Testing disgust- and shame-based safe water and handwashing promotion in urban Dhaka, Bangladesh

Raymond Guiteras, University of Maryland

Kaniz Jannat, International Centre for Diarrheal Disease Research, Bangladesh

David I. Levine, University of California, Berkeley

Tom Polley, University of California, Berkeley

Grantee Final Report

Accepted by 3ie: February 2015

Note to readers

This impact evaluation has been submitted in partial fulfilment of the requirements of grant OW2.037 issued under Open Window 2. This version is being published online as it was received. A copy-edited and formatted version will be produced in the near future. All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, David I. Levine at levine@haas.berkeley.edu

.

Suggested citation: Guiteras, R, Jannat, K, Levine D and Polley, T (2015). *Testing disgust- and shame-based safe water and handwashing promotion in Dhaka, Bangladesh, 3ie Grantee Final Report.* New Delhi: International Initiative for Impact Evaluation (3ie). Funding for this impact evaluation was provided by 3ie's donors, which include UKaid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided on the 3ie website.

Acknowledgements

The authors gratefully acknowledge financial support from the Bill and Melinda Gates Foundation, the Clausen Center for International Business and Policy at U.C. Berkeley Haas School of Business, and the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN). The views expressed in this article are not necessarily those of 3ie or its members.

Contents

1 Abstract

Authors: Raymond Guiteras, University of Maryland; KanizJannat, icddr,b; David I. Levine*, University of California, Berkeley, levine@haas.berkeley.edu; Tom Polley, University of California, Berkeley

Background and research question

Objective: Test whether behaviour change messages designed to elicit disgust and shame can promote treating drinking water and hand washing with soap in low income urban housing compounds more effectively than classic public health messages based on germs.

Measure willingness to pay for compound-level chlorine dispensers using new group-versions of the Becker-DeGroot-Marschak(Becker, Degroot, and Marschak 1964) procedure.

Measure the effect of providing low-cost compound-level soapy bottles on handwashingbehavior.

1.1 Methods

Participants We studied households living within compounds in slums of Dhaka, Bangladesh. Compounds are clusters of households, typically located around a small courtyard, sharing a common toilet, water source and cooking facilities. We identified five communities around Dhaka with poor water quality, high incidence of water-borne disease, and high population density: Mohammedpur, Mirpur, Badda, Khilgaon and Bashabo. Compounds were eligible if they had between 6 and 18 Households, a shared water source, space to hold a compound meeting in or nearby, no other large water or hygiene interventions going on at this time, and primarily spoke Bengali.

Interventions All intervention compounds received a free trial of a chlorine dispenser followed by a sales meeting where they could choose to subscribe for the following year. Half the intervention compounds received a traditional behaviour change message focusing on germs and health, while half received messages designed to elicit disgust that untreated drinking water had shit in it, and fear of shame if they did not treat drinking water. Orthogonal to message assignment, two-thirds of compounds also received a soapy bottle that could be refilled with water and inexpensive laundry soap and left near the latrine or water source. Compounds receiving the soapy bottle received additional behavior change messages that emphasized either germs and health (for the compounds with similar water messages) or that hands that only rinse (and do not use soap) after leaving the latrine still have shit on them, and that this is shameful.

Outcomes For water treatment our primary outcomes are H2S tests for bacterial contamination of drinking water stored in the home; treatment with chlorine, as measured by chlorine residual tests of drinking water; and willingness to pay to subscribe to the chlorine dispenser. For handwashing our primary outcomes are direct observation of handwashing during a structured observation (during the free trial) and a cleanliness score of hands (as observed by enumerators).

Randomization We stratified compounds by neighbourhood and size. We used Atkinson's D-a optimal method to allocate treatments sequentially as compounds were enrolled(Atkinson 1982; Atkinson 1999).

1.2 Results

Numbers randomized. We randomized 650 compounds, 215 to control, 220 to the standard intervention arm and 215 to the disgust & shame intervention arm.

Recruitment is completed.

Numbers analyzed Due to the destruction of a few compounds, 214 and 210 compounds received each message. By the endline we had 205 and 203 compounds in each group.

Outcome Usage rates of the chlorine dispenser were low. During the free trial there wasdetectable chlorine in household drinking water within 24 hours of reported treatment at 8% of the homes.At the end of the free trial, only about a fourth of households agreed to participate in the auction, where compounds stated their willingness to pay to subscribe to the chlorine dispenser. Among this subset, the mean willingness to pay was about $0.10 / month per household where the auction was at the household level, and $0.83 per month per compound when the auction reported compound-level collective willingness to pay.

Adding our handwashing messages to either study arm and providing the soapy water bottle increased the share of handwashing stations that also had soap from 18% at baseline to 43% at endline (seven months post-intervention). During structured observation during the free trial, about 9% washed both hands with soap after toileting in study arms without the handwashing intervention, and 13% (P = 0.03) in arms with the handwashing intervention.

There is no consistent evidence that disgust and shame improved demand for or usage of the chlorine dispenser or soapy bottle relative to a traditional behavior change message focused on germs and health.

1.3 Conclusions

The chlorine dispenser is not popular in Dhaka, though a niche market may exist for a small share of compounds. The soapy bottle, in contrast, has very low cost and holds promise for increasing handwashing in other settings where households share water sources or latrines.

Our messages designed to elicit disgust and fear of shame did not reach most of the compound members (especially men) and also did not lead to measurable changes in self-reported attitudes or observed social interactions. Thus, it is unsurprising these messages also did not lead to larger changes in behavior than traditional messages emphasizing germs and health. It remains unclear if a more sustained set of messages would lead to the hypothesized changes in attitudes and behavior.

Trial registration: PR-11024, NCT02037243, Disgust and Shame Based Safe Water and Handwashing Promotion. Name of the register is US National Institute of Health (NIH).

FundingThe authors gratefully acknowledge financial support from the Bill and Melinda Gates Foundation, the Clausen Center for International Business and Policy at U.C. Berkeley, and the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN). We have no conflicts of interest.

2 Introduction

Diarrheal disease is one of the two leading killer diseases in the world(WHO 2008). An estimated 2.2 million children under the age of 5 years die from diarrheal disease each year. Most of these deaths are from middle- and low-income countries. Improvements in sanitation, water quality, and hygiene could reduce the burden of diarrheal diseases by about one-fourth(Pruss-Ustun, Bonjour, and Corvalan 2008). Specifically, both water treatment with dilute chlorine solution and handwashing with soap reduce self-reported diarrhea(Clasen et al. 2007; Ejemot-Nwadiaro et al. 2008; Rabie and Curtis 2006;

Luby et al. 2004; Luby et al. 2005; Luby et al. 2011a). However, even when enabling conditions are present, including access to soap and water and high baseline knowledge of handwashing, handwashingbehavior is difficult to change (Research Brief Vietnam; WSP 2012).Study participants in an evaluation of a large-scale intervention in Bangladesh were observed to wash both hands with soap or ash at only 1.7% of all handwashing opportunities at the promoted key times (ICDDR,B 2008).

Fear of disease alone does not change people's hygiene behavior significantly(van der Pligt 1996); interventions to date have not been able to provide substantial evidence of behavior change to make the interventions scalable(Gupta et al. 2008; Sallis, Owen, and Fisher 2008). Interventions need to provide more conclusive evidence of behaviour change in relation to handwashing practice with soap and water treatment. In addition to the traditional health messages based on increasing knowledge of health risks and germ transmission, other ways to motivate behavior need evaluation.

Research into the determinants of handwashing and sanitation behaviors has identified disgust and social affiliation as key motivators for behavior change(Curtis, Danquah, and Aunger 2009). In developed countries, studies using disgust and shame messages have shown that both are more effective than traditional health messages(Porzig-Drummond et al. 2009a; Judah et al. 2009). We hypothesized that incorporating disgust and shame messages in safe water and hand hygiene interventions in a developing country setting could increase the effectiveness of the interventions. To rigorously test this hypothesis, we *conducted a randomized controlled trial in urban Dhaka, the capital of Bangladesh.*

This report describesthe development of the disgust and shame intervention that started with community assessment of practice, beliefs and attitudes of the people. It includes the summary report of the qualitative assessment done during the implementation phase of the intervention to assess its effectiveness. We measured the differences in change in behaviour between the two intervention groups, one that received the traditional health messages and the other that received the disgust and shame messages. We analyse how the intervention affected willingness to pay--a crucial outcome for understanding if safe water interventions can be financially sustainable with little or no donor or government support. We discuss the results and provide recommendations for potential use of these messages.

2.1 Background and objectives

2.1.1 Scientific background and explanation of rationale

An estimated 2.2 million children under the age of 5 years die from diarrheal disease each year(WHO 2008). Children who survive multiple episodes of diarrhea and enteric infections commonly develop tropical enteropathy, an inflammatory disorder of the intestines that compromises nutrient absorption (Haghighi and Wolf 1997). Repeated episodes of diarrhea and chronic tropical enteropathy in early childhood often reduce growth and cognitive function, and impair school performance (Alderman, Hoddinott, and Kinsey 2006; Checkley et al. 2008; Lorntz et al. 2006; Niehaus et al. 2002; Petri et al. 2008). There Is some evidence that the result can be lower income later in life(Boissiere, Knight, and Sabot 1985). Thus, repeated episodes of childhood diarrhea and enteric infection may exact a long-run toll, perpetuating a cycle of poverty and ill health.

Most of the burden of diarrheal disease is preventable with better sanitation, water quality, and hygiene(Ezzati et al. 2003; Pruss-Ustun, Bonjour, and Corvalan 2008). As noted above, both water treatment with dilute chlorine solution and handwashing with soap reduce self-reported diarrhea(Clasen et al. 2007; Ejemot-Nwadiaro et al. 2008; Rabie and Curtis 2006).For handwashing with soap one meta-analysis found a mortality risk reduction factor of 48%(Cairncross et al. 2010).The effectiveness of point-of-use water treatment is less certain given that four blinded studies found no significant diarrhea risk reduction; still the much larger body of contrary evidence led them to estimate a risk reduction factor of 17%(Cairncross et al. 2010). These encouraging results come from small-scale studies that couple intense interventions and low-cost (typically free) handwashing and water treatment supplies(Khan 1982; Stanton and Clemens 1987; Quick et al. 1999; Chiller et al. 2006). Unfortunately, the behavior change approaches used in these studies are too costly to apply at larger scale.

In contrast, large-scale water or hygiene interventions have either not been rigorously evaluated or have resulted in limited uptake(Clasen 2009; Scott et al. 2008; Luby et al. 2008; Olembo et al. 2004; Luby, Halder, et al. 2009). The result is that both treating drinking water(Rosa and Clasen 2010) and handwashing with soap(ICDDR,B 2008) remain uncommon among the world's poor.

2.1.2 Current practices and barriers at the individual level

In our setting, few residents of Dhaka treat their drinking water. According to a recent national survey only 3.7% of urban dwellers reported filtering their water, 0.2% reported using chlorine, and 15.1% reported boiling their water(Rosa and Clasen 2010). Even these low self-reported rates likely overestimate the proportion of households that treat their water. For example, in an earlier evaluation of a point-of-use water treatment intervention in Bangladesh, 21% of households reported using the device, but only 4% were subsequently observed to be regular users(Gupta et al. 2008).

When households did boil, it was time-consuming. Most compounds we studied had communal (and unmetered) gas burners, so boiling also often requires waiting for access to one of the burners. In other compounds, boiling requires costly fuel.

Handwashing with soap is also uncommon in Bangladesh(ICDDR,B 2008). Fieldworkers observed handwashing practices in 10 housing compounds in a low income community in urban Dhaka(ICDDR,B 2010). Only 26% of study subjects washed their hands with soap after defecation and only 30% after cleaning a child's anus who had defecated. None washed their hands before preparing food in 301 instances, and fewer than 1% of persons washed their hands with soap before eating(Luby, Kadir, et al. 2009).

The Health Belief Model assumes that health behaviors occur when there is perceived susceptibility to an adverse outcome, high perceived benefit of adopting a new behavior,

lower perceived barriers to adopting the behavior, cues to action, and perceived self-efficacy (that is, the person believes they can correctly perform the behavior).

The perceived susceptibility to cholera, diarrhea, or dysentery is variable among low-income urban residents in Dhaka. Although residents note that such diseases occur within their families and communities, they also see diarrhea is a part of normal childhood development.

Most people living in low-income communities Bangladesh do not perceive substantial benefit to adopting household water treatment and more regular handwashing with soap. The lack of perceived benefits is consistent with their experience. Water and hands, even when contaminated with microscopic pathogens, usually appear clean. Most people in Bangladesh drink untreated water and do not regularly wash their hands with soap, yet they are not usually sick. In one study from Bangladesh only 21% of people believed that their drinking water sometimes made their family ill(Gupta et al. 2008).

Residents of low-income communities in Bangladesh perceive important barriers to household water treatment and regular handwashing with soap. The most immediate of these barriers is cost. Residents of low-income urban communities in Dhaka have difficulty covering basic housing and nutritional expenses for their family. In other low-income settings, there has been little demand for and limited willingness to pay for products that improve drinking water quality(Kremer et al. 2008; Ashraf, Berry, and Shapiro 2010). Although hand soap is present in most low- and middle-income households in Dhaka (92%), household residents commonly express concern about wasting soap(Luby and Halder 2008).

Inconvenience and the time required are also barriers. In low-income settings in Dhaka, toilet facilities and a water supply are typically shared by several families. Community soap is not available, and a person's own soap is typically not conveniently accessible after toileting. In rural Bangladesh people who lived in households that had soap or water available at their most convenient place to wash hands were twice as likely to wash their hands with soap after contact with feces compared to households that did not have soap or water present at a handwashing station(Luby, Halder, et al. 2009). Similarly, people in Dhaka who have a place to wash hands that includes soap and water have less respiratory illness (Luby and Halder, 2008). These findings are consistent with studies of handwashing in hospitals that concluded that hand hygiene was improved when equipment and supplies were present that made it easier to wash hands(Kaplan and McGuckin 1986).

Furthermore, handwashing instructions promoted by organizations may be overly complex, and may themselves become another perceived barrier. Handwashing instructions used by non-governmental and international organizations range from simple 3-step instructions (rinse, lather, rinse) to far more complex instructions (rinse, lather for 30 seconds or as long as it takes to sing "happy birthday", lather up past wrists, clean between fingers and under nails, rinse, dry with a clean towel or air dry). Complex handwashing instructions may pose a particular barrier for school-aged children and for busy caregivers of young children.

Similarly, prominent complaints among users of some point-of-use water treatment approaches include the time required and inconvenience of using the system(Gupta et al. 2008; Luby et al. 2008). Indeed, the broader scholarship on behavior change reinforces the importance of creating an environment that facilitates the desired behaviour (Sallis, Owen, and Fisher 2008).

There are currently few cues for action to encourage household water treatment or handwashing with soap in low-income communities in Dhaka. Both behaviors are uncommon enough that they are not a part of the standard cultural norm. Although cleanliness and purity (including washing hands with water alone) are important concepts within the culture, washing hands with soap – especially in association with food preparation or eating – is not generally connected to these concepts of cleanliness

and purity(Zeitlyn and Islam 1991). These behaviors were not typically taught in early childhood nor have they been acquired as a habit(Curtis, Danquah, and Aunger 2009). Because these behaviors are not part of the usual social landscape, people do not think about them and are not reminded to practice them either by their environment or by their family, friends or neighbors.

We are not aware of any formal assessments of community residents' self-efficacy to regularly wash hands with soap and to treat drinking water among low-income communities in Bangladesh, but we anticipate that with the substantial perceived barriers and the introduction of new technologies, explicit attention to self-efficacy of the target community will be important to achieve high levels of adoption and sustained behavior change.

The failure to achieve widespread adoption of these behaviors has inspired much research into the individual factors that affect health behavior change. A meta-analysis of 18 health interventions based on the Health Belief Model found that increased levels of perceived susceptibility and perceived severity of an illness were not associated with increased preventative behaviors as the model predicts; however, increasing perceived benefits and reducing perceived barriers of preventative behaviors was associated with safer behavior(Carpenter 2010).

2.2 Emotional and social motives

Unfortunately, substantial research implies that overcoming the information and cost constraints emphasized by the Health Belief Model often leads to only modest increases in preventive behaviors(Albert, Luoto, and Levine 2010; Luoto et al. 2011, 26–28). Thus, it is likely successfulinterventions will need to go beyond basic health education about germs and disease and include social forces(Madden, Ellen, and Ajzen 1992),andpsychological, sociological and economic theories of norms and self-identity all posit that, in addition to forces emphasized in the Health Belief Model,behavior change is based on normative beliefs and subjective norms(Akerlof and Kranton 2000).

Curtis, Danquah and Aunger(2009) highlight the concept of disgust and the need for affiliation with others as examples of emotional and cultural motivations that may lead to safer health behaviors. Following their suggestion, we focus on disgust and shame as our additional motivators.

2.2.1 Disgust

The Oxford English Dictionary defines disgust as "a feeling of revulsion or strong disapproval aroused by something unpleasant or offensive." Disgust at fecal contamination appears to be an evolved psychological system for protecting organisms from infection(Curtis, Barra, and Aunger 2011), making it one of the very few universal human norms. Curtis, Danquah and Aunger(2009)identify disgust both as a motivation for handwashing and as a potential tool for increasing the effectiveness of handwashing promotion.

We know of four studies that combine disgust components with hand hygiene interventions: two at-scale developing country interventions and two studies that tested the effect of including the concept of disgust in hand hygiene promotion to individuals. The first study was a full scale national campaign conducted in Ghana using a television advertising campaign. The focus was the transfer of fecal contamination from a mother to her child following a visit to the toilet. Assessment of this intervention suggested that reported handwashing after toileting increased by 13%, and increased before eating by 41%(Ram 2009). The second was a social marketing campaign in Burkina Faso which used the concept of disgust in components of its intervention and found that handwashing after using the toilet increased by 16%(Sallis, Owen, and Fisher 2008). The third study, in Sydney, Australia comprised two parts. The first part tested whether subjects who watcheda brief (3-minute) video-based intervention using disgust and education improved hand hygiene more than a comparison group viewing a video

including only education alone. The second part examined whether the findings from the first study could be replicated in the field. Disgust-based interventions were significantly better at promoting hand-hygiene than education alone(Porzig-Drummond et al. 2009b). The fourth study tested the effect of various text-only messages on soap use in highway service station restrooms. Disgust-based messages proved most effective for males, increasing soap use by 9.8%. For females disgust messages increased use by 5% (not statistically significant)(Judah et al. 2009).

These studies show promise for the potential of disgust to increase handwashing with soap. We know of no studies that show whether disgust can also motivate point-of-use water treatment.

2.2.2 Shame

Definitions of shame are less agreed upon than those of disgust; we will use the definition given by sociologist Thomas Scheff: a large family of emotions that involve the feeling of a threat to the social bond(Scheff 2000). He explains that even if people rarely feel shame, they frequently correct their behaviors to avoid it.1

The definition of *affiliation* used by Curtis, Danquah and Aunger(2009)captures one aspect of shame: "Being a good member of society by joining in and by doing what everyone else is perceived to be doing." They argue this desire for group affiliation "is an important motive for handwashing. This helps ensure membership in the social group. Conformity with local social norms is known to be a powerful driver of behavior" (Curtis, Danquah, and Aunger 2009; Verplanken 2006).

In Muslim societies such as Bangladesh the desire for affiliation has specific implications for handwashing. The Muslim religion emphasizes the importance of rinsing hands after defecation and before prayer(Zeitlyn and Islam 1991).Consistent with our approach, there is qualitative evidence that people in Bangladesh rinse their hands primarily to become pure and to be thought of as a clean person(Zeitlyn and Islam 1991). At the same time, community residents do not have a strong sense that people they know want them to wash their hands with soap (Zeitlyn, 1991 p. 523) or want them to treat their drinking water.

Unfortunately, because handwashing with soap and water treatment with chlorine are not established norms in the communities we study, there is a risk that shame or affiliation motives will discourage handwashing with soap and water treatment. People may be ostracized for attempting to be "too clean" (Pedersen, Keithly, and Brady 1986). However, for handwashing with soap, we believe there is a low risk of ostracization because handwashing without soap and general cleanliness are already highly valued; thus, the addition of soap would not be extremely different. For water treatment there may be greater risk of ostracization.

Curtis, Danquah and Aunger also note that avoiding loss of status (another aspect of our definition of shame) can also motivatehandwashing with soap. People want to enhance their social status, and being seen as clean can lead to admiration and respect. However, concerns with status concerns also pose a risk of discouraging safe behaviors. For example, if handwashing is considered a behavior only practiced by those with much higher social status, then low-status individuals who wash hands might be seen as "trying to get above themselves" (Pedersen, Keithly, and Brady 1986).

1This definition differs from that found in some of the psychology literature which uses shame for feelings related to status or social rank (Fessler 2004). Under that definition, the tendency to conform to social norms is a separate domain of emotion. Here however, because the tendency to optimize social rank and the tendency to conform are both means of avoiding social sanction, Scheff's definition fits our setting well.

If shame is an important motivator and if handwashing is associated with conforming to a community norm, then handwashing will be more common when observers are present. A number of studies have demonstrated this effect in a variety of settings. For example, in rich nations, people in public restrooms(Pedersen, Keithly, and Brady 1986; Munger and Harris 1989; Drankiewicz and Dundes 2003; Judah et al. 2009) and doctors in hospitals (Pittet et al. 2004) are much more likely to wash their hands if they know someone is observing them. In fact, in one study an observer in hospitals increased hand-washing by substantially more than did education regarding germs and transmission (Pittet et al. 2004). Closer to our population, a study in rural Bangladesh that placed inconspicuous motion detectors in soap found that handwashing with soap increased by 40% when an observer who had promoted handwashing was present (PK Ram, personal communication, 2012).

2.3 Description of Intervention and Theory of Change

2.3.1 Theory of change: Methods of water treatment and handwashing

Given the high costs in time and money of effective water treatment, we posit that a chlorine dispenser at the tap will lower the costs (in time and money) for effective water treatment. Similarly, we posit that providing a soapy water bottle will lower the costs (in time and money) for effective handwashing with soap. (We explain our interventions in detail in section 3.4.7 below)

The Health Belief Model(Carpenter 2010) and many related rational theories imply that lower costs will increase uptake of behaviors that improve health. In addition, the visual salience of the chlorine dispenser and soapy water bottle can act as reminders to engage in the safe behaviors.

These optimistic implications regarding water treatment are offset by the fact that many people complain about the smell and taste of chlorinated water, particularly when they are unused to the smell.

These optimistic implications regarding handwashing require that the soapy bottle be refilled with a few cents' worth of laundry soap.

2.3.2 Theory of change: Standard public health messages

The Health Belief Model(Carpenter 2010) posits that safe health behaviors will improve when people understand the risks of not performing the safe behaviors.

Based on the Health Belief Model, our standard message explained that there are dangerous germs on hands after rinsing and in drinking water from the community tap. We emphasized the costs of these diseases in terms of health risks, lower schooling for children, and the financial costs for treatment of severe diarrhea. We explained the importance of water treatment with the chlorine dispenser and handwashing with soap. We created the most vivid and memorable messages possible within our budget (for example, we had no video).

We posit that these messages will encourage the belief that untreated water is unsafe and hands not washed with soap are unsafe. Thus, we posit this treatment will increase rates of water treatment with chlorine and handwashing with soap.

At the same time, our own previous research(Albert, Luoto, and Levine 2010; Luoto et al. 2011, 26–28) and the literature cited above (Carpenter 2010) indicate that messages about health risks usually do not motivate large changes in preventive behaviors. Thus, we expect only modest increases in safe behaviors.

2.3.3 Theory of change: Engaging disgust and shame

For this intervention our key assumptions are:

1. People in this community perceive eating feces is disgusting
2. People do not want their neighbors to see them doing disgusting things, as they will feel shame and fear social sanction and loss of status.
3. People do not currently think that there is fecal contamination on hands after rinsing with water alone, nor in drinking water from the community tap.

Shame will be mobilized largely in settings where people care about the opinion of the observer. Thus, our auxiliary assumptions include that there is sufficient social cohesiveness and ability to observe unsafe behaviors to mobilize norms of shame, and that our intervention has sufficiently long-lasting effects to make the new norms persist. We recognize these auxiliary assumptions may not hold.

Most residents of Dhaka grew up elsewhere and have moved to their current compounds. If people have weak ties within a compound, or they recently moved into the compound or expect to move away soon, many people may not care strongly about what their neighbors think of them. (It is also possible that newcomers are especially sensitive to their neighbors' views.)

Our theory of change requires people care about what their neighbors think of them, because sanctioning takes time, can be unpleasant, and risks further unpleasant confrontations. Social sanctions may include correction and rebuke from the observer and criticism, mockery or ostracism from both the observer and from other neighbors who hear about the behavior from the observer. Sanctioning to enforce norms is more likely if there is a "meta-norm" that people in this community sanction norm-breakers. With that meta-norm in place, someone who sanctions a norm breaker receives social approval, while failing to sanction can lead to risks of lower status and incurring sanctions oneself.

Similarly, our theory of change requires that people believe their safe water and hygiene behaviors are frequently observable by their neighbors who have taken up these new norms. Thus, we must have trained a sufficiently high share of the compound. To the extent that only high-status people can sanction other high-status people, we must also have trained enough high-status people.

Our final assumptions are that people will retain knowledge from the training and will anticipate their neighbors have retained knowledge from the training.

If these assumptions hold, then teaching people there is fecal contamination after washing hands with water alone and in untreated drinking water from the community tap will mobilize their feelings of disgust. We explained (with a vivid demonstration) how feces came into contact with community tap water, starting from open defecation and leaking latrines, to sewers and into leaky pipes. (Bengali lacks a word corresponding to "crap" and we wanted more emotional engagement than if we used the words corresponding to "feces' or "poop." Thus, we used a vulgar word that corresponds to the English word "shit.") We explained that shit remains on hands after defecation unless wash them with soap. We also explained that if they treat drinking water with chlorine or a filter and wash hands with soap, they remove the dangerous and disgusting contamination that comes from shit.

We also emphasized that their *neighbors'* failure to treat drinking water or failure to wash hands with soap spreads shit and disease to *their* family. We used role-playing and vivid stories to emphasize that if neighbors see someone serving untreated water or

failing to wash hands with soap they will feel disgusted. We emphasized that the compound can work together to ensure families do not eat shit.

Because we taught many neighbors from their compound at the same time, participants knew that their neighbors also know that is fecal contamination on hands after rinsing and in drinking water from the community tap.

Our theory of change posits participants perceivedtheir neighbors would be disgusted at hands not washed with soap or untreated water. Thus, participants would feel shame and would fear social sanction if their neighbors observed them not using soap after defecation and not treating drinking water.

All of these shame-based implications are amplified if an person perceived his or her behavior was likely to be observed, especially if the observed person cared about the opinion of their neighbors and perceived the observers can influence the group, that the observers had higher status relative to the person, and that the observers felt strongly about the importance of water treatment and handwashing with soap.

Fear of sanctions may encourage more participants to treat their drinking water and remind their children treat water and to drink treated water. Further, participants who treat water may have encouraged neighbors to treat water and to avoid or sanction neighbors who drink or serve untreated water.

Treating the water, in turn, will lead to higher rates of detectable chlorine and lower rates of bacterial contamination detected with H2S tests. Disgust and fear of shame will also increase willingness to pay to subscribe to the water treatment product.

Almost identical hypotheses hold for unwashed hands. In that case, fear of sanctions will lead participants to wash hands with soap.

2.4 Specific objectives or hypotheses

Hypotheses:

1. Residents of communities which have participated in an intervention emphasizing disgust and shame related to unsafe water procurement behaviors will:
    a. be more likely to regularly drink treated water than residents of communities which have received a standard public health intervention.
    b. have higher willingness to pay for water treatment products than residents of communities which have received a standard public health intervention.
2. Residents of communities which have participated in an intervention emphasizing disgust and shame related to unsafe hand hygiene behaviors will...
    a. be more likely to regularly wash hands with soap at key times than residents of communities which have received a standard public health intervention.

Specific Objectives:

1. Develop an intervention adjusted in the local context using disgust and shame eliciting messages to promote treating drinking water and hand washing with soap in low income urban housing compounds.
2. Develop a new group version of the Becker-DeGroot-Marschak (BDM) procedure to measure willingness to pay of compound members for shared hardware.
3. Develop a new survey instrument to measure behavioral determinants of hand washing and water treatment like disgust and shame or social pressure.
4. Identify new methods for measuring hand washing and water treatment behavior.
5. Compare the effectiveness of the disgust- and shame-based interventions with standard public health interventions.
6. Measure demand for compound-basedliquid chlorine dispensers, and differences in demand between thestandard health treatment and the disgust and shame treatment.
7. Measure the effective of providing low-cost compound-level soapy bottles on handwashingbehavior.

3 Methods

3.1 Trial design

We randomly assigned two-thirds of eligible, consenting compounds (435 compounds) to a treatment arm. We assigned the other third (216 compounds) to a control group. Control compounds were randomized and data collection remained pending for a later period with availability of funds. Moreover, our objective was to assess the effectiveness of the disgust and shame messages over standard public health messages not over a control population. We compared the outcomes between the standard and disgust intervention arms.

3.2 Participants

3.2.1 Settings and locations where the data were collected

This study was conducted in compounds in slums of Dhaka, Bangladesh. Compounds are clusters of households, typically located around a small courtyard, sharing a common toilet, water source and cooking facilities. This setting was chosen for the following reasons:

- Poor water quality
- High incidence of water-borne disease (diarrhea, cholera)

- High population density and multiple households sharing water sources allow for improve potential scalability / sustainability of dispenser intervention
- Public visibility of water collection and handwashing habits make social norms a potentially powerful tool for behavior change
- Experience with chlorine dispenser intervention in Dhaka, as an arm of the Introduction of Cholera Vaccine in Bangladesh (ICVB) study

Against these advantages, the setting had the following disadvantages:

- Unfamiliarity with chlorine, possible taste barrier
- Heavily subsidized natural gas (zero marginal cost) available for boiling
- High population mobility leading to number of compound residents exposed to treatment decreasing over time
- Groups from disparate backgrounds and not necessarily subject to mutual social influence

Six communities around Dhaka were identified as most compatible with our study, on the basis of poor water quality, high incidence of water-borne disease, and high population density: Mohammedpur, Mirpur, Badda, Korail, Khilgaon and Bashabo. Of these, Korail was excluded, primarily because sanitary conditions were so poor that (a) handwashing and chlorination of household water alone would not improve health and (b) appeals to disgust and shame would not be effective, since fecal matter was essentially unavoidable, even with consistent handwashing and household water treatment.

3.2.2 Visit 1 and eligibility criteria for participants

Within the chosen field sites, enumerators searched for compounds which match all the essential criteria and match at least 2 of the preferred criteria and which are located at least 75 footsteps2apart from each other.

The essential compound criteria were:

- Between 6 and 18 Households
- Shared water source
- Physical space exists to hold a compound meeting here or nearby
- No other large water or hygiene interventions going on at this time
- Primarily Bengali language

The preferred criteria were:

- Use of water source is visible to others
- Shared kitchen
- Shared toilets visible from common area

Enumerators asked managers of eligible compounds for written consent to participate in the study. If they did not agree, the enumerators moved to the next suitable compound. (See participant flow in Figure 3)

If the compound manager signed the consent form, then enumerators recorded stratification data (listed in Section3.3 below) and performed baseline data collection (see Section 3.5)

3.3 Assignment to treatment

We used a stratified, cluster-randomized design, with the compound as the cluster. Sample compounds were divided into four strata based on compound size and presence of gas burners connected to the municipal gas supply. Compound size was chosen

2 Footsteps were determined to be the most intuitive unit of distance for our FRAs to use.

because social dynamics could be very different in small versus large compounds. Gas was chosen because compounds with gas had a low-cost alternative to chlorination (that is, boiling).

Strata:

1. Gas & Fewer than 8 households per compound
2. Gas & 8 households per compound or more
3. No Gas & Fewer than 8 households per compound
4. No Gas & 8 households per compound or more

Within strata, compounds were assigned to one of eight distinct study arms according to the following ratios.

1/6th: Disgust & Shame – Hand washing & Water treatment – Collective BDM

1/6th: Standard Health – Hand washing & Water treatment – Collective BDM

1/6th: Disgust & Shame – Hand washing & Water treatment – Individual BDM

1/6th: Standard Health – Hand washing & Water treatment – Individual BDM

1/12th: Disgust & Shame – Water treatment only – Collective BDM

1/12th: Standard Health – Water treatment only – Collective BDM

1/12th: Disgust & Shame – Water treatment only – Individual BDM

1/12th: Standard Health – Water treatment only – Individual BDM

Hand washing arms were given double weight to assure sufficient statistical power to test the effect of the disgust and shame treatment on take up of the hand washing technology.

Compounds were assigned to treatment or control immediately after giving consent. Enumerators called, or sent an SMS text message to the field office staff providing stratification covariates (gas status and compound size). The field office staff then used a pre-printed randomization list along with the covariates provided to assign treatment or control. If treatment was assigned, the enumerator conducted a baseline survey and scheduled a marketing visit. If control was assigned, the enumerator notified the compound that we would return to collect data.

The randomization list used by the field office staff consisted of four separate lists, one for each stratum, of permutations of two treatments and one control ((T,T,C),(T,C,T),(C,T,T)). Here is an image of the top of two of the four lists:

> 8 Households			
gas		no gas	
assign ment	compound_id	assign ment	compound_id
C		C	
T		T	
T		T	
T		T	
T		T	
C		C	

Further assignment of treatment compounds to treatment arm took place 2 to 3 days after visit one. Enumerators provided a list of newly enrolled compounds along with relevant covariates to the office research assistant daily. Every two to three days, the research assistant entered the list into our randomization computer program which randomized the order of the compounds, then computed the optimal treatment

allocation. The randomization program uses Atkinson's optimal sequential allocation algorithm(Atkinson 1982; Atkinson 1999). The final allocation of compounds, by treatment and stratum, is shown in table 30, section 9.

3.4 Interventions

3.4.1 Visit 2: Marketing meeting: Promotion and free trial

Intervention specialists conducted a three-hour integrated behavior change/product promotion program in the compound from September 2011 through December 2011.The content of the program, which varied by treatment arm, was developed during piloting and is briefly described below under "Meeting Content and Hardware".

At the end of the program, compounds were offered a four-month free trial of the hardware specific to their study arm (see "Meeting Content and Hardware").

3.4.2 Visit 3: Reminder visits

Roughly two weeks after the first meeting (September 2011 to December 2011), compounds received a reminder visit. Intervention specialists met with available individuals,reiterated messages from the first meeting, discussed concerns with the new equipment, and helped troubleshoot problems. A second reminder visit was conducted at a three and half month interval (December 2011 to March 2012).

3.4.3 Visit 4: Sales mechanism coaching

3.4.4 Roughly four months after the first meeting(January 2012 to April 2012), intervention specialists visited compounds to set up a time and date for the second compound meeting and to prepare the compound members for the sales procedure (described below). Intervention specialists met with residents to explain the process and held a practice willingness-to-pay elicitation procedure for a real good (detergent). This meeting was conducted to make the compound members familiar with the bidding process. Since most of the participants were female and individual decision making was difficult in most cases, we let them now possible five prices in the coaching meeting so that they can discuss with their spouse before bidding in the actual auction meeting. Visit 5: Sales offer (BDM)

Within a week or less of the coaching meeting, at the end of the extended free trial,intervention specialists returned to treatment compounds to conduct the sales meeting (January 2012 to April 2012).

The intervention specialists first conducted a brief (30-minute) refresher of the behavior change and product promotion program from the first meeting. The intervention specialists then offered the compound an opportunity to purchase one year of use of the dispenser, including maintenance and regular resupply of chlorine. The sales offer was made via a modified Becker-DeGroot-Marschak (BDM) mechanism(Becker, Degroot, and Marschak 1964), adapted to a group decision (described in detail in Sections 3.6 and 3.7).

3.4.5 Fee Collection visits

Every month for a year after the sales meeting (February 2012 to January 2013) and quarterly for next one year (March 2013 to April 2014), staff visited compounds who subscribed to the hardware to collect fees and record compoundphysicalobservation (detailed in Section 3.7 below).

3.4.6 Hardware

We distributed water treatment hardware at all participating compounds.

- One wall mounted liquid chlorine dispenser
- Two 15L reservoirs

- Two plastic stools (Figure 1)

We asked compound members to treat the reservoir water with liquid chlorine from the dispenser, with three turns for each 15L of water. For convenience we also marked one of their own vessels for five litres of water and instructed to add one turn of liquid chlorine from the dispenser. We advised residents to share the responsibility for filling the reservoirs. Hardware were removed from the compounds if they lost the auction or decided not to take further subscription.

At two-thirds of the compounds (selected at random) we also distributedhandwashing hardware, consisting of:

- a soapy water bottle for handwashing. The soapy water bottle is a plastic bottle containing 30g of detergent dissolved in 1.5L of water (Figure 2).The soapy water bottle is a plastic bottle white in color with icddr,b logo printed on it. A two taka ($0.03) sachet (30g) of detergentis needed to make 1.5L of soapy water. 1.5L lasts for about two weeks for a five-member family. We distributed 2-3 sachets per households per month, usually to the compound manager.
- During the free trial we provided free detergent (2 sachets monthly per household). If a compound subscribed to the chlorine dispenser, we continued to provide free detergent. If the compound did not subscribe, they retained the plastic bottle but we stopped delivering free detergent.

3.4.7 Meeting content

All compounds:

All 435 compounds were assigned to receive water treatment promotion interventions, whether standard or disgust- and shame-based. Two-thirds were randomly selected also to receive handwashing promotion. 142 compounds received disgust and shame messages for handwashing promotion and 149 compounds received standard messages.

Standard public health intervention:

The standard public health intervention meetings weremodelled after pre-existing high quality water treatment and/or hand washing interventions, in particular those in use by ICDDR,B. The content, communicated primarily as a presentation accompanied by flip charts, included explaining and demonstrating how germs can enter our bodies via untreated water and unwashed hands, how they can make us sick and lead to death, and how these risks can be reduced by practicing safe water and hand hygiene behavior.3

Disgust- and shame-based intervention:

Disgust and shame based intervention messages were developed through rigorous piloting. Before development of the intervention messages we surveyed the community to understand the existing practice of handwashing and water treatment. We also gathered information on what is considered to be disgusting by the community people, norms related to hygiene practice, and social dynamics related to hygiene. We developed various intervention messages and field tested each of the items individually. We assessed the effectiveness of the items quantitatively through a semi-structured questionnaire and also performed qualitative assessment through face-to-face interview. We selected the most effective and feasible messages to create a complete set of package.

3 Copies of the intervention presentation are at http://faculty.haas.berkeley.edu/levine/courtyards/.

The disgust and shame meetings contained similar explanations and demonstrations of contamination mechanisms and risk. While it mentioned germs and disease, the message placed most emphasis on the presence of fecal matter in or on contaminated objects and neighbors' role in spreading fecal matter to their families. The intervention specialists used the more vivid and harsh local term equivalent to "shit" to elicit a stronger reaction from participants. They communicated that we sometimes unknowingly serve shit to our family by not washing hands with soap or by not treating our drinking water. They emphasized how shit can spread between people, especially neighbors, to encourage people to care about others'behaviors and what others think of their behaviors. The disgust and shame intervention was intended to be more vivid than the standard intervention, and used different methods of communication in addition to flip charts, such as demonstrations, role play and storytelling. We demonstrated how shit gets into our drinking water and how shit can remain on hands if they are not washed with soap. We designed a special container to show how water becomes contaminated through leaky pipes, which we named the "Disgust box demonstration". We used UV light and UV powder to show invisible dirt or shit on hands when not washed with soap. We engaged the participants in role plays to illustrateneighbors' role in spreading shit.4. The messages were delivered in a courtyard meeting. The meetings were scheduled at a convenient time of the compound members when maximum participants were present. The first promotional meeting were long and took more than an hour. Reminder visits were shorter in small groups with 3-4 households at a time. 1-3 small group meetings were done for the reminder messages.

3.4.8 Changes in the protocol

Chlorine concentration and water reservoirs

Several iterations of piloting were conducted to optimize the hardware. The first piloting phase used a single 40L bucket and a chlorine concentration of 2.5%. The buckets are binstypicallyused for bathing or washing clothes. We customized them for water storage by attaching a tap at the bottom side. The goal was simplicity: one bucket (both for drinking water and handwashing among HW treatment compounds) and a single turn of the dispenser wheel to provide the proper dosage of chlorine.However, pilot users found the 40L bucket difficult to refill (due to its weight) and often too large for the available space. Furthermore, the 2.5% concentration gave a very strong smell during dosing and users reported that it took a long time for the strong chlorine taste to diminish. As a result, usage was very low: there were zero users out of fourteen pilot compounds attwo weeks follow-up.

A second round of piloting used two 15L reservoirs and a 0.3025% chlorine concentration. This lower concentration required three turns of the dispenser wheel to treat 15L of water. The size and weight of the reservoirs were more accepted by users, it was durable, it looked more like a water vessel, and it was safer for water storage because it had a firm lid.

In spite of the lower concentration of chlorine, in the part of this pilot phase users continued to report the odor and taste to be too strong to use.

To address this problem, in the main study we implemented a gradual increase in chlorine concentration. Water quality in Dhaka slums is varies according to season, with the poorest quality during the monsoon season (roughly June through September). We started our intervention well in advance of the monsoon season, and used a 1/3 strength dose (approximately 0.1008%), which was adequate to treat water during that season. After four months, we increased the concentration to approximately 0.2343%. Finally, we increased the concentration again to full strength (0.3025%) about a year after the

4 Copies of the intervention presentation are at http://faculty.haas.berkeley.edu/levine/courtyards/.

first promotion meeting. This phase-in process was intended to allow users to acclimate slowly to the smell and taste.

Added a reminder visit before the sales mechanism

We initially planned to combine teaching the complex sales mechanism and refreshing the behavior change messages into one meeting. In fact, teaching and practicing the sales mechanism took longer than expected. As a result, we added a separate refresher session at which intervention specialists revised the promotional message. The intervention specialists also checked the hardware and made any necessary repairs or modifications.

3.5 Data collection

3.5.1 Visit 1: Baseline data collection

After agreeing to signed consent, we collected the baseline data for the compound: two household surveys, including a rapid physical observation, and four additional household rapid physical observations (see Section3.7 below).

Household representatives who participated in baseline data collection were randomly selected from the pool of compound members present. The Field Research Assistant (FRA) asked the available members of the household who was responsible for collecting water and invited that person to participate. If the invited person declined, the FRA repeated the random draw and invitation until six consenting compounds were selected. The first two households received the questionnaire and household physical observation form and the last four received only the household physical observation.

3.5.2 Visit 6: Midline data collection

Three and half months after baseline data collection and first promotinal meeting, enumerators collected midline surveys at two households and household physical observation forms at six households (see section3.7). Enumerators attempted to interview the same households as before, but if they were not present after returning to the compound for a second attempt, then enumerators randomly selected new households from among those who were living in the compound at the time of the first promotion meeting and that had adults present.

3.5.3 Visit 7: Structured observation

Two months after the intervention began, staff carried out a 5-hour structured observation of handwashing and water treatment behavior in all compounds (see section 3.5.3). The observer introduced him/herself and asked for verbal consent from the compound manager (written consent was taken during enrolment). The observer selected a place to sit from where s/he could observe the handwashing station, latrine and the cooking space. S/he was allowed to move a few steps from his/her position to observe an event but was not allowed to follow anyone. The structured observation instrument had three sections: handwashing event, water handling events, and social interaction event. We recorded eight items for handwashing: whether both hands were cleaned, use of soap, hand drying methods, location of the handwashing station, water source and using the soapy water bottle. For water events we looked for use of chlorine for treating water, drinking chlorinated water, and use of chlorine dispenser and the reservoirs. For social events we observed any interaction between neighbors related to hygiene or water treatment.

3.5.4 Visit 8: Endlinedatacollection

Seven months after baseline data collection and first promotional meeting, Enumerators enumerated endline surveys at two households and household physical observation forms at six households using procedures similar to those at midline (see Section3.7 below).

Survey data was collected at three time points; at baseline, 3.5month midline and 7month endline. Rapid observation or spot check was done for 21 times from the baseline until end of two years of fee collection that is the end of the study.

3.5.5 Fee collection visits

Fee collection visits started about five months after the first promotional meeting with the subscriber compounds only. 106 compound own the auction and fee collection was done from these compounds. We assessed handwashing and water treatment behavior of the compound members using spot checks, hand inspection and chlorine residual test of the household drinking water. All measurements were done only at subscriber households. In the first year the payment visit was made monthly. In the second year we made quarterly visits to the compound to collect the payment and supply chlorine stock. A reminder phone call was made before each visit. We added the chlorine residual test and hand inspectionafter about 7 months of payment collection.

The spot check was performed from the first payment collection through the 24-month payment collection visit. We have 1 to 15 rounds of spot checks per compound and 8 to 12 rounds of chlorine testing and hand inspection per compound.

Among 424 enrolled compounds, payment collection started with 106 compounds that won the auction and made their first payment. Five compounds missed spot check data because this data collection was introduced after they had quit paying. By the end of 12 months, 37 compounds refused to continue subscription and 69 compounds continued for the second year. Another 36 dropped out during the second year of payment collection and 33 compounds made the last payment. Thus, data from fee collection visits are from self-selected compounds and self-selected households within those compounds. 15 fee collection visits were made to each compound during two years of follow up period.

3.5.6 Controls datacollection

During the initial selection of compounds we selected a matched group that we hoped would act as a control group. This set of compounds received no intervention. Due to funding constraints, we did not collect data on this potential control group during the main study. Eight months after the endline (April 5 to July 9, 2012) we took water samples from these other compounds under control sub-study (September 19, 2012 to January 23, 2013).

3.6 Outcomes

We are interested in the effect of our intervention on (a) the proportion of people who drink treated water; (b) the proportion of people who wash their hands with soap after fecal contact and before eating; and (c) willingness to pay for one year of use of the chlorine dispenser, including regular supply of chlorine and maintenance visits. Because the compounds that chose to subscribe to the service were a selected sample, in (a) and (b) we focused on measures recorded during the free trial period, before the subscription decision was made. We inferred changes in these proportions based on direct observation and multiple proxy measurements outlined below.5

3.6.1 Water treatment

Primary outcomes:

- Water quality, as measured by H2S tests for bacterial contamination of source drinking water and of drinking water stored in the home.
- Treatment with chlorine, as measured by chlorine residual tests of drinking water by color wheels.

Secondary outcomes:

- Self-report of water treatment behaviour (van de Mortel 2008). Self-reported water treatment is often higher than true water treatment, but is correlated with true treatment(Luoto et al. 2011).
- Physical condition of water treatment station: Present, usable, filled.
- Structured observation of water treatment (boiling or chlorination).

3.6.2 Willingness to pay (WTP)

Our final measure of how the interventions affected water treatment was willingness to pay for a compound-based chlorine dispenser. We measured collective and individual willingness to pay (WTP) using two variants of the Becker-DeGroot-Marschak (BDM) mechanism, an incentive-compatible real-money sales mechanism(Becker, Degroot, and Marschak 1964). We randomly assigned compounds to one of two variants:

- collective BDM, in which the compound collectively determined its WTP;
- individual BDM, in which individual households determined their own willingness to contribute to the public good.

In both cases, the classic BDM, in which a single individual or household is offered a single, private good, is modified to provide incentives for true revelation of maximum WTP in its given context. We included the two variants of BDM to observe how well the compound could solve its collective action problem (measured in collective BDM) and how much heterogeneity there was in individual WTP within compound (measured in individual BDM). Both procedures are described in greater detail in Section 3.7.5below.

3.6.3 Handwashing

Because there are no well-established, field-practical measures of handwashing with highly established validity, we relied on several proxies:

- cleanliness score observed in direct observations of hands. The score is one when all the three parts of the hand (palm, fingers and nail bed) showed no visible dirt or unclean appearance. The score is zero when any of the part was visibly dirty or showed unclean appearance.

5We did not measure health outcomes because their connection with these behaviors is fairly well established and such measurement would require a larger and more costly study.

- % hands observed washed with soap at 3 key events we mentioned (among # of key events we mentioned observed during 5 hours of structured observation per compound)

We also have three secondary measures of handwashing:

- Soap was present near the latrine, handwashing station or kitchen
- Used soap when we asked respondent to show how she washes hands after defecation
- % of self-reported handwashing at last key events (after defecating, before eating, and, if respondent has a child, after cleaning a child's anus)

Most of these maternal measures have weak correlations with predicting child diarrhea(Luby et al. 2011b).

3.7 Measures[6]

3.7.1 Household survey

Households were surveyed at baseline (less than a week before installation of hardware for the free trial), midline (three to four months after the baseline, which was just prior to the end of the free trial) and endline (roughly 7 months after the baseline, roughly three months after the end of the free trial). We interviewed two households per compound; we tried to reach the same household in all survey rounds. If any household had migrated we randomly selected another household from the compound. The questionnaires for each wave (baseline, midline and endline) were largely the same excepting corrections made during implementation and small changes for administering at different times. The questionnaire measured self-reported usage, reports of neighbors' usage, knowledge and practice pertaining to handwashing and safe water, perceptions of risk and severity of diarrhea, norms/ beliefs,feelings,andreactions related to disgust and shame, social networking, and demographic information.

3.7.2 Household physical observation

At the baseline, midline, and endline and during structured observation, we performeda Household Physical Observation. We randomly selected two households that received the survey. Four other households per compound were again selected randomly for physical observation. In each visit we tried to reach the original sample but if they were not available at that time another household were selected randomly. In order to achieve six adult and six child samples per compound, adults and children were not matched in some of the compounds. Sometimes we had to take multiple adult or multiple child sample from the same household when enough distinct households were not present at the time of visit.

At each of the six households we collected:

- Chlorine residual testing of stored drinking water (not taken at baseline)
- Hand cleanliness inspections of child (see scoring sheet in Appendix 4)
- Handwashing demonstration by motherorcaregiver (or by any adult available, if caregiver not present).
    - We asked her to demonstrate how she washes her hands after defecation. We evaluated her performance as outlined in Appendix 4.
- Drinking water quality viaH2S testing

3.7.3 Compound rapid observation

6Copies of the surveys and other data collection instruments are at http://faculty.haas.berkeley.edu/levine/courtyards/.

Compoundrapidobservationstook place during all compound visits and consisted of identifying the location andusage status of handwashing and water treatment hardware.

3.7.4 Structured observation

After nearly two months of the free trial, the field team conducted the 5 hour structured observation in all the enrolled compounds (25 October 2011-19 March, 2012). Field workers, 80% of whom were female, took consent from 418 compounds; six compounds had beendemolishedbeforethis follow-up.

Field workers explained they were visiting to observe daily household activities. They stayed in the compound from 7:00am to 12:00pm. This was a culturally acceptable hour for visitors and the typical time for a range of personal hygiene and food preparation behaviors.

Using a pre-tested instrument, field workers noted handwashingbehavior at key times: before preparing food, before eating or feeding a child, after defecating and after cleaning a child's anus who had defecated. They observed all available household members and noted whether they washed their hands, and whether they used water, with or without soap, and/or ash for food and defecation related events. Field workers also noted water treatment behaviors. Hand hygiene and water treatment behaviors were assessed using the score sheet in Appendix 6.

3.7.5 Willingness to pay

We measured compound members' willingness to pay for a monthly subscription to the shared chlorine dispenser and storage reservoir. To elicit willingness to pay, we developed two group-level adaptations of the Becker-DeGroot-Marschak (BDM) mechanism. One adaptation measures compound members' collective willingness to pay, the other measures willingness to pay of individual households within a compound. Compounds were randomly assigned one of the two BDM methods.

As both procedures are complex, we held a coaching meeting one week prior to the sales meeting in each compound at which we explained the BDM procedure and carried out a real-money practice for some laundry detergent. The sales meeting was roughly 48 weeks after the free trial began.

At the *individual* group's coaching meeting, the procedure was explained as:

> The full cost of the chlorine dispenser, if you were to pay for it yourself, would be 35 taka ($0.44) per household, per month. But to encourage people to use this facility, icddr,b is paying most of the cost and is providing it to you at a much lower monthly subscription fee. As it is not possible to offer all compounds the same low subscription fee, to be fair to all we will draw the price from a lottery. There are six possible prices that could be drawn from the lottery: 20, 15, 10, 7, 5, 2 taka ($0.25 to $0.03) per household per month.
>
> It is very important that you think and discuss within your household, about these possible prices—20, 15, 10 , 7, 5 and 2 taka per household per month—and decide what prices your household would agree to pay as a monthly subscription fee.
>
> On the day of the lottery, the icddr,b staff person will first ask you each individually if you would agree to the prices I just mentioned first, then will reveal the lottery price. If you and all of your neighbors have already agreed to pay that price, then you will keep the facility and the lottery price will be your monthly subscription fee, to be paid by each of the households in your compound. If you or any of your neighbors have not agreed to that price, then ICDDRB will remove the facility from your compound.

We then explained why respondents' best choice was to reveal their true maximum willingness to pay (among the prices listed), as their agreement on a price did not affect

the price which they would be offered. During and after the practice BDM procedure the icddr,b staff further highlighted the negative consequences of expressing a willingness to pay that was higher or lower than their own. After the practice BDM procedure with laundry detergent, we assessed respondents' understanding of the procedures and explained further when necessary.

We provided each household with a price sheet (piece of paper with the six prices listed) and a pencil and asked them to put a check next to prices they would pay. To assist members who couldn't read, we provided colored pictures of currency equalling the taka amounts corresponding to each number. To allow households privacy while stating their willingness to pay, we asked that participants not peek at others price sheets, and instructed participants to fold their sheets twice before handing them to the enumerator. Enumerators brought extra price sheets to allow participants to change their stated willingness to pay.

At the dispenser sales meeting we repeated the instructions from the coaching meeting,thenallowed participants to fill out and submit their price sheets. As a basic understanding check, we asked the group whether they believe their stated willingness to pay can influence the price they will get in the lottery. We explained again when necessary.We asked respondents whether they had considered if absent members would use and pay for the dispenser. We reminded them that regardless of this expectation, they should only be considering their own household's willingness to pay and gave them a chance to submit a new maximum price if they had taken absent households into account.

We then announced the highest price that everyone had agreed to pay (that is, the minimum stated maximum willingness to pay). As a final understanding check, we asked the group to recall what would happen if the lottery price were greater than, less than or equal to this price. If there was confusion or disagreement with this outcome, we encouraged discussion between members and allowed each household to change their maximum WTP again before finally revealing the lottery price.

The procedure was slightly different for the *collective* group procedure, as the participants chose a total price, not a price per household. At the *collective* group's coaching meeting, the procedure was explained as:

> The full cost of the chlorine dispenser, if you were to buy it yourself, would be 350 taka per month. But to encourage people to use this facility, icddr,b is paying most of the cost and providing it to you at a much lower monthly subscription fee. As it is not possible to offer all compounds the same low subscription fee, to be fair to all we will draw the price from a lottery. There are six possible prices that could be drawn from the lottery: 200, 150, 100, 75, 50 and 25 taka per month.
>
> It is very important that you think and discuss among yourselves about these possible prices—200, 150, 100, 75, 50 and 25 taka per month—and decide what prices your group would agree to pay as a monthly subscription fee before the lottery meeting.
>
> You should also work out how you are going to share the total subscription fee among yourselves. If all compound members would share the total fee, the per-household subscription fee will be lower, but if only some of you share, then the per-household subscription fee will be higher.
>
> On the day of the lottery, the icddr,b staff person will ask you if you would agree to the prices I just mentioned first, then will reveal the lottery price. If you (as a group) have already agreed to pay that price, then you will keep the facility and that lottery price will be your monthly subscription fee. If your group has not agreed to that price, then icddr,b will remove the facility from your compound.

As in the individual coaching meeting, we then explained why it was best to reveal one's true maximum willingness to pay, we conducted a practice BDM with laundry detergent, and assessed understanding.

At the dispenser sales meeting we repeated these instructions, allowed the group a moment to discuss, then asked the group which of the six prices their compound would agree to pay. We asked if they had considered absent members willingness to pay, but we didn't discourage it, as it was their collective willingness that was of interest. Additionally, we asked how they planned to divide the costs, equally, or unequally and why.

Before revealing the lottery price, we performed the same basic understanding checks: we asked if they believed their stated price could influence the lottery and we asked them to recall the hypothetical outcome given a lottery price greater, less than or equal to their price.

3.7.6 Payment data

After the sales meeting we measured continued enrolment and maintenance of the safe water (or safe water and handwashing) station on a monthly basis.

3.7.7 Qualitative data collection

During implementation our qualitative team irregularly visited sub-groups of our sample to gather data on how the intervention was going and to solicit ideas for improving the intervention and data collection.

3.8 Statistical methods

When comparing the effects of the traditional health message with the disgust and shame message we measured the primary and secondary outcomes for the water treatment behavioratall compounds received the intervention. For the handwashingoutcomes we analyzed the two third of the sample who received the handwashing intervention.

The unit of data collection was either compound, households or individual, depending on the measure.

For most outcomes we had baseline data. Thus, we compared across study arms either trends from baseline to 3.5 months after the first promotion meeting (still during the free trial) or trends from baseline to 7 months (when only some compounds still subscribed to the chlorine dispenser). We adjusted the p-value for Difference in Differences using GEE for compound level clustering.

We carried out structured observation once at 2 months of first promotion. For this cross-sectional measure we used GEE to calculate adjusted risk ratios.

To study the effects of the soapy bottle intervention, we compared the compounds who received no handwashing intervention and those who received handwashing interventions (pooling both health messages).

Winner compounds were analyzed as a subgroup to see changes among the subscribers at seven-month follow up.

3.9 Attendance

Figure 4, in section 10 shows attendance rates at each visit.

We randomly selected two households for survey at baseline. In the next visits during midline and endline we tried to reach the same household. If any household was absent we revisited the compound. If the household is migrated we replaced it with a randomly selected household preferably among those who were present at the first promotion meeting. For physical observation we randomly selected six households and tried to

reach them in the subsequent follow up. For absent or migrated households we replaced them with randomly selected new household.

3.9.1 Baseline characteristics

The compounds averaged 9 households and just over 2 toilets, almost always shared. The households were poor, but not the very poorest. For example, almost all had electricity, almost all cooked with gas (95%), 90% had a mobile phone, and 73% had a TV. At the same time, only 20% had a refrigerator, only 1% had a motorcycle, and the vast majority (93%) shared a latrine. Two-thirds of household heads had lower secondary education or more, and 20% had attended or graduated upper secondary (high school). The typical household had monthly income of 8000-12000 taka ($110-$160) for a household averaging 5 members.

90% said they washed with soap after defecating. At the same time, a much lower 68% said they washed with soap after the last time they defecated, the same percentage (67%) who washed using soap when we asked them to demonstrate how they washed after toileting. Both rates may be higher than actual, due to experimenter demand effects (where respondents give replies they believe the enumerator is expecting).

3.9.2 Balance checks

Table 30 shows balance on stratification covariates. Allocation ratios are nearly perfect except for in the no gas strata, where there were too few compounds to allocate all treatments evenly.Tables 1 through 6 show balance at baseline over the disgust and shame treatment.Two of 37 covariates differed significantly at the 5% level. Given that all other related covariates did not differ significantly, we dismiss this as spurious correlation.There is no evidence against ourallocationprocedure working as would true randomization.

4 Results

4.1 Baseline: Setting and Subjects

The study sites were low income communities of urban Dhaka, the capital of Bangladesh. Participants lived in compounds which are clusters of households. The households typically share a cooking area with a few gas burners, a water source, and a few latrines. Mobility was high in the compounds, with most residents born outside of Dhaka and many stating they did not expect to live in the compound for much longer.

Descriptive statistics on compounds are in (Table 2). The average compound size was 9 households per compound. Descriptive statistics on subject households are provided in (Table 3).Most households had cement floors (95%), and a corrugated iron roof (95%). Most of the compounds had improved shared latrine (standard intervention 91% and disgust intervention 94%) and used natural gas for cooking (95%). Only one fifth of the household heads completed primary education. More than 60% of households had a total monthly income of 4000-12000 taka ($54-$162). Descriptive statistics on household assets are provided in (Table 4). Most of the households had working mobile phones (90%) and televisions (73%). The respondents were almost always (93%) adult females, with a mean age of 29 years (SD ±8yrs). While we did not collect a household roster, other studies in similar neighborhoods found the typical household had 5-6 members(SHEWA, B data, May 2008).

More than 90% of the compounds were supplied with piped water to the courtyard (Table 5). Almost two-thirds of the compounds reported always treating water before drinking (65%). At the same time, almost three-quarters of the household drinking water were contaminated within 48 hours of H2S test (72%).

Sixty seven percent claimed to have used soap to wash hands after the last time they defecated, although only 18% mentioned handwashing their usual times to wash with included before eating, after defecation and after cleaning child anus (the three key times promoted in the intervention) (Table 14). Enumerators observed both soap and water available at the handwashing station (that is, at the public water source) in 17% of compounds. Most of the respondents said after defecation and before eating are their usual key times of handwashing (Table 15).About two-thirds of the respondents used soap during the handwashing demonstration at baseline (Table 15). Soap was used in the demonstration, even when not observed at the water source, because residents kept soap in their homes.

4.2 Effect of standard vs. disgust treatment

4.2.1 Effect on water treatment

We measured antecedents of water treatment including self-reported feeling of disgust at untreated water and observed social interactions related to water treatment (or its lack).

We measured water treatment using the following indicators:

- Self-reported treatment of drinking water
- Spot check for use of the water storage vessels and boiling practice
- Willingness to pay for the water treatment product (liquid chlorine)
- Detection of residual chlorine in household drinking water
- Observed water treatment behavior through structured observation
- Determination of water contamination using H2S test

Self-reported feeling of disgust at untreated water was similar at baseline (63% at standard and 61% at disgust compounds, Table 10). These self-reported rates rose 10 percentage points among the standard intervention arm and a marginally statistically significantly higher 17 percentage points in the disgust intervention arm (double difference P = 0.07). This marginally higher rate of self-reported disgust was not maintained, as both rates ended up about 70% (71% standard arm, 70% disgust arm, P on double difference = 0.82).

Self-reported perceived disgust did not lead to social interactions supporting safe water treatment. Social interactions related to water treatment (e.g., praise for treatment or criticism for not treating) was very rare, occurring in only 3 of 2739 cases in the standard arm and 6 of 2346 water gathering events in the disgust arm (Table 12, p = 0.22).

The rate of self-reported "always treat drinking water" rose from a baseline level near 65% for both groups to 83% for both groups at the midline(Table 7).This rate then dropped part-way back to 73% at the standard intervention arm and 67% at the disgust intervention arm. None of the double differences are statistically significant (all P > 0.4).

While the rate of self-reported treatment rose, the rate we observed boiling occurring during our spot checks declined a bit. (Recall this is a compound-level measurement, so not directly comparable to the household level measures.) At baseline our enumerators saw people boiling water at 28% of compounds by the midline and endline those rates had fallen by more than half at both standard intervention arm and 67% at the disgust intervention arm households. (Recall these changes over time could just be seasonal or other factors, not due to the intervention). None of the double differences across study arms approach statistical significance (all P > 0.3).

At the two-month follow-up there wasdetectable chlorine in household drinking water within 24 hours of reported treatment at 8% of the homes (Table 11). This low rate is consistent with (but ever lower than) low uptake in previous distributions of free trials of

chlorine point-of-use water products in Dhaka (e.g. 10% inLuoto et al. (2011). The rate at disgust homes is not quite as low (12%), with the 4 percentage point advantage for the disgust group marginally significant (P = 0.09).

The slightly higher usage in disgust compounds was not sustained. At the 3.5 month midline, still during the free trial of the dispenser, we asked about whether stored water in the home had been treated with the chlorine(Table 12a).Rates of self-reported use were the same (25%) in both study arms. Among self-reported users, we detected chlorine at identical rates (33%) in both study arms. Thus, there was detectable chlorine at only 8% of homes (=25% * 33%), which was the same in both study arms.

At the 7-month follow-up we asked about whether stored water in the home had been treated with the chlorine in compounds that were subscribing to the hardware. Rates of self-reported use were very similar (40% standard, 41% disgust arm, P = 0.85 adjusted for clustering) in both study arms. Among self-reported users, we detected chlorine a bit more often at disgust arm homes (50%) than standard arm homes (42, P=0.80%). This result is not too encouraging, as rates of detectable are also similar if we compare all subscribing homes, not just those that self-reported use (17%, 34/198, in standard vs. 21%, 48/225, in disgust arms, P = 0.59 adjusting for clustering).

Moving to the intention to treat framework, we can think of the potential number of tested homes at the endline as being similar to that at the midline. (Ideally we would adjust for compounds that were destroyed, but such adjustments would not change these results.) In that case self-reported use was about 7.9% in the standard arm (83/1046), which was not statistically distinguishable from the 9.2% (91/992) in the disgust arm (P = 0.16 without adjusting for clustering). Similarly, shares with detectable chlorine were also close (3.9%, 41/1046, in the standard arm and 4.6%, 46/992, P = 0.21 without adjusting for clustering).

We can measure both chlorine and boiling during the structured observation of water treatment, which took place 2 months after the intervention began (Table 8). Of all cases of gathering water or drinking water directly, only 5% of the standard intervention arm used chlorinated water from the reservoir or added chlorine to untreated water. The rate was not only very slightly higher for the disgust intervention arm (6%, difference P= 0.16). If we add in boiling, the rate of effective water treatment for the standard arm was 15% while it was 19% for the disgust arm (difference P = 0.11).

The water was contaminated (as measured by the H2S test at 48 hours) at 72% of households at baseline (Table 11). This level fell to 60% at standard arm and 61% at disgust arms at the 3.5 month follow-up (difference P = 0.91). Unfortunately, the improvement in water quality was not sustained in either arm, with contamination rates rising to 78% (control) and 77% (disgust) arms after seven months (difference P = 0.47). The contamination rate was significantly less in the winner compounds (65% vs 82%) in both arms (P=<0.01, Table 12b)

The increase of water contamination at 7 month could have several reasons. We saw that in the disgust intervention arm compound members started using chlorine at 3.5 months which reduced water contamination at this point. In 7 month follow up during spot check we found disgust treatment compounds were still treating water through boiling. Residual chlorine level decreased at this time and water contamination increased. This may be due to incomplete boiling, improper storage of boiled water, and poor water-handling practices. Other explanations could be cleaner supply water during

monsoon led more infrequent treatment. It is also possible increased availability of gas for cooking may have increased (evidently ineffective) boiling.7

We collected chlorine residual data within subscribing households of compounds that won the auction and continued payments. Within this self-selected group almost all of the chlorination hardware appeared to be recently used in both study arms (95% of 560 spot checks in the standard intervention, 93% of 541 spot checks in the disgust intervention, P = 0.28). The rates of positive chlorine tests in the water of subscribing homes were also almost identical across the study arms (73% with detectable chlorine in 1388 spot checks in the standard arm, and 74% of 1274 checks in the disgust arm, p=0.77).

4.2.2 Effect on willingness to pay

Nearly half (43%) of compounds dropped out of the program and returned their dispenser before the sales meeting. The dropout rate was similar for compounds in the standard and disgust treatment arms (54% vs. 60%, P = 0.135) as well as for compounds in the collective and individual groups (57% vs. 57%).

When a compound had a sales meeting, only 37% of households attended on average (compounds are composed of 6 to 16 households, median = 9). Again, these shares were similar for the standard and disgust treatment arms (39% vs. 36%, P = 0.305) and the collective and individual groups (35% vs. 40%, P = 0.108). Essentially everyone who attended the sales meeting was a self-reported user (94%). Moreover, a large majority (74%) of self-reported regular users attended the sales meeting.

We estimated that a small business or NGO running at scale could promote and distribute chlorine dispensers and visit monthly to replenish chlorine and collect fees at a break-even cost of 200 to 300 taka ($2.50 to $3.50) per compound per month (we use the market exchange rate of March 1, 2012). Among compounds with a sales meeting, willingness to pay for the collective sales meeting had a mean of 68 taka ($0.83) per month (SD = 45). Only 3% of compounds with a sales meeting and 2% of all compounds were willing to pay $2.50 per month. These rates were, again, similar for compounds in the standard and disgust treatment arms (3% vs. 3%).

The mean willingness to pay of households in the individual group was 8.2 taka ($0.10) per household per month (SD = 4.9). The weakest link (that is, lowest willingness to pay among those participating) had a mean of 7.8 taka (SD = 5.1). Both results are similar for households in the standard and disgust treatment arms (8.9 vs. 7.6 for all households, P = 0.070; 8.4 vs. 7.3 for weakest link, P = 0.197).

See tables 24 – 29 in Section 9 for more sales meeting results.

4.2.3 Effect on subscription to the chlorine dispenser

We sold subscriptions to the dispensers at a highly subsidized prices from 25 to 150 taka ($0.30 to 1.80) per month. At those prices, 21% of the compounds subscribed to dispensers (89 / 420). An additional 4% (17 compounds) agreed to subscribe at the sales meeting, but never made the first payment. The subscription rate was similar for compounds in the standard and disgust treatment arms (22% vs. 21%, P = 0.585) as well as for compounds in the collective and individual group procedures (19% vs. 22%, P = 0.164).

Dropout after subscription was fairly low, 82% of those compounds that made a first payment were still subscribing at the 10 month mark (73/89). About 7% of payments

---

7Find that water quality in Dhaka is worse during floods : Sirajul Islam, M., et al. "Faecal contamination of drinking water sources of Dhaka city during the 2004 flood in Bangladesh and use of disinfectants for water treatment." *Journal of applied microbiology* 103.1 (2007): 80-87.

arrived incomplete or late. The 17% of surviving compounds that were subscribers six months after subscriptions began (and 10 months after the intervention began) paid an average of 43 taka ($0.50) per month.

At six months the subscription rate was nearly identicalfor compounds in the standard and disgust treatment arms (17% vs. 16%, P = 0.780) as well as for compounds in the collective and individual groups (17% vs. 17%, P = 0.982).

At the seven-month mark, in compounds that subscribed to the dispensers, 41% of households self-reported using chlorine and among those households fewer than half had detectable chlorine. Thus, 18% of homes in subscribing compounds had detectable chlorine, which is less than 4% of the entire treatment sample.

4.2.4 Effect on handwashing

In addition to the handwashing proxy measures discussed above, we also measured self-reported feeling of disgust at hands not washed with soap and observed social interactions related to handwashing with soap. Among compounds that received the handwashing intervention, feeling of disgust at hands not washed with soap was nearly universal at baseline (95% for both standard and disgust compounds). This rate increased to 100% at the 7-month follow-up at both standard and disgust compounds. The ceiling effect implies that no difference in trends would be detectable (Table 13).

Self-reported perceptions of disgust did not translate into frequent social interaction. We observed social interactions related to handwashing (or its lack) at less than 1% of the opportunities (that is, the sum of toileting events, food preparation, and eating). If our observation period covered half of toilet use and was a representative day, then our results imply a typical compound would have about 5 interactions per week and a typical person would give or receive feedback related to handwashing about five times per year. These interactions occurred at similar rates at the standard and disgust treatment arms (P = 0.63, Table 17).

Self-reported use of soap during all 3 promoted key times were the same for both arms at baseline (19 %). The increase was 8 percentage points larger among disgust compounds at 3.5 months (p = 0.04). Similarly at 7 months, the percent increase was 9 points larger (p = 0.02) for the disgust compounds (Table 14).

Availability of soap and soapy water at the handwashing station increased over time in both treatment arms. The increase was slightly greater in the standard treatment arm, but the gap was not statistically significant (48 vs 43 p=0.5, Table 15).

The soapy water bottle was slightly more regularly in use in the disgust treatment arm (7 %, p = 0.17) when spot checks included other locations such as kitchens, courtyards, and latrines, in addition to handwashing stations and checked for evidence of recent use (Table 15).

About 20% of the adult respondents at baseline had visibly clean hands (for all of palm, finger tips and nails; 20% for standard arm and 21% for disgust arm). This rate increased to 26% for the standard arm and 27% for the disgust arm at the midline (3 ½ months later). The double difference is tiny and not statistically significant. The rate of clean hands then fell slightly to 24% for the standard arm and rose slightly to 30% for the disgust arm; the double difference (relative to baseline) of +5 percentage point relative increase in disgust compounds is marginally statistically significant (P = 0.10).

We have the most confidence in our structured observation of handwashing, which took place 2 months after the intervention began (Table 16). As with the other measures, there is no large or statistically significant difference in observed rates of handwashing between the standard andthe disgust arms. For example, people washed both hands with soap after toileting 11% of the time in the standard intervention arm and 12% at the disgust intervention arm (P = 0.53).

4.3 Effect of the handwashing treatment on handwashing

We measured multiple proxies for handwashing:

- Spot checks of the availability of soap and water at the handwashing station
- Self-reported use of soap and water at the promoted key times
- Observation of use of soap during handwashing demonstration
- Hand inspection to determine cleanliness of hands
- Hand wipes on a white paper
- Observed handwashing practice through structured observation during promoted key times

We compared shares or trends in those in compounds that had the handwashing training and the soapy water bottle added to the regular intervention to compounds that received only the standard or the disgust water treatment intervention.

During the free trial the soapy water appeared to be in use at about 59% of compounds where we distributed it (3.5 month follow-up). The distribution of the soapy bottle drove rates of either soap or our soapy bottle to rise from about 16% at baseline to 59% at the 3.5 month follow-up (still during the free trial) and remained at 44% at the seven-month follow-up (Table 14).

When the free trial ended, only 44% of handwashing compounds had the soapy bottle present at the handwashing station (7-month follow-up). Most all of these compounds (65%) were subscribers to the chlorine deliveries. Recall that the free delivery of detergent continued only for this subset of compounds; others retained the bottle but had to provide their own detergent. Although the soapy bottle provides a potentially low-cost means to keep soap near the handwashing station. Few (38%) non-subscribing compounds continued to use when the free trial ended. Nevertheless, the incidence of soap near the handwashing station remained higher than at compounds that never received the handwashingintervention (31%), even for non-subscribers.

Self-reported use of soap during promoted key times were consistently more frequent (about 8 percentage point increase in both follow ups) in the disgust treatment arm than the standard treatment arm (p=0.05) (Table 18). Self-reported washing hands after last defecation was similar at baseline for compounds that would and would not receive the handwashing intervention (68%).During the free trial (at 3.5 months) this share rose 7percentage points more at handwashing compounds (81% no handwashing intervention vs. 89% handwashingcompounds p=0.07). By the endline, when most compounds no longer subscribed to the chlorine (so no longer received free detergent), some of this advantage was no longer present (82% without the handwashing and 87% of handwashing compounds reported washing with soap, p=0.37).

The use of soap in the handwashing demonstration was significantly higher in the handwashing arm at both follow up visits; the increase was 9 percentage points (p=0.001) at 3.5 months and 6 percentage points (p=0.03) at seven months (Table 19).

Our enumerators observed hand cleanliness (palm, fingers and nails) of 1-3 people per household in 5-6 households per compound. The rate of adults having both hands that appeared clean was 21% in baseline and rose to 27% at the 3.5 month midline and 30% at the 7-month endline (Table 166).The increase in the share of respondents with clean hands was almost identical for the compounds with and without the handwashing intervention, so the double differences were small and nowhere close to statistically significant.

We have the most confidence in our structured observation of handwashing, which took place 2 months after the intervention began (Table 20). In the arms without the

handwashing intervention (but had either the standard message or disgust and shame message about safe water), 9% washed both hands with soap after toileting. This rate rose to 13% for those with the handwashing intervention (P = 0.03).The increase was similar in magnitude for washing hands after cleaning a child's anus (from 23 to 32%), but the smaller sample size left the result statistically insignificant (P =0.14). We found no effect of the handwashing intervention on washing hands before eating, which remained below ½% in all study arms.

In summary, almost all adults in Bangladesh claim to wash hands after defecation and about 2/3 use soap when we ask them to demonstrate how they wash after toileting. We observed rates after toileting closer to 9% in our arms without the handwashing intervention. While many visits to the latrine involve only urination, this rate may be increased over a true baseline because we had given a safe water intervention in these compounds (which may have also increased hygiene awareness and knowledge) and we had an observer present (even if often inconspicuous). Thus, we find (as have others) that many poor Bangladeshis do not wash with soap after defecation. The handwashing training plus provision of soapy water bottle raised handwashing by modest amounts but had no effect on observably clean hands.

4.4 Willingness to pay in collective vs. individual sales meeting

Among compounds with a sales meeting, willingness to pay for the collective sales meeting had a mean of 68 taka ($0.83) per month (SD = 45). The mean willingness to pay of households in the individual group was 8.2 taka ($0.10) per household per month (SD = 4.9). The weakest link (that is, lowest willingness to pay among those participating) had a mean of 7.8 taka (SD = 5.1).

In a weakest link decision, the compound's total obligation is the lottery price times the number of households in the compounds (not just the number of households attending the sales meeting). Given the typical number of households per compound (median = 9), the weakest link willingness to pay from the individual sales meeting is similar in magnitude to the compound-level willingness to pay of the collective sales meeting (7.8 taka * 9 households = 70.2 taka).

At the same time, we asked households to consider only their own WTP, assuming their neighbors also agreed. This bidding rule was incentive compatible in the weakest-link format if all neighbors were present. In that extreme case, our results would show lower per household WTP for the individual group (8.2 taka mean) than for the collective group (20.5 taka mean if we divide collective willingness to pay by the number of households who agreed to pay prior to the meeting). However, our suggestion to ignore neighbors is not always incentive compatible when neighbors are absent. If attendees know absent members will not pay, then stating their full willingness to pay puts them at risk of agreeing to cover some payment for the absent members as well.

More than half (57%) of bidders in individual sales meeting stated they *did* take their expectation of absent members' willingness to pay into account while bidding (contrary to our instructions). Thus, we cannot tell if the individual group truly had lower per household willingness to pay.

The collective sales meeting could increase collective willingness to pay if those with high willingness to pay agree to pay more than those who have (or claim to have) lower WTP. In fact, 89% of participants in collective compounds stated they plan to share the cost equally among payers; thus, the ability to have unequal costs probably did not increase WTP by much.

In results not shown, the hand washing and water treatment intervention compounds had similar willingness to pay as those with only the water treatment intervention (65 vs. 74 taka in the collective group, P = 0.309; 7.9 vs. 7.7 taka in the individual group, P = 0.833).

Though some households may have deliberately not followed instructions in the individual group, considering absent members willingness to pay, our other measures show a high level of understanding of the BDM process. When attendees were asked as a group whether their stated willingness to pay could affect the lottery price they received, 81% of compounds members correctly agreed, no. When asked about hypothetical lottery results, 95% of compounds answered correctly.

See tables 24 – 29 in Section 9 for more sales meeting results.

4.5 Summary of Qualitative Findings

4.5.1 Effects of the disgust and shame intervention

Where the disgust and shame messages were delivered most of the respondent mentioned the demonstration of disgust box that made them aware how there might have shit in their household water. Most of them stated that they would feel disgusted observing the plastic poop and the process how shit could be mixed with their drinking water. A substantial share of the participants also said the flipchart's picture of leaky pipe passed through the drain made them changes their behavior regarding water treatment. Participants who received the disgust and shame message also mentioned preventing disease, following the example of other users, and freeing themselves from the hassle of boiling water as the reasons for drinking chlorinated water.

A focus group participant said:

> *"I felt disgusted when they [the ICDDR,B presenters] had shown how shit [gue] could be mixed with water…If someone defecates on the leaky water supply pipe and if it rains then the rainy water will make the shit mix with the supplied water. Thus water becomes dirty. That makes me think to drink this [chlorinated] water.*

In qualitative interviews a few months after the presentation, most of the respondent recalled the disgust box demonstration, describing it as *plastikergu* [plastic shit],*rubberergu*[rubber shit],*borotob* [big vessel], *dibbaarplastikergu* [small pot and plastic poop], etc. A good number of respondents also recalled the 'leaky pipe' shown in the flip chart that showed how their household water could be contaminated.

A respondent from an area where disgust and shame messages were delivered mentioned the presenters

> *told us there may be shit in our water. Sanitary pipe may have link with the pipe of water supply. They had shown with a plastic poop how it could be mixed with our water. We generally could not see it.*

After discussing previous water treatment practices respondents were asked how they felt after the promotion meeting. The respondents who would drink raw water directly from their compound's water source expressed their feelings about their previous practices using various terms. Those from the area that received disgust and shame messages the term disgusting [*grinnalaga*] was the most commonly uttered followed by the term dirty work [*nungrakaj*], doing wrong [*bhulkaj*], feeling uneasy or embarrassed [*shokshoklaga*], unfair [*bisri*], and religiously forbidden work [*haram kaj*].

One respondent from an area receiving disgust and shame messages said:

> *Now I feel embarrassed [shokshok] that I would drink dirty water before. After seeing the plastic poop [demonstration of disgust box] I never think of drinking raw water. ---Remembering that I feel disgust.*

A respondent from another compound said:

I feel bad thinking that I would drink shit mixed water that was religiously forbidden [haram]. Everybody will acknowledge that shit is the dirtiest thing in the world. I must feel disgust now [thinking that].

4.6 Comparison to true controls

During the initial selection of compounds we selected a matched group that we hoped would act as a control group. This set of compounds received no intervention.

Due to funding constraints, we did not collect data on this potential control group during the main study. Several months after the endline (April 5 to July 9, 2012) we took water samples from these other compounds (September 19, 2012 to January 23, 2013). Due to seasonal effects, the results are not directly comparable to our endline for the treatment sample; nevertheless, they are interesting.

The main study showed 78% of water samples were contaminated at 7 months of intervention (as measured by the H2S test at 48 hours), and these samples were largely drawn during the monsoons when water quality typically is at its worst in Dhaka. Among the winners this contamination was 64% and among the loser compounds the contamination was 82% ($p<0.01$).Several months later, 85% of the H2S test for the comparison group were contaminated. Thus, our compounds that did not subscribe and the control group had similar rates of contamination, while the subscribing compounds had somewhat lower rates of contamination than the control group.

4.7 Results up to two years

At the end of one year, rates of continued subscription were similar for the standard and the disgust arms (71% and 60% of those who ever subscribed, $P = 0.21$). During the end of the second year about half the remaining compounds dropped out. Rates of subscription remained similar by study arm (29% of standard arm who ever subscribed remained and 33% of disgust arm, $P = 0.60$).

Among the subset of compounds that subscribed to the chlorine dispenser after the free trial, the handwashing intervention had a lasting effect. Only 11% of the 383 spot checks over two years period at compounds without the handwashing intervention had both soap and water at the handwashing location, while 89% of the 740 spot checks at compounds with the handwashing intervention had both soap and water present ($P = 0.05$, clustering by compound to adjust for repeated measures). There is some evidence the soapy bottle increased handwashing. At spot checks enumerators inspected hands of an adult and of a child (15 years or younger). Only 35% of the 421 times an enumerator inspected an adult's hands was the report that both hands were clean (that is, no visible dirt in the finger tips, palm or nails). That rate rises to 45% of 829 hand inspections at compounds with the handwashing intervention ($P < 0.001$). For children (15 years and under) rates were not different (both 0.30, $n = 325$ with and 648 without the handwashing intervention).

5 Discussion

5.1.1 The chlorine dispenser

Overall, usage rates of the chlorine dispenser were low. During the free trial there was detectable chlorine in household drinking water within 24 hours of reported treatment at 8% of the homes (Table 13, at the two-month follow-up). This low rate is consistent with low uptake in previous distributions of free trials of chlorine point-of-use water products in Dhaka (10%, Luoto, *et al.*, 2011). Thus, even with free distribution and the two interventions we tested, chlorine dispensers are not appealing products in this setting.

Only about a fourth of households participated in the auction. Their mean willingness to pay was about $0.10 / month for the individual auction or $0.87 per month for the

compound-level collective auction. As noted above, true willingness to pay might be higher or lower than we measured. Regardless of any mismeasurement, with the level of marketing in this intervention the market for compound-level chlorine dispensers in Dhaka is a fairly small one. With that said, roughly 27% of compounds expressed willingness to pay of $1.20 per month or more, implying a niche market exists.

5.1.2 The soapy water bottle

Adding our handwashing messages to either study arm and providing the soapy water bottle increased the share of handwashing stations that also had soap from 18% at baseline to 43% at endline (seven months post-intervention). Regardless of their message, in the arms without the handwashing intervention about 9% washed both hands with soap after toileting. (These data are from the structured observation about 2 months post-intervention, during the free trial.) This rate rose to 13% for those with the handwashing intervention (P = 0.03). Thus, there is evidence that providing the soapy water bottle and our relatively modest behavior change intervention can improve handwashing by a meaningful – but still modest – amount.

The messages appeared useful, as rates of access to soap and self-reported handwashing remained slightly higher even when we no longer provided free detergent for the soapy water bottle.

5.1.3 Disgust and shame

Our qualitative evidence suggested the disgust and shame intervention was powerful for many who attended the marketing meeting. However, most compound residents did not attend the intervention and as a result were not directly exposed to the intervention. Importantly, very few men observed the full presentation. Furthermore, our qualitative evidence suggests that women cannot typically put social pressure on men to conform to norms of water treatment and handwashing with soap.

The shame and disgust intervention did not increase either self-reported feelings that untreated water is disgusting or self-reports that hands not washed with soap were disgusting. Given the lack of success in changing perceptions, it is unsurprising that there is no consistent evidence that disgust and shame improved demand for or usage of the chlorine dispenser or soapy bottle relative to a traditional behavior change message focused on germs and health.

5.2 Limitations and generalizability

5.2.1 Internal validity

Each of our measures has limitations.

For both water treatment and handwashing

- *Structured observation* was visible to participants. Although the observer did not state that water treatment and handwashing were the objectives of the observation, having an observer may have increased participants' water treatment and handwashing with soap.
- *Self-reports of water treatment and of handwashing* are subject to experimenter demand effects. That is, we told people treating water with chlorine and handwashing with soap are important. We then asked them if they treated water and washed with soap. Many respondents may reply "Yes" due to the common inclination to tell people what they want to hear.

  Consistent with this hypothesis, after three and half months of intervention during midline survey 96% of the respondents say they wash with soap after defecating. At the same time, a lower 89% say they washed with soap after the

last time they defecated, which was also the share (87%) who washed using soap when we asked them to demonstrate how they wash after toileting.

In structured observation, only 11% of respondents washed at least one had with soap after toileting. This rate is not strictly comparable to rates of self-reported or demonstrated handwashing after defecation, as many visits to the latrine may have been solely for urination. At the same time, participants may have increased their handwashing in response to being observed. In short, it is plausible that there is substantial over-statement of handwashing with soap using self-reports or when people demonstrate for an audience that they believe expects handwashing with soap.

- *Direct observation of the status of the chlorine station and soapy water bottle*can confirm lack of use, but are only indirect measures of water treatment or of handwashing with soap. That is, the chlorine station is useful for water treatment, but it is neither necessary (as people can boil) nor sufficient (as people need not use the chlorine station). Similarly, having the soapy water bottle is useful for handwashing with soap, but it is neither necessary (as people can wash with soap from their homes) nor sufficient (as people need not use the soapy water bottle).

Willingness to pay

While in principle BDM is incentive compatible under modest assumptions (Horowitz 2006), in practice many factors can lead to a divergence between BDM valuation and the subject's true maximum willingness to pay(Berry, Fischer, and Guiteras 2012). Necessary conditions for incentive compatibility include:subjects understand how the complex procedure operates, believe they will not interact again with this seller (so that a high stated willingness to pay will not raise prices in the future), think hard about the incentives and do not just follow the rule of thumb "Understate your true willingness to pay or it will be used against you," and do not over-bid just to "see the price in the envelope." All of these assumptions hold imperfectly in our setting.

The group setting adds new complications.

- In the group BDM, households have an incentive to understate their true WTP if they believe it will lower their share of the compound's total WTP. {Something about equal sharing rule}
- In the individual BDM, we let the "weakest link" of each compound determine the compound's WTP. When some neighbors are not present, those stating a WTP may reasonably fear that they will end up paying more than they have agreed to. For example, if only half of a compound of 10 households is present and the other half have told everyone they will pay nothing, then a stated willingness to pay of 20 taka per month per household (which implies 200 taka per compound) will lead to a bill of *40* taka for the 5 households who stated a price they would pay. Thus, a reasonable respondent might report only half his or her true willingness to pay, knowing the participating families will end up paying double shares. Although we instructed respondents to ignore this issue, 57% stated they did, in fact, take into account the willingness to pay of those families not represented at the sales meeting.

For handwashing

Direct observation of hands is an imperfect proxy of washing with soap, largely because in these settings most hands become dirty very quickly. Error can also arise if different observers rate hands differently.

For water treatment

- *Self-reported knowledge* about the importance of treating drinking water had a ceiling effect: essentially all reported knowledge at baseline.

- *Detectable free residual chlorine*can under-state true chlorination if all the free chlorine had reacted with the storage container or contaminants in the water. Our primary measure, detectable chlorine, can over-state *effective* water treatment if some people under-dose their water (perhaps to reduce the taste and smell of chlorine). Our secondary measure, correctly dosed chlorine, checks for this possibility. Detectable chlorine can also understate effective water treatment because some people boil their water (and a few may use an effective filter).
- *Chlorine* can predict low self-reported diarrhea due to politeness or due to omitted variables such as wealth or concern for health that both increase chlorine use and reduce diarrhea (or increase self-reported diarrhea, conditional on true illness).
- *The water reservoir* can be absent but households can still use chlorine if the households chlorinate directly to their own container.
- *H2S test* can over-state human fecal contamination because bacteria from other sources such as cow feces (for example) can also lead to a positive H2S test. Our suspicion is that most bacterial contamination in urban Dhaka that leads to a positive H2S test comes from human feces, but we have no direct evidence on this point.

Attrition

Attrition among compounds in this study was very low, 4% (see Figure 5). Most (11 of 16) attrition at the compound level was due to compounds being destroyed. This factor is not correlated to treatment (p=1).

At the household level attrition was 51%. However no correlation is found with the treatment group (p=1).

Implications for internal validity

Fortunately, most of these sources of error are uncorrelated with the treatment arm, with a few exceptions. Tothe extent the disgust-and-shame treatment led to increased sensitivity to acting "appropriately" when observed, and if participants thought icddr,b staff viewed using chlorine and handwashing with soap as "appropriate" behaviors, both structured observation and self-reported behaviors may over-state the effectiveness of the disgust-and-shame treatment.

Given the consistently small effects we find, and the fact that the estimated effects are not strongest on these outcomes, we believe these potential sources of bias have little effect.

5.2.2 External validity

- Having water treatment be a compound-wide club good (that is, an excludable public good), like our chlorine dispenser, is unusual for water treatment
- Handwashing and water treatment station were usually somewhat visible, which might increase effects of shame.
- High mobility, high expected mobility, and low concern about the views of their neighbors might decrease shame

5.3 Interpretation

5.3.1 Implications for research and evaluation

Measuring handwashing

Structured observation reported a 12% of those exiting the latrine washed hands with soap. This rate is much lower than the rate of self-reported washing with soap after defecating (96% at midline)or the share that demonstrated use of soap when asked to demonstrate how they wash hands after defecation (87% in midline). Some of the gap is presumably because many visits to the latrine were solely for urination and rates of

handwashing with soap are presumably lower after urination than defecation. At the same time, participants may have increased their handwashing in response to being observed. Thus, as many others have observed(Stanton et al. 1987; Biran et al. 2008; Cousens et al. 1996; Manun'Ebo et al. 1997), there appears to be substantial upward bias in self-reported handwashing. Thus, we recommend that objective measures based on structured observation, motion detectors in soap, etc. be the primary measure for evaluating changes in hand washing practices in handwashing interventions.

Measuring willingness to pay

The Becker-DeGroot-Marschak (BDM) procedure is a fairly standard way to measure willingness to pay in a (largely) incentive-compatible fashion. Incentive compatibility fails if people do not understand the instructions, if they do not believe the BDM procedure will be carried out as described, if they use a rule of thumb to understate true WTP, or if they fear repeated interaction (so current stated WTP will lead to higher future prices).

Our two implementations of the group-level BDM were novel, as far as we know. Both procedures seemed fairly effective.

The "weakest-link" sales method can lead to under-reporting of true WTP if participants fear their neighbors or icddr,b will somehow use their high stated WTP against them. It also leads to bidding below (or perhaps above) their true willingness to pay if some participants know their neighbors will bid zero, so there is no chance of the purchase occurring. At the same time, stated WTP can exceed true WTP if some participants with high WTP offer to pay the share of those with low (or zero) WTP.

The group auction is not incentive-compatible at the individual level, as households had an incentive to falsely state a low WTP hoping to free ride on the higher WTP of their neighbors. This incentive is weakened in compounds with a heuristic of equal payment per household. The group-level procedure has the potential to increase WTP relative to a weakest-link procedure if those with high WTP either pressure those with low demand to join in or if they take on an above-average share of payments. Thegroup-level bid also permitted compounds to decide their own rules for adjusting for household size. At the same time, the transaction costs of agreeing to any rule other than equal sharing may have lowered group WTP in this auction.

Both auctions suffered from absences, as it was unclear how to interpret bids when many compound households had no representative at the meeting. Any future implementations will need to ensure essentially all participants submit information on their WTP.

5.3.2 Implications for policy

It is plausible that an intervention that repeated the message more often and that reached a higher share of the compound might have been more effective. At the same time, we have no consistent evidence in favor of that hypothesis.

5.3.3 Implications for theory

We engaged in one long presentation and 2short presentations of the behavior change material. Although the disgust and shame material tested well, the intervention was not sufficient to create feelings of disgust about untreated water or hands not washed with soaps. Thus, we have clear evidence that changing attitudes requires more effort than our budget permitted.

The failure to change attitudesleaves open the question of whether the underlying theory is correct. It remains to be seen if an intervention that successfully mobilizes disgust and fear of shame can lead to sustained increases in healthful water and hygiene behaviors.

5.3.4 Implications for project managers

Multiple visits are necessary to change attitudes – and presumably even more effort is needed to change behavior.

The soapy bottle was a very low-cost and effective means of promoting handwashing, especially where multiple households share a single hand-washing station.

5.3.5 Implications for chlorine dispenser manufacturers

The recommended chlorine concentration is calibrated to the poorest water-quality season of the year. In settings where chlorine demand (the amount of chlorine required to make water safe) varies by season, it may make sense to market chlorine-based products in less-dangerous seasons and provide (or recommend) a lower chlorine concentration at first, gradually adjusting the concentration depending on the season. Starting with a somewhat lower concentration should permit many users to acclimate to the smell and taste of the chlorine.

6 Other information

6.1 Protocol

The trial protocol is at http://faculty.haas.berkeley.edu/levine/courtyards/.

6.2 Ethics statement

In a courtyard meeting participants were briefed on the details of the study and afforded the opportunity to ask questions and receive answers. Enumerators obtained informed written consent from the compound manager prior to inclusion in the study. For follow-up visits we used verbal consent for each of the procedures like structured observation, rapid observation or hand inspection.The qualitative team obtained written consent from the compound manager for their activities separately. This study was reviewed and approved by the Ethical Review Committee at icddr,b.

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had access to all the data in the study.

7 References

Akerlof, George A, and Rachel E Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics* 115 (3): 715–753. http://www.jstor.org/stable/2586894.

Albert, Jeff, Jill Luoto, and David I. Levine. 2010. "End-User Preferences for and Performance of Competing POU Water Treatment Technologies Among the Rural Poor of Kenya." *Environmental Science & Technology* 44 (12) (June 15): 4426–4432. doi:10.1021/es1000566. http://pubs.acs.org/doi/abs/10.1021/es1000566.

Alderman, Harold, John Hoddinott, and Bill Kinsey. 2006. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers* 58 (3) (July 1): 450–474. doi:10.1093/oep/gpl008. http://oep.oxfordjournals.org/content/58/3/450.

Ashraf, Nava, James Berry, and Jesse M. Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review* 100 (5): 2383–2413. doi:10.1257/aer.100.5.2383. http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.5.2383.

Atkinson, Anthony C. 1982. "Optimum Biased Coin Designs for Sequential Clinical Trials with Prognostic Factors." *Biometrika* 69 (1): 61–67. http://www.jstor.org/stable/2335853.

———. 1999. "Optimum Biased-coin Designs For Sequential Treatment Allocation With Covariate Information." *Statistics in Medicine* 18 (14): 1741–1752.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak. 1964. "Measuring Utility by a Single-response Sequential Method." *Behavioral Science* 9 (3) (January 1): 226–232. doi:10.1002/bs.3830090304. http://dx.doi.org/10.1002/bs.3830090304.

Berry, James M., Greg M. Fischer, and Raymond P. Guiteras. 2012. "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana". Working Paper. http://r4d.dfid.gov.uk/Output/191387/Default.aspx.

Biran, A., T. Rabie, W. Schmidt, S. Juvekar, S. Hirve, and Valerie A. Curtis. 2008. "Comparing the Performance of Indicators of Hand-washing Practices in Rural Indian Households." *Tropical Medicine & International Health* 13 (2): 278–285. doi:10.1111/j.1365-3156.2007.02001.x. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-3156.2007.02001.x/abstract.

Boissiere, M., J. B. Knight, and R. H. Sabot. 1985. "Earnings, Schooling, Ability, and Cognitive Skills." *The American Economic Review* 75 (5) (December 1): 1016–1030. doi:10.2307/1818642. http://www.jstor.org/stable/1818642.

Cairncross, Sandy, Caroline Hunt, Sophie Boisson, Kristof Bostoen, Valerie A. Curtis, Isaac CH Fung, and Wolf-Peter Schmidt. 2010. "Water, Sanitation and Hygiene for the Prevention of Diarrhoea." *International Journal of Epidemiology* 39 (suppl 1) (April 1): i193–i205. doi:10.1093/ije/dyq035. http://ije.oxfordjournals.org/content/39/suppl_1/i193.abstract.

Carpenter, Christopher J. 2010. "A Meta-Analysis of the Effectiveness of Health Belief Model Variables in Predicting Behavior." *Health Communication* 25 (8) (November 30): 661–669. doi:10.1080/10410236.2010.521906. http://dx.doi.org/10.1080/10410236.2010.521906.

Checkley, William, Gillian Buckley, Robert H. Gilman, Ana MO Assis, Richard L. Guerrant, Saul S. Morris, Kåre Mølbak, Palle Valentiner-Branth, Claudio F. Lanata, and Robert E. Black. 2008. "Multi-country Analysis of the Effects of Diarrhoea on Childhood Stunting." *International Journal of Epidemiology* 37 (4) (August 1):

816–830. doi:10.1093/ije/dyn099. http://ije.oxfordjournals.org/content/37/4/816.

Chiller, Tom M., Carlos E. Mendoza, M. Beatriz Lopez, Maricruz Alvarez, Robert M. Hoekstra, Bruce H. Keswick, and Stephen P. Luby. 2006. "Reducing Diarrhoea in Guatemalan Children: Randomized Controlled Trial of Flocculant-disinfectant for Drinking-water." *Bulletin of the World Health Organization* 84 (1) (January): 28–35. http://www.who.int/bulletin/volumes/84/1/28.pdf.

Clasen, Thomas. 2009. "Scaling up Household Water Treatment Among Low-income Populations." *World Health Organization: Geneva*. http://whqlibdoc.who.int/hq/2009/WHO_HSE_WSH_09.02_eng.pdf.

Clasen, Thomas, Wolf-Peter Schmidt, Tamer Rabie, Ian Roberts, and Sandy Cairncross. 2007. "Interventions to Improve Water Quality for Preventing Diarrhoea: Systematic Review and Meta-analysis." *BMJ* 334 (March): 782–791. doi:10.1136/bmj.39118.489931.BE. http://www.bmj.com/content/334/7597/782.full.

Cousens, S., B. Kanki, S. Toure, I. Diallo, and Valerie A. Curtis. 1996. "Reactivity and Repeatability of Hygiene Behaviour: Structured Observations from Burkina Faso." *Soc Sci Med* 43 (November): 1299–308. 8913000. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8913000.

Curtis, Valerie A., Mícheál de Barra, and Robert Aunger. 2011. "Disgust as an Adaptive System for Disease Avoidance Behaviour." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1563) (February 12): 389–401. doi:10.1098/rstb.2010.0117. http://rstb.royalsocietypublishing.org/content/366/1563/389.

Curtis, Valerie A., L. O. Danquah, and R. V. Aunger. 2009. "Planned, Motivated and Habitual Hygiene Behaviour: An Eleven Country Review." *Health Education Research* 24 (August): 655–73. 19286894. PubMed. doi:10.1093/her/cyp002. http://www.ncbi.nlm.nih.gov/pubmed/19286894.

Drankiewicz, Danielle, and Lauren Dundes. 2003. "Handwashing Among Female College Students." *American Journal of Infection Control* 31 (2) (April): 67–71.

Ejemot-Nwadiaro, Regina I, John E Ehiri, Martin M Meremikwu, and Julia A Critchley. 2008. "Hand Washing for Preventing Diarrhoea." In *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004265.pub2/abstract.

Ezzati, Majid, Stephen Vander Hoorn, Anthony Rodgers, Alan D Lopez, Colin D Mathers, and Christopher JL Murray. 2003. "Estimates of Global and Regional Potential Health Gains from Reducing Muliple Major Risk Factors." *The Lancet* 362 (9380) (July 26): 271–280. doi:10.1016/S0140-6736(03)13968-2. http://dx.doi.org/10.1016/S0140-6736(03)13968-2.

Fessler, Daniel M.T. 2004. "Shame in Two Cultures: Implications for Evolutionary Approaches." *Journal of Cognition and Culture* 4 (2): 207–262. doi:10.1163/1568537041725097.

Gupta, S. K., M. S. Islam, R. Johnston, Pavani K. Ram, and Stephen P. Luby. 2008. "The Chulli Water Purifier: Acceptability and Effectiveness of an Innovative Strategy for Household Water Treatment in Bangladesh." *The American Journal of Tropical Medicine and Hygiene* 78 (June): 979–84. 18541780. http://www.ajtmh.org/content/78/6/979.abstract.

Haghighi, P, and P L Wolf. 1997. "Tropical Sprue and Subclinical Enteropathy: a Vision for the Nineties." *Critical Reviews in Clinical Laboratory Sciences* 34 (4): 313–341. doi:10.3109/10408369708998096.

Horowitz, John K. 2006. "The Becker-DeGroot-Marschak Mechanism Is Not Necessarily Incentive Compatible, Even for Non-random Goods." *Economics Letters* 93 (1). Economics Letters: 6–11. http://ideas.repec.org/a/eee/ecolet/v93y2006i1p6-11.html.

ICDDR,B. 2008. "Handwashing Behavior in Rural Bangladesh". Health and Science Bulletin 6(3). ICDDR,B. http://www.icddrb.org/what-we-do/our-library/journal-of-health-population-and-nutrition/forthcoming-articles/cat_view/52-publications/10042-icddrb-periodicals/10048-health-and-science-bulletin-bangla-and-english/10271-vol-6-no-3-english-2008/11786-handwashing-behavior-in-rural-bangladesh.

———. 2010. "Piloting Interventions to Promote Home Water Treatment and Handwashing in Dhaka". IRCCR,B Research Protocol PR-09048. ICDDR,B.

Judah, Gaby, Robert Aunger, Wolf-Peter Schmidt, Susan Michie, Stewart Granger, and Valerie A. Curtis. 2009. "Experimental Pretesting of Hand-Washing Interventions in a Natural Setting." *American Journal of Public Health* 99 (S2) (October 1): S405–S411. doi:10.2105/AJPH.2009.164160. http://dx.doi.org/10.2105/AJPH.2009.164160.

Kaplan, L M, and M McGuckin. 1986. "Increasing Handwashing Compliance with More Accessible Sinks." *Infection Control: IC* 7 (8) (August): 408–410. http://www.ncbi.nlm.nih.gov/pubmed/3638277.

Khan, Moslem Uddin. 1982. "Interruption of Shigellosis by Hand Washing." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 76 (2): 164–168. doi:10.1016/0035-9203(82)90266-8. http://www.sciencedirect.com/science/article/pii/0035920382902668.

Kremer, Michael, Clair Null, Edward Miguel, and Alix Peterson Zwane. 2008. "Trickle down: Diffusion of Chlorine for Drinking Water Treatment in Kenya." http://www.stanford.edu/group/SITE/archive/SITE_2008/segment_1/papers/de_giorgi_SITE_Trickle_Down.pdf.

Lorntz, Breyette, Alberto M. Soares, Sean R. Moore, Relana Pinkerton, Bruce Gansneder, Victor E. Bovbjerg, Helen Guyatt, Aldo M. Lima, and Richard L. Guerrant. 2006. "Early Childhood Diarrhea Predicts Impaired School Performance." *The Pediatric Infectious Disease Journal* 25 (6) (June): 513–520. doi:10.1097/01.inf.0000219524.64448.90. http://journals.lww.com/pidj/Abstract/2006/06000/Early_Childhood_Diarrhea_Predicts_Impaired_School.9.aspx.

Luby, Stephen P, Mubina Agboatwalla, Daniel R Feikin, John Painter, Ward Billhimer, Arshad Altaf, and Robert M Hoekstra. 2005. "Effect of Handwashing on Child Health: a Randomised Controlled Trial." *The Lancet* 366 (9481) (July 16): 225–233. http://linkinghub.elsevier.com/retrieve/pii/S0140673605669127.

Luby, Stephen P, Amal K Halder, Tarique Huda, Leanne Unicomb, and Richard B Johnston. 2011a. "The Effect of Handwashing at Recommended Times with Water Alone and with Soap on Child Diarrhea in Rural Bangladesh: An Observational Study." *PLoS Medicine* 8 (6): e1001052. doi:10.1371/journal.pmed.1001052.

Luby, Stephen P., Mubina Agboatwalla, John Painter, Arshad Altaf, Ward L Billhimer, and Robert M Hoekstra. 2004. "Effect of Intensive Handwashing Promotion on Childhood Diarrhea in High-risk Communities in Pakistan: a Randomized Controlled Trial." *JAMA: The Journal of the American Medical Association* 291 (21) (June 2): 2547–2554. doi:10.1001/jama.291.21.2547.

Luby, Stephen P., and Amal K. Halder. 2008. "Associations Among Handwashing Indicators, Wealth, and Symptoms of Childhood Respiratory Illness in Urban Bangladesh." *Tropical Medicine and International Health* 13 (June): 835–44. 18363587. doi:10.1111/j.1365-3156.2008.02074.x. http://www.ncbi.nlm.nih.gov/pubmed/18363587.

Luby, Stephen P., Amal K. Halder, Tarique M. N. Huda, Leanne Unicomb, and Richard B. Johnston. 2011b. "Using Child Health Outcomes to Identify Effective Measures of Handwashing." *The American Journal of Tropical Medicine and Hygiene* 85 (5) (November 1): 882–892. doi:10.4269/ajtmh.2011.11-0142. http://www.ajtmh.org/content/85/5/882.

Luby, Stephen P., Amal K. Halder, Carole Tronchet, Shamima Akhter, Abbas Bhuiya, and Richard B. Johnston. 2009. "Household Characteristics Associated with Handwashing with Soap in Rural Bangladesh." *The American Journal of Tropical Medicine and Hygiene* 81 (5) (November 1): 882–887. doi:10.4269/ajtmh.2009.09-0031. http://www.ajtmh.org/content/81/5/882.abstract.

Luby, Stephen P., A. K. Kadir, Y. Sharker, S. Islam, Yeasmin F., W.A. Brooks, and L. Unicomb. 2009. "A Pilot Study to Improve Hand Hygiene in a Low Income Urban Community in Dhaka, Bangladesh."

Luby, Stephen P., C. Mendoza, B. H. Keswick, T. M. Chiller, and R. M. Hoekstra. 2008. "Difficulties in Bringing Point-of-use Water Treatment to Scale in Rural Guatemala." *The American Journal of Tropical Medicine and Hygiene* 78 (March): 382–7. 18337330. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18337330.

Luoto, Jill, Nusrat Najnin, Minhaj Mahmud, Jeff Albert, M. Sirajul Islam, Stephen Luby, Leanne Unicomb, and David I. Levine. 2011. "What Point-of-Use Water Treatment Products Do Consumers Use? Evidence from a Randomized Controlled Trial Among the Urban Poor in Bangladesh." *PLoS ONE* 6 (10) (October 20): e26132. doi:10.1371/journal.pone.0026132. http://dx.doi.org/10.1371/journal.pone.0026132.

Madden, T. J., P. S. Ellen, and I. Ajzen. 1992. "A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action." *Personality and Social Psychology Bulletin* 18 (1) (February 1): 3–9. doi:10.1177/0146167292181001. http://psp.sagepub.com/cgi/doi/10.1177/0146167292181001.

Manun'Ebo, M., S. Cousens, P. Haggerty, M. Kalengaie, A. Ashworth, and B. Kirkwood. 1997. "Measuring Hygiene Practices: a Comparison of Questionnaires with Direct Observations in Rural Zaïre." *Tropical Medicine & International Health* 2 (11): 1015–1021. doi:10.1046/j.1365-3156.1997.d01-180.x. http://onlinelibrary.wiley.com/doi/10.1046/j.1365-3156.1997.d01-180.x/abstract.

Munger, Kristen, and Shelby J. Harris. 1989. "Effects of an Observer on Hand Washing in a Public Restroom." *Perceptual and Motor Skills* 69 (3) (December): 733–734. doi:10.2466/pms.1989.69.3.733. http://www.amsciepub.com/doi/abs/10.2466/pms.1989.69.3.733.

Niehaus, Mark D., Sean R. Moore, Peter D. Patrick, Lori L. Derr, Breyette Lorntz, Aldo A. Lima, and Richard L. Guerrant. 2002. "Early Childhood Diarrhea Is Associated with Diminished Cognitive Function 4 to 7 Years Later in Children in a Northeast Brazilian Shantytown." *The American Journal of Tropical Medicine and Hygiene* 66 (5) (May 1): 590–593. http://www.ajtmh.org/content/66/5/590.

Olembo, Lynette, FAD Kaona, Mary Tuba, and Gilbert Burnham. 2004. "Safe Water Systems: An Evaluation of the Zambia CLORIN Program". USAID Environmental Health Project. http://www.ehproject.org/PDF/Others/Zambia%20Report%20Format.pdf.

Pedersen, Darhl M., Sheila Keithly, and Karie Brady. 1986. "Effects of an Observer on Conformity to Handwashing Norm." *Perceptual and Motor Skills* 62 (1) (February): 169–170. doi:10.2466/pms.1986.62.1.169. http://www.amsciepub.com/doi/abs/10.2466/pms.1986.62.1.169.

Petri, William A., Mark Miller, Henry J. Binder, Myron M. Levine, Rebecca Dillingham, and Richard L. Guerrant. 2008. "Enteric Infections, Diarrhea, and Their Impact on Function and Development." *The Journal of Clinical Investigation* 118 (4) (April 1): 1277–1290. doi:10.1172/JCI34005. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2276781/.

Pittet, Didier, Anne Simon, Stéphane Hugonnet, Carmen Lúcia Pessoa-Silva, Valérie Sauvan, and Thomas V. Perneger. 2004. "Hand Hygiene Among Physicians: Performance, Beliefs, and Perceptions." *Annals of Internal Medicine* 141 (1) (July 6): 1–8. doi:10.7326/0003-4819-141-1-200407060-00008. http://dx.doi.org/10.7326/0003-4819-141-1-200407060-00008.

Porzig-Drummond, Renata, Richard Stevenson, Trevor Case, and Megan Oaten. 2009a. "Can the Emotion of Disgust Be Harnessed to Promote Hand Hygiene? Experimental and Field-based Tests." *Social Science & Medicine* 68 (6) (March): 1006–1012. doi:10.1016/j.socscimed.2009.01.013.

———. 2009b. "Can the Emotion of Disgust Be Harnessed to Promote Hand Hygiene? Experimental and Field-based Tests." *Social Science & Medicine* 68 (6) (March): 1006–1012. doi:10.1016/j.socscimed.2009.01.013. http://www.sciencedirect.com/science/article/pii/S0277953609000136.

Pruss-Ustun, Annette, Sophie Bonjour, and Carlos Corvalan. 2008. "The Impact of the Environment on Health by Country: a Meta-synthesis." *Environmental Health* 7 (February 25): 7. doi:10.1186/1476-069X-7-7. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2276491/.

Quick, R. E., L. V. VENCZEL, E. D. MINTZ, L. SOLETO, J. APARICIO, M. GIRONAZ, L. HUTWAGNER, et al. 1999. "Diarrhoea Prevention in Bolivia through Point-of-use Water Treatment and Safe Storage: a Promising New Strategy." *Epidemiology &Infection* 122 (01): 83–90. http://journals.cambridge.org/article_S0950268898001782.

Rabie, Tamer, and Valerie A. Curtis. 2006. "Handwashing and Risk of Respiratory Infections: a Quantitative Systematic Review." *Tropical Medicine & International Health* 11 (3) (March 1): 258–267. doi:10.1111/j.1365-3156.2006.01568.x. http://dx.doi.org/10.1111/j.1365-3156.2006.01568.x.

Ram, Pavani K. 2009. "Practical Guidance for Measuring Handwashing Behavior." *World Bank Water and Sanitation Program.*

Rosa, Ghislaine, and Thomas Clasen. 2010. "Estimating the Scope of Household Water Treatment in Low- and Medium-Income Countries." *The American Journal of Tropical Medicine and Hygiene* 82 (2) (February 1): 289–300. doi:10.4269/ajtmh.2010.09-0382. http://www.ajtmh.org/content/82/2/289.abstract.

Sallis, J. F., N. Owen, and E.B. Fisher. 2008. "Ecological Models of Health Behavior." In *Health Behavior and Health Education*, edited by K. Glanz, B.K. Rimer, and K. Viswanath, 4th ed., 465–485. San Francisco: Josey-Bass.

Scheff, Thomas J. 2000. "Shame and the Social Bond: A Sociological Theory." *Sociological Theory* 18 (1): 84–99. doi:10.1111/0735-2751.00089. http://onlinelibrary.wiley.com/doi/10.1111/0735-2751.00089/abstract.

Scott, Beth E., Wolf P. Schmidt, Robert Aunger, Nana Garbrah-Aidoo, and Rasaaque Animashaun. 2008. "Marketing Hygiene Behaviours: The Impact of Different Communication Channels on Reported Handwashing Behaviour of Women in Ghana." *Health Education Research* 23 (3) (June 1): 392–401. doi:10.1093/her/cym056. http://her.oxfordjournals.org/content/23/3/392.abstract.

Stanton, Bonita F., and John D. Clemens. 1987. "An Educational Intervention for Altering Water-Sanitation Behaviors to Reduce Childhood Diarrhea in Urban Bangladesh. II: a Randomized Trial to Assess the Impact of the Intervention on Hygienic Behaviors and Rates of Diarrhea." *American Journal of Epidemiology* 125 (2) (February 1): 292–301. http://aje.oxfordjournals.org/content/125/2/292.

Stanton, Bonita F., John D. Clemens, K. M. Aziz, and M. Rahman. 1987. "Twenty-four-hour Recall, Knowledge-attitude-practice Questionnaires, and Direct Observations of Sanitary Practices: a Comparative Study." *Bull World Health Organ* 65: 217–22. 3496987. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3496987.

Van de Mortel, Thea F. 2008. "Faking It: Social Desirability Response Bias in Self-report Research." *Australian Journal of Advanced Nursing* 25 (4): 40–48. http://www.ajan.com.au/vol25/vol25-4.pdf.

Van der Pligt, Joop. 1996. "Risk Perception and Self-Protective Behavior." *European Psychologist* 1 (1) (January 1): 34–43. doi:10.1027/1016-9040.1.1.34. http://dx.doi.org/10.1027/1016-9040.1.1.34.

Verplanken, Bas. 2006. "Beyond Frequency: Habit as Mental Construct." *British Journal of Social Psychology* 45 (3): 639–656. doi:10.1348/014466605X49122. http://onlinelibrary.wiley.com/doi/10.1348/014466605X49122/abstract.

WHO. 2008. "The Global Burden of Disease: 2004 Update." World Health Organization. http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf.

Zeitlyn, S., and F. Islam. 1991. "The Use of Soap and Water in Two Bangladeshi Communities: Implications for the Transmission of Diarrhea." *Reviews of Infectious Diseases* 13 Suppl 4 (March): S259–64. 2047648. PubMed. http://www.ncbi.nlm.nih.gov/pubmed/2047648.

8 Appendix: Sample Size and Power Calculations

These sample size calculations were performed using the Optimal Design software from the William T. Grant Foundation, available at http://hlmsoft.net/od/.

Our primary outcomes of interest were the proportion of people who drink treated water and the proportion of people who wash their hands with soap after fecal contact and before eating. Direct measures of water and hygiene behaviors (in the absence of an observer) are not available to us, thus we used structured observation in conjunction with proxy measures. Willingness to pay for water treatment hardware was also of primary interest for generating demand curves, but its role as a secondary outcome for evaluating water treatment behavior requires a larger sample so it was included below in the water treatment section.

Chlorine residual test of stored drinking water

Proportion with chlorine residual in Arm 1: 0.278

Using data from the Gates-CBC 10-5 pilot, we determined that given a standard public health intervention, the proportion with residual chlorine should be around 0.278 with a plausible interval from 0.03 to 0.6

Proportion with chlorine residual in Arm 2: 0.358

Using data from Najnin et al, we estimated that at 10 percentage point increase in usage would correspond to an 8 percentage point increase in the proportion with residual chlorine. This is our minimum practically significant effect size.

Sample size needed: **202 to 380compounds** per treatment arm given **6 and 2 households** per compound

Setting significance to 0.05, we find that to measure this increase in proportions with 80% power we need 380 compounds with 2 households each or 202 compounds per arm with 6 households each.

Use soap when demonstrating handwashing

Proportion who use soap when demonstrating handwashing in Arm 2: 0.54

Here we set the proportion to 10 percentage points greater than that in Treatment arm 1. This is our minimum practically significant effect size.

Sample size needed: **186compounds** per treatment arm given **2 households** per compound

Setting significance to 0.05, we find that to measure this increase in proportions with 80% power we need 186 compounds with 2 households each.

Under reasonable assumptions our other outcomes required smaller sample sizes to find meaningfully large effects.

Using the largest needed sample size—202 compounds per arm for Chlorine residual testing—and comparing sample configurations for cost effectiveness, we have chosen to enrol 420 compounds total (210 in standard treatment and 210 in disgust and shame). Given that the largest sample size needed for detecting a difference in handwashingbehavior is 138 compounds per arm—for mother's handwashing demonstrations—we have decided to enrol 280 compounds (140 per arm), two-thirds of the total sample in the handwashing with soap combined intervention.

9 Tables

Table 1: Characteristics of respondents

Indicators	Values	Standard Intervention Arm N=428 *n (%)	Disgust Intervention Arm N=420 *n (%)	p-value
Sex	Female	400 (93)	385 (92)	0.35
Marital status	Married	415 (97)	400 (95)	0.20
	†Others	13 (3)	20 (5)	
		Mean (SD)	Mean (SD)	
Age of the respondent		28.4 (7.9)	28.8 (9.10)	0.46

*n=Number of respondents within indicator, N=Total number of respondents by arm
†Others include divorce, widow and unmarried

Table 2: Compound-level Descriptive Statistics

Indicator	Standard Intervention Arm Mean (SD)	Disgust Intervention Arm Mean (SD)	P-value
Number of households per compound	9.24 (2.8)	9.21 (2.6)	0.92
Number of toilets per compound	2.38 (0.94)	2.35 (1)	0.74
Under 5 years of age children per household	0.73 (0.78)	0.73 (0.74)	0.96
# of compounds	214	210	

Table 3: Characteristics of Households

Indicators	Values	Standard Intervention Arm N=428 households *n (%)	Disgust Intervention Arm N=420 households *n (%)	P-value
Material of the wall	Finished wall (cement/ brick) Rudimentary wall (tin/ wood/ mud/stick)	344 (80) 84 (20)	306 (73) 114 (27)	0.04
Material of the floor	Finished floor (cement/ concrete Rudimentary (earth/mud/dung/sand/wood)	406 (95) 22 (5)	399 (95) 21 (5)	0.94
Material of the roof	Finished roof (cement/concrete) Rudimentary (corrugated iron/tin/wood)	21 (5) 407 (95)	21 (5) 399 (95)	0.96
Fuel used for cooking	Natural gas Wood/ charcoal/ Kerosene	407 (95) 21 (5)	398 (95) 22 (5)	0.85
Type of latrine at home used by adults	**Improved Unimproved **Improved (own) Improved (shared)	398 (93) 27 (6) 37 (9) 391 (91)	397 (95) 27 (6) 27 (6) 393 (94)	0.51 0.96 0.36

Level of education of the household head	No education	127 (30)	144 (34)	0.17
	Primary	213 (50)	205 (49)	-
	Secondary	60 (14)	44 (10)	-
	Higher secondary	19 (4)	17 (4)	-
	Graduation and above	9 (2)	10 (2)	-
Monthly household income (‡taka)	<4000	5 (1)	6 (1)	0.76
	4000-8000	114 (27)	115 (27)	-
	8000-12000	161 (38)	172 (41)	-
	>12000	148 (35)	127 (30)	-

*n=Number of household within indicator, N=Total number of household by arms
‡1 USD = 74 taka (at the time of data collection during baseline)
** Improved means presence of slab, water seal and super structure

Table 4: Household Assets

Variables	Standard Intervention Arm N=428 *n (%)	Disgust Intervention Arm N=420	p-value
Bicycle	22 (5)	28 (7)	0.35
Motorcycle	5 (1)	1 (0.2)	0.11
Rickshaw for passengers or cargo	7 (2)	15 (4)	0.08
Working radio/Cassette Player/CD player	53 (12)	75 (18)	0.03
Working Television/VCD	317 (74)	308 (73)	0.81
Working computer	16 (4)	10 (2)	0.25
Working mobile phone	383 (89)	390 (93)	0.08
Working refrigerator	90 (21)	74 (18)	0.21
Bed/Chouki	398 (93)	399 (95)	0.22
Sofa set	19 (5)	20 (5)	0.82
Working sewing machine	63 (15)	52 (12)	0.32
Locally-made holder for clothing	212 (50)	188 (45)	0.16
Blanket	402 (94)	383 (91)	0.13
Almira/Wardrobe (wood or metal)	304 (71)	286 (68)	0.35
Electricity	425 (99)	419 (100)	0.33

*n=Number of household within indicator, N= Total number of households by arm

Table 5: Water treatment behaviour (self-report, spot check)

Indicators	Standard Intervention Arm % (n/N)	Disgust Intervention Arm % (n/N)	**P-value**
Piped water source to the courtyard	92 (394/428)	91 (382/420)	0.65
Always treat water before drinking (self-report)	66 (283/428)	63 (265/420)	0.45
Water treatment methods (self-report, totals can add to > 100% because respondents can give > 1 reply) Boiling Filtration Filter	 65 (280/428) 40 (172/428) 4 (19/428)	 64 (268/428) 39 (162/428) 3 (12/428)	 0.68 0.69 0.24
During spot check observed someone boiling water in the compound, and at least one said it was for drinking	28 (37/131)	26 (33/126)	0.71
H2S test of household drinking water shows contamination within 48 hours of test	72 (771/1071)	72 (741/1027)	0.87

- Multiple response were allowed for water treatment methods
- Six households were randomly selected for H2S test.H2S test was done with the household drinking water to detect fecal contamination. Report was recorded after 24 hours and 48 hours. Contamination at 48 hours was higher than 24 hours.

Table 6: Feeling of disgust related to unsafe water following disgust intervention (self report)

Indicators	Standard Intervention Arm			Disgust Intervention Arm			DID§ (baseline vs 3.5m)	DID§ (baseline vs 7m)
	Baseline	3.5m follow up	7m follow up	Baseline	3.5m follow up	7m follow up		
All compounds	N=428 % (n)	N=416 % (n)	N=410 % (n)	N=420 % (n)	N=410 % (n)	N=406 % (n)		
Feeling of disgust at untreated water †	63 (270)	73 (302)	71 (292)	61 (256)	78 (319)	70 (283)	8 (p 0.07)	1 (p 0.82)
Strongly agrees with the statement "tap water could make me sick"	47 (200)	36 (149)	32 (133)	48 (203)	31 (129)	29 (118)	-6 (p 0.20)	-5 (p 0.28)

† Survey sample two households per compound per survey wave.

- Statistical tests adjust for clustering at compound level

- [§] DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline). P values adjusted for clustering at compound level when multiple data collected from one compound.

Table 7: Water treatment behavior (self-report from survey, spot check)

Indicators	Standard Intervention Arm			Disgust Intervention Arm			DID[§] (3.5 m)	DID[§] (7 m)
	Baseline % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)	Baseline % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)		
Always treat water before drinking (self-report) [†]	66 (283/428)	83 (347/416)	73 (301/410)	63 (265/420)	83 (341/410)	67 (274/406)	3 (p = 0.4)	-2 (p = 0.5)
Both the reservoirs are regular in use (spot check) [††]	0	29 (60/207)	12 (24/205)	0	33 (67/205)	10 (21/203)	4 (p = 0.5)	-1 (p = 0.7)
Both the reservoirs treated with chlorine (self-report) [††]	0	28 (57/207)	11 (23/205)	0	32 (66/205)	10 (21/203)	5 (p = 0.4)	-1 (p = 0.8)
During the spot check[††] someone was boiling water and said it was for drinking	28 (37/131 compounds)	14 (29/207)	9 (18/205)	26 (33/126)	6 (13/205)	11 (22/203)	-6 (p = 0.3)	4 (p = 0.5)

[†] Survey sample two households per compound per survey wave.

[††] Spot check one per compound per visit

[§] DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline)

- Self-reports are based on 2 surveys per compound. Spot checks are compound-level observations. The reservoirs were considered regular in use with the following criteria; water within reservoirs was reduced to some extent, the reservoirs were not empty and water found inside them , the reservoirs were found being used
- DID adjusted for clustering at compound level
- N for spot check of boiling was lower at baseline as we added this observation after some baseline date were collected.

Table 8: Observed water treatment (structured observation)

Indicators	Standard Intervention Arm	Disgust Intervention Arm	ARR
	2 months % (n/N)	2 months % (n/N)	
% of [‖]water drinking or [¶]gathering using chlorinated water, chlorinating water, or boiling water	15 (417/2819)	19 (452/2436)	1.18 (0.96, 1.46) p 0.11
% of water drinking or gathering either previously chlorinated water or chlorinating water	5 (130/2819)	6 (157/2436)	1.30 (0.90, 1.87) p 0.16

- [‖]Water drinking means when any compound members were seen drinking water. If the treatment status of that water could be confirmed it was coded accordingly. If not then the event was not included in the analysis
- [¶]Water gathering means when anyone found collecting water from the source or from the reservoir we provided for water storage. We observed for subsequent treatment and use of water and coded accordingly
- This index was created using water collection, treatment and drinking events related to boiling and/or chlorination
- N=total water collection (drinking or gathering) events observed
- n= collection of chlorinated water + treatment by boiling/chlorine + drinking boiled or chlorinated water
- Confidence intervals adjusted for clustering by compound

Table 9: Drinking practice during structured observation

Drinking event	Standard treatment N=343	Disgust treatment N=382	ARR[†](95% CI)	P-value
	% (n*)	% (n*)		
Drank untreated water	38 (129)	33 (125)	1.03 (.89, 1.20)	0.61
Drank any treated[‡] water	62 (214)	67 (257)		
Of which: Drank chlorinated	31 (108)	36 (138)	1.09 (.81, 1.47)	0.55

*n=Number of event within indicator, N=Total number of drinking events observed
[†]Adjusted Relative Risk (ARR) for compound, reference is standard treatment group
[‡] Treated with any method-chlorine, boiling, tablets, filter, sieving

- Table shows drinking event by intervention arms
- This event was coded when the observer seen someone drinking water. If treatment status was confirmed it was coded accordingly, if treatment status was not sure it was coded as "treatment status unknown". This category is not included in the denominator. Two types of missing data are coded here, "not applicable" means when no drinking event observed another is "treatment status unknown".
- % is presented in columns
- 66% standard treatment and 63% of the disgust treatment compounds self-reported treating drinking water at baseline survey

Table 10: Water treatment practice during structured observation

Water treatment event	Standard treatment N=2739	Disgust treatment N=2346	ARR[†](95% CI)	P-value
	n* (%)	n*(%)		
Treated source water with any method[‡]	349 (13)	375 (16)	1.17 (.93, 1.47)	0.17
Treated source water with chlorine	53 (2)	68 (3)	1.37 (.83, 2.27)	0.21
Treated source water by boiling	281 (10)	291 (12)	1.18 (.91, 1.52)	0.19

*n=Number of event within indicator, N=Total number of water treatment event observed
[†]Adjusted Relative Risk (ARR) for compound, reference is standard treatment group
[‡] Any method includes chlorine, boiling, tablets, filter, sieving

- Table shows the water treatment practice by arms
- Subsequent treatment of the collected source water was observed
- Event for chlorinated water collection is not included in the denominator
- % is shown in column
- Increased frequency of treatment is noted in the disgust treatment arm however the difference is not statistically significant when adjusted for cluster
- 96% standard treatment and 99% of the disgust treatment compounds self-reported treating drinking water by boiling at baseline survey
- 59% standard treatment and 60% of the disgust treatment compounds self-reported treating drinking water by sieving at baseline survey

Table 11: Water treatment behaviour (by chemical tests)

Indicators	Standard Intervention Arm			Disgust Intervention Arm			Disgust – Standard Arm		
	2 m follow up % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)	2 m follow up % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)	2m	3.5m	7m
Chlorine residual test: Any detectable chlorine in household drinking water within 24 hours of reported treatment	8 (96/ 1135)	8 (85/ 1044)	6 (36/ 581)	12 (133/ 1124)	8 (83/ 992)	8 (48/ 598)	3 (p 0.09)	0.1 (p 0.94)	1 (p 0.50)
	Baseline % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)	Baseline % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)	DID (3.5 months)		DID (7 months)
H2S test: Household drinking water shows contamination within 48 hours of test	72 (768/1067)	60 (630/1046)	78 (787/1007)	72 (740/1025)	61 (607/991)	77 (775/1005)	0.5 (p 0.86)		-2 (p 0.53)

[§] DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline). P values adjust for clustering by compound.

- [†]Concentration of the supplied sodium hypochlorite solution was 0.1008% during 2 months and 3.5 months follow up and 0.2343% during 7 months follow up

- During endline chlorine test was done only in the winner compounds where chlorine hardware was present. We tested all self-reported user households, one non-user and the source water. Source water was tested in the loser compounds as well
- H2S test was done in all compounds with or without the presence of chlorine hardware
- [§] DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline). P values adjusted for clustering at compound level when multiple data collected from one compound.

Table 12a: Chlorine residual and H2S test results

Indicators	3.5 months follow up All 413 compounds		ARR	7 months follow up ONLY Among 106 subscriber compounds		ARR
	Standard Intervention Arm % (n/ N)	Disgust Intervention arm % (n/ N)		Standard Intervention Arm % (n/ N)	Disgust Intervention arm % (n/ N)	
Self-reported users of liquid chlorine (tried to get observations at up to 6 observations per compound at 3.5 months, all [up to 8] users per compound at 7 months)	25 (260/1046)	25 (249/992)	0.89	42 (83/196)	40 (91/227)	0.75
Presence of any detectable chlorine among self-reported use within last 24 hours	33 (86/260)	33 (83/249)	0.78	41 (34/83)	51 (46/91)	0.68
H2S test shows no contamination after 48 hours in presence of detectable chlorine among self-reported use within last 24 hours	18 (47/260)	20 (49/249)	0.67	20 (17/83)	31 (28/91)	0.40

Table 12b: Water contamination by winning status

Indicators	Standard Intervention Arm % (n/ N)		ARR	Disgust Intervention arm % (n/ N)		ARR
	Winner	Loser		Winner	Loser	
Water contaminated after 48 hours of H2S test	66 (154/234)	82 (230/279)	0.001	64 (167/259)	82 (270/331)	0.000

Standard compound 115; 2-6 observations

Disgust compounds 136; 2-12 observations

Table 12: Social interaction promoting soap and water treatment (structured observation)

Indicators	Standard Intervention Arm	Disgust Intervention Arm	p-values
	2 months follow-up % (n./ N)	2 months follow-up % (n./ N)	
Opportunities to treat drinking water that involved reminders to treat, praise for treating, or criticism for not using treating	0.1 3/2739	0.3 6/2346	0.22

- N=total water gathering event observed
- P value not adjusted for clustering

Table 13: Feeling of disgust related to handwashing and water following disgust intervention (self-report)

Indicators	Standard Intervention Arm			Disgust Intervention Arm			DID[§] (3.5m)	DID[§] (7m)
Handwashing compound	Baseline N=288 % (n)	3.5m follow-up N=280 % (n)	7m follow-up N=276 % (n)	Baseline N=280 % (n)	3.5m follow-up N=270 % (n)	7m follow-up N=268 % (n)		
Feeling of disgust at hands not washed with soap [†]	95 (273)	99 (276)	100 (275)	95 (267)	97 (263)	100 (268)	-2 (p 0.45)	-0.1 (p 0.94)

[†] Survey sample two households per compound per survey wave.

- Statistical tests adjust for clustering at compound level
- [§] DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline). P values adjusted for clustering at compound level when multiple data collected from one compound.

Table 14: Knowledge and practice of handwashing with soap (self-report)

Indicators	Standard Intervention Arm			Disgust Intervention Arm			DID[§] (3.5m)	DID[§] (7m)
	Baseline N=288 % (n)	3.5m follow up N=280 % (n)	7m follow up N=276 % (n)	Baseline N=280 % (n)	3.5m follow up N=270 % (n)	7m follow up N=268 % (n)		
Mentioned handwashing after defecation as usual time to wash with soap	85 (244)	94 (264)	97 (264)	86 (241)	97 (263)	94 (253)	2 (0.55)	-2 (0.48)
Mentioned all three key times* as usual time of handwashing with soap	18 (52)	20 (57)	26 (72)	19 (53)	28 (76)	32 (86)	7 (0.06)	8 (0.04)
# of other key times mentioned[†] as usual time of handwashing with soap	0.4 (1)	1 (4)	3 (8)	0.4 (1)	2 (5)	9 (25)	1 (p 0.2)	5 (0.01)
Washed both hands with soap in all three key times* (self-reported last event)	9 (26)	11 (31)	14 (40)	7 (20)	16 (42)	18 (49)	4 (p 0.10)	4 (p 0.2)
Self-reported washing hands after last defecation	67 (192)	89 (249)	89 (246)	70 (195)	89 (241)	84 (225)	-1 (0.77)	-6 (0.10)

* The three key times promoted during the intervention were: after defecation, after cleaning a child's anus, and before eating

† Other key times include before preparing food, before feeding child and after cleaning child feces (not promoted during intervention)

§ DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline)

- N is the household number. Two households were interviewed from each compound. % and frequencies are for households but DIDs are adjusted for clustering at compound level
- Only handwashing households are included in the denominator. Drop out compounds are excluded, drop out households are not considered as the data is summarized for compounds

Table 15: Observed handwashing behaviour (spot check, demonstration, inspection)

Indicators	Standard Intervention Arm			Disgust Intervention Arm			DID[§] (baseline vs 3.5m)	DID[§] (baseline vs 7m)
	Baseline % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)	Baseline % (n/N)	3.5m follow up % (n/N)	7m follow up % (n/N)		
Soapy water appeared to be in regular use (spot check): bottle was half empty, enough lather was present within the bottle stating recent use, or seen being used	N/A	68 (95/140)	16 (22/138)	N/A	69 (93/135)	23 (31/134)	1 (p 0.88)	7 (p 0.17)
Soap and water or filled soapy water bottle present at the handwashing place (spot check)	16 (14/87)	61 (85/140)	48 (66/138)	18 (15/83)	59 (79/135)	43 (58/134)	-4 (p 0.6)	-7 (p 0.5)
Used soap to wash both hands while demonstrating handwashing after defecation (handwashing demonstration)	68 (484/712)	85 (593/694)	85 (505/593)	67 (453/681)	89 (579/654)	87 (509/588)	5 (0.11)	3 (0.38)
Adult both hands appear clean[‡] (hand inspection)	20 (142/712)	26 (177/694)	24 (141/595)	21 (145/681)	27 (175/657)	30 (181/596)	-0.4 (p 0.9)	5 (p 0.10)

[‡]Clean is defined when no visible dirt was seen on any of the three parts of the hands (palm, finger tips and nails)

[§] DID is the percentage points difference-in-differences = (Disgust arm at 3.5 or 7 month follow up – Standard arm at same follow-up) – (Disgust arm at baseline – Standard arm at baseline). P values adjusted for clustering at compound level when multiple data collected from one compound.

Table 16: Observed handwashing behaviour (structured observation, all households in the compound)

Indicators	Standard Intervention Arm	Disgust Intervention Arm	ARR
	2months % (n/N)	2 months % (n/N)	
Washed both hands with soap before eating	0.4 (6/1598)	0.1 (2/1420)	0.38 (0.07, 2.01) p 0.26
Washed both hands with soap after cleaning child anus	30 (39/128)	34 (48/140)	1.04 (0.68, 1.59) p 0.84
Washed both hands with soap after toileting	11 (274/ 2407)	12 (266/ 2188)	1.08 (0.84, 1.39) p 0.53
Washed both hands with soap after toileting (adult)	13 (221/1747)	12 (200/1604)	1.01 (0.77, 1.31) P 0.93
Washed both hands with soap after toileting (3-15 years)	8 (53/660)	11 (66/584)	1.4 (0.93, 2) P 0.10
Soap/ soapy water and water available at the handwashing station	79 (112/141)	80 (111/139)	(p 1.00)

- Only the key times promoted in the intervention are shown
- Confidence intervals adjusted for clustering by compound
- Soap and soapy water was more available at the handwashing station as because the observation was done very early in the morning unlike the other observations

Table 17: Social interaction promoting handwashing with soap (structured observation)

Indicators	Standard Intervention Arm	Disgust Intervention Arm	p-values
	2 months follow-up % (n./ N)	2 months follow-up % (n./ N)	
Opportunities to use soap and water that involved reminders to use soap, praise for using soap, or criticism for not using soap	0.7 26/3789	0.6 22/3680	0.63

- N=total toileting event observed
- P value not adjusted for clustering

Effect of handwashing intervention on handwashing

Table 18: Knowledge of handwashing with soap (self-report, two households per compound)

	Baseline		3.5m follow up		7m follow up		Baseline vs 3.5m DID[§] (95%CI) (P-values)	Baseline vs 7m DID[§] (95%CI) (P-values)
	Without HW intervention N=280 % (n)	With HW intervention N=568 % (n)	Without HW intervention N=276 % (n)	With HW intervention N=550 % (n)	Without HW intervention N=272 % (n)	With HW intervention N=544 % (n)		
Mentioned handwashing after defecation as usual time of handwashing with soap								
	85 (238)	85 (485)	93 (257)	96 (527)	95 (259)	95 (517)	10 (7, 13) (0.000)	9 (6, 13) (0.000)
Mentioned all three key times* as usual time of handwashing with soap								
	24 (66)	18 (105)	25 (70)	24 (133)	30 (81)	29 (158)	4 (0.34)	4 (0.30)
Washed both hands with soap in all three key times* (last key event)								
	7 (19)	8 (46)	7 (20)	13 (73)	18 (49)	16 (89)	5 (0.12)	3 (0.37)
Washed both hands with soap after last defecation								
	68 (190)	68 (387)	81 (224)	89 (490)	82 (224)	87 (471)	7 (0.07)	4 (0.37)

Notes: Three key times include: after defecation, after cleaning child anus and before eating (promoted during intervention)

[§] DID is the percentage points difference-in-differences = (handwashing intervention compounds– no handwashing intervention compounds in standard/disgust treatment arm at 3.5 or 7 month follow up) – (handwashing intervention compounds – no handwashing intervention compounds in standard/disgust treatment arm at baseline)

“Without handwashing intervention”arecompounds who received only the standard or the disgust water treatment intervention

Table 19: Observed handwashing behaviour (spot check, demonstration, inspection)

	Baseline		3.5m follow up		7m follow up		Baseline vs 3.5m DID[§] (95%CI) (P-values)	Baseline vs 7m DID[§] (95%CI) (P-values)
	Without HW intervention % (n/N)	With HW intervention % (n/N)	Without HW intervention % (n/N)	With HW intervention % (n/N)	Without HW intervention % (n/N)	With HW intervention % (n/N)		
Soap or soapy water bottle and water present at the handwashing place (spot check)								
	29 (25/86)	18 (30/170)	16 (22/138)	59 (161/275)	32 (44/136)	44 (121/272)	54 (40, 61) (0.000)	24 (8, 39) (0.003)
Used soap or soapy water to wash both hands while demonstrating handwashing after defecation								
	69 (471/682)	67 (937/1393)	80 (550/688)	87 (1172/1348)	81 (477/588)	86 (1014/1181)	9 (3, 14) (0.001)	6 (0.5, 11) (0.03)
Adult both hands appear clean[‡] on inspection								
	22 (152/682)	21 (287/1393)	25 (173/690)	26 (352/1351)	27 (163/593)	27 (322/1191)	2 (-3, 8) (0.37)	1 (-5, 7) (0.72)

- For spot check N= common handwashing stations in compounds received handwashing intervention; for demonstration/ hand inspection N= total number of demonstration/ hand inspection.
- [§] DID is the percentage points difference-in-differences = (handwashing intervention compounds– no handwashing intervention compounds in standard/disgust treatment arm at 3.5 or 7 month follow up) – (handwashing intervention compounds – no handwashing intervention compounds in standard/disgust treatment arm at baseline)

Table 20: Observed handwashing behaviour (structured observation, all households in the compound)

Indicators	Without any HW intervention	With HW Intervention	ARR (95% CI) (p value)
	% (n/N)	% (n/N)	
Washed both hands with soap before eating	0.3 (5/1480)	0.3 (8/3018)	0.77 (0.20, 2.88) (0.69)
Washed both hands with soap after cleaning child anus	23 (31/134)	32 (87/268)	1.38 (0.92, 2.09) (0.12)
Washed both hands with soap after toileting	7 (154/2130)	12 (540/4595)	1.66 (1.27, 2.15) (0.00)
Washed both hands with soap after toileting (adult)	8 (124/1547)	13 (421/3351)	1.60 (1.22, 2.09) (0.001)
Washed both hands with soap after toileting (3-15 years)	5 (30/583)	10 (119/1244)	1.91 (1.19, 3.09) (0.007)

- Only the key times promoted in the intervention are shown
- N=total number of specific event observed
- ARR adjusted for clustering at compound level

Table 23: Payments data (# of compounds)

Payment visit	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	16th	20th	24th
Full paid	76	78	77	72	73	75	76	71	73	73	73	69	51	41	29
Partial due	11	1	3	3	4	1	1	1	2	1	1	0	6	6	4
Full due	2	8	5	5	1	1	0	4	0	0	0	0	0	0	0
Refused	17	19	21	26	28	29	29	30	31	32	32	37	49	59	73
Visit due	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total =106	106	106	106	106	106	106	106	106	106	106	106	106	106	106	106

Source: Payment data as of April 30, 2014. Payment visits were monthly for the first year and quarterly for the second year. Data are out of 420 compounds that were in baseline and not destroyed by the time of the sales meeting.

Table 24: How does participation in the sales meeting compare across treatment groups?[a]

	All			Standard treatment			Disgust & Shame treatment		
	compounds enrolled.	compounds that had a sales meeting.	compounds that purchased dispenser.	compounds enrolled.	compounds that had a sales meeting.	compounds that purchased dispenser.	compounds enrolled.	compounds that had a sales meeting.	compounds that purchased dispenser.
Proportion of the total sample (n of proportion 'n/N', N=435)									
All compounds	**1.00** (435)	0.57 (248)	0.13 (58)	**0.50** (217)	0.27 (117)	0.06 (27)	**0.50** (218)	0.30 (131)	0.07 (31)
Compounds assigned to the Collective Auction	**0.50** (217)	0.29 (124)	0.06 (24)	**0.25** (109)	0.14 (59)	0.03 (11)	**0.25** (108)	0.15 (65)	0.03 (13)
Compounds assigned to the Individual Auction	**0.50** (218)	0.29 (124)	0.08 (34)	**0.25** (108)	0.13 (58)	0.04 (16)	**0.25** (110)	0.15 (66)	0.04 (18)
Average proportion of households in each compound that attended the sales meeting[b,c]									
All compounds	-	0.37	0.29	-	0.39	0.32	-	0.36	0.26
Compounds assigned to the Collective Auction	-	0.35	0.28	-	0.37	0.36	-	0.33	0.21
Compounds assigned to the Individual Auction	-	0.40	0.30	-	0.41	0.29	-	0.39	0.30
Average proportion of households in each compound that were reported as users at the sales meeting[d,c]									
All compounds	-	0.50	0.45	-	0.54 *	0.50	-	0.46 *	0.40
Compounds assigned to the Collective Auction	-	0.47	0.42	-	0.50	0.48	-	0.44	0.38
Compounds assigned to the Individual Auction	-	0.53	0.47	-	0.58	0.52	-	0.49	0.42
Average proportion of sales meeting attendees in each compound who also attended the marketing meeting[e,c]									
All compounds	-	0.83	0.84	-	0.80	0.76	-	0.85	0.90
Compounds assigned to the Collective Auction	-	0.83	0.79	-	0.82	0.68	-	0.83	0.88
Compounds assigned to the Individual Auction	-	0.83	0.87	-	0.79	0.82	-	0.87	0.91
Average proportion of sales meeting attendees in each compound who were female[f,c]									
All compounds	-	0.94	0.98	-	0.96	0.98	-	0.92	0.98
Compounds assigned to the Collective Auction	-	0.95	0.98	-	0.99	1.00	-	0.91	0.96
Compounds assigned to the Individual Auction	-	0.93	0.98	-	0.93	0.96	-	0.92	1.00

** For each row group, difference between the figures for All standard compounds that had a sales meeting and All disgust and shame compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a * appears to the right of the two figures.*

† For each row group, difference between the figures for All collective compounds that had a sales meeting and All individual compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a † appears to the right of the two figures.

a. Values are given for the sub-sample defined by the row and column headings.

b. For each compound, the number of households who had a representative attend the sales meeting is divided by the total number of households in the compound. The mean of these proportions is then calculated for each sub-

c. Compounds were composed of 6 to 16 households (median: 9 households)

d. At the start of the sales meeting, participants were asked to report on the user status of all members of the compound (present or not). For each compound, the number of households reported as users is divided by the total number of households in the compound. The mean of these proportions is then calculated for each sub-sample.

e.At the start of the sales meeting, each present household representative was asked whether she or he also attended the initial marketing meeting. For each compound, the number of households with representatives who answered yes to this question is divided total number of households participating in the sales meeting. The mean of these proportions is then calculated for each sub-sample.

f.At the start of the sales meeting, each present household representatives gender is recorded. For each compound, the number of households with female representatives present is divided total number of households participating in the sales meeting. The mean of these proportions is then calculated for each sub-sample.

Table 25: How does willingness to pay compare across treatment groups?[a]

	All		Standard treatment		Disgust & Shame treatment	
	compounds that had a sales meeting.	compounds that purchased the dispenser.	compounds that had a sales meeting.	compounds that purchased the dispenser.	compounds that had a sales meeting.	compounds that purchased the dispenser.
Mean group willingness to pay, from COLLECTIVE sales meeting (SD)	68 (45)	72 (44)	68 (38)	77 (44)	69 (51)	67 (45)
Mean per household willingness to pay, from COLLECTIVE sales meeting (SD)[b]	22.4 (15.2)	21.0 (15.3)	21.6 (16.4)	19.5 (19.7)	23.2 (14.2)	22.4 (11.1)
Mean individual weakest link willingness to pay, from INDIVIDUAL sales meeting (SD)[c]	7.8 (5.1)	10.0 (5.1)	8.4 (5.3)	10.3 (5.5)	7.3 (4.9)	9.7 (5.0)
Proportion of compounds with HIGHEST possible willingness to pay, 200 taka (N)[d]						
All compounds	0.03 (248)	0.10 (58)	0.03 (117)	0.11 (27)	0.03 (131)	0.10 (31)
Compounds assigned to the Collective Sales meeting	0.02 (124)	0.04 (24)	0.02 (59)	0.00 (11)	0.02 (65)	0.08 (13)
Compounds assigned to the Individual Sales meeting (weakest link)[c]	0.05 (124)	0.15 (34)	0.05 (58)	0.19 (16)	0.05 (66)	0.11 (18)
Proportion of compounds with LOWEST possible willingness to pay, 25 taka (N)[d]						
All compounds	0.42 (248)	0.10 (58)	0.42 (117)	0.11 (27)	0.43 (131)	0.10 (31)
Compounds assigned to the Collective Sales meeting	0.48 (124)	0.17 (24)	0.51 (59)	0.18 (11)	0.45 (65)	0.15 (13)
Compounds assigned to the Individual Sales meeting (weakest link)[c]	0.37 (124)	0.06 (34)	0.33 (58)	0.06 (16)	0.41 (66)	0.06 (18)
Proportion of compounds with ZERO willingness to pay (N)[d]						
All compounds	0.01 (248)	0.00 (58)	0.01 (117)	0.00 (27)	0.02 (131)	0.00 (31)
Compounds assigned to the Collective Sales meeting	0.01 (124)	0.00 (24)	0.00 (59)	0.00 (11)	0.02 (65)	0.00 (13)
Compounds assigned to the Individual Sales meeting (weakest link)[c]	0.02 (124)	0.00 (34)	0.02 (58)	0.00 (16)	0.02 (66)	0.00 (18)
Proportion of compounds with 100% of households stating a positive willingness to pay (N)[d,e]						
All compounds	0.07 (248)	0.02 (58)	0.08 (117)	0.04 (27)	0.06 (131)	0.00 (31)
Compounds assigned to the Collective Sales meeting (100% agreed prior)	0.08 (124)	0.04 (24)	0.10 (59)	0.09 (11)	0.06 (65)	0.00 (13)
Compounds assigned to the Individual Sales meeting (100% attended and bid, more strir	0.06 (124)	0.00 (34)	0.05 (58)	0.00 (16)	0.06 (66)	0.00 (18)
Proportion of compounds who changed their willingness to pay after understanding questions (N)[d,f]						
All compounds	0.08 (248)	0.12 (58)	0.07 (117)	0.11 (27)	0.08 (131)	0.13 (31)
Compounds assigned to the Collective Sales meeting	0.03 (124) †	0.00 (24)	0.00 (59)	0.00 (11)	0.06 (65)	0.00 (13)
Compounds assigned to the Individual Sales meeting (weakest link)[c]	0.12 (124) †	0.21 (34)	0.14 (58)	0.19 (16)	0.11 (66)	0.22 (18)
Mean % change in willingness to pay for compounds who changed (SD)[f,g]						
All compounds	65 (72)	81 (38)	49 (104)	63 (25)	76 (37)	94 (43)
Compounds assigned to the Collective Sales meeting	57 (19)	0 (0)	0 (0)	0 (0)	57 (19)	0 (0)
Compounds assigned to the Individual Sales meeting (weakest link)[c]	67 (81)	81 (38)	49 (104)	63 (25)	87 (42)	94 (43)

** For each row group, difference between the figures for All standard compounds that had a sales meeting and All disgust and shame compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a * appears to the right of the two figures.*

† For each row group, difference between the figures for All collective compounds that had a sales meeting and All individual compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a † appears to the right of the two figures.

a. Figures are calculated using final willingness to pay unless otherwise stated (that is the last wtp stated, after changes if any). Willingness to pay figures are in Bangladeshi Taka.

b. Before eliciting willingness to pay, participants of the collective sales meeting are asked how many households have agreed to pay for the dispenser (regardless of whether they are present or absent). Collective willingness to pay is divided by the number of households who agreed to pay to obtain the per household willingness to pay.

c. In the individual sales meeting, the highest price that all present members agree to pay, which is also the lowest of everyone's maximum stated willingness to pay, is known as the "weakest link" willingness to pay. It is this price that is compared to the lottery price to determine whether the compound will keep the dispenser.

d.The proportion is calculated for the sub-sample defined by the row and column headings. The denomenator is the total number of compounds within the sub-sample, N, which is given for each value.

e. In the collective group, we don't measure directly if every individual states a positive willingness to pay, but we ask the group how many households have agreed to pay. If everyone has agreed to pay, we consider this 100% stating a positivie willigness to pay for this figure. In the individual group, individuals must be present in order to state their willingness to pay. Thus, in order to have 100% stating a positive willingness to pay, 100% must be present. Thus, the comparison is not perfect with the collective group, as there may have been households who wanted to pay but couldn't attend the meeting.

Table 26: How do individual housholds' willingness to pay compare across treatment groups?[a]

	All		Standard treatment		Disgust & Shame treatment	
	compounds who had the Auction meeting.	compounds who won at the Auction meeting.	compounds who had the Auction meeting.	compounds who won at the Auction meeting.	compounds who had the Auction meeting.	compounds who won at the Auction meeting.
Mean per household willingness to pay, from COLLECTIVE sales meeting (SD)[b]	20.5 (12.7)	21.0 (15.3)	18.6 (15.8)	19.5 (19.7)	22.0 (9.5)	22.4 (11.1)
Mean individual weakest link willingness to pay, from INDIVIDUAL sales meeting (SD)[c]	7.5 (4.9)	10.0 (5.1)	7.9 (5.2)	10.3 (5.5)	7.1 (4.7)	9.7 (5.0)
Mean individual willingness to pay, from INDIVIDUAL sales meeting (SD)[d]	8.2 (4.9)	10.7 (5.0)	8.9 (5.0)	11.1 (5.5)	7.6 (4.6)	10.4 (4.7)
Proportion of households with HIGHEST possible willingness to pay, 20 taka (N)[e]	0.08 (190)	0.18 (84)	0.10 (88)	0.24 (38)	0.06 (102)	0.13 (46)
Proportion of households with LOWEST possible willingness to pay, 2 taka (N)[e]	0.15 (190)	0.01 (84)	0.15 (88)	0.03 (38)	0.16 (102)	0.00 (46)
Proportion of households with ZERO willingness to pay (N)[e]	0.00 (190)	0.00 (84)	0.00 (88)	0.00 (38)	0.00 (102)	0.00 (46)
Proportion of households who changed willingness to pay after understanding question	0.17 (190)	0.23 (84)	0.22 (88)	0.21 (38)	0.13 (102)	0.24 (46)
Mean % change in willingness to pay for households who changed (SD)[f,g]	81 (31)	76 (31)	83 (26)	72 (18)	78 (39)	79 (39)
Mean coefficient of variation of households' initial willingness to pay within compound[f]	0.14	0.16	0.13	0.09	0.15	0.21
Mean coefficient of variation of households' final willingness to pay within compound[f,h]	0.07	0.06	0.06	0.03	0.07	0.09

** For each row group, difference between the figures for All standard compounds that had a sales meeting and All disgust and shame compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a * appears to the right of the two figures.*

a. Due to a data entry mistake, individual households' willingness to pay from the first 40% of individual group compounds visited has not yet been entered. Results in this table (collective and individual) are derived from the subset of compounds that were visited during the period of time for which we have individual households willingness to pay entered. Household willingness to pay data is entered for 190 households from 72 compounds. Figures are calculated using final willingness to pay unless otherwise stated (that is the last wtp stated, after changes if any). Willingness to pay figures are in Bangladeshi Taka.

b. Before eliciting willingness to pay, participants of the collective sales meeting are asked how many households have agreed to pay for the dispenser (regardless of whether they are present or absent). Collective willingness to pay is divided by the number of households who agreed to pay to obtain the per household willingness to pay.

c. In the individual sales meeting, the highest price that all present members agree to pay, which is also the lowest of everyone's maximum stated willingness to pay, is known as the "weakest link" willingness to pay. It is this price that is compared to the lottery price to determine whether the compound will keep the dispenser.

d. Mean of all households' willingness to pay (not just weakest link) from the individual group.

e.The proportion is calculated for the sub-sample of households in the individual group whose compounds are in the category defined by the column heading. The denomenator is the total number of households within the sub-sample, N, which is given for each value.

f. After first obtaining participants' willingness to pay, but before moving to the lottery, respondents are asked questions to test their understanding of the BDM procedure. After they answer, the correct answer is explained again and respondents are given the opportunity to change their willingness to pay.

g. The average of first and final willingness to pay is used as the denomenator to calculate % change.

h. The coefficient of variation of households' willingness to pay is calculated for each compound, the mean of these compound figures is given (note, over all compounds, not just those whose members changed bids).

Table 27: How does participation and willingness to pay compare across the handwashing treatment?

	All			Water Treatment Only			Water Treatment & Handwashing		
	compounds enrolled.	compounds that had a sales meeting.	compounds that purchased dispenser.	compounds enrolled.	compounds that had a sales meeting.	compounds that purchased dispenser.	compounds enrolled.	compounds that had a sales meeting.	compounds that purchased dispenser.
Proportion of the total sample (n of proportion 'n/N', N=435)[a]									
All compounds	**1.00** (435)	0.57 (248)	0.13 (58)	**0.33** (144)	0.20 (89)	0.05 (21)	**0.67** (291)	0.37 (159)	0.09 (37)
Compounds assigned to the Collective Auction	**0.50** (217)	0.29 (124)	0.06 (24)	**0.17** (72)	0.09 (40)	0.02 (9)	**0.33** (145)	0.19 (84)	0.03 (15)
Compounds assigned to the Individual Auction	**0.50** (218)	0.29 (124)	0.08 (34)	**0.17** (72)	0.11 (49)	0.03 (12)	**0.34** (146)	0.17 (75)	0.05 (22)
Mean group willingness to pay, from COLLECTIVE sales meeting (SD)	-	68 (45)	71.88 (44)	-	74 (47)	78 (51)	-	65 (44)	68 (41)
Mean "individual" willingness to pay, from COLLECTIVE sales meeting (SD)[b]	-	22.4 (15.2)	21.0 (15.3)	-	23.6 (12.0)	21.7 (13.9)	-	21.8 (16.6)	20.7 (16.6)
Mean individual weakest link willingness to pay, from INDIVIDUAL sales meeting (SD)[c]	-	7.8 (5.1)	10.0 (5.1)	-	7.7 (5.1)	9.4 (5.8)	-	7.9 (5.1)	10.3 (4.9)

† For each row group, difference between the figures for All Water treatment only compounds that had a sales meeting and All water treatment and handwashing compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a † appears to the right of the two figures.

a. Values are given for the sub-sample defined by the row and column headings.

b. Before eliciting willingness to pay, participants of the collective sales meeting are asked how many households have agreed to pay for the dispenser (regardless of whether they are present or absent). Collective willingness to pay is divided by the number of households who agreed to pay to obtain the "indivindual' willingness to pay.

c. In the individual sales meeting, the highest price that all present members agree to pay, which is also the lowest of everyone's maximum stated willingness to pay, is known as the "weakest link" willingness to pay. It is this price that is compared to the lottery price to determine whether the compound will keep the dispenser.

Table 28: How was the cost shared within compounds

	All		Standard treatment		Disgust & Shame treatment	
	compounds who had the Auction meeting.	compounds who won at the Auction meeting.	compounds who had the Auction meeting.	compounds who won at the Auction meeting.	compounds who had the Auction meeting.	compounds who won at the Auction meeting.
Proportion of compounds with the COLLECTIVE auction who	(N= 124)	(N= 24)	(N= 59)	(N= 11)	(N= 65)	(N= 13)
Will permit only paying households to use the dispense	0.47	0.46	0.37 *	0.36	0.55 *	0.54
Will try to recruit more households to pay	0.88	0.79	0.85	0.82	0.91	0.77
Expect all or most non-payers to try to use	0.10	0.21	0.10	0.27	0.11	0.15
Expect all or most absent households to contribute	0.30	0.46	0.31	0.55	0.29	0.38
Have absent members who already agreed to pay	0.46	0.50	0.54	0.55	0.38	0.46
Plan to share the cost equally among payers	0.89	0.92	0.93	1.00	0.85	0.85
Proportion of compounds with the INDIVIDUAL auction who	(N= 124)	(N= 34)	(N= 58)	(N= 16)	(N= 66)	(N= 18)
Will permit only paying households to use the dispense	0.34	0.53	0.31	0.63	0.36	0.44
Will try to recruit more households to pay	0.57	0.91	0.52	0.81	0.62	1.00
Expect all or most non-payers to try to use	0.12	0.26	0.14	0.44	0.11	0.11
Expect all or most absent households to contribute	0.64	0.68	0.66	0.63	0.62	0.72

** For each row group, difference between the figures for All standard compounds that had a sales meeting and All disgust and shame compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a * appears to the right of the two figures.*

a. Figures are based on majority answer given by participating households at sales meeting. Questions were asked after obtaiing households' initial willingness to pay, but before giving them a final chance to change.

Table 29: Did respondents understand the BDM method?[a]

	All compounds who had the Auction meeting.	Standard treatment compounds who had the Auction meeting.	Disgust & Shame treatment compounds who had the Auction meeting.
Proportion of compounds where all or most members correctly believed their bids cannot affect the lottery price (N)			
All compounds	0.81 (248)	0.80 (117)	0.82 (131)
Compounds assigned to the Collective Auction	0.83 (124)	0.85 (59)	0.82 (65)
Compounds assigned to the Individual Auction	0.79 (124)	0.76 (58)	0.82 (66)
Proportion of compounds who answered both hypothetical lottery questions correctly (N)[b]			
All compounds	0.95 (248)	0.97 (117)	0.93 (131)
Compounds assigned to the Collective Auction	0.94 (124)	0.98 (59)	0.91 (65)
Compounds assigned to the Individual Auction	0.95 (124)	0.95 (58)	0.95 (66)
Proportion of compounds whose members reported they did not take their expectation of absent members' willingness to pay into account while bidding, as we instructed (N)			
Compounds assigned to the Individual Auction (only collected in individual)	0.43 (124)	0.33 (58) *	0.52 (66) *

** For each row group, difference between the figures for All standard compounds that had a sales meeting and All disgust and shame compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a * appears to the right of the two figures.*

† For each row group, difference between the figures for All collective compounds that had a sales meeting and All individual compounds that had a sales meeting was tested for statistical significance (ttest for means, binomial test for proportions). If the difference was shown to be significant at the 5% level, a † appears to the right of the two figures.

a. Figures are based on majority answer to understanding questions given by participating households at sales meeting. Questions were asked after obtaiing households' initial willingness to pay, but before giving them a final chance to change. Proportions are calculated for the sub-sample defined by the row and column headings. The denomenator is the total number of compounds within the sub-sample, N, which is given for each value.

b. Before giving participants a final chance to change their willingness to pay, we asked them two hypothetical lottery questions: one, "What will happen if the lottery price is greater than your stated willingness to pay?" and two, "What will happen if the lottery price is less than or equal to your stated willingness to pay?".

Table 30: Final Allocation of Compounds

Study Arm	Intended Allocation %	Allocation by Stratum: One Gas, Small n	One Gas, Small % of stratum	Two Gas, Large n	Two Gas, Large % of stratum	Three No Gas, Small n	Three No Gas, Small % of stratum	Four No Gas, Large n	Four No Gas, Large % of stratum	**TOTAL SAMPLE n**	**TOTAL SAMPLE % of total**
Disgust & Shame – Hand washing & Water treatment – Collective BDM	16.7	20	15.5	51	17.1	0	0.0	1	16.7	**72**	**16.6**
Standard Health – Hand washing & Water treatment – Collective BDM	16.7	24	18.6	47	15.7	0	0.0	2	33.3	**73**	**16.8**
Disgust & Shame – Hand washing & Water treatment – Individual BDM	16.7	18	14.0	53	17.7	1	100.0	2	33.3	**74**	**17.0**
Standard Health – Hand washing & Water treatment – Individual BDM	16.7	25	19.4	46	15.4	0	0.0	1	16.7	**72**	**16.6**
Disgust & Shame – Water treatment only – Collective BDM	8.3	12	9.3	24	8.0	0	0.0	0	0.0	**36**	**8.3**
Standard Health – Water treatment only – Collective BDM	8.3	12	9.3	24	8.0	0	0.0	0	0.0	**36**	**8.3**
Disgust & Shame – Water treatment only – Individual BDM	8.3	11	8.5	25	8.4	0	0.0	0	0.0	**36**	**8.3**
Standard Health – Water treatment only – Individual BDM	8.3	7	5.4	29	9.7	0	0.0	0	0.0	**36**	**8.3**
Total compounds per column / Percentage of total sample:		**129**	**29.7**	**299**	**68.7**	**1**	**0.2**	**6**	**1.4**	**435**	**100.0**

10 Figures

Figure 1: The chlorine dispenser hardware

Figure 2: Soapy Water Bottle

Figure 3: Subject Flow - Enrollment to Allocation

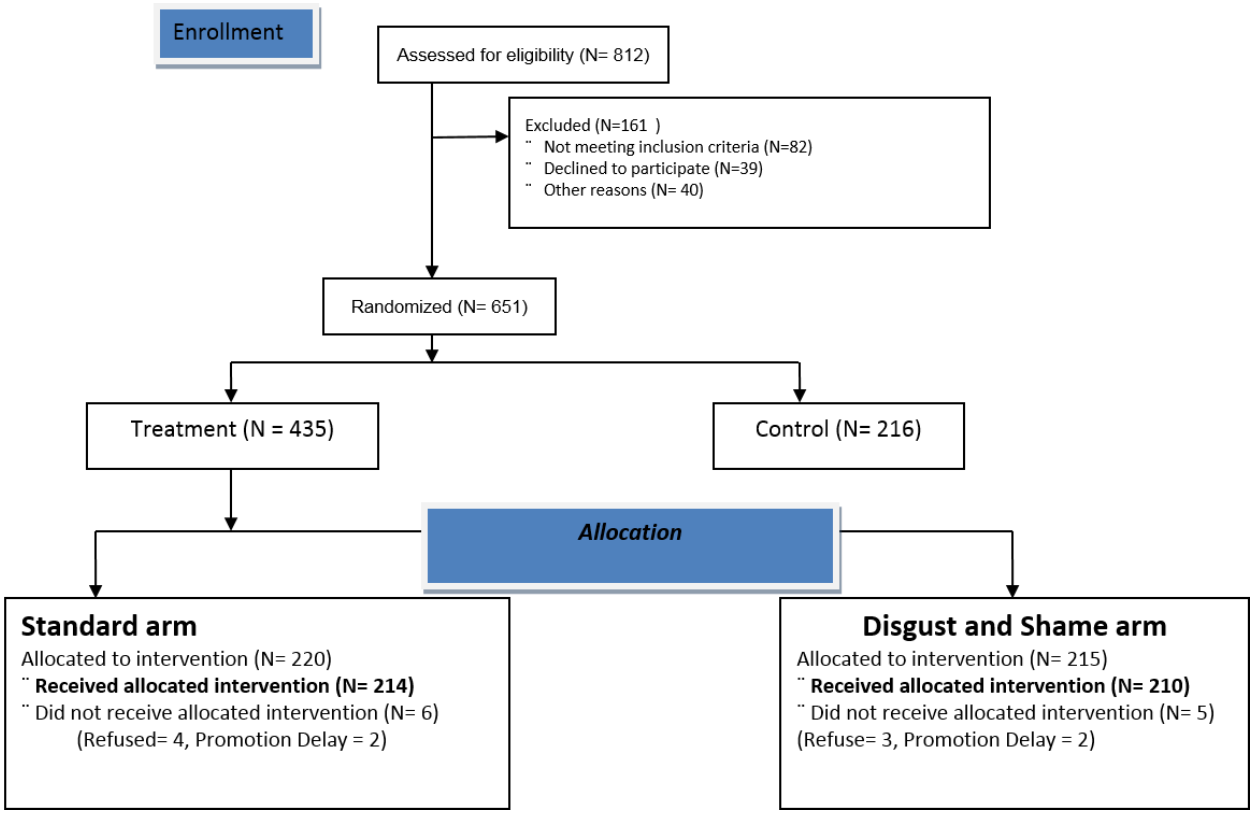
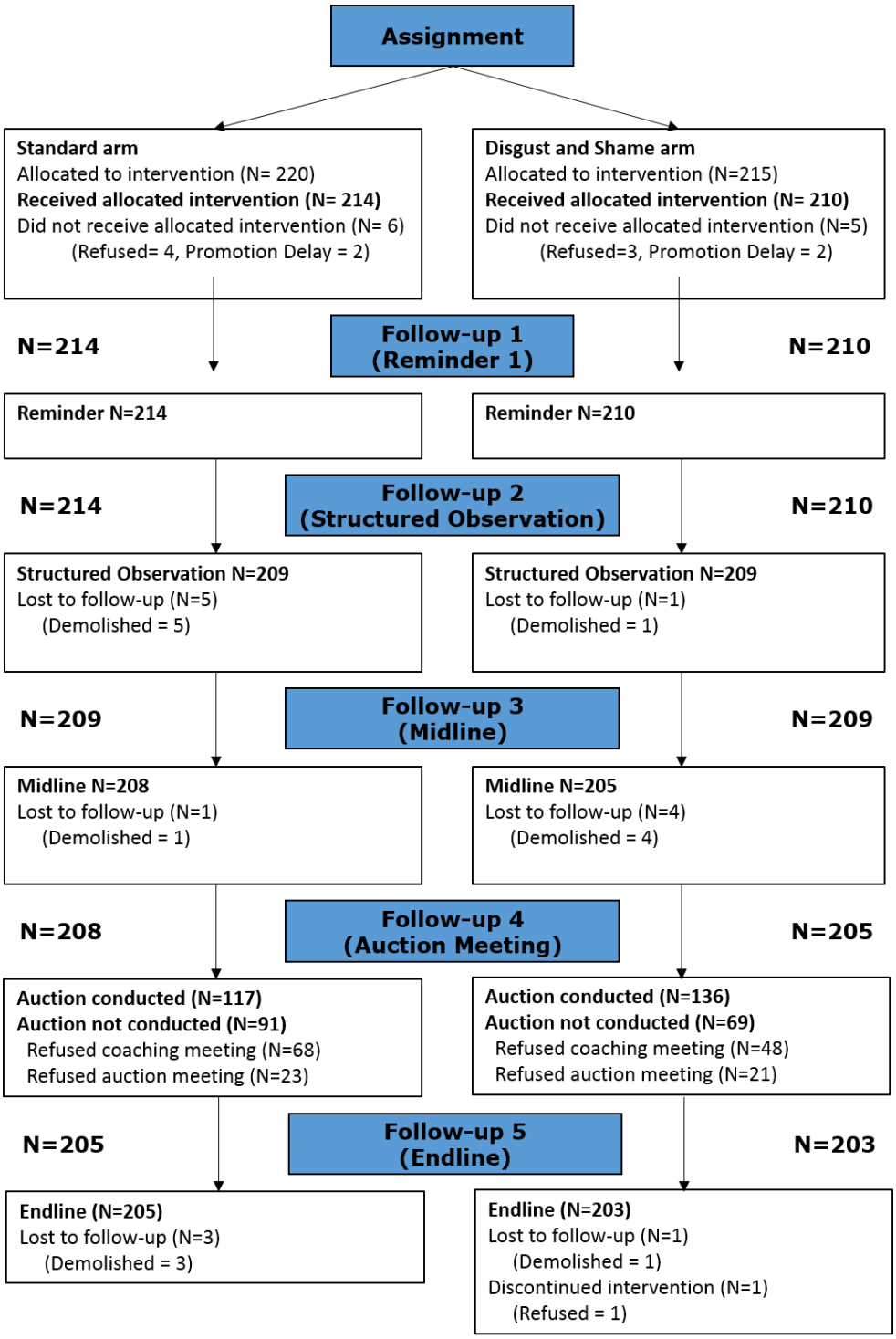
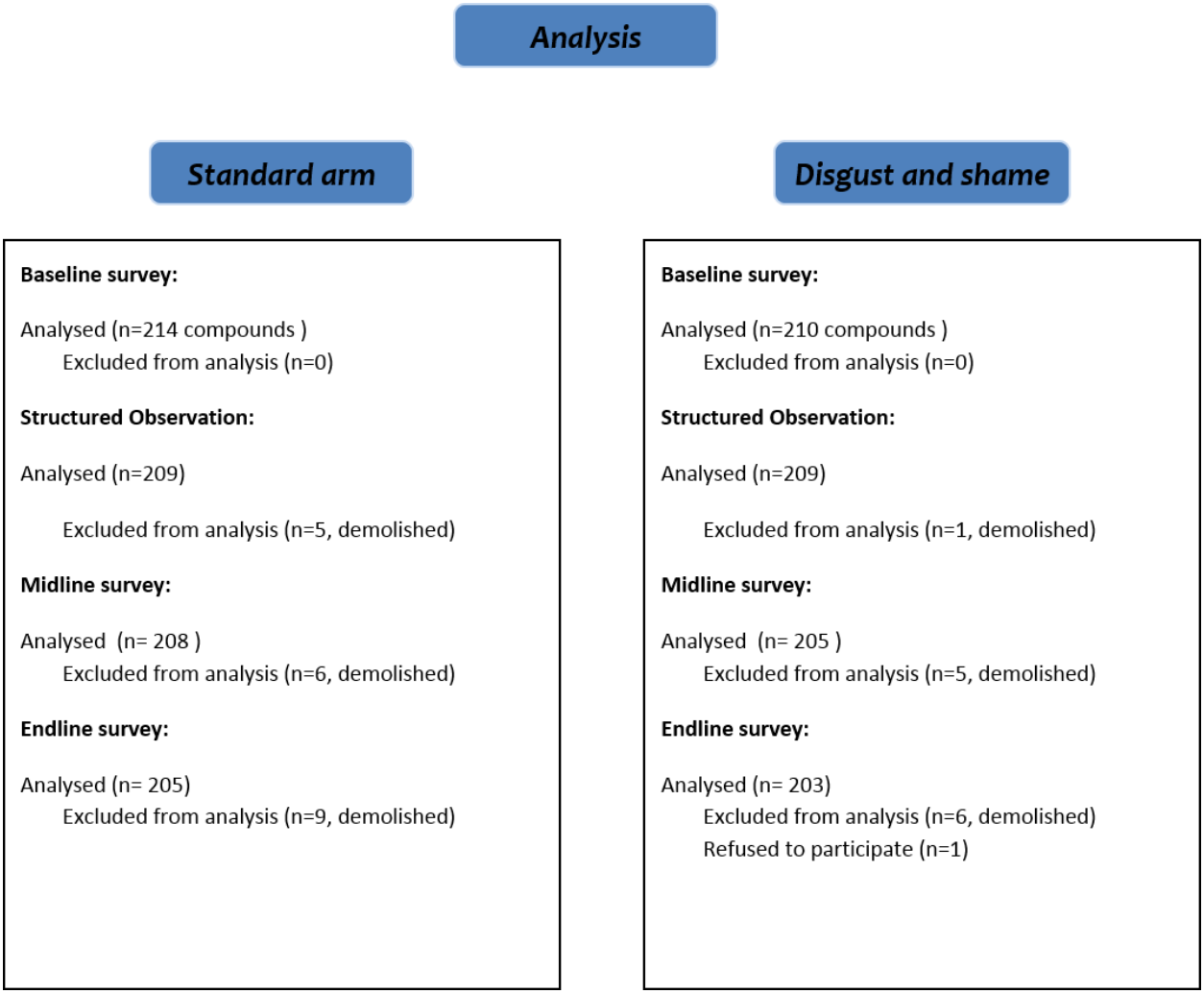


Figure 4: Subject Flow - Assignment to Endline



Note: Enrolment: September 2011, Reminder1: September 2011; Structured observation: October 2011 (during free trial); Midline: January 2012 (during free trial); Auction meeting: January 2012; Endline: April 2012

Figure 5: Subject Flow - Analysis



See note to Figure 4.

Figure 6: Subscription rates by treatment (% of compounds subscribing)