Kate, Can Simple Psychological Interventions Increase Preventive Health Investment? (April 2019). NBER Working Paper No. w25731, Available at SSRN: https://ssrn.com/abstract=3368021.

Dupas, P. (2014), Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment. Econometrica, 82: 197-228. https://doi.org/10.3982/ECTA9508.

Jalan, Jyotsna and Somanathan, E. (2008), The importance of being informed: Experimental evidence on demand for environmental quality, Journal of Development Economics, 87, issue 1, p. 14-28, https://EconPapers.repec.org/RePEc:eee:deveco:v:87:y:2008:i:1:p:14-28.

Anett John, Kate Orkin, Can Simple Psychological Interventions Increase Preventive Health Investment?, Journal of the European Economic Association, Volume 20, Issue 3, June 2022, Pages 1001–1047, https://doi.org/10.1093/jeea/jvab052.

Storey, J. D. (2002). A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(3), 479-498.

APPENDICES

Appendix 1: Power Calculation Scenarios

Outcome variable: Binary variable measuring self-reported intent to take up household water testing.

Effect Size *h*	**Power**	**Sample Size per group**	**Sample Size for Two Groups** (1T + 1C)	**Sample Size for Three Groups** (2T + 1C)
0.15	80%	698	1396	2094
0.2	80%	392	784	1176
0.3	80%	175	350	525

Note on effect Size:

- The effect size estimate is from a somewhat comparable study: 3ie Bangladesh study[3] was a disgust information intervention at the community level, where the outcome (measured at the individual level) was the *intent to use water treatment methods*:
    - Effect Size (calculated Cohen's h)[4] observed in the study: 0.128
        - Increase from 65% to 73%

[3] Guiteras, R, Jannat, K, Levine D and Polley, T, 2015. Testing disgust and shame-based safe water and handwashing promotion in Dhaka, Bangladesh, 3ie Impact Evaluation Report 29.

[4] Effect Size Calculations for Proportions: https://rdrr.io/cran/pwr/man/ES.h.html

Appendix 2: Water Taste-Or-Test Game: Show Cards

GAME ROUND 1 - SHOW CARD

	Actions		Earnings
Option 1			*Rs. 200*
Option 2			*Rs. 170*
			*Rs. 70*
Option 3			*Rs. 0*

GAME ROUND 2 - SHOW CARD

	Actions		Earnings
Option 1			*Rs. 200*
Option 2			*Rs. 180*
			*Rs. 80*
Option 3			*Rs. 0*

GAME ROUND 3 - SHOW CARD

	Actions		Earnings
Option 1			*Rs. 200*
Option 2			*Rs. 190*
			*Rs. 90*
Option 3			*Rs. 0*

Appendix 3: Water Quality Test Report - Format

TDS Test

Your TDS result is _____, refer to the table below to understand its meaning

Levels	Inference	
Less than 50	Unacceptable	✖
50-150:	Acceptable range	✓
150-250:	Healthiest range	✓✓
250-500:	Acceptable range	✓
500-900	Not acceptable.	✖

Chlorine Test

Your Chlorine result is _____, refer to the table below to understand its meaning.

Colour	Levels	Inference	
	0.0	Unhealthy	✖
	0.2	Healthy	✓
	0.5	Healthy	✓
	1.0	Healthy	✓
	1.5	Healthy	✓
	2.0	Unhealthy	✖
	3.0	Unacceptable	✖ ✖
	5.0	Unacceptable	✖ ✖

Bacteria Test

Your H2S result is _____, refer to the table below to understand its meaning

Colour	Inference	
	Unacceptable	✖
	Acceptable	✓