Experiential Learning in Health

Agha Ali Akram, Gabriella Fleischman, Reshmaan Hussam, and Akib Khan *

November 1, 2024

Abstract

This study examines how to sustain and scale the adoption of new health behaviors. We engage 1800 households in peri-urban Pakistan in a field experiment on water chlorination. Relative to control households who receive no intervention, households are randomized into receiving chlorine tablets only, also receiving daily incentives for chlorine use, or also participating in an experiential learning program. In the latter, community health workers help caregivers observe salient signals about the efficacy of chlorine tablets by visually tracking child diarrhea rates relative to control households before and after chlorine distribution for six months. While monetary incentives generate higher chlorination than experiential learning in the short run, these effects quickly dissipate while those in the learning arm remain, leading to significant reductions in diarrhea and improvements in child anthropometry. These sustained effects among learning households are driven entirely by those who also have a neighbor in the learning arm. Households *not* in the learning arm exhibit no difference in behavior by whether they have a neighbor in the learning arm. Our results suggest that experiential and social learning not only complement, but *rely*, on each other. Trust, built through information sources that participants are familiar with and have experienced as credible, emerges as the key mechanism driving this complementarity. We rule out a variety of alternative explanations, such as mimicry, social network formation, shifts in conversation content, and changes in social norms. The welfare implications are significant: ITT (TOT) estimates suggest that learning households with learning neighbors exhibit a 0.12 SD (0.52 SD) increase in an index of their child anthropometrics after one year. These results have direct implications for how public health information campaigns should be administered and scaled.

*Akram: Mathematica, Fleischman: Harvard Kennedy School, gfleischman@g.harvard.edu, Hussam: Harvard Business School, Khan: Stockholm School of Economics, akib.khan@hhs.se.

1 Introduction

How to, not only generate, but *sustain* and *scale* the adoption of new technologies or behaviors is a question central to economic theory and policy. The literature in technology adoption identifies learning – both how to use a good and what the returns to use are – as central to adoption (Bennett et al., 2018; Corno, 2014; Hanna et al., 2014; Arrow, 1962; de Janvry et al., 2016; Foster and Rosenzweig, 1994; Conley and Udry, 2010). We focus on the process by which people learn and consider the implications of this process for both sustaining and scaling the use of health technologies.

There are three primary channels through which people may learn about health technologies: experts (such as doctors, health workers, or scholars), laypeople (such as those in one's personal networks or popular information from non-experts), and the self (through personal experience). In environments with social and economic inequality, reliance on experts to disseminate information may exacerbate inequity given the variation in trust of experts (Alsan and Wanamaker, 2018; Alsan and Eichmeyer, 2024; Lowes and Montero, 2021; Martinez-Bravo and Stegmann, 2022; Banerjee et al., 2023), access to experts (Dussault and Franceschini, 2006), and access to well-informed networks (Banerjee et al., 2019; Calónico et al., 2023; Chen et al., 2022). An increasing body of evidence further documents that individuals may underweight information coming from other people, but react strongly to personal experiences (Conlon et al., 2022; Malmendier and Nagel, 2015; D'Acunto et al., 2021; Simonsohn et al., 2008). This study asks: what role can data-informed experiential learning play in sustainably altering health behavior? And what do social complementarities in experiential learning imply about scale?

To be sure, there are numerous reasons to be skeptical of learning through one's personal experience, particularly in the health domain. Health is a high-dimensional problem that is subject to random shocks and reliant on many inputs. These inputs are sometimes within an individual's control (for example, compliance with medical treatment, nutrition, and exercise), but oftentimes outside an individual's control (for example, pollution or traffic accidents) and potentially unobservable (for example, pathogens). Furthermore, health outcomes are noisy and rare occurrences that are difficult to observe precisely. Misattributing health inputs to health outcomes could bias beliefs and lead to suboptimal health behavior. Precisely for these reasons, the medical community relies on clinical trial evidence to inform medical and behavioral recommendations. However, individuals might more readily *trust* information they have observed. This source of learning also has fewer distributional implications, as the ability to observe is unconstrained by access to experts or social marginalization.

We seek to understand the role of experiential learning in changing health behaviors and the conditions under which experiential learning may be most effective. We conduct a field experiment with 1800 households in the peri-urban slums of Karachi, Pakistan. Our experiment has four arms: control households, who receive no intervention; chlorine households, who receive free chlorine tablets; incentivized households, who receive tablets and small daily financial incentives for chlorine use; and learning households, who receive tablets and a visual tool to

help track their household-specific health returns to their chlorination behavior. Incentive and learning tool provisions are temporary, but we continue providing chlorine tablets (to all treatment households) and tracking chlorine presence in drinking water and child health for over 12 months following intervention withdrawal. A comparison of the incentive and learning arms allows us to disentangle habit formation from learning, as both may raise the consumption stock of chlorine use in the short run, but only the latter can alter the perceived health returns to use. A comparison of both of these with the chlorine arm allows us to disentangle the degree to which access to chlorine is a barrier to use. And a comparison of chlorine with pure control offers us a useful policy benchmark against which to measure the relative health impacts of incentives and learning.

Our learning arm builds on the intervention of Akram and Mendelsohn (2021) (hereafter, AM). They developed a visual tool that plots the health returns to water treatment by (1) utilizing household-specific data on the frequency of diarrhea among treated children, (2) providing a community-specific benchmark against which to estimate health returns to chlorination, and (3) generating multiple observations over time per household before and after chlorine tablet provision during the peak diarrhea months of the year, facilitating learning from repeated data and trends in health. In a sense, households are offered the opportunity to perform their own "difference-in-differences" exercise. The information is presented through a simple visual tool (henceforth referred to as the Info-tool) in the form of a colorful bar chart that plots days of diarrhea for one's own treated household relative to the control households in the neighborhood, twice each month for six months: three months prior to chlorine tablet provision and three months following provision. Designed through extensive piloting both in AM and for the present study, the tool is easily usable and comprehensible to the low-literacy and low-numeracy households in our sample. AM, who also work in peri-urban Karachi, provide all households in their study with free chlorine tablets. They randomize the Info-tool treatment at the neighborhood-block level. Fifteen months after ending the Info-tool treatment, they find a 197% increase in the rate of chlorine detection in Info-tool households' water (from 31% to 86%, with no change among the control group).

The present study utilizes the same Info-tool intervention in our experiential learning arm and implements the study in peri-urban slums nearby, but not overlapping with, those of AM eight years after their intervention was completed. One key difference is that, while they randomize the Info-tool at the community level, we do so at the individual (household) level. As such, our learning arm serves as a type of replication of the previous study, with important differences (beyond our addition of the pure control and the incentives arm, both essential for disentangling underlying mechanisms). This enables us to see whether our results replicate under a new methodology, a different time, and a comparable but not identical context.

Returning to our study, we find that all three interventions (chlorine, incentives, and learning) significantly increase chlorine presence in drinking water for as long as we provide chlorine tablets, approximately fourteen months. While incentive households perform better than learn-

ing households during the period of incentive and Info-tool provision (the "short-run"), learning households outperform incentive households thereafter, with child diarrhea rates dropping the most for learning households following the withdrawal of interventions. Notably, however, differently from AM, we cannot rule out equality in average chlorination rates between the learning and chlorine-only arms in the long run. Recall that, while AM saturate the Info-tool treatment at the neighborhood level, we randomize the learning treatment at the household level. This motivates an exercise to test for learning spillovers.

We consider two possible ways in which there might be learning spillovers: experiential learning may be infectious in the social network, wherein non-learning-arm households are more likely to use chlorine when they have learning arm neighbors, diluting our ability to detect treatment effects; or experiential learning and social learning may be complements, wherein learning arm households are more likely to use chlorine when they have other learning arm neighbors, meaning that we weakened the potential impact of the treatment by randomizing at the individual level, thereby reducing the number of learning arm neighbors a learning household has. This distinction is important for policy. If experiential learning is infectious, the optimal policy would be to seed experiential learning to network-central individuals who can then "infect" others with this information. If experiential and social learning are complements, meaning that participants are most likely to change their behavior when they receive *both* information from their own experiences and from social connections, the optimal policy would be to geographically saturate experiential learning.

We find that experiential and social learning are complements. Participants in the learning arm who also have a neighbor in the learning arm use chlorine at a significantly higher rate than any other group. We find that this complementarity in learning is in fact *essential to sustained use*: the higher rates of utilization that we observe among participants in the learning arm fade out after one quarter *unless* they have a neighbor also in the learning arm. Participants outside of the learning arm behave non-differentially across access to a neighbor in the learning arm: in other words, they do not react to having a learning arm neighbor. Taken together, these results suggest that, while individuals do alter their behavior based on their own experience, this behavior change does not persist unless they are also able to observe, and *trust the experience of*, others in their social setting who are doing so as well. In our setting, experiential and social learning *rely* on one another.

The consequences of learning are substantial: one year after the withdrawal of interventions, children of caregivers in the learning arm exhibit the largest improvement in child health. The treatment leads to a 0.24 SD ($p = 0.025$) increase in an index of child anthropometrics among any participants who use chlorine (as instrumented by randomization into treatment), a 0.33 SD ($p = 0.005$) increase among *learning arm* participants who use chlorine, and a 0.51 SD ($p = 0.002$) increase among learning arm participants who use chlorine *and have a learning arm neighbor*.

To discipline our exploration of the learning process that Info-tool households may be engag-

ing in, we outline a model of experiential learning in which participants, when forming their beliefs about the efficacy of a health technology, weigh health signals differently according to how the health signal was observed: more-salient signals are weighed more heavily than less-salient signals, and signals observed through a more-trusted method are weighed more heavily than signals observed through less-trusted methods. We interpret the Info-tool as a technology that improves salience, wherein individuals place more weight on the health signals they receive when recording their health status than those that they observe but do not record. The model embeds a complementarity between ones' own signals and others' signals by assuming that individuals have higher levels of trust in signal observation methods that they have experienced. Consequently, if a participant has experience using the Info-tool, she more readily trusts information from somebody else whose information was likewise acquired through the Info-tool.

Our model poses belief-updating in response to health signals as the mechanism driving changes in chlorination behavior. To test this mechanism, we first check whether individuals are responding to health signals at all. Using a machine learning algorithm on a host of baseline variables, we identify which households are *predicted* to experience health improvements during the intervention period. We find that, during the three months that incentives and the Info-tool interventions are ongoing, the effects of the learning intervention on chlorination are 31% ($p = 0.029$) greater among these households who are predicted to improve relative to their non-predicted-improved counterparts within the Info-tool arm, suggesting that participants indeed respond to the information that they observe in their environment.[1] This separation does not exist for the same predicted-improved households in the chlorine-only or incentives treatment arms, suggesting that the pattern we identify is unlikely to be due to selection.[2]

Next, we check if learning arm participants are responding to their learning arm-neighbors' health signals. We find that the spillover is almost entirely explained by learning arm neighbors whose diarrhea rate was ex-ante predicted to improve. Learning arm participants do not respond differentially when they have access to *any* neighbor who was predicted to improve, bolstering the assertion that the spillover is driven by learning arm neighbors' health signals, rather than access to other attributes of people who are ex-ante predicted to have improved health.

We offer two caveats. First, while we model the learning arm as a program that facilitates participants in observing health signals by providing *salient* signals, it may instead work by leading to higher rates of early adoption which then provides *more* signals, or creates habit formation via building greater 'consumption stock' of chlorination.[3] We are able to rule out

[1] This effect continues but diminishes over the course of the fourteen months, likely because the other Info-tool households are also learning over time.

[2] Reassuringly, participants who are predicted *ex-ante* to observe null or negative health signals during the facilitated learning experience are *no less likely* to chlorinate their water than other participants outside this treatment arm, indicating that the potential for misattribution to lead to negative belief formation about the health technology is not realized.

[3] Salient signals may work more effectively if participants have limited attention (Hanna et al., 2014). However, if participants can attend to and learn from many signals, these signals will be more likely to converge to the truth, which will decrease the potential for misattribution.

this alternative channel by comparing the learning arm to the incentives arm. Incentives for chlorine use should lead to high rates of early adoption *without* helping participants observe more salient health signals, so any technology adoption mechanism that relies on early adoption should be realized in this treatment arm. Indeed, participants in the incentives arm use chlorine tablets at higher rates in the short-term, but they immediately revert to a lower rate of chlorine use after the withdrawal of incentives. This suggests that facilitating participants in observing health returns works by providing *salient*, not merely numerous, signals.

Second, we do not find evidence that participants in the learning arm are more likely to believe that chlorine is an effective technology (everyone rates chlorine as "very effective," and responses are likely subject to experimenter demand). This is not unlike Hussam et al. (2023), which finds that a hygiene information intervention significantly alters behavior but does not alter priors about the returns to proper hygiene. Instead, we find in our setting that learning arm participants are more likely to say that they are motivated to use chlorine to achieve a standard of health for their family. Furthermore, learning arm participants are more likely to believe that other learning arm participants would know more about child health than participants in the chlorine or incentives groups. While nearly all households can express that chlorine is a useful and effective technology, the Info-tool appears to make that information trustworthy and thereby motivates action.

As detailed in the paper, we are able to rule out a variety of competing mechanisms as explanations for the social learning we observe, including mimicry, social network formation, changing content of conversations (rather than changing beliefs about the content of the conversations), and changing social norms.

We highlight three central contributions of this study. First, we add to the literature on technology adoption and behavioral change by proposing a model of learning in which individual experience and social learning are complementary. In our model, this complementarity is a key underlying mechanism by which repeated exposure to, or experience with, a good or behavior translates into long-term adoption. We generate evidence for this model using a randomized experiment that suggests that actors are sensitive to *how* information is acquired and more readily trust information whose source they can understand. This may contribute to why people more readily trust laypeople with health information than experts (Alsan and Eichmeyer, 2024). Furthermore, we show that the learning treatment changes behavior without changing underlying beliefs about the efficacy of chlorine, adding to a small but growing literature which shows that information that individuals possess is not the only, and potentially not the primary, process through which behavior changes. Lab-in-the-field evidence suggests that an ownership effect over information that an individual herself generates might be important in determining if this information stimulates behavioral change (Conlon et al., 2022). We document this pattern outside the lab in a setting with high-stakes decision-making, but depart from Conlon et al. (2022) by showing that, under the right conditions, people can "take ownership" of information that comes from other people.

Second, we operate in a setting with significant policy implications for scaling education and information campaigns in preventive health. The returns to preventive health behaviors, such as handwashing with soap, using clean cookstoves or bednets, or treating water, are particularly challenging to perceive the returns to. Existing literature on using social networks for information dissemination has focused on seeding information. This literature assumes that everyone can learn from social connections if the information reaches them, but seeding omits the value of the experiences of the person *receiving* the information. Our evidence suggests that, for some types of learning, these assumptions could seriously undermine the potential for individuals to learn from disseminated information. In our setting, personal experience is essential to learning from others, and having others to learn from is essential to acting on information from personal experiences. *Own experience cannot be seeded*, suggesting that there might be large efficacy costs in seeding certain types of information, rather than saturating it.

Finally, we are one of few papers in the learning-through-adoption literature that can definitively rule out habit formation as a confounding mechanism to learning, two behavioral processes that are often unknowingly conflated in existing work (see Appendix Section 8.6.1 for a detailed review of the existing literature).

Our paper proceeds as follows: Section 2 provides the details of the experimental design, sample, and data; Section 3 presents a model of complementarities in experiential and social learning; Section 4 presents overall results on chlorine use, and spillover results on chlorine use; Section 5 discusses mechanisms; Section 6 presents results on child health; and Section 7 concludes.

2 Experimental Design

We conducted a randomized controlled trial in Ibrahim Hyderi, a peri-urban neighborhood of Karachi, Pakistan. We identified 1,800 eligible households through a census. Households were eligible if there was at least one child between the ages of 6 months and 5 years old at baseline, and if the caregiver was generally home during the day. In all but one case the caregiver was a female adult, typically the mother of the children. Furthermore, we screened households on several dimensions of water usage to ensure that chlorine would be useful and effective for improving children's health.[4]

[4]Households were excluded if they did not store drinking water in a separate vessel from the water for other uses, if children drank from a separate vessel from adults, if the vessel's capacity was less than 10 liters (to ensure there would be enough water to avoid over-chlorination), or if children frequently drank bottled water (see Figure 15 for a sample vessel). Of the households we approached, 93.6% were eligible based on these criteria, and 80% of the ineligible households were ineligible on the basis that nobody was home, nobody could regularly be home in the future, or there was no child in the correct age range in the household. Only nineteen households (less than 1% of all households approached) were ineligible based on water infrastructure.

2.1 Baseline Survey, Randomization, and Balance

Our baseline survey of these 1,800 eligible households ran from mid-May to early June of 2022. We collected data on chlorine exposure, chlorine knowledge, diarrhea prevalence,[5] and child anthropometrics.[6] As expected, there was not a single household whose water vessel tested positive for chlorine residual, as 98% of households reported that chlorine tablets were not available in local markets. Awareness of chlorine tablets in this community was likewise low: Twenty percent of the sample reported having heard of chlorine tablets, but forty percent of this sample could not properly identify what chlorine tablets were used for. Despite low awareness, eighty-seven percent reported that they would consider using chlorine tablets after enumerators explained the purpose of chlorine for purifying water.

We electronically randomized households to treatment groups between the baseline visit and the first biweekly visit. Eligible households were individually randomized into the four 450-person experimental arms as follows:

Control (C): Pure control (no offer of free chlorine tablets; chlorine testing)

Treatment 1 (T1): Free chlorine tablets and chlorine testing (Chlorine Only)

Treatment 2 (T2): Free chlorine tablets and chlorine testing + Info-tool (Info-tool)

Treatment 3 (T3): Free chlorine tablets and chlorine testing + financial incentive (Incentives)

Treatment groups are balanced on most, but not all, baseline measures. In Tables 12 and 13, we compare each treatment group with the pure control group. Children from T1 are reported to have fewer diarrhea days than the control group both when considering a binary measure for whether or not any child in the household had diarrhea in the past fourteen days (Table 12), or when considering both a binary measure and the number of diarrhea-days at the child level (Table 13). However, there are no significant differences in child anthropometrics, nor do we see lower levels of diarrhea among T1 households in the following visit (before households

[5]We asked respondents to describe their perception of "motions", or diarrhea. Then, we described diarrhea to respondents as loose or watery stools. For respondents who were not sure what classifies as diarrhea, we also shared a visualization of stool types called the Bristol stool chart. Respondents then reported how many days each child under five years old experienced diarrhea in the past fourteen days. Diarrhea rates were higher than expected during the baseline period. Average rates of diarrhea in May are close to 3% of days per child in the literature, whereas households in our sample experienced diarrhea on 5.9% of child-days (Luby et al., 2006). There was an extreme heatwave ongoing in Karachi during the baseline period, which we believe explains the high rates of diarrhea (Xu et al., 2014). Diarrhea rates converged to the expected rate in the following month, without any increase in chlorine use.

[6]Anthropometric outcomes were only recorded at baseline and endline – child weight and mid-upper-arm circumference at baseline and endline, and child height and age at endline only. We calculated the children's weight as the difference between the weight of the mother alone and the weight of the mother while holding the child. For children older than 6 months, we also collected mid-upper-arm circumference. There were 338 households for whom we did not collect anthropometric data in the baseline visit due to issues with the measurement equipment. As such, we collected anthropometric data for these households in the following visit. We did not reveal treatment status to households until the first Phase 1 visit, and enumerators were blinded to treatment status during the baseline visit, so there is no reason to believe that treatment would differentially affect child anthropometrics between the baseline and the following visit. Furthermore, we visited these households at the beginning of the following visit to minimize the time between the end of the baseline survey and the time of measurement for any anthropometric outcomes that we consider as baseline measures.

had received any treatment interventions), indicating these differences are likely spurious.

		Phase 1		Phase 2		
		Round 1	Round 2			
	Month 1	Pre-Treatment: Months 2-4	Treatment: Months 5-7	Post-Treatment: Months 8-18	Month 19	Month 24
Chlorine	Baseline Survey		Chlorine Distribution	Chlorine Distribution	Endline Survey 1	Endline Survey 2
Incentives			Chlorine Distribution + Incentives			
Info-Tool		Info-tool Training	Chlorine Distribution + Info-tool			
Control						
		Twice-monthly visits for child health data	Twice-monthly visits for child health and chlorine presence data	Once-monthly visits for child health and chlorine presence data		

Figure 1: Study Timeline

2.2 Phase 1: Biweekly Visits

Phase 1, during which we introduced the behavioral interventions and visited households every two weeks, consisted of two rounds. Throughout Round 1, households were aware of their treatment status and we trained T2 households on how to use the Info-tool chart. In Round 2, we distributed chlorine tablets to all three treatment groups and delivered the incentives treatment to T3 participants, while T2 households continued to use the Info-tool chart. Figure 1 lays out the study timeline and key events.

2.2.1 Round 1 of Phase 1

Throughout Round 1 of Phase 1 (June to August 2022), households were visited by surveyors every two weeks for a short round of data collection. These surveys were standardized and involved a few key components: measuring child diarrhea days over the past two weeks, testing for the presence of chlorine residual, and training caregivers in the T2 group on how to use the Info-tool chart.[7] Since chlorine distribution had not yet begun, and chlorine tablets are not widely available in local markets, ten percent of households were randomly selected for chlorine testing during every other visit during Round 1 of Phase 1. Chlorine testing in this period was

[7]In the first biweekly survey, we also collected data on household bargaining power, as we were unable to incorporate this into our first baseline visit. Since we ask caregivers questions about decision-making power before revealing treatment status, we consider this as a baseline measure of household bargaining power. There were no differences across treatment groups in being the sole decision-maker, or a part of decision-making, in issues of child health, household purchases, or household visits. We also collected anthropometric data for 338 households for whom we were not able to collect anthropometric data in the baseline visit.

simply to ensure that knowledge of treatment status did not affect people's ability to procure chlorine tablets on their own.

In the first Phase 1 visit, we revealed treatment status to caregivers. We explained each treatment group to caregivers and told them that a lottery would determine their treatment status. We explained that the lottery would run through the enumerator's tablet, and that whichever number would appear on the screen would determine their treatment status. These numbers were pre-loaded so as to deliver the treatment status we pre-assigned. After treatment status was revealed, we reminded participants about what being a part of their treatment group would entail.

Info-tool

To activate the 'learning' mechanism, participants in T2 received the Info-tool chart on the first visit of Phase 1.[8] The Info-tool chart was a simple pen-and-paper chart that allowed caregivers to track their children's diarrhea (Figure 13). Each chart consisted of two bars to represent each month: one bar in which caregivers filled in a square for each child-day with diarrhea (for example, if two children had diarrhea on the same day, the caregiver would fill in two boxes), and one bar in which the enumerator filled in a benchmark diarrhea rate. The benchmark diarrhea rate was calculated in two ways: initially, by using the fourteen-day diarrhea rate in the pure control group across all children, multiplied by the number of children in the household; and subsequently, by using the monthly incidence of childhood diarrhea from the epidemiological literature (Luby et al., 2006). When we used the diarrheal incidence in the Control group, the benchmark was updated daily so as to reflect the past fourteen days of data collection from the pure control households. We switched from using the diarrhea rate in our Control group to the diarrhea rate in the epidemiological literature out of concern that the incidence of diarrhea in the Control group would be impacted by a reduced disease environment due to the presence of treated households in their community. Enumerators clearly explained to participants that the benchmark rate was the average rate of diarrhea from people in the community who did not use chlorine to purify their water.

This simple, visual paper-and-pencil tool was borrowed from Akram and Mendelsohn (2021), which was comfortable for low-literacy-and-numeracy caregivers to use. In this way, T2 caregivers were able to compare their children's diarrheal incidence with the average diarrheal incidence in households that did not use chlorine tablets, both before they were offered chlorine tablets (months 1-3 of the intervention) and after (months 4-6 of the intervention).

[8]The study was non-blinded with respect to study subjects and treatment administrators (CHWs). Enumerators were not informed by the research team on the specific treatment status of the households at baseline. While all enumerators were hired through our partner organization, Interactive Research and Development (IRD), we recruited a different team for the endline survey to ensure objectivity, and these enumerators were not informed of the participants' treatment status before beginning the survey. However, participants were not blind to their own treatment status, so it is possible that they discussed their treatment status with the enumerators and effectively rendered the endline survey non-blinded. Community health workers used electronic tablets and smartphones to collect data.

2.2.2 Round 2 of Phase 1

Round 2 of Phase 1 began in late August 2022 and continued for three months. Throughout this round, we continued to visit households every 2 weeks and commenced the distribution of chlorine tablets to all T1, T2, and T3 households. We continued to collect information on diarrhea prevalence and tested water for the presence of chlorine in all households. We tested for chlorine residual every other visit (once per month).

Incentives

At the beginning of Round 2, we also began the incentive treatment with T3 households. This treatment, which should build consumption stock, was designed to test learning against habit formation as the primary mechanism. Caregivers were offered tokens redeemable for child and household goods if they could show empty chlorine tablet wrappers as proof of usage. The chlorine tablets came individually wrapped and participants were instructed to save the empty wrappers in a small pouch provided for them. Each daily reward for proper chlorine use was equal to approximately 5 US cents (with 'proper use' calibrated to the household's pre-intervention water consumption). To hold income effects constant, we also gave comparable products to participants in the remaining groups – *unconditioned* on chlorine tablet wrappers – as a token of appreciation for participating in our surveys. Since T3 households could redeem tokens for household goods that had varying values, we implemented a lottery to determine the value of the unconditional gifts that non-T3 households received.

Info-tool

At the end of Round 2, we aggregated the monthly Info-tool chart statistics across three-month intervals and presented the aggregated data to T2 participants visually (Figure 14). As such, T2 participants were able to visualize their own children's diarrhea rate relative to the average diarrhea rate among households that did not use chlorine in the three-month interval before chlorine tablet distribution, and make the same visual comparison for the three-month interval following chlorine tablet distribution: in effect, a difference-in-differences.

Throughout Round 2, we conducted random unscheduled audits where CHWs tested water for the presence of chlorine residual. These audits were conducted to ensure that households were not chlorinating their water in anticipation of our bi-weekly visits and to increase confidence that our scheduled monthly chlorine tests serve as a good proxy for chlorine use throughout the month (audit visits are discussed in more detail in Appendix Section 8.7).

2.3 Phase 2: Monthly Visits

Phase 2 began in late November 2022 and continued through October 2023. In this phase, we visited households once per month. We continued to distribute chlorine tablets and test water for the presence of chlorine residual but ceased provision of incentives or gifts, no longer informed Info-tool participants of their own children's diarrhea rate relative to the benchmark

diarrhea rate, and stopped providing Info-tool sheets or helping Info-tool participants track their children's diarrhea rates.

2.4 Endline Visits

The first endline survey was completed in December 2023 through January 2024. In this round, we collected all of the measures from our baseline: diarrhea rates; presence of chlorine residual; child anthropometrics, including weight, mid-upper arm circumference, and height; and household bargaining power. We additionally elicited willingness-to-pay for chlorine tablets using a take-it-or-leave-it offer with a randomized price.[9] Finally, we collected detailed social network data at endline to understand potential mechanisms for spillovers, including directly serving neighbors chlorinated water (through meal-sharing, for example), and indirectly through conversations about the project and lessons learned through the Info-tool.

We were able to find 87% of households for the endline survey. Incentives households were the most likely to attrit (14.9%), and were 4.2 percentage points more likely to attrit than Control ($p < 0.05$; Table 12).[10] While this selective attrition is a concern, none of our primary analyses at endline rely on a comparison between Incentives and Control. Our primary analyses compare Info-tool with Control or the Chlorine Only group, each of whose endline attrition is non-differentiable from the other.

We conducted a final short followup survey in June 2024, motivated by three reasons: (1) to provide participants with chlorine tablets for the summer months, the period in which diarrhea rates peak, (2) to monitor stockpiling of chlorine tablets, and (3) to ask additional survey questions to disentangle mechanisms.[11]

Taken together, we administered 26 surveys per household: a baseline survey, 11 rounds of bi-weekly visits in months 1-6 of the intervention (Phase 1), 12 monthly visits in months 7-18 of the intervention (Phase 2), one endline visit immediately after Phase 2, and a second endline visit four months later. Our primary outcome, chlorine residual in drinking water, was recorded in nineteen of these visits (baseline; bi-weekly visits 6, 8, 10, and 11; and every visit thereafter). We randomly tested households for chlorine residual in months 1-3 to ascertain that people never attained and used chlorine before we began distribution.

[9] After providing respondents with a one-month supply of chlorine tablets, free of charge, we offered them a second month's supply of chlorine at a randomized price (market price, a 29% subsidy, or a 53% subsidy).

[10] Note, however, that the endline Incentives respondents do not statistically differ from their counterparts in the other arms (Tables 38 and 39).

[11] We sought to monitor stockpiling because there was a sharp drop-off in rates of chlorine detection in the last two rounds of surveying (the final survey of Phase 2, and the first endline survey), which was precisely when we began to tell respondents that the trial was nearly complete and that we would soon cease to provide free chlorine tablets. Since we ended the survey during the winter, the time of year when diarrhea rates are lowest, stockpiling chlorine tablets for the summer would have been a rational response from participants. We find some evidence of this: although only 3% of respondents reported that they still had chlorine tablets remaining in this final followup survey, 44% of respondents reported that they ran out of their chlorine tablets later than they should have if they were chlorinating their water daily, and for 24% of households the discrepancy was by greater than one month.

2.5 Outcomes

Our primary outcomes are chlorine use and child health. Our measure for chlorine use is a binary indicator for the presence of chlorine residual in drinking water. After we started distributing free chlorine tablets, CHWs tested for chlorine residual in each household's drinking water vessel every month using a simple test strip that, when dipped into a small cup of water, turns shades of blue depending on the chlorine concentration in the water (the minimum amount of chlorine we can detect with our test strips is 0.5 parts per molecule). The presence of chlorine residual is an objective measure of chlorine tablet use that cannot be manipulated. However, chlorine residual only presents itself if the chlorine tablet was added to the water in the past twenty-four hours, meaning we may not capture a treatment effect for households who imperfectly chlorinate their water.

Our measures of child health are self-reported diarrhea incidence and child anthropometrics. Diarrheal incidence is measured by the number of days of diarrhea across all children under five years old, which we collect in every visit. While we do collect data on the frequency of childhood diarrhea among households, we do not use diarrhea as our only measure of child health because it is subject to measurement and misclassification errors. Firstly, our interventions, particularly Info-tool, are likely to change the reporting of diarrhea in children, since we explicitly encourage participants to track diarrhea. Thus, we may capture a treatment effect on reporting, rather than a treatment effect on actual incidence of diarrhea. Secondly, the presence of non-infectious diarrhea and asymptomatic infection make diarrhea presence a poor outcome for measuring the effectiveness of water and sanitation interventions (Watson et al., 2022).

Frequent diarrhea can result in poor absorption of nutrients, leading to sub-optimal physical development in children. As such, we supplement diarrhea incidence with child anthropometrics, including height-for-age, weight-for-age, weight-for-height, and mid-upper-arm-circumference-for-age, which we collected at endline among children under 80 months old[12].

3 Model of Complementarities in Experiential and Social Learning

We present a model of learning whose key insight is to show how a complementarity arises between experiential and social learning. Crucially, we depart from standard models of learning by allowing beliefs about information to depend on how that information is acquired, and allowing trust in information-acquisition technologies to depend on personal experience with those technologies.

[12]This is because, in baseline, we measured children who were under 5 years or 60 months old.

Net information in the model:

We adapt a model of naive DeGroot learning that allows for repeated learning from the self as well as social learning. As with a standard naive DeGroot model, caregivers interpret information from each person in the network as new, and do not internalize that two network connections' information may have come from a common source. We depart from the standard DeGroot model by allowing caregivers to draw from multiple observations from their own experiences. Thus, rather than averaging their own and their neighbors beliefs in each period, they average all of their own and the signals that their neighbors send to them across all periods. Similarly, caregivers do not consider common source shocks, such as common health shocks, that may lead to serial correlation between their own observations, or spacial correlation with other people's observations about the technology.

Let there be J people in a community. An individual aggregates signals from her own experiences and from information sent by people in her network to form beliefs about the efficacy of chlorine in improving children's health. Individual i's belief about the efficacy of chlorine in time T, $\hat{S}_i^T$, is:

$$\hat{S}_i^T = \frac{\sum_{t=0}^{T}\sum_{j=1}^{J} \alpha_{ij}^t \cdot a_{ij}^t \cdot s_{ij}^t}{\sum_{t=0}^{T}\sum_{j=1}^{J} a_{ij}^t}$$

where s_{ij}^t is the information that individual j sends to individual i in time t (one's own signal is s_{ii}^t), α_{ij}^t is the weight that individual i gives signal s_{ij}^t, and $a_{ij}^t \in \{0,1\}$ is a binary variable that indicates if i received a signal from j in time t. We let $a_{ii}^t = 1$ if individual i adopts chlorine in period t. In other words, an individual will not interpret her observations about her environment, such as the presence or absence of her children's health symptoms, as information about chlorine unless she uses chlorine in that time period.

We define weights assigned to each signal $\alpha_{ij}^t \in [0,1]$ as the caregiver's believed probability that a signal is accurate.

$$\alpha_{ij}^t = \alpha_{ij}^t(\gamma_j^t, \omega_{\gamma_j,i}^t, \cdot)$$

where γ_j^t is the technology that sender j uses to observe signals in time t, and $\omega_{\gamma_j,i}^t$ is receiver i's time t trust in the γ_j^t observation technology. Furthermore, we will assume that i's trust in γ_j^t depends on her experience with that technology:

$$\omega_{\gamma_j,i}^t = \omega_{\gamma_j,i}^t(\gamma_i^{t-1}, \gamma_i^{t-2}, ..., \gamma_i^0, \cdot)$$

We assume that other reasons for i to be skeptical towards a signal from j in time t are exogenous.

Thus, caregiver i's net information is the average of the signals she has received from her own observations over time and the information she has received from others, weighted by her beliefs about the probability of accuracy of those signals. The probability of the accuracy of those signals will depend on how the signal was observed, and the caregiver's experience with

observing signals in the same way.

She will adopt chlorine in time t if $\hat{S}_i^t > c_i^t$, where c_i^t is her cost associated with taking chlorine in time t. We model our caregiver as a passive observer of health signals, rather than an active experimenter who makes adoption decisions with the purpose of learning. Thus, her adoption decision depends only on her contemporaneous belief about the efficacy of chlorine relative to the cost, and she does not optimize adoption decisions in order to aggregate more health signals and learn.

Adoption in t=0:

Note that adoption in any given period will only happen if an individual has adopted chlorine already, has learned about chlorine already, or if $c_i^t \leq 0$. In this setting, the technology is new and it is unlikely that individuals have previous experience with the technology before the experimenter arrives (at baseline, only 12.7% of respondent said that they had heard of chlorine before and could name that its purpose was to purify water). We allow the information sent by the Community Health Worker (CHW) to enter into the information the caregiver receives from the network. That is, the CHW shares information in the network such that $a_{i,chw} = 1, \forall i$. The CHW always sends a signal that $s_{i,chw}^t > c_i^t$ (the CHW will always inform the caregiver that chlorine is efficacious). Then, a caregiver will only adopt chlorine when we first distribute chlorine if her intial beliefs that the probability the CHW sent an accurate signal, $\alpha_{i,chw}^0$, is sufficiently high, or if she has costs associated with taking chlorine c_i^t close to zero. It is likely that $\alpha_{i,chw}^0$ is high, considering that the CHW represents a trusted NGO and had already visited households for three months regularly prior to distributing chlorine. It is also likely that c_i^t is close to zero in the first period, since the largest reported cost associated with taking chlorine is an unpleasant taste if the chlorine dose is not titrated just right, which caregivers would not yet have experienced.

Interventions:

All chlorine groups: We reduce the cost of taking chlorine c_i^t in all time periods, through free distribution and free delivery.

Incentives: We further reduce the cost of taking chlorine c_i^t in the Incentives group for the first 3 months of chlorine distribution by providing monetary incentives for use.

Info-tool: We change the signal observation technology γ_i^t in *only* the first 3 months of chlorine distribution. Consequently, we increase trust in signals sent by other Info-tool participants, $\omega_{\gamma_j,i}^t | j \in IT, \forall t > 3$ where sender j informs receiver i that she observed her signal through the Info-tool.

Comparative statics:

Shock to c_i^t for Incentives:

Prediction 1: Increased contemporaneous adoption relative to Chlorine Only

A negative shock to c_i^t will increase the probability in any period that $\hat{S}_i^t > c_i^t$, leading to increased contemporaneous chlorine adoption. This effect will be immediate from the first period of chlorine distribution, but will fade out as soon as the CHW stops providing chlorination incentives.

Shock to γ_i^t for Info-tool:

Prediction 2: Increased adoption relative to Chlorine Only among individuals with larger early-period health signals

Changing γ_i^t leads to a contemporaneous increase in the weight placed on health signals α_{ii}^t. We should expect to see higher use in the Info-tool group than in the Incentives or Chlorine groups immediately after the end of the behavior interventions. After the CHW stops assisting participants in the Info-tool, and stops providing gifts to the Incentives groups, costs are the same for all three groups. However, Info-tool participants will still incorporate the higher weights that they placed on the health signals from the early part of the trial into her beliefs about chlorine's efficacy. Thus, we should see higher average use immediately after the behavioral interventions end.

Since individuals learn from short-term health signals, we should expect to see heterogeneity by the sign on individual contemporaneous health signals, and the Info-tool group should be the most sensitive to small differences in health signals. However, placing heavier weights on early health signals is not the only way in which Info-tool participants differ from the other groups. Info-tool participants also place heavier weights on signals from certain other people, described below, which may supersede heterogeneity by individual early health signals. As such, heterogeneity by early health signals is suggestive of learning through observation, but the absence of heterogeneity by health signals is not dispositive.

Shock to $\omega_{\gamma_j,i}^t | j \in IT$ for Info-tool:

Prediction 3: Increased contemporaneous adoption relative to Chlorine Only among individuals with social connections who were also in the Info-tool group

As soon as an Info-tool participant has used the Info-tool, which is a shock to γ_j for the Info-tool technology j, this is also a shock to $\omega_{\gamma_j,i}^t$, when the technology j is the Info-tool. This, in turn, is a shock to α_{ij}^t, when the signal observation technology is the Info-tool, even when the signal comes from somebody else. Thus, Info-tool caregivers with access to another network connections who was in the Info-tool group will also respond to those network connections' health signals, while other groups will not. Furthermore, Info-tool participants should be

responsive to the sign of the signals they receive from other Info-tool social network connections. We allow network connections to share net information or period-specific information, and for informants to share information in any period.

4 Results: Chlorination

4.1 Overall Average Treatment Effects

We begin by plotting chlorine use by treatment groups across the study period (Figure 2). The short-run is defined as Phase 1 (weeks 12-20). The medium-run spans the first three months of Phase 2 (weeks 22-32). The long-run spans the remainder of Phase 2 (weeks 36-72).

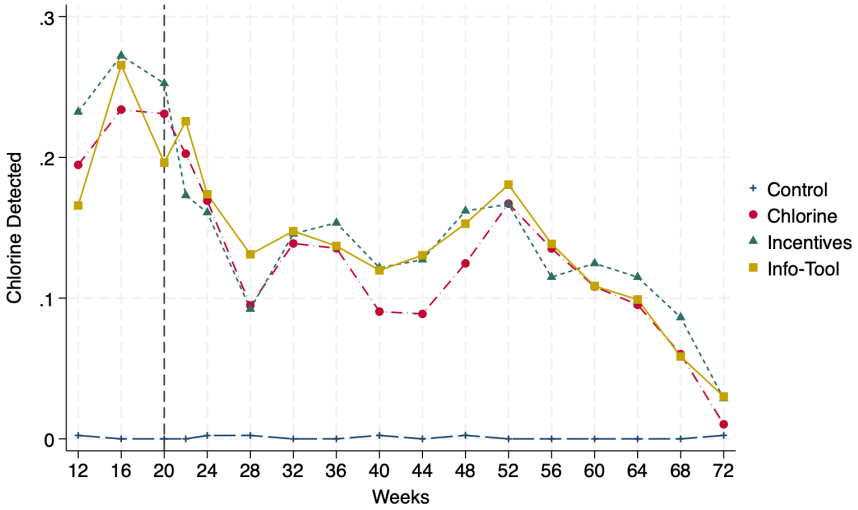


Figure 2: Raw Rates of Chlorine Detection

Free chlorine distribution leads to a large increase in chlorination in all treated groups. As we will see in Section 6, there are large decreases in reported child diarrhea and improvements in child anthropometrics in all groups, indicating that free chlorine distribution alone, even with low levels of perfect compliance, can have substantive impacts on child health. On aggregate, the impacts of the behavioral interventions are small in magnitude and at times statistically insignificant. We highlight two important patterns: first, during the short-run, Incentives households chlorinate more than those in Info-tool or Chlorine. Immediately following incentive withdrawal (week 20) however, chlorination rates for incentive households plummet, while the decay among Info-tool households is slower. Chlorination patterns across the treatment groups appear, on average, to follow one another thereafter.

On average, chlorine residual is detected in the water of over 20% of the treatment households during the short-run (weeks 12-20). In the medium run (weeks 22-32, or over the three months post-intervention), the detection rate declines, especially among Incentives households. Info-tool households, in contrast, chlorinate more than the other two intervention groups throughout this period. In the succeeding period, however, Incentives households and Info-tool household

converge around week 32, with households in Chlorine finally converging to their chlorination rate in week 52 (eight months after the end of interventions). Finally, there is a gradual decline in chlorination rates across all groups in the last four months of the study period.

Our main empirical specification is the following:

$$Y_{i,s} = \beta_0 + \beta_1 T1_i + \beta_2 T2_i + \beta_3 T3_i + Y_{i0} + X_{i0} + \delta_s + \gamma_b + \epsilon_{i,s}, \tag{1}$$

where $Y_{i,s}$ is an outcome for individual i measured in survey-round s; $T1_i$, $T2_i$, and $T3_i$ are binary variables representing Chlorine Only, Info-tool, and Incentive group participants, respectively; Y_{i0} is the baseline measurement of the outcome (whenever available); X_{i0} is a vector of unbalanced baseline covariates and other covariates that we select using the double-lasso method proposed by Urminsky et al. (2016); δ_s denotes survey-round fixed effects; and γ_b are block fixed effects (geographic units of stratification). We cluster standard errors at the household level.[13]

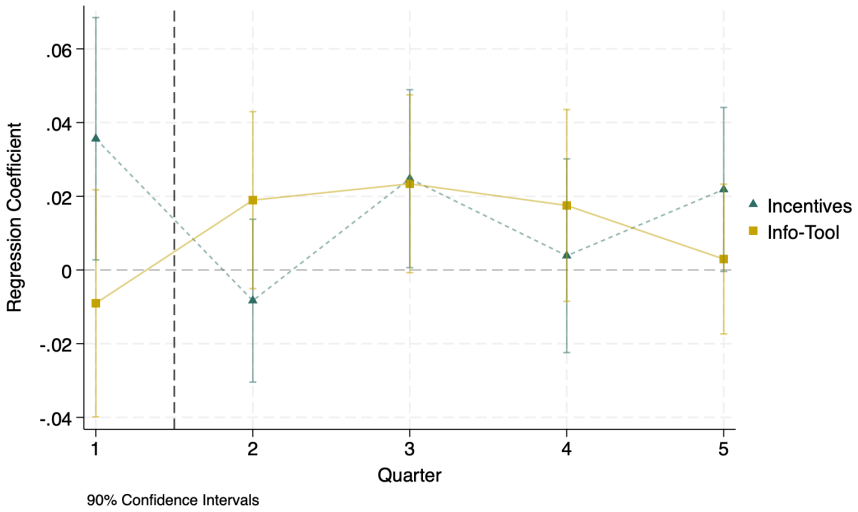


Figure 3: Regression Coefficients of Chlorine Detection (Omitted: Chlorine Only)

Figure 3 shows the regression results for chlorination using the panel dataset from household visits (also reported in Appendix Table 15). Consistent with the observations from the raw chlorination plots, during the short-run period (weeks 12-20, or Quarter 1), households assigned to Incentives outperform those in Info-tool or Chlorine, at a rate just shy of conventional levels of statistical significance (p-value = 0.110). In the immediate post-intervention period (weeks 22-32, or Quarter 2), households in the Info-tool group chlorinate at a higher level than Incentives and Chlorine; however, contrary to the findings in AM, the difference in treatment effects between Info-tool and Chlorine is not statistically distinguishable (p-value = 0.290).

[13]If a participant could not be found or refused to let us test their water in any given survey round, we consider this observation as an attriter and drop it from the analysis.

4.2 Learning Spillovers

In our model, individuals update their beliefs about health technologies when they receive signals from themselves *and others*. Motivated by this theory, we investigate the effects of learning spillovers. Learning spillovers are also a natural potential explanation for why we are not able to replicate the large positive Info-Tool treatment effects that AM find, given that the key difference between the two studies is that we randomize the treatment at the individual level, and AM randomizes the Info-tool treatment at the neighborhood-block level.

Households are, on average (median), 13.6 (10.4) meters from the closest other participant; 16.0 (12.1) meters from the closest participant in any chlorine treatment group; and 27.5 (22.7) meters from the closest Info-tool participant. While caregivers are surrounded by other participants, only 43% of participants have a nearby neighbor (someone within twenty meters) in the Info-tool group. When we recenter this measure to account for endogenous neighborhood density (following Borusyak and Hull (2023)), only 35% of participants are exposed to a nearby neighbor in the Info-tool group by random variation in the treatment assignment (Appendix Table 14 demonstrates that this measure is balanced across treatment arms).

We compute exposure to Info-Tool neighbors in the AM data using the same spillover definition. In that experiment, 49% of Info-tool participants were randomly exposed to other Info-tool participants by variation in the treatment assignment, while 2% of Chlorine Only participants were.[14] This means both that, in our trial, non-Info-tool participants are *more* exposed to other Info-tool participants than in AM, and that Info-tool participants are *less* exposed to other Info-tool participants. If there is a learning spillover from Info-tool participants onto non-Info-tool neighbors, our individual-level randomization would conceal Info-tool treatment effects; and if there is a learning spillover from Info-tool participants onto other Info-tool neighbors, the treatment would be less powerful in our setting.

Motivated by our theoretical framework of learning and the contextual differences between our study design and AM, we now test for learning spillovers from Info-tool participants onto their neighbors. We find that the Info-tool treatment has powerful spillovers, but these spillovers are only absorbed by other Info-tool participants. In other words, there is an Info-tool-to-Info-tool spillover.

First, we define our spillover measure. Because randomization was conducted at the individual level, individuals are randomly exposed to Info-tool neighbors. We consider somebody as being in the "spillover sample" if they have at least one Info-tool neighbor within an r-meter radius of where they live. First, we determine the number of Info-tool participants within each radius $r \in \{20(20)200\}$ meters. Next, we recenter these measures to purge our estimates of omitted variable bias that may arise from neighborhood density or other endogenous environmental factors, following Borusyak and Hull (2023). We construct $AnyT2_i^r$, a binary measure that indicates if the recentered number of Info-tool participants within radius r of individual i

[14] There were no Incentives or Pure Control groups in AM.

is greater than 0. We interpret this variable as indicating if participants were randomly more exposed to an Info-tool neighbor than they would be in expectation based on endogenous spatial factors. Following Egger et al. (2022), we select the R that minimizes the Schwarz Bayesian Information Criterion (BIC) for each $r \in \{20(20)200\}$ meters. Then, $\sum_{r=20}^{R} \theta_r(\overline{AnyT2_i^r})$ is our estimate of learning spillovers. For all of our specifications, the smallest radius (20 meters) minimizes the Schwarz BIC. The spillover and non-spillover samples are balanced on observable baseline variables, across the full sample (Appendix Table 34), and within each individual treatment group (Appendix Tables 35, 36, 37).

We test for the presence of learning spillover using the following specification:

$$\begin{aligned} Y_i =& \theta_0 + \theta_1 T1_i + \theta_2 T2_i + \theta_3 T3_i + \sum_{r=20}^{R} \theta_{4,r} AnyT2_i^r \\ &+ \sum_{r=20}^{R} \theta_{5,r} AnyT2_i^r \times T1_i + \sum_{r=20}^{R} \theta_{6,r} AnyT2_i^r \times T2_i + \sum_{r=20}^{R} \theta_{7,r} AnyT2_i^r \times T3_i \\ &+ Y_{i0} + X_{i0} + \gamma_{1,b} + \epsilon_i \end{aligned}$$

Then $\theta_{6,r}$ is our estimate of Info-tool to Info-tool spillovers. We use $\theta_{4,r}$, $\theta_{5,r}$ and $\theta_{7,r}$ to rule out spillovers from Info-tool onto other groups.

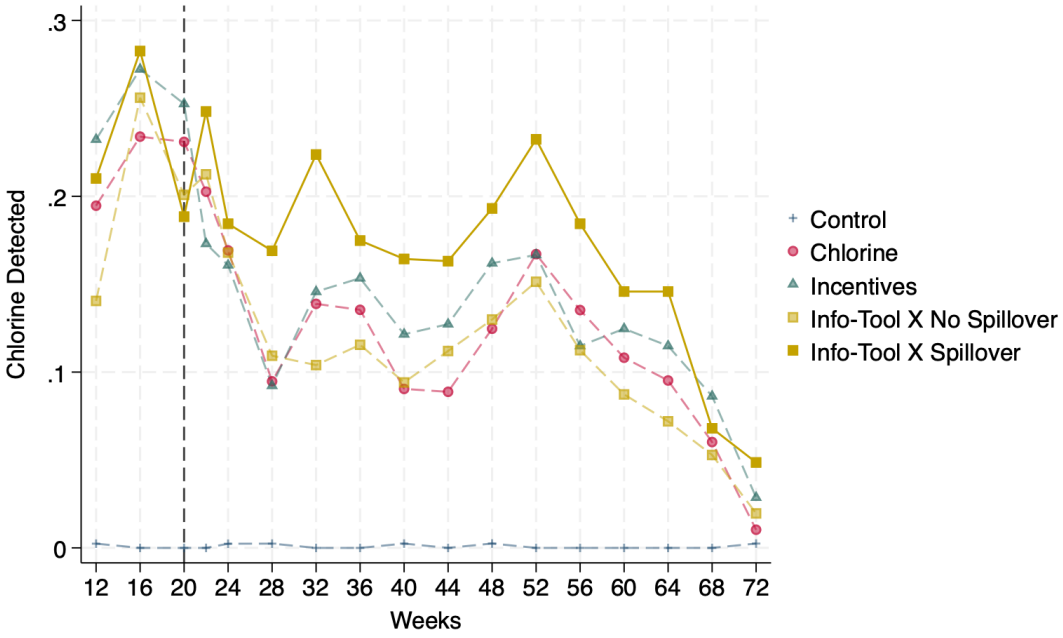


Figure 4: Raw Rates of Chlorine Detection
where "Spillover" is exposure to any Info-tool neighbor

While the presence of Info-tool participants within twenty meters does indeed increase chlorine detection, this is only true for *other Info-tool participants*. Figure 4 plots raw chlorination rates across these groups, while Table 1 and Figure 5 present the regression analogs. This pattern emerges in the medium-run (the quarter after behavioral treatments end), and persists through

Figure 5: Regression Coefficients of Chlorine Detection (Omitted: Pure Control)
where "Spillover" is exposure to any Info-tool neighbor

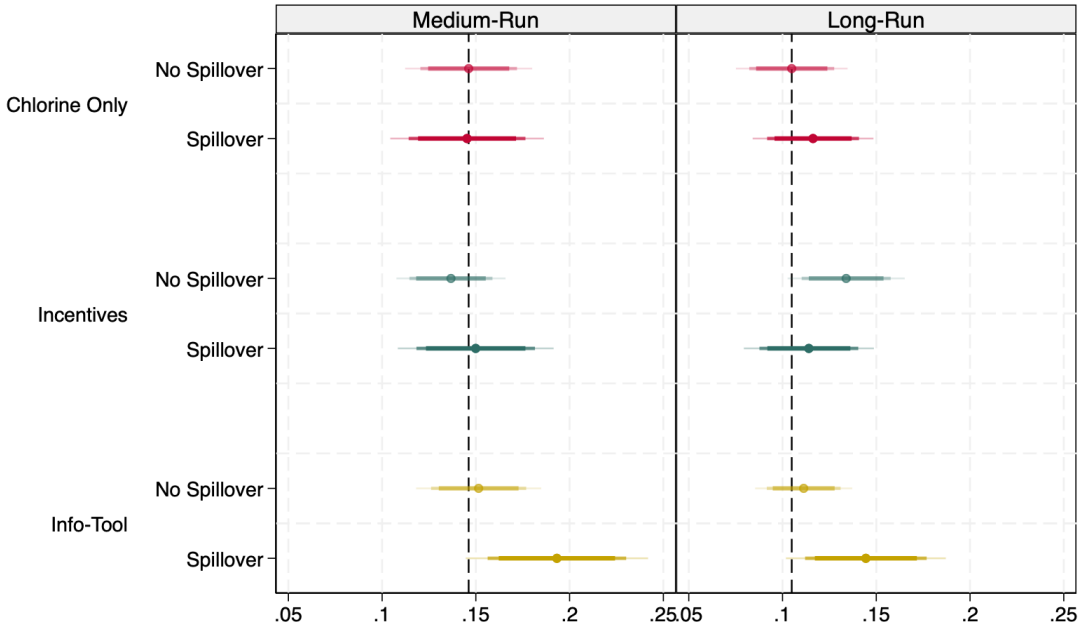


Table 1: Chlorine Detection (Household-Survey Panel)
where "Spillover" is exposure to any Info-tool neighbor

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × No Spillover	0.219***	(0.018)	0.146***	(0.013)	0.105***	(0.012)
Chlorine × Spillover	0.219***	(0.024)	0.145***	(0.016)	0.116***	(0.012)
Incentives × No Spillover	0.248***	(0.019)	0.137***	(0.011)	0.134***	(0.012)
Incentives × Spillover	0.259***	(0.024)	0.150***	(0.016)	0.114***	(0.013)
Info-Tool × No Spillover	0.202***	(0.017)	0.151***	(0.013)	0.111***	(0.010)
Info-Tool × Spillover	0.212***	(0.022)	0.193***	(0.019)	0.144***	(0.017)
Observations	4711		6354		14066	
P-values:						
Chlorine: No Spillover = Spillover	0.982		0.966		0.452	
Incentives: No Spillover = Spillover	0.705		0.491		0.228	
Info-Tool: No Spillover = Spillover	0.713		0.063		0.063	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

the end of the study. In the medium-run, Info-tool participants with another Info-tool participant neighbor chlorinate 27% more often than participants with no Info-tool neighbor, and the difference is statistically distinguishable at conventional levels. There are no meaningful

differences in use (statistically or in magnitude) by this spillover measure in either the Chlorine Only or Incentives groups. In the long-run, Info-tool participants with another Info-tool neighbor still chlorinate 30% more frequently. Again, there are no differences in the Chlorine Only or Incentives groups.

On aggregate across the study, Chlorine participants, Incentive participants, and Info-tool participants without another Info-tool participant in a twenty meter radius chlorinate 11.9 percent of the time – whereas Info-tool participants with at least one other Info-tool participant in a twenty-meter radius chlorinate 15 percent of the time. We conduct two tests for robustness. First, as a placebo test, we replicate our analysis to test for spillovers from Chlorine Only or Incentives participants onto their neighbors. There is no time period in which we find spillover treatment effects from Chlorine Only participants or Incentives participants onto any treatment group (Appendix Tables 17 and 18). Second, we reanalyze the AM data to test for within-treatment Info-Tool spillover treatment effects in Appendix Section 8.2, and find comparable effects.

5 Mechanisms

In our model, the Info-tool changes individuals' behavior, which teaches them that the Info-tool is a trustworthy signal. Then, Info-tool participants are primed to learn from other Info-tool health signals. This motivates us to first discuss potential mechanisms to explain why the Info-tool changes individual behavior. Next, we discuss potential mechanisms to explain why there is a complementarity in experiential and social learning.

5.1 Info-tool Individual Effects

Why does a simple pen-and-paper chart lead to changes in health behaviors? We posit that the Info-tool works by making health signals more salient, which leads beliefs to be more sensitive to positive health signals. The Info-tool provides individuals with information about their children's health, and health in the community, before and after chlorine distribution. Individuals might update their beliefs about the efficacy of chlorine, or they might update their beliefs about their children's risk profile.

Alternatively, the Info-tool might lead participants to adopt chlorine early on at a higher rate than they would without the Info-tool. Long-term adoption could then be explained by early adoption leading to habit formation, or intertemporal complementarities in chlorine use. We use the Incentives arm to rule out stories of chlorine use related to higher rates of early adoption.

5.1.1 Beliefs

We show three pieces of evidence that jointly suggest that belief-updating is the driving mechanism. Info-tool stimulates learning by drawing the participants' attention to differences in children's health, before and after chlorine use, as compared to the average among community

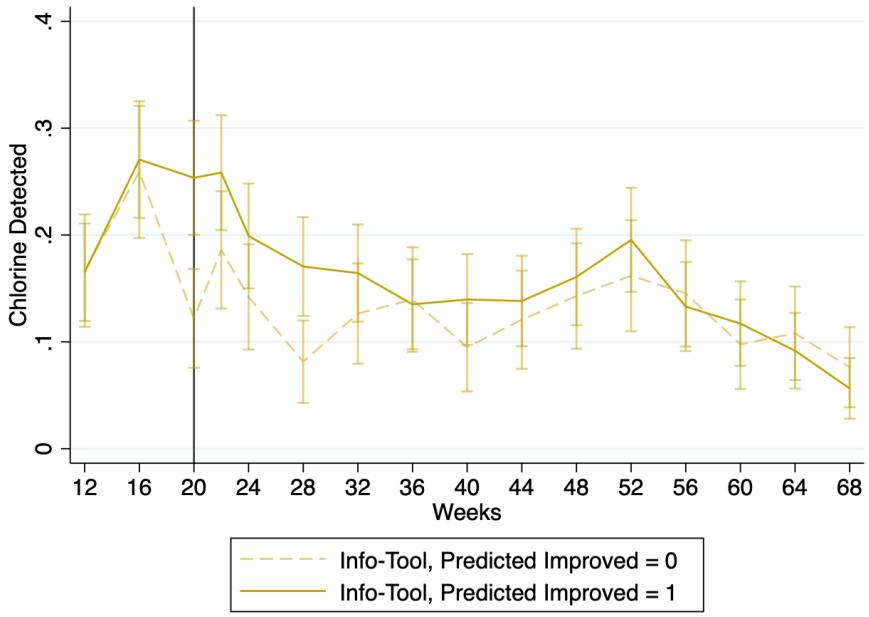


Figure 6: Info-tool Chlorine Detection by Predicted Health Improvement

members who do not use chlorine. While, on average, the Info-tool participants' diarrhea rates decrease relative to the benchmark, this decline is heterogeneous depending on the underlying health and other environmental factors. First, we show that Info-tool participants' chlorine use is heterogeneous by the sign on their health signals (the expected rate of decline in diarrhea). Next, we show that Info-tool participants have different stated beliefs than other groups: they report a different set of motivating reasons for using chlorine, and they report different beliefs about the types of people who are most knowledgeable about child health.

Responsiveness to Health Signals

First, we show that Info-tool participants who are predicted to experience improvements in their diarrhea rate during the time period where they are tracking their data – or, Info-tool participants who we expect to have a positive sign on the health signals they observe in their environment and place higher weights on – chlorinate at a higher rate immediately after the Info-tool treatment ends (Table 2 and Figure 6). This effect is very valuable in the short-run: predicted-improved Info-tool participants chlorinate 31% more often than not-predicted-improved Info-tool participants in the short-run (p-value = 0.029). In the medium- and long-run, this effect becomes smaller and noisier. However, predicted-improved and not-predicted-improved Info-tool participants' chlorination rates converge not because people in the predicted-improved sample are learning less, but because people in the not-predicted-improved sample are learning more from their Info-tool neighbors, as evidenced by Appendix graph 12.

Actual health will be correlated with treatment adoption. To understand how treatment responds to health signals that are exogenous to treatment status, we use a predicted measure of health improvement. First, we define "health improvement" as the number of days of diarrhea in the three months before chlorine adoption minus the number of days of diarrhea in the first

three months of chlorine adoption. The first three months of chlorine adoption is the relevant period of time to analyze health signals, because this is the time in which Info-tool participants receive salient information about their health. We consider someone to be "predicted improved" if the predicted diarrhea rate improvement is above the median.

We use lasso with the Pure Control sample to construct a measure of predicted health improvement (as defined above), using variables collected before treatment status was revealed to predict this measure (all variables from the baseline survey and the first half of the first follow-up survey). To avoid bias that can arise from endogenous stratification, we use the leave-one-out procedure proposed by Abadie et al. (2018).

To test for individual learning from one's own health signals in the Info-tool group, we use the following specification:

$$\begin{aligned} Y_i =& \theta_0 + \theta_1 T1_i + \theta_2 T2_i + \theta_3 T3_i + \theta_4 \hat{I}_i \\ &+ \theta_5 \hat{I}_i \times T1_i + \theta_6 \hat{I}_i \times T2_i + \theta_7 \hat{I}_i \times T3_i \\ &+ Y_{i0} + X_{i0} + \gamma_{1,b} + \epsilon_i \end{aligned}$$

where $\hat{I}_i = 1$ if an individual was predicted to improve (continuous predicted health improvement is above the median). Then θ_6 is our object of interest. We use θ_4, θ_5, and θ_7 to test if there are differences in long-term use between those predicted to improve and those not predicted to improve when the information is not made salient.

We find that Info-tool participants whose health was predicted to improve in the first three months of chlorine distribution use chlorine at a higher rate than Info-tool participants whose health was not predicted to improve. In the short run, predicted-improved Info-tool participants use chlorine 31% more often than Info-tool participants who are not predicted to improve. The two groups only diverge in their chlorine use in the third month, by which time they would have received and been able to respond to health signals that the Info-tool makes salient. There are no differences between predicted-improved and not-predicted-improved participants in any other treatment group in the short-run.

In the quarter immediately after behavioral treatments end, the predicted-improved Info-tool respondents continue using chlorine at a higher rate than not predicted improved Info-tool participants (23% higher rate of use, Table 2). However, we are not able to rule out these differences at traditional measures of statistical significance (p-value = 0.135). [15]

Stated Beliefs

Because the health signals that individuals observed are endogenous to personal characteris-

[15] In this quarter, differences in use by predicted-improvement in the Chlorine Only group also emerge. This could be explained by some learning that happens through observation without the Info-tool, or by selection. For example, perhaps the characteristics of people who are predicted to improve also inclines them more towards longer-term use without any learning. While these patterns fade in both groups in the long-run, the magnitude of the difference remains slightly larger in the Info-tool group.

Table 2: Chlorine Detection (Household-Survey Panel)
by predicted health improvement
Omitted group: Chlorine × Not Improved

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × Improved	0.033	(0.029)	0.035*	(0.021)	0.013	(0.016)
Incentives × Not Improved	0.045	(0.028)	0.002	(0.018)	0.022	(0.015)
Incentives × Improved	0.059**	(0.029)	0.019	(0.019)	0.027*	(0.016)
Info-Tool × Not Improved	-0.019	(0.026)	0.020	(0.019)	0.013	(0.015)
Info-Tool × Improved	0.034	(0.028)	0.053**	(0.022)	0.030**	(0.015)
Observations	3463		4689		10472	
P-values:						
Incentives: Not Improved = Improved	0.662		0.365		0.719	
Info-Tool: Not Improved = Improved	0.049		0.135		0.281	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). Improved is a binary indicator for if the participant's predicted improvement in health after the beginning of chlorine distribution was above the median. See Section 6 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

tics, we take these results as highly suggestive, but not conclusive, of Info-tool participants responding to health signals. Thus, we ask participants a series of questions at endline to elicit their beliefs about chlorine.

We ask respondents several questions to interrogate three ways in which the Info-tool could have changed participants' beliefs in a way that would lead them to use chlorine at higher rates: (1) Info-tool increases the probability that participants believe that chlorine is effective; (2) Info-tool increases the probability that participants believe that their disease environment poses a risk to children's health; and (3) Info-tool operates on a sub-conscious level, such as by causing women to feel more engaged in their children's health. We find no evidence of mechanisms (1) or (2) by endline, and some evidence of mechanism (3).[16] It is possible that channel (1) and (2) beliefs updated immediately after the Info-tool treatment but then converged with time, with social learning superseding individual learning, such that we do not observe differences in beliefs at endline. Furthermore, all households reported high levels of belief in the need for chlorine and its efficacy, which is potentially driven by experimenter demand.

[16] At endline, we find no differences across treatment groups in beliefs about how many child-days of diarrhea a hypothetical household would experience after using chlorine relative to before using chlorine; no differences in the number of child-days of diarrhea participants believe their household would experience in the absence of chlorine, in either the summer or winter seasons; no differences in reporting "unclean water" as a primary cause of illness in the household prior to using chlorine tablets (unprompted); and no differences in how high they rank unclean water as a primary source of child illness prior to using chlorine tablets, relative to other potential sources of illness (prompted).

Although at endline households in all treatment groups non-differentially report that they believe chlorine to be an effective technology, and unclean water to be a source of sickness, Info-tool participants are more likely to report that attempting to achieve health for the family is the *motivating reason* for why they use chlorine.[17] Info-tool households rated "to attain a certain level of health for my family" as a stronger motivating factor for using chlorine, controlling for their ratings of other potential motivating factors. Although almost everyone believes that chlorine is an effective technology, and that they live in a risky disease environment where that technology is needed, the Info-tool gives people the motivation to act upon this information. Thus, our evidence suggests that Info-tool might operate on a sub-conscious level.

Why does Info-tool motivate people to act upon something that, at the surface, everybody knows? To understand if trust is the key mechanism, we reminded all participants about the three treatment groups in our study: people receiving chlorine only, people receiving chlorine as well as gifts in exchange for using chlorine, and people receiving chlorine as well as a pen and paper chart with which they tracked their children's diarrhea rates. We asked participants to imagine a neighbor in each of these three groups and asked: who among these three neighbors do you think knows more about children's health? Participants in the Info-tool group were 93% more likely to list the Info-tool respondent than Pure Control households (who were the second most likely to consider Info-tool respondents as the most knowledgeable), controlling for a binary indicator for listing one's own treatment group. While all participants are able to repeat the information that they have learned throughout the study – that their children are exposed to unclean water which makes them sick, and chlorine tablets can purify their water and make their children healthier – the Info-tool makes this information trust-worthy and actionable.

5.1.2 Early Adoption

Our Incentives arm builds on Hussam et al. (2022), in which persistence in handwashing behavior is engendered through exogenous short-run financial incentives to handwash in West Bengal.[18] Our Incentives arm serves as a parallel intervention: households are incentivized to chlorinate their water daily, receiving tokens that can be exchanged for household goods for each day of empty chlorine wrappers they present to enumerators. Should water chlorination, which is a repeated act performed at the same time and place each day, indeed be habit

[17] At endline, we asked caregivers a series of questions to understand their motivation behind using chlorine (question formulation adapted from Tremblay et al. (2009)). We read several statements describing reasons why someone might use chlorine, and asked women to rate on a scale from 1 to 7 how true that statement was for them during the times when they have ever used chlorine. The statements described a motivation related to health ("Because using treatments like chlorine tablets help me to attain a certain level of health for my family"), habit formation ("Because it has become a fundamental part of my routine"), income generation ("Because it could allow me to earn money"), and intrinsic motivation ("Because I derive pleasure from trying new things", "Because I want to be very good at taking care of my family, otherwise I would be very disappointed", and "For the satisfaction I experience when I am successful at doing difficult tasks").

[18] Hussam et al. (2022) finds that households who experienced larger health improvements (either across weeks or in aggregate) from the intervention did not exhibit differentially greater persistence in handwashing, and therefore attribute the long run behavior change to habit formation rather than learning.

Table 3: Endline Stated Beliefs and Motivations
Omitted group: Chlorine Only

	(1)		(2)	
	Info-Tool Knows Most		Motivation: Health	
Incentives	0.068**	(0.028)	0.088	(0.062)
Info-Tool	0.297***	(0.027)	0.191***	(0.062)
Observations	1080		1049	
P-values:				
Incentives = Info-Tool	0.000		0.095	

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

forming, then this exogenous increase in initial consumption stock via financial incentives will activate the intertemporal complementarities in use, thereby generating long-run use even after incentives are withdrawn. Alternatively, a higher rate of use may lead households to acquire more signals in a shorter period of time than in the other two groups, leading participants to learn from their own experience how the tablets improve their children's health. We do not attempt to distinguish between habit formation and learning through rapid accumulation of health signals in the Incentives arm. Instead, we use the Incentives arm to rule out either of these two mechanisms relative to the role of salient signals.

We observe that Incentives arm households use chlorine at a higher rate than either of the Chlorine or Into-Tool arms during the behavioral treatment period. In order to ensure that Incentives households are actually chlorinating at a higher rate, rather than chlorinating just on the days where the enumerator is present, we do not directly incentivize the water chlorination test (household incentives are tied to presenting empty chlorine tablet wrappers), and we conduct unscheduled audit tests. In these audit tests, we detected chlorine at a higher rate in Incentives than in Info-tool. Furthermore, we detected chlorine at a much higher rate during the audit visits than during the regularly scheduled visits (45% detection rate in the Incentives arm during audit visits), further suggesting that Incentives participants were not more likely to use chlorine when they expected that we would come test their water than at other times.

Since the Incentives arm built up a higher stock of chlorine use in the first three months of chlorine distribution, any theory of sustained behavioral change relating to early adoption will be born out in the long term in the Incentives arm. In the three months immediately after the behavioral interventions ended, the Incentives arm reverted to the same rate of use as the

Chlorine Only group, while the Info-tool group chlorinated 16% more often in the medium-run, the quarter after behavioral interventions ended, than the Incentives group ($p = 0.095$; we cannot, however, rule out equality between the Info-Tool group and the Chlorine Only group during this time period). Thus, the mechanism that helped the Info-tool sustain chlorine adoption in the months proceeding the behavioral interventions cannot be explained by habit formation or any story related to early adoption.[19] For a more detailed discussion on the roles of learning and habit formation, see Appendix 8.6.1.

5.2 Info-tool-to-Info-tool Spillovers

Why might spillovers only flow from one Info-tool participant to another? We hypothesize that Info-tool participants are able to learn from one another more readily than from people who do not use the Info-tool because they have experience with the same learning technology, and therefore have higher levels of trust in information – and apply heavier weights to the health signals – sent from other Info-tool participants. Thus, Info-tool participants with an Info-tool neighbor should update their beliefs to a greater extent than Info-tool participants without Info-tool neighbors.

First, we show that the spillover is completely explained by Info-tool neighbors whose diarrhea rate was predicted to improve, indicating that Info-tool participants only learn from their Info-tool neighbors who received positive signals. We will then provide evidence to suggest that Info-tool participants' stated beliefs are more likely change when they have an Info-tool neighbor, and evidence to show that Info-tool changes the type of information that individuals find trustworthy. In Appendix Section 8.8, we consider three competing explanations, none of which we find any evidence for: (1) Info-tool encourages interaction between Info-tool participants, who then become more likely to mimic one another's behaviors, (2) Info-tool encourages more conversations between participants, or different conversation content, rather than differential beliefs about the content of those conversations, and (3) Info-tool changes participants beliefs about the social norms around water chlorinating in their community.

Responsiveness to Health Signals

If Info-tool participants are able to learn from their Info-tool neighbors information about the efficacy of chlorine to a greater degree than other participants are, then the spillover should be explained by cases where information shared between two participants is likely about chlorine being highly efficacious. To test if the spillover treatment effect can be explained by positive signals from neighboring Info-tool participants we limit $AnyT2_i^r$ to only be equal to 1 if there

[19] Average rates of chlorine use in the Incentives and Info-tool groups converge at 32 weeks. Notably, the Info-Tool participants in the Spillover sample continue to chlorinate at a higher rate. This could be due to forgetting in the Info-tool arm without an Info-Tool neighbor, or it could be due to some catching-up in the Incentives group (cyclical habit formation, learning over a different time horizon in the absence of salient signals, etc.). We do not attempt to disentangle potential reasons why Incentives and Info-tool eventually converge on average because, ultimately, the higher rate of use that Info-tool experienced, though temporary, resulted in greater child health improvements. Furthermore, the Info-Tool group in the Spillover sample continued to chlorinate at a higher rate after the two groups converged on average.

is any Info-tool participant within radius r of participant i *who was predicted to see their health improve*. When we change the definition of the spillover from *any Info-tool neighbor* to *any predicted-improved Info-tool neighbor* the spillover estimates remain almost unchanged (comparing estimates in Table 4 with esimates in Table 1, estimates are 88% of the magnitudes of the spillover treatment effect with the broader definition of spillovers).[20]

It is possible that participants who are predicted to improve have characteristics that make them more likely to influence their neighbors, regardless of treatment status or whether they share information that can be interpreted as a positive signal about chlorine. To rule out this possibility, we conduct a placebo test where we test heterogeneity by the total number of predicted-improved treated neighbors. In this placebo test, there are no differences in use by exposure to this spillover measure in the short-run or medium-run, in any group (Appendix Table 19).

Stated Beliefs

To corroborate these findings, we show that Info-tool participants are *all* more likely to believe that another Info-tool participant is more trust-worthy with regards to child health than another study participant, but that only Info-tool participants who randomly have another Info-tool participant neighbor are able to utilize this potential trust-worthy information source.

Indeed, we find that Info-tool participants were 93% more likely to list an Info-tool respondent as the most knowledgeable about child health, controlling for a binary indicator for listing one's own treatment group. This is not explained by participants with access to an Info-tool spillover – in fact, participants *without* access to an Info-tool spillover are more likely than their Info-tool spillover counterparts to rate other Info-tool participants as the most knowledgeable (104% increase, $p < 0.05$) – indicating that personal experience with the Info-tool itself is enough to change participants' trust in Info-tool as an information source. Then, Info-tool participants are uniquely primed to trust health signals from other Info-tool participants. Only when an individual has access to these signals, by being randomly assigned an Info-tool neighbor, are their beliefs able to further update.

Info-tool participants rated attaining health as a greater motivating reason for using chlorine than both Incentives and Chlorine participants, relative to other possible motivating reasons.

[20]In the long-run, a *reverse* pattern emerges in the Incentives group, where people without access to a predicted-improved treated neighbor use chlorine at a higher rate. While we do not have a prediction for this result, these participants were more likely to report at endline that they had shared their tablets with others. Incentives participants in the spillover sample are in an environment where many others have a higher demand for chlorine than they do. We do not have any evidence as to why the Incentives group is more responsive to this higher aggregate demand than the Chlorine Only group, but we speculate that it could be due to two possibilities. Firstly, Incentives participants are used to receiving income for their chlorine tablets, and so they might have sold their tablets. Only one person reports that they sold chlorine tablets at endline (the rest say they shared them for free), but it could be that participants do not wish to admit they sold tablets or forgot. Secondly, Incentives participants may wish to share chlorine tablets out of equity concerns. Although the participants not in the Incentives group did receive gifts via a lottery, it could be that a feeling of inequity arises between the Incentives group and other households because they were chosen for a group where receipt of gifts was within the caregivers' control and explicitly linked to the treatment assignment.

Table 4: Chlorine Detection (Household-Survey Panel)
where "Spillover" is exposure to any Info-tool neighbor predicted to improve
Omitted group: Chlorine × No Spillover

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × Spillover	-0.031	(0.033)	0.013	(0.023)	0.032*	(0.017)
Incentives × No Spillover	0.032	(0.024)	-0.008	(0.015)	0.033**	(0.014)
Incentives × Spillover	0.013	(0.031)	0.005	(0.024)	0.006	(0.016)
Info-Tool × No Spillover	-0.017	(0.022)	0.010	(0.017)	0.015	(0.012)
Info-Tool × Spillover	-0.018	(0.029)	0.047**	(0.024)	0.044**	(0.019)
Observations	3463		4689		10472	
P-values:						
Incentives: No Spillover = Spillover	0.541		0.569		0.115	
Info-Tool: No Spillover = Spillover	0.971		0.123		0.115	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) whose health was predicted to improve than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined. See Section 6 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Endline Stated Beliefs and Motivations
where "Spillover" is exposure to any Info-tool neighbor
Omitted group: Chlorine × No Spillover

	(1)		(2)	
	Info-Tool Knows Most		Motivation: Health	
Chlorine × Spillover	-0.097**	(0.046)	-0.096	(0.093)
Incentives × No Spillover	0.053	(0.034)	0.053	(0.077)
Incentives × Spillover	0.000	(.)	0.055	(0.090)
Info-Tool × No Spillover	0.316***	(0.035)	0.114	(0.078)
Info-Tool × Spillover	0.168***	(0.045)	0.229**	(0.091)
Observations	1080		1049	
P-values:				
Incentives × No Spillover = Incentives × Spillover	0.124		0.982	
Info-Tool × No Spillover = Info-Tool × Spillover	0.010		0.209	

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

We consider motivation as a component of beliefs that participants have about the efficacy of chlorine. Our model predicts that beliefs should update more among Info-tool participants who receive signals about the efficacy of chlorine from other Info-tool participants, but that participants in other groups should not respond to exposure to Info-tool neighbors. Indeed, we see that the treatment effect on the motivation score that participants give "health" is twice as large in the spillover sample in the Info-tool group, as compared to Info-tool participants who are not in the spillover sample. These treatment effects are not statistically distinguishable and therefore are not conclusive on their own, but it is encouraging that the patterns we see are aligned with our model predictions. Chlorine Only and Incentives participants do not rate health as a motivating factor differentially by access to spillovers from the Info-tool group.

6 Child Health

6.1 Diarrhea

We find substantial impacts of water chlorination on child health. Diarrhea rates over time offer a lens into how the various interventions impact children's health dynamically over the course of the experiment. Treatment group participants report 34% (Chlorine Only), 28% (Info-Tool), and 39% (Incentives) reductions in diarrhea in the short-run relative to the Pure Control group

(Appendix Table 16). These relative reductions mirror relative rates of chlorine use during this time period. Reductions in diarrhea rates continue to mirror relative levels of chlorine use in the medium- and long-run, with the Incentives groups experiencing the smallest reductions in diarrhea prevalence (12% reduction in the medium-run and 19% reduction in the long-run, neither statistically distinguishable from the Pure Control group); and the Info-Tool group experiencing the largest reductions in diarrhea prevalence (39% reduction in the medium-run, and 34% reduction in the long-run, both statistically distinguishable from the Pure Control group at conventional levels of significance). More details on measurement and results on child diarrhea are reported in Appendix Section 8.9.

As a self-reported measure, children's diarrhea rate is subject to respondent bias; that it is directly connected to the substance of the Info-tool also leaves room for concerns of endogeneity in outcome reporting. As an objective and therefore our preferred measure of health, we measure child anthropometrics at endline in an index that combines four standardized measures: weight-for-age (WAZ), weight-for-height (WHZ), height-for-age (HAZ), and MUAC-for-age (mid-upper arm circumference-for-age) z-scores.

6.2 Child Anthropometrics

6.2.1 ITT Results: Any Treatment

To analyze the impact of our treatment on child anthropometrics, we use the following specification:

$$Y_c, h = \beta_0 + \beta_1 T_h + X_{h0} + \gamma_b + \epsilon_{c,h}$$

where $Y_{c,h}$ is health outcome Y for child c in household h, T_h is treatment status of household h, and X_{h0} are household-level baseline control variables. We cluster standard errors at the household level for all child-level specifications.

Table 6 reports the ITT effects of *any* treatment on the various measures of child anthropometry, whereas Table 8 reports the treatment-specific estimates. Following Anderson (2008), we create a summary index combining height-for-age, weight-for-height, weight-for-age, and MUAC-for-age z-scores. For households in any of the three treatment arms, the index increases by over 7% of a standard deviation ($p < 0.05$).[21]

6.2.2 IV Results: Effects of Water Purification

Treatment compliance was not perfect – we only were able to detect chlorine in the water of our treated participants 13% of the time across the whole study, and 47% of treatment participants said that they were using chlorine to treat their water at endline. To understand the effect

[21] For households receiving any of the three treatments, the estimates for height-for-age and weight-for-height are close to zero whereas those for weight-for-age and MUAC-for-age are 0.10 and 0.05 (albeit not significant at conventional levels).

Table 6: ITT Impacts on Endline Child Health (Pooled Treatments)

	(1) Index (Anthro-pometry)	(2) Height-for-Age	(3) Weight-for-Height	(4) Weight-for-Age	(5) MUAC-for-Age
Any Treatment	0.074**	0.015	-0.003	0.114*	0.055
	(0.034)	(0.066)	(0.081)	(0.065)	(0.054)
Observations	2616	2371	2439	2492	1954
Endline Control Mean	-0.019	-1.773	-0.291	-1.407	-1.453

Standard errors in parentheses
Any treatment is an indicator for not being in the Control group (ever receiving chlorine). Child-level cross-section of the endline survey. Standard errors are clustered at the household level. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, neighborhood block fixed effects, and lasso-selected controls. Indeces are constructed using following Anderson (2008).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: IV Impacts on Endline Child Health (Instrument: Any Chlorine Treatment)

	(1) Index (Anthro-pometry)	(2) Height-for-Age	(3) Weight-for-Height	(4) Weight-for-Age	(5) MUAC-for-Age
Boils, Bleaches, or Chlorinates Water	0.243**	0.048	-0.011	0.380*	0.201
	(0.113)	(0.208)	(0.261)	(0.213)	(0.184)
Observations	2616	2371	2439	2492	1954
Endline Control Mean	-0.019	-1.773	-0.291	-1.407	-1.453
Weak-IV robust F statistic	119.09	126.68	120.85	121.56	95.24
C-statistic p-value	0.025	0.499	0.410	0.127	0.054

Standard errors in parentheses
Any treatment (i.e., received chlorine) is an instrument for if the respondent reported at endline that she boils, bleaches, or chlorinates her water. Child-level cross-section of the endline survey. Standard errors are clustered at the household level. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, and neighborhood block fixed effects. Indeces are constructed using following Anderson (2008).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of purified water on child health, we use an instrumental variables specification. We consider compliers to be any individual who at endline reported that they boil, bleach, or chlorinate their water. While this is a self-reported measure of compliance, we think this is a better measure of purified water for two reasons: (1) our objective measure of chlorine residual detection is an underestimate and changes with time, and (2) participants may have learned about the negative effects of impure water through the study and changed their water treatment behavior in some way *other* than chlorinating their water. Indeed, we find that Control participants are more likely to boil their water at endline. If we consider treatment assignment to be an instrument for using chlorine to purify water, then the exclusion restriction does not hold because the treatment also affects participant's use of other water-treatment methods. If we instead consider treatment to be an instrument for using any effective water purification technique, which includes boiling, then higher rates of boiling in the Control group is not a threat to the exclusion restriction and the Control participants who boil their water can be considered always-takers.

Table 7 reports the results of an IV exercise where we use the assignment to any treatment as an instrument for whether the household reported to be boiling, bleaching, or chlorinating water at endline. Forty-eight percent of the sample report using one of these effective water purification technologies at endline. We find that effective water purification leads to a 0.24 SD increase in the index of child anthropometrics ($p < 0.05$).

6.3 Heterogeneity by Treatment

6.3.1 ITT Results

Table 8 shows that the improvements in anthropometric outcomes are consistently higher for households in the Info-tool group. Furthermore, the treatment effects are more than twice the magnitude among Info-tool households who have another Info-tool neighbor, relative to Info-tool households without access to this spillover (Table 9). Comparing only the spillover-sample Info-tool participants with the rest of the sample, there was a 0.12 SD increase in the index of child anthropometry.

6.3.2 IV Results

The IV results may also be heterogeneous with respect to the specific treatment group and access to an Info-tool neighbor. Our instrumented measure (self-reported use of boiling, bleach, or chlorine for water purification) is self-reported and is an extensive margin measure of using an effective water purification method, which does not imply perfect compliance.[22] Furthermore, even with perfect compliance, individuals who use effective water purification methods at endline may have purified their water at different rates throughout the study; since health is a stock, this historical use matters. However, we cannot use a cumulative measure, such as the number of visits in which we detected chlorine in participants' water, because these mea-

[22] Among the people who reported that they chlorinate their water at endline, we only detected chlorine in the water of 3.8%.

Table 8: ITT Impacts on Endline Child Health (Separated by Treatment)

	(1) Index (Anthro-pometry)	(2) Height-for-Age	(3) Weight-for-Height	(4) Weight-for-Age	(5) MUAC-for-Age
Chlorine-Only	0.081**	0.019	0.009	0.082	0.026
	(0.039)	(0.084)	(0.099)	(0.081)	(0.064)
Incentives	0.032	0.021	-0.095	0.069	0.062
	(0.042)	(0.078)	(0.101)	(0.079)	(0.065)
Info-Tool	0.108***	0.006	0.073	0.189**	0.080
	(0.039)	(0.080)	(0.101)	(0.080)	(0.066)
Observations	2616	2371	2439	2492	1954
Endline Control Mean	-0.019	-1.773	-0.291	-1.407	-1.453
P-values:					
Chlorine = Incentives	0.191	0.984	0.317	0.872	0.567
Chlorine = Info-Tool	0.453	0.867	0.525	0.192	0.387
Incentives = Info-Tool	0.045	0.840	0.108	0.136	0.768

Standard errors in parentheses
Child-level cross-section of the endline survey. Standard errors are clustered at the household level. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, neighborhood block fixed effects, and lasso-selected controls. Indeces are constructed using following Anderson (2008).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Child Index of Anthropometry (ITT): where "Spillover" is exposure to any Info-tool neighbor

	(1) Ommitted Group: Pure Control		(2) Ommitted Group: Everyone Else	
Chlorine × No Spillover	0.097**	(0.045)		
Chlorine × Spillover	0.056	(0.056)		
Incentives × No Spillover	0.029	(0.045)		
Incentives × Spillover	0.020	(0.054)		
Info-Tool × No Spillover	0.085*	(0.044)		
Info-Tool × Spillover	0.163***	(0.057)	0.120**	(0.060)
Observations	2616		2616	
P-values:				
Chlorine: No Spillover = Spillover	0.494			
Incentives: No Spillover = Spillover	0.876			
Info-Tool: No Spillover = Spillover	0.202			

Standard errors in parentheses
Child-level cross-section of the endline survey. Standard errors are clustered at the household level. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, neighborhood block fixed effects, and lasso-selected controls. The index is constructed using following Anderson (2008) from the following variables: WAZ, WHZ, HAZ, and Muac-for-age. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: IV Impacts on Endline Child Health:
where "Spillover" is exposure to any Info-tool neighbor

	(1) Instrument: Any Treatment	(2) Instrument: Info-Tool	(3) Instrument: Info-Tool × Spillover
Boils, Bleaches, or Chlorinates Water	0.243**	0.332***	0.514***
	(0.114)	(0.126)	(0.186)
Observations	2616	2616	2616
Endline Control Mean	-0.017	0.023	0.031
First-Stage Coefficient	0.310	0.328	0.310
Weak-IV robust F statistic	119.090	85.947	34.817
C-statistic p-value	0.025	0.005	0.002

Standard errors in parentheses
Child-level cross-section of the endline survey. Standard errors are clustered at the household level. Any treatment (i.e., received chlorine) is an instrument for if the respondent reported at endline that she boils, bleaches, or chlorinates her water. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, neighborhood block fixed effects, and lasso-selected controls. The index is constructed using following Anderson (2008) from the following variables: WAZ, WHZ, HAZ, and Muac-for-age. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

sures will not be excludable if the treatments changed water purification methods aside from chlorinating, for which we don't have cumulative measures.

The Info-tool treatment, and especially being in the Info-tool group with a neighbor in the Info-tool group, led to a higher rate of cumulative use of chlorine throughout the study. Therefore, it is likely that someone in these groups who reports purifying her water at endline had a higher rate of historical use than other treated participants, and that the effect of being someone who reports purifying her water at endline has a different effect on child health. To understand if there is heterogeneity in the causal effect of water purification on child health, we test three sets of compliers by varying our instrument.

Column (1) of Table 10 replicates column (1) of Table 7: assignment to any treatment group is the instrument for using an effective water purification method, and the compliers are any participants in the Chlorine Only, Incentives, or Info-tool group who were induced to report using an effective water purification method due to random assignment into any treatment group. In Table 10 column (2), the instrument is an indicator for being in the Info-tool group, and the compliers are individuals who were induced to report using an effective water purification method due to random assignment into the Info-tool group (the regression controls for being in the Incentives or Chlorine-Only groups). The estimate is 37% higher (0.332 SD increase in child anthropometrics), suggesting that the higher intensity of chlorine use that we detect among Info-tool throughout the study period translates into greater improvements in child health. Finally, in Table 10 column (3), the instrument is an indicator for being in the Info-tool group *and* having random exposure to an Info-tool neighbor, and the compliers are individuals who were induced to report using an effective water purification method due to

Table 11: Benchmarking Child Health Estimates

Intervention	Paper	HAZ	WHZ	WAZ	Muac-for-age
Chlorine: Average	Table 7	0.048	-0.011	0.380*	0.201
Chlorine: Saturated Info-tool	Appendix Table 32	-0.096	0.604	0.807**	-0.424
Handwashing	Hussam et al. (2022)	0.272	–	0.203	0.078
Hygiene	Bennett et al. (2018)	0.290	–	0.270	–
Nutrient Supplements	Sazawal et al. (2013)	0.180	–	0.030	–
	Soofi et al. (2022)	0.290	0.050	0.260	–
ECD	Bos et al. (2024)	-0.024	0.230	0.137	–

randomly being in the Info-tool group and having an Info-tool neighbor (the regression controls for being in the Incentives or Chlorine-Only groups, and for being in the spillover sample). The estimate of water purification on child health is 112% higher than the any-chlorine-group treatment effect, and 55% higher than the average Info-tool treatment effect (0.514 SD increase in child anthropometrics).[23]

These heterogeneous results further point towards full saturation of the Info-tool treatment as the optimal policy. Comparing our results on anthropometrics with results from other recent programs in South Asia, our results on weight-for-age are significantly larger than those from handwashing, hygiene, nutrient supplements, or early childhood education programs (Table 11). We do not see any results on height-for-age, unlike the handwashing, hygiene, and nutrient supplement programs.

7 Conclusion

We offer three policy takeaways from this study. The first directly addresses the question of scale. When experiential and social learning are complementary, there is high added value in saturating a learning or information treatment. It is important that policymakers not assume that these interventions will work similarly if disseminated sparsely within a population, for example, by treating or seeding knowledge among a specific sub-population.

Second, we add to the literature documenting the necessity of trust in information that individuals receive about health in order for it to impact behavior, a phenomenon with specific implications for historically marginalized communities. We operate in a setting where people lack formal consumer protection and have recent experience with insincere and politicized health campaigns (Martinez-Bravo and Stegmann, 2022).[24] This adds to a body of research showing that, while the identity of someone delivering information is important for generating trust, *how* they deliver that information is likewise critical. As outsiders in these communities, researchers may not have the relevant knowledge and skills to craft the messages that will be

[23]Table 32 reports the corresponding estimates for the four components of the index.

[24]It is important to note that information coming from neighbors was effective above and beyond the information coming from the Community Health Workers, most of whom live in the same community as the respondents.

most effective. Allowing community members to experience and learn for themselves, and then craft their own messages, might be a more effective information campaign strategy.

Third, we add to a small but growing body of evidence that the underlying mechanism behind behavioral change or technological adoption is more complex than a mere shift in explicit knowledge about the returns to the behavior or technology (Hussam et al., 2022; Conlon et al., 2022); rather, in our context, learning appears to be most effective when it is action-based and similarly experienced by others in one's network.

Our results speak to public health programs across a wide geographic space. The take-up of water purification products is low in many developing countries. In Pakistan, only 0.3% of the population report usage of chlorine tablets, with 7.1% adopting any purification technology (Pakistan DHS 2017-18). The cost and availability of such tablets does not appear to be the only barrier: free provision induces only partial usage. For example, free delivery of chlorine solution in Kenya yielded a usage rate of only 34% (Dupas et al., 2016). Many features of the study context are characteristic of low- and middle-income countries across South Asia and Sub-Saharan Africa (Supply and Programme, 2014). Inadequate public WASH infrastructure and low use of cheap point-of-use water purification technology, such as chlorine tablets, lead to environments with water contamination and high diarrhea prevalence. Full subsidization of these technologies has not been successful in bringing about substantive increases in take-up (Akram and Mendelsohn (2021) in Pakistan; Dupas et al. (2016) in Kenya). The materials that we use in our intervention are reasonable to use outside of an experiment and in other contexts. The Info-tool is a simple and cheap pencil-and-paper intervention that low-literacy-and-numeracy individuals can easily use. Community health workers are a common feature of health systems across low- and middle-income countries, and our field protocols fit into the typical health worker workflow; by integrating into existing workflows, this intervention does not impose added human resource burdens on institutions working in this space. As such, this learning tool is potentially fairly easily scalable.

Aside from point-of-use chlorination for water purification, for what types of technologies might there be a complementarity between experiential and social learning? Our model suggests that the best candidates are technologies that are costly to adopt, require persistent use, and produce observable yet noisy treatment effects. Technologies that require persistent use provide users with many opportunities for adoption and observation. However, with sufficiently noisy treatment effects and sufficiently high costs, experiential learning may need reinforcement. Examples of health technologies that meet these criteria include: sanitation (e.g., hand-washing and proper treatment of food), mental health care practices (e.g., therapy and meditation), lifestyle changes (e.g., diet and exercise), and annual vaccinations (e.g., influenza and COVID-19). Examples of non-health technologies that meet these criteria include: agricultural practices, childcare practices, and learning or study habits.[25] A promising avenue for future research is to

[25]List and Uchida (2024) finds within-treatment (but no cross-treatment) spillover effects of early childhood education on cognition, driven by social network effects. While they do not specify what activities or behaviors drive these cognitive gains, it is conceivable that students learn behaviors that promote cognitive development

test for within-treatment spillover effects of programs encouraging take-up of technologies that, according to our criteria, are candidates for experiential and social learning complementarities.

in preschool, which the students reinforce in one another if they are in classrooms together in kindergarten and grade school.

References

Abadie, Alberto, Matthew M. Chingos, and Martin R. West, "Endogenous Stratification in Randomized Experiments," *Review of Economics and Statistics*, 2018, *C* (4), 567–580.

Aggarwal, Shilpa, Rebecca Dizon-Ross, and Ariel D. Zucker, "Incentivizing behavioral change: The role of time preferences," *NBER Working Paper No. w27079*, 2020.

Akram, A. A. and R. Mendelsohn, "Diaries to Increase the Adoption of Chlorine Tablets for Water Purification by Poor Households," *Water Economics and Policy (WEP)*, 2021, *7* (2), 1–34.

Allcott, Hunt and Todd Rogers, "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation," *American Economic Review*, 2014, *104* (10), 3003–3037.

Alpízar, Francisco, Maria Bernedo Del Carpio, Paul J. Ferraro, and Ben S. Meiselman, "Exposure-enhanced goods and technology disadoption," *Working Paper.*, 2022.

Alsan, Marcella and Marianne Wanamaker, "Tuskegee and the Health of Black Men," *The Quarterly Journal of Economics*, 2018, *133* (1), 407–455.

_ **and Sarah Eichmeyer**, "Experimental Evidence on the Effectiveness of Nonexperts for Improving Vaccine Demand," *American Economic Journal: Economic Policy*, 2024, *16* (1), 394–414.

Anderson, Michael L., "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Arrow, Kenneth, "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 1962, *29* (3), 155–173.

Banerjee, Abhijit, Abhijit Chowdhury, Jishnu Das, Jeffrey Hammer, Reshmaan Hussam, and Aakash Mohpal, "The Market for Healthcare in Low Income Countries," *Working Paper*, 2023.

_ **, Arun G. Chandrasekhar, Esther Duflo, and Matthew O Jackson**, "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials," *The Review of Economic Studies*, 2019, *86* (6), 2453–2490.

Becker, Gary S and Kevin M Murphy, "A theory of rational addiction," *Journal of political Economy*, 1988, *96* (4), 675–700.

Bennett, Daniel, Asjad Naqvi, and Wolf-Peter Schmidt, "Learning, Hygiene and Traditional Medicine," *The Economic Journal*, 2018, *128* (612), F545–74.

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 2023, *91*, 2155–2185.

Bos, Johannes M., Abu S. Shonchoy, Saravana Ravindran, and Akib Khan, "Early childhood human capital formation at scale," *Journal of Public Economics*, 2024, *231*.

Calónico, Sebastian, Rafael Di Tella, and Juan Cruz Lopez del Valle, "The Political Economy of a "Miracle Cure": The Case of Nebulized Ibuprofen and its Diffusion in Argentina," *NBER Working Paper 31781*, 2023.

Caro-Burnett, Johan, Judith A. Chevalier, and Ahmed Mushfiq Mobarak, ""Is Habit a Powerful Policy Instrument to Induce Prosocial Behavioral Change?," *Cowles Foundation Discussion Papers*, 2021, *2600*.

Celhay, Pablo A., Paul J. Gertler, Paula Giovagnoli, and Christel Vermeersch, "Long-Term Effects of Temporary Incentives on Productivity in Medical Care Clinics," *American Economic Journal: Applied Economics*, 2015, *11* (3), 92–127.

Chen, Yiqun, Petra Persson, and Maria Polyakova, "The Roots of Health Inequality and the Value of Intrafamily Expertise," *American Economic Journal: Applied Economics*, 2022, *14* (3), 185–223.

Conley, Timothy G. and Chris Udry, "Learning about a New Technology: Pineapple in Ghana," *The American Economic Review*, 2010, *100* (1), 35–69.

Conlon, John, Manvika Mani, Gautam Rao, Matthew Ridley, and Frank Schilbach, "Not Learning from Others," *Revise Resubmit, Econometrica*, 2022.

Corno, Lucia, "Learning (or Not) in Health-Seeking Behavior: Evidence from Rural Tanzania," *Economic Development and Cultural Change*, 2014, *63* (1), 27–72.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina-Tejeiro, and Michael Weber, "Exposure to Grocery Prices and Inflation Expectations," *Journal of Political Economy*, 2021, *129*, 1615–1639.

de Janvry, Alain, Karen Macours Macours, and Elisabeth Sadoulet, "Learning for adopting: Technology adoption in developing country agriculture," *FERDI*, 2016.

Duhigg, Charles, *The power of habit: Why we do what we do in life and business*, Vol. 34, Random House, 2012.

Dupas, Pascaline, Vivian Hoffman, Michael Kremer, and Alix Peterson Zwane, "Targeting Health Subsidies through a Nonprice Mechanism: A Randomized Controlled Trial in Kenya," *Science*, 2016, *353* (6302), 889–895.

Dussault, Gilles and Maria Cristina Franceschini, "Not enough there, too many here: understanding geographical imbalances in the distribution of the health workforce," *Human Resources for Health*, 2006, *4* (12).

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker, "General equilibrium effects of cash transfers: experimental evidence from Kenya," *Econometrica*, 2022, *90* (6), 2603–2643.

Facchini, Gabriel, "Forgetting-by-not-doing: The case of surgeons and cesarean sections," *Health Economics*, 2022, *31* (3), 481–495.

Foster, Andrew D. and Mark R. Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *The Journal of Political Economy*, 1994, *103* (6), 1176–1209.

Halm, Ethan A., Clara Lee, and Mark R. Chassin, "Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature," *Annals of Internal Medicine*, 2002, *137* (6).

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein, "Learning Through Noticing: Theory and Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 2014, *129* (3), 1311–1353.

Hussam, Reshmaan, Atonu Rabbani, Giovanni Reggiani, and Natalia Rigol, "Rational Habit Formation: Experimental Evidence from Handwashing in India," *American Economic Journal: Applied Economics*, 2022, *14* (1), 1–41.

_ , Kailash Pandey, Abu Shonchoy, and Chikako Yamauchi, "Translating Information into Action," *Revise and resubmit, American Economic Review*, 2023.

Kremer, Michael, Stephen P Luby, Ricardo Maertens, Brandon Tan, and Witold Wiecek, "Water Treatment And Child Mortality: A Meta-Analysis And Cost-effectiveness Analysis," Technical Report, National Bureau of Economic Research 2023.

List, John A. and Haruka Uchida, "Here Today, Gone Tomorrow? Toward an Understanding of Fade-Out in Early Childhood Education Programs," *NBER Working Paper 33027*, 2024.

Lowes, Sara and Eduardo Montero, "The Legacy of Colonial Medicine in Central Africa," *American Economic Review*, 2021, *111* (4), 1284–1314.

Luby, Stephen P, Mubina Agboatwalla, John Painter, Arshad Altaf, Ward Billhimer, Bruce Keswick, and Robert M. Hoekstra, "Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomised controlled trial," *The American Journal of Tropical Medicine and Hygiene*, 2006, *11* (4), 479–489.

Malmendier, Ulrike and Stefan Nagel, "Learning from Inflation Experiences," *The Quarterly Journal of Economics*, 2015, *131* (1), 53–87.

Martinez-Bravo, Monica and Andreas Stegmann, "In Vaccines We Trust? The Effects of the CIA's Vaccine Ruse on Immunization in Pakistan," *Journal of the European Economic Association*, 2022, *20* (1), 150–186.

Neal, David, Jelena Vujcic, Orlando Hernandez, Wendy Wood, J Vujcic, O Hernandez, and W Wood, "The science of habit: creating disruptive and sticky behavior change in handwashing behavior," *Washington DC, USA. USAID/WASHplus Project*, 2015.

Quick, Robert E., Akiko Kimura, Angelica Thevos, Mathias Tembo, Isidore Shamputa, Lori Hutwagner, and Eric Mintz, "Diarrhea Prevention through Household-Level Water Disinfection and Safe Storage in Zambia," *The American Journal of Tropical Medicine and Hygiene*, 2002, *66* (5), 584–589.

Quick, Robert E, LV Venczel, ED Mintz, L Soleto, J Aparicio, M Gironaz, L Hutwagner, K Greene, C Bopp, K Maloney et al., "Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy," *Epidemiology & Infection*, 1999, *122* (1), 83–90.

Reller, Megan E, Carlos E Mendoza, M Beatriz Lopez, Maricruz Alvarez, Robert M Hoekstra, Christy A Olson, Kathleen G Baier, Bruce H Keswick, and Stephen P Luby, "A randomized controlled trial of household-based flocculant-disinfectant drinking water treatment for diarrhea prevention in rural Guatemala," *American Society of Tropical Medicine and Hygiene*, 2003, *69* (4), 411–419.

Royer, Heather, Mark Stehr, and Justin Sydnor, "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company," *American Economic Journal: Applied Economics*, 2015, *7* (3), 51–84.

Sazawal, Sunil, AKM Ahsan Habib, Usha Dhingra, Arup Dutta, Pratibha Dhingra, Archana Sarkar, Saikat Deb, Jahangir Alam, Asmaul Husna, and Robert E. Black, "Impact of micronutrient fortification of yoghurt on micronutrient status markers and growth – a randomized double blind controlled trial among school children in Bangladesh," *BMC Public Health*, 2013, *13* (514).

Scharf, Rebecca J, Mark D DeBoer, and Richard L Guerrant, "Recent advances in understanding the long-term sequelae of childhood infectious diarrhea," *Current infectious disease reports*, 2014, *16*, 1–7.

Simonsohn, Uri, Niklas Karlsson, George Loewenstein, and Dan Ariely, "The tree of experience in the forest of information: Overweighing experienced relative to observed information," *Games and Economic Behavior*, 2008, *62*, 263–286.

Soofi, Sajid Bashir, Gul Nawaz Khan, Shabina Ariff, Yasir Ihtesham, Mahamadou Tanimoune, Arjumand Rizvi, Muhammad Sajid, Cecilia Garzon, Saskia de Pee, and Zulfiqar A. Bhutta, "Effectiveness of nutritional supplementation during the first 1000-days of life to reduce child undernutrition: A cluster randomized controlled trial in Pakistan," *The Lancet Regional Health – Southeast Asia*, 2022, *4*, 1–11.

Supply, WHO/UNICEF Joint Water and Sanitation Monitoring Programme, *Progress on drinking water and sanitation: 2014 Update*, World Health Organization, 2014.

Tremblay, Maxime A., Céline M. Blanchard, Sara Taylor, Luc G. Pelletier, and Martin Villeneuve, "Work Extrinsic and Intrinsic Motivation Scale: Its Value for Organizational Psychology Research," *Canadian Journal of Behavioural Science*, 2009, *41* (4), 213–226.

UNICEF and WHO, "Progress on Sanitation and Drinking Water: 2015 Update and MDG Assessment," *Geneva, 80 pages*, 2015.

Urminsky, Oleg, Christian Hansen, and Victor Chernozhukov, "Using double-lasso regression for principled variable selection," *Working Paper*, 2016.

Watson, Samuel I, Ryan TT Rego, Timothy Hofer, and Richard J Lilford, "Evaluations of water, sanitation and hygiene interventions should not use diarrhoea as (primary) outcome," *BMJ Global Health*, 2022, *7* (5), e008521.

Wellsjo, Alexandra Steiny, "Simple Actions, Complex Habits: Lessons from Hospital Hand Hygiene," Technical Report, Working paper 2021.

WHO, "Diarrhoeal disease factsheet 2017," 2017.

Xu, Zhiwei, Yang Liu, Zongwei Ma, Ghasem (Sam) Toloo, Wenbiao Hu, and Shilu Tong, "Assessment of the temperature effect on childhood diarrhea using satellite imagery," *Scientific Reports*, 2014, *4* (5389).

8 Appendix

8.1 Baseline Balance

Table 12: Baseline Balance: Adult Outcomes

Variable	(1) Control Mean/SE	(2) Chlorine-tablets Mean/SE	(3) Incentives Mean/SE	(4) Info-tool Mean/SE	(5) Total Mean/SE	T-test Difference (1)-(2)	(1)-(3)	(1)-(4)
Any Child Had Motions	0.283	0.356	0.304	0.291	0.309	-0.073**	-0.021	-0.008
	(0.021)	(0.023)	(0.022)	(0.021)	(0.011)			
Reported Highest Diarrhea in Summer	0.836	0.844	0.825	0.820	0.831	-0.008	0.011	0.016
	(0.017)	(0.017)	(0.018)	(0.018)	(0.009)			
Number of Children <5	1.496	1.517	1.479	1.507	1.499	-0.021	0.017	-0.011
	(0.034)	(0.038)	(0.031)	(0.037)	(0.017)			
Has Heard of Chlorine	0.188	0.194	0.193	0.240	0.204	-0.006	-0.005	-0.052**
	(0.018)	(0.019)	(0.019)	(0.020)	(0.009)			
Would consider using chlorine	0.869	0.880	0.887	0.878	0.878	-0.010	-0.017	-0.008
	(0.016)	(0.015)	(0.015)	(0.015)	(0.008)			
Enumerator Observes Dirt in Water	0.164	0.160	0.173	0.156	0.163	0.003	-0.009	0.008
	(0.017)	(0.017)	(0.018)	(0.017)	(0.009)			
Reports Dirt in Water	0.752	0.753	0.743	0.751	0.750	-0.001	0.009	0.001
	(0.020)	(0.020)	(0.021)	(0.020)	(0.010)			
Boils, Bleaches, or Chlorinates Water	0.148	0.131	0.131	0.147	0.139	0.017	0.017	0.002
	(0.017)	(0.016)	(0.016)	(0.017)	(0.008)			
Strains or Filters Water	0.633	0.657	0.619	0.633	0.635	-0.024	0.014	-0.001
	(0.023)	(0.022)	(0.023)	(0.023)	(0.011)			
Believes Chlorine is for Water Purification	0.188	0.198	0.220	0.218	0.206	-0.010	-0.031	-0.030
	(0.018)	(0.019)	(0.020)	(0.019)	(0.010)			
Caretaker Asked about Chlorine Test	0.142	0.176	0.140	0.144	0.150	-0.034	0.002	-0.003
	(0.016)	(0.018)	(0.016)	(0.017)	(0.008)			
Attrited (endline)	0.106	0.136	0.149	0.113	0.126	-0.030	-0.042**	-0.007
	(0.015)	(0.016)	(0.017)	(0.015)	(0.008)			
N	452	449	451	450	1802			
F-test of joint significance (F-stat)						0.977	0.748	0.430

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Standard errors are robust. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 13: Baseline Balance: Child Outcomes

	(1) Control		(2) Chlorine-tablets		(3) Incentives		(4) Info-tool		T-test Difference		
Variable	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	(1)-(2)	(1)-(3)	(1)-(4)
Child Weight (kg)	672	10.461	675	10.417	664	10.621	670	10.624	0.044	-0.159	-0.163
	[451]	(0.111)	[448]	(0.109)	[448]	(0.131)	[449]	(0.111)			
Child MUAC (cm)	676	14.504	678	14.428	666	14.474	675	14.546	0.076	0.030	-0.041
	[452]	(0.048)	[449]	(0.042)	[450]	(0.049)	[450]	(0.050)			
Number of Motion Days	676	0.761	679	0.928	667	0.886	672	0.779	-0.167*	-0.125	-0.019
	[452]	(0.067)	[449]	(0.077)	[451]	(0.080)	[450]	(0.074)			
Child Had > 0 Motion Days	676	0.214	680	0.278	667	0.226	675	0.230	-0.063***	-0.012	-0.015
	[452]	(0.016)	[449]	(0.019)	[451]	(0.016)	[450]	(0.017)			

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are clustered at variable HHID. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 14: Treatment Balance: Spillover

	(1) Spillover Sample		(2) Spillover Sample	
Any Treatment Group	0.027	(0.026)		
Chlorine Only			0.026	(0.032)
Incentives			0.027	(0.031)
Info-Tool			0.028	(0.032)
Observations	1802		1802	
P-values:				
Incentives = Chlorine Only			0.962	
Info-Tool = Chlorine Only			0.928	
Info-Tool = Incentives			0.967	

Standard errors in parentheses
Each observation is at the household level. This regression include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

8.2 Akram and Mendelsohn (2021) Reanalysis

We reanalyze the data from Akram and Mendelsohn (2021), computing the same measure of spillovers that we compute in our data. We find that Info-tool participants with another Info-tool participant within twenty meters accept chlorine 17% more often than in other households (Akram and Mendelsohn (2021) used chlorine tablet acceptance as their primary outcome throughout the trial, and only tested water for chlorine residual in the final endline visit.)

Similar to our study, the spillover and non-spillover samples diverge most starkly soon after the period ends where the participants are using the Info-tool, and after they see the cumulative bar chart (Appendix Figure 7). Also similar to our study, the spillover and non-spillover samples converge towards the end of the trial.

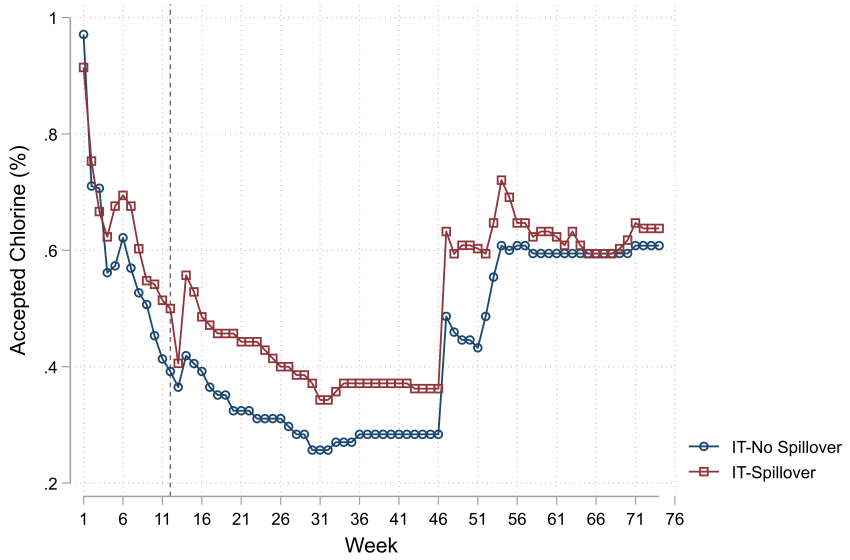


Figure 7: Akram and Mendelsohn (2021) Reanalysis
Rates of Chlorine Acceptance
where "Spillover" is exposure to any Info-tool neighbor

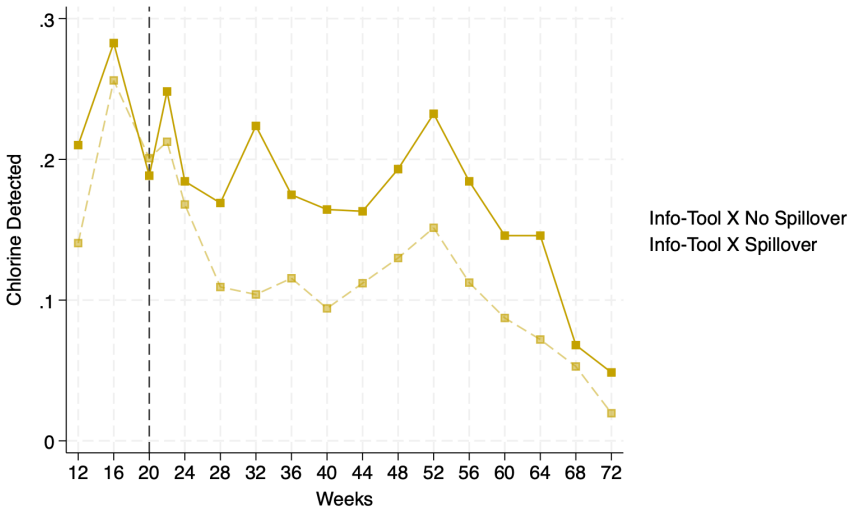


Figure 8: Our Study
Rates of Chlorine Detection
where "Spillover" is exposure to any Info-tool neighbor

8.3 Additional Figures and Tables

Table 15: Chlorine Detection (Household-Survey Panel)

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine Only	0.222***	(0.015)	0.151***	(0.011)	0.109***	(0.009)
Incentives	0.253***	(0.015)	0.143***	(0.010)	0.127***	(0.010)
Info-Tool	0.208***	(0.013)	0.166***	(0.011)	0.123***	(0.009)
Observations	4711		6354		14066	
P-values:						
Chlorine = Incentives	0.110		0.574		0.116	
Chlorine = Info-Tool	0.458		0.290		0.219	
Incentives = Info-Tool	0.016		0.095		0.712	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Child-Days of Diarrhea (Household-Survey Panel)

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine Only	-0.176***	(0.062)	-0.061	(0.045)	-0.084***	(0.032)
Incentives	-0.203***	(0.058)	-0.028	(0.046)	-0.052	(0.033)
Info-Tool	-0.143**	(0.063)	-0.094**	(0.043)	-0.095***	(0.031)
Observations	7872		6354		14066	
Control Mean	0.515		0.240		0.277	
P-values:						
Chlorine = Incentives	0.591		0.385		0.305	
Chlorine = Info-Tool	0.571		0.348		0.704	
Incentives = Info-Tool	0.248		0.070		0.147	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a continuous measure of the total child-days of diarrhea that the household reported over the preceding two weeks (aggregated across all children). All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, the total number of children in the household under five in that survey round, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

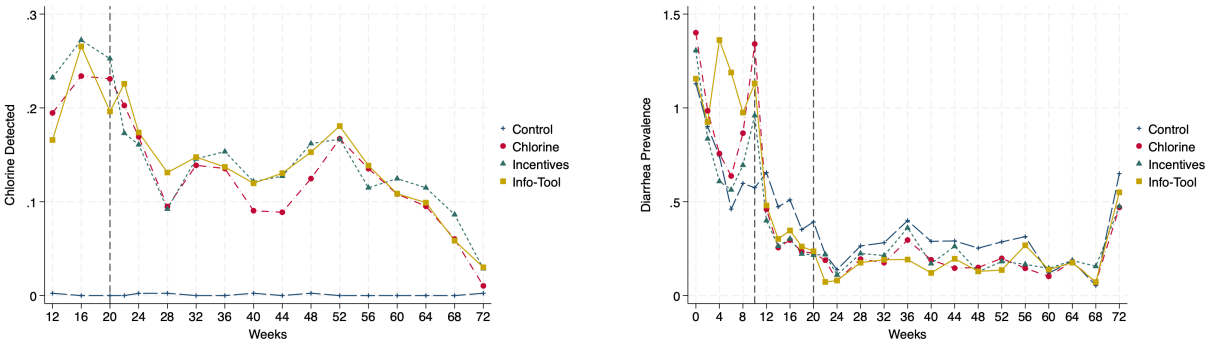


Figure 9: Raw Average: Chlorine Detection and Diarrhea Prevalence

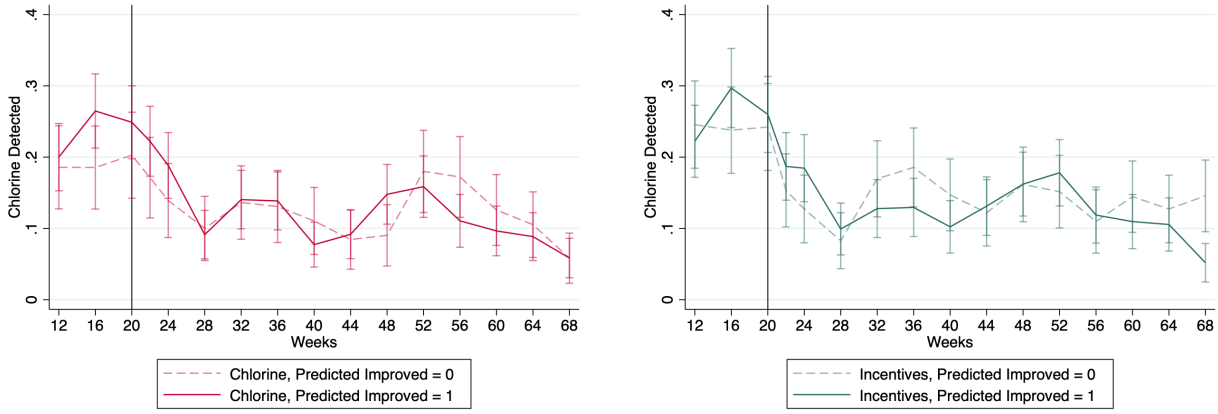


Figure 10: Heterogeneity: Short-Run Health Improvements

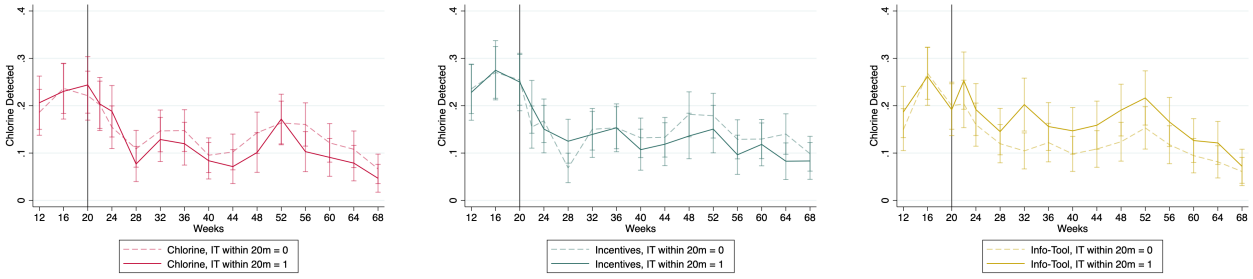


Figure 11: Heterogeneity: Info-tool Neighbors

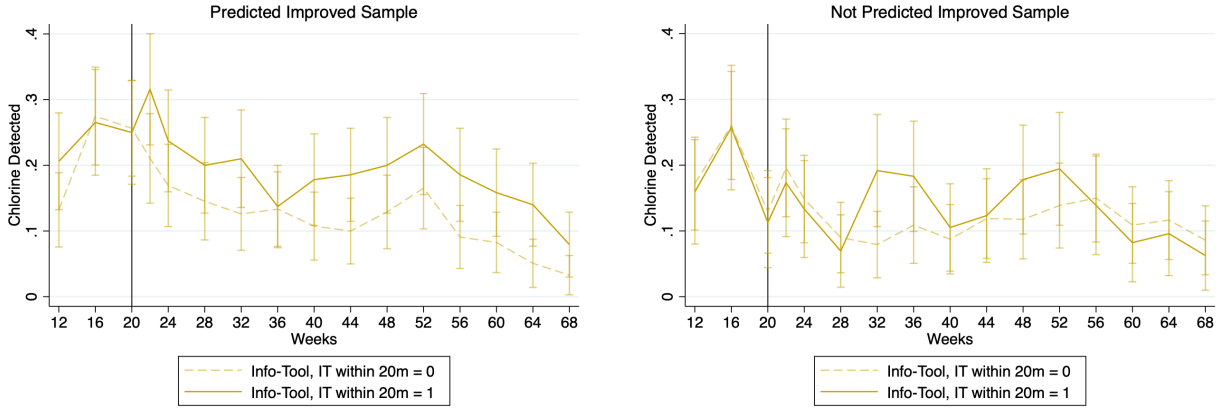


Figure 12: Heterogeneity: Info-tool Neighbors

Table 17: Placebo Test – Chlorine Detection (Household Survey Panel)
where "Spillover" is exposure to any Chlorine Only neighbor
Omitted group: Chlorine × No Spillover

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × No Spillover	0.209***	(0.018)	0.134***	(0.013)	0.108***	(0.011)
Chlorine × Spillover	0.237***	(0.026)	0.169***	(0.017)	0.112***	(0.017)
Incentives × No Spillover	0.250***	(0.019)	0.136***	(0.012)	0.131***	(0.013)
Incentives × Spillover	0.255***	(0.024)	0.151***	(0.016)	0.121***	(0.015)
Info-Tool × No Spillover	0.212***	(0.017)	0.167***	(0.014)	0.127***	(0.012)
Info-Tool × Spillover	0.194***	(0.022)	0.167***	(0.018)	0.117***	(0.015)
Observations	4711		6354		14066	
P-values:						
Chlorine: No Spillover = Spillover	0.364		0.102		0.831	
Incentives: No Spillover = Spillover	0.873		0.450		0.616	
Info-Tool: No Spillover = Spillover	0.515		0.997		0.598	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to a Chlorine Only neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Placebo Test – Chlorine Detection (Aggregate Specification)
where "Spillover" is exposure to any Incentives neighbor
Omitted group: Chlorine × No Spillover

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × No Spillover	0.212***	(0.018)	0.141***	(0.013)	0.101***	(0.011)
Chlorine × Spillover	0.231***	(0.026)	0.154***	(0.017)	0.123***	(0.018)
Incentives × No Spillover	0.264***	(0.019)	0.148***	(0.012)	0.135***	(0.013)
Incentives × Spillover	0.231***	(0.024)	0.130***	(0.015)	0.112***	(0.015)
Info-Tool × No Spillover	0.194***	(0.016)	0.164***	(0.014)	0.124***	(0.012)
Info-Tool × Spillover	0.223***	(0.024)	0.172***	(0.019)	0.125***	(0.016)
Observations	4711		6354		14066	
P-values:						
Chlorine: No Spillover = Spillover	0.555		0.571		0.293	
Incentives: No Spillover = Spillover	0.290		0.346		0.242	
Info-Tool: No Spillover = Spillover	0.327		0.716		0.962	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to an Incentives neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Placebo Test – Chlorine Detection (Household-Survey Panel)
where "Spillover" is exposure to any treated neighbor predicted to improve
Omitted group: Chlorine × No Spillover

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × Spillover	0.007	(0.029)	0.001	(0.021)	0.014	(0.016)
Incentives × No Spillover	0.059*	(0.032)	-0.011	(0.021)	0.042**	(0.019)
Incentives × Spillover	0.026	(0.027)	-0.004	(0.020)	0.014	(0.015)
Info-Tool × No Spillover	-0.013	(0.027)	0.009	(0.023)	0.012	(0.016)
Info-Tool × Spillover	0.000	(0.027)	0.025	(0.022)	0.030*	(0.016)
Observations	3463		4689		10472	
P-values:						
Incentives: No Spillover = Spillover	0.275		0.720		0.107	
Info-Tool: No Spillover = Spillover	0.624		0.454		0.250	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to any treatment neighbor (someone within 20m) who was predicted to have their health improve than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined. See Section 6 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Endline Social Networks

	(1) Number in Social Network	(2) Number Discuss Health	(3) Number Discussed Water Purification
Incentives	-0.052	-0.022	0.114**
	(0.076)	(0.054)	(0.045)
Info-Tool	-0.119	-0.004	0.005
	(0.076)	(0.054)	(0.045)
Observations	1116	1116	1116
P-values:			
Incentives = Info-Tool	0.380	0.728	0.015

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Endline Social Networks
where "Spillover" is exposure to any Info-tool neighbor

	(1) Number in Social Network	(2) Number Discuss Health	(3) Number Discussed Water Purification
Chlorine × Spillover	0.116	-0.065	0.132**
	(0.128)	(0.080)	(0.067)
Incentives × No Spillover	-0.016	-0.050	0.148***
	(0.096)	(0.068)	(0.056)
Incentives × Spillover	0.000	-0.037	0.184***
	(.)	(0.079)	(0.066)
Info-Tool × No Spillover	-0.045	0.004	0.037
	(0.096)	(0.068)	(0.056)
Info-Tool × Spillover	-0.131	-0.081	0.081
	(0.127)	(0.079)	(0.066)
Observations	1116	1116	1116
P-values:			
Incentives × No Spillover = Incentives × Spillover	0.866	0.872	0.590
Info-Tool × No Spillover = Info-Tool × Spillover	0.590	0.288	0.514

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 22: Mimicry of Other Behaviors

	(1) Take-up of Other Programs: Vitamin A	(2) Take-up of Other Programs: Deworming	(3) Take-up of Other Programs: Iron	(4) Number of Friends Believes Uses Same Savings Technology
Chlorine Only	0.030	0.004	0.012	-0.001
	(0.031)	(0.023)	(0.010)	(0.057)
Incentives	0.066**	0.001	0.010	0.006
	(0.031)	(0.022)	(0.010)	(0.057)
Info-Tool	0.084***	0.004	0.016	-0.113**
	(0.031)	(0.023)	(0.010)	(0.057)
Observations	1512	1512	1512	1503
P-values:				
Chlorine = Incentives	0.247	0.906	0.890	0.909
Chlorine = Info-Tool	0.085	1.000	0.688	0.051
Incentives = Info-Tool	0.565	0.906	0.588	0.039

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Mimicry of Other Behaviors
where "Spillover" is exposure to any Info-tool neighbor

	(1) Take-up of Other Programs: Vitamin A	(2) Take-up of Other Programs: Deworming	(3) Take-up of Other Programs: Iron	(4) Number of Friends Believes Uses Same Savings Technology
Chlorine × No Spillover	0.069*	-0.001	0.013	0.040
	(0.038)	(0.026)	(0.011)	(0.065)
Chlorine × Spillover	-0.041	0.013	0.009	-0.073
	(0.053)	(0.031)	(0.013)	(0.079)
Incentives × No Spillover	0.046	0.017	0.001	0.031
	(0.038)	(0.025)	(0.011)	(0.065)
Incentives × Spillover	0.100*	-0.026	0.027**	-0.036
	(0.052)	(0.031)	(0.013)	(0.079)
Info-Tool × No Spillover	0.102***	0.015	0.011	-0.088
	(0.038)	(0.026)	(0.011)	(0.065)
Info-Tool × Spillover	0.048	-0.016	0.023*	-0.158**
	(0.052)	(0.031)	(0.013)	(0.079)
Observations	1512	1512	1512	1503
P-values:				
Chlorine × No Spillover = Chlorine × Spillover	0.094	0.683	0.751	0.187
Incentives × No Spillover = Incentives × Spillover	0.402	0.205	0.073	0.433
Info-Tool × No Spillover = Info-Tool × Spillover	0.405	0.353	0.409	0.412

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Endline Social Norms

	(1) Believes Neighbor Would Accept Chlorinated Water	(2) Number of Friends Believes Chlorinates
Chlorine Only	0.082***	0.038*
	(0.030)	(0.021)
Incentives	0.139***	0.065***
	(0.030)	(0.021)
Info-Tool	0.098***	0.066***
	(0.030)	(0.021)
Observations	1503	1342
P-values:		
Chlorine = Incentives	0.061	0.200
Chlorine = Info-Tool	0.601	0.178
Incentives = Info-Tool	0.176	0.953

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 25: Endline Social Norms
where "Spillover" is exposure to any Info-tool neighbor

	(1) Believes Neighbor Would Accept Chlorinated Water	(2) Number of Friends Believes Chlorinates
Chlorine × No Spillover	0.079**	0.039*
	(0.037)	(0.024)
Chlorine × Spillover	0.088*	0.034
	(0.051)	(0.029)
Incentives × No Spillover	0.157***	0.065***
	(0.037)	(0.024)
Incentives × Spillover	0.109**	0.065**
	(0.050)	(0.029)
Info-Tool × No Spillover	0.060	0.058**
	(0.037)	(0.024)
Info-Tool × Spillover	0.164***	0.081***
	(0.050)	(0.029)
Observations	1503	1342
P-values:		
Chlorine × No Spillover = Chlorine × Spillover		
Incentives × No Spillover = Incentives × Spillover	0.444	0.987
Info-Tool × No Spillover = Info-Tool × Spillover	0.100	0.477

Standard errors in parentheses
Each observation is at the household level. The outcome for column (1) is an indicator for if the respondent chose a hypothetical Info-tool respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Incentives participant). This regression controls for an indicator for if the respondent chose someone from within their own treatment group. The outcome for column (2) is a rating between 1 and 7 for how true the following statement felt: I use chlorine to achieve a standard of health for my family. This regression controls for the average motivation score the respondent gave across all motivation questions, and the order of questions (randomized). Both specifications include neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

8.4 Main Tables: Robustness to Alternate Specification

For robustness, we also estimate the following specification using an aggregated version of a given outcome:

$$Y_i = \beta_0 + \beta_1 T1_i + \beta_2 T2_i + \beta_3 T3_i + NV_i + X_{i0} + \delta_s + \gamma_b + \epsilon_i, \tag{2}$$

where Y_i represents the aggregated outcome measurements over relevant visits (e.g., total number of visits with chlorine presence detected in the water), and NV_i captures the number of visits with non-missing outcome measurements. All of our tables are replicated using this specification in the Appendix.

Table 26: Chlorine Detection: Aggregate
where "Spillover" is exposure to any Info-tool neighbor

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × No Spillover	0.641***	(0.056)	0.541***	(0.051)	0.886***	(0.099)
Chlorine × Spillover	0.641***	(0.069)	0.561***	(0.062)	0.951***	(0.119)
Incentives × No Spillover	0.743***	(0.056)	0.506***	(0.051)	1.093***	(0.098)
Incentives × Spillover	0.744***	(0.068)	0.528***	(0.061)	0.932***	(0.118)
Info-Tool × No Spillover	0.616***	(0.057)	0.543***	(0.051)	0.920***	(0.099)
Info-Tool × Spillover	0.632***	(0.069)	0.716***	(0.062)	1.193***	(0.120)
Observations	1599		1690		1690	
P-values:						
Chlorine: No Spillover = Spillover	0.995		0.774		0.613	
Incentives: No Spillover = Spillover	0.987		0.733		0.208	
Info-Tool: No Spillover = Spillover	0.835		0.011		0.037	

Standard errors in parentheses
Each observation is at the household level. The outcome is the aggregate of the total number of times that chlorine was detected across all visits in the specified time period. All specifications include a control for the number of visits that the household was surveyed during the specified time period, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

8.5 Baseline Balance: Spillover and Endline Sample

Table 27: Chlorine Detection: Aggregate
by predicted health improvement

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × Improved	0.089	(0.082)	0.150**	(0.073)	0.153	(0.132)
Incentives × Not Improved	0.138*	(0.081)	0.008	(0.072)	0.168	(0.130)
Incentives × Improved	0.168**	(0.082)	0.079	(0.074)	0.231*	(0.132)
Info-Tool × Not Improved	-0.044	(0.081)	0.088	(0.073)	0.105	(0.131)
Info-Tool × Improved	0.101	(0.083)	0.180**	(0.074)	0.269**	(0.134)
Observations	1188		1261		1261	
P-values:						
Incentives: Not Improved = Improved	0.709		0.332		0.631	
Info-Tool: Not Improved = Improved	0.082		0.217		0.223	

Standard errors in parentheses
Each observation is at the household level. The outcome is the aggregate of the total number of times that chlorine was detected across all visits in the specified time period. All specifications include a control for the number of visits that the household was surveyed during the specified time period, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). Improved is a binary indicator for if the participant's predicted improvement in health after the beginning of chlorine distribution was above the median. See Section 6 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 28: Chlorine Detection: Aggregate
where "Spillover" is exposure to any Info-tool neighbor predicted to improve

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × Spillover	-0.069	(0.093)	0.037	(0.082)	0.260*	(0.146)
Incentives × No Spillover	0.112*	(0.065)	-0.038	(0.059)	0.244**	(0.106)
Incentives × Spillover	0.025	(0.091)	0.014	(0.082)	0.042	(0.148)
Info-Tool × No Spillover	-0.032	(0.067)	0.025	(0.061)	0.093	(0.108)
Info-Tool × Spillover	-0.045	(0.089)	0.166**	(0.079)	0.377***	(0.142)
Observations	1188		1261		1261	
P-values:						
Incentives: No Spillover = Spillover	0.337		0.533		0.172	
Info-Tool: No Spillover = Spillover	0.880		0.079		0.049	

Standard errors in parentheses
Each observation is at the household level. The outcome is the aggregate of the total number of times that chlorine was detected across all visits in the specified time period. All specifications include a control for the number of visits that the household was surveyed during the specified time period, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20m) whose health was predicted to improve than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined. See Section 6 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 29: Chlorine Detection: Aggregate
where "Spillover" is exposure to any treatment neighbor predicted to improve

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine × Spillover	0.015	(0.082)	0.010	(0.074)	0.145	(0.133)
Incentives × No Spillover	0.186**	(0.087)	-0.039	(0.079)	0.367***	(0.141)
Incentives × Spillover	0.069	(0.080)	-0.022	(0.073)	0.095	(0.130)
Info-Tool × No Spillover	-0.049	(0.088)	0.008	(0.080)	0.109	(0.142)
Info-Tool × Spillover	0.009	(0.082)	0.101	(0.074)	0.247*	(0.132)
Observations	1188		1261		1261	
P-values:						
Chlorine: No Spillover = Spillover	0.858		0.891		0.273	
Incentives: No Spillover = Spillover	0.154		0.816		0.043	
Info-Tool: No Spillover = Spillover	0.494		0.224		0.316	

Standard errors in parentheses
Each observation is at the household level. The outcome is the aggregate of the total number of times that chlorine was detected across all visits in the specified time period. All specifications include a control for the number of visits that the household was surveyed during the specified time period, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months). The Spillover Sample is defined as participants who were randomly more exposed to any treatment neighbor (someone within 20m) who was predicted to have their health improve than they would be in expectation based on endogenous spatial factors. See Section 5 for a detailed explanation of how the measure is defined. See Section 6 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table 30: Chlorine Detection: Aggregate

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine Only	0.641***	(0.050)	0.578***	(0.046)	0.933***	(0.090)
Incentives	0.743***	(0.049)	0.546***	(0.046)	1.086***	(0.089)
Info-Tool	0.622***	(0.050)	0.648***	(0.046)	1.060***	(0.089)
Observations	1599		1661		1653	
P-values:						
Chlorine = Incentives	0.039		0.492		0.091	
Chlorine = Info-Tool	0.706		0.137		0.161	
Incentives = Info-Tool	0.015		0.029		0.770	

Standard errors in parentheses
Each observation is at the household level. The outcome is the aggregate of the total number of times that chlorine was detected across all visits in the specified time period. All specifications include a control for the number of visits that the household was surveyed during the specified time period, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 31: Diarrhea: Aggregate

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine Only	-0.892***	(0.278)	-0.264*	(0.157)	-0.820***	(0.265)
Incentives	-1.006***	(0.276)	-0.103	(0.156)	-0.462*	(0.262)
Info-Tool	-0.720***	(0.279)	-0.334**	(0.156)	-0.792***	(0.262)
Observations	1599		1661		1653	
Control Mean	2.505		0.949		2.348	
P-values:						
Chlorine = Incentives	0.681		0.308		0.177	
Chlorine = Info-Tool	0.544		0.660		0.916	
Incentives = Info-Tool	0.306		0.143		0.209	

Standard errors in parentheses
Each observation is at the household level. The outcome is a continuous measure of the total child-days of diarrhea that the household reported over the preceding two weeks (aggregated across all children). All specifications include a control for the number of visits that the household was surveyed during the specified time period, neighborhood block fixed effects, unbalanced baseline controls, the number of children under five the household had at baseline, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 32: IV Impacts on Endline Child Health (Instrument: Info-tool × Spillover) where "Spillover" is exposure to any Info-tool neighbor

	(1) Index (Anthro-pometry)	(2) Height-for-Age	(3) Weight-for-Height	(4) Weight-for-Age	(5) MUAC-for-Age
Boils, Bleaches, or Chlorinates Water	0.514***	-0.096	0.604	0.807**	-0.424
	(0.184)	(0.379)	(0.521)	(0.402)	(0.337)
Observations	2616	2371	2439	2492	1954
Endline Control Mean	-0.019	-1.773	-0.291	-1.407	-1.453
Weak-IV robust F statistic	34.82	33.67	33.49	35.77	26.95
C-statistic p-value	0.002	0.983	0.449	0.057	0.384

Standard errors in parentheses
Any treatment (i.e., received chlorine) is an instrument for if the respondent reported at endline that she boils, bleaches, or chlorinates her water. Child-level cross-section of the endline survey. Standard errors are clustered at the household level. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, and neighborhood block fixed effects. Indeces are constructed using following Anderson (2008).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 33: Chlorine Acceptance

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Incentives	0.066***	(0.023)	0.077***	(0.024)	0.059**	(0.028)
Info-Tool	0.015	(0.024)	0.010	(0.026)	0.016	(0.029)
Observations	5545		4609		10455	
Chlorine-only Mean	0.794		0.756		0.740	

Standard errors in parentheses
Each observation is a household-level survey visit. The outcome is a binary indicator for whether or not the respondent accepted free chlorine during the survey visit. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Medium-run is defined as the next three months of chlorine distribution (immediately after the behavioral interventions ended); and Long-run is defined as the remainder of the trial (nine months).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 34: Baseline Balance: Spillover Sample

Variable	(1) No Spillover Mean/SE	(2) Spillover Sample Mean/SE	(3) Total Mean/SE	T-test Difference (1)-(2)
Any Child Had Motions	0.295	0.332	0.309	-0.036
	(0.013)	(0.019)	(0.011)	
Reported Highest Diarrhea in Summer	0.817	0.856	0.831	-0.039
	(0.011)	(0.014)	(0.009)	
Number of Children <5	1.491	1.514	1.499	-0.023
	(0.020)	(0.033)	(0.017)	
Has Heard of Chlorine	0.212	0.188	0.204	0.024
	(0.012)	(0.015)	(0.009)	
Would consider using chlorine	0.872	0.890	0.878	-0.019
	(0.010)	(0.012)	(0.008)	
Enumerator Observes Dirt in Water	0.164	0.162	0.163	0.002
	(0.011)	(0.014)	(0.009)	
Reports Dirt in Water	0.750	0.750	0.750	-0.000
	(0.013)	(0.017)	(0.010)	
Boils, Bleaches, or Chlorinates Water	0.143	0.133	0.139	0.010
	(0.010)	(0.013)	(0.008)	
Strains or Filters Water	0.620	0.662	0.635	-0.042
	(0.014)	(0.019)	(0.011)	
Believes Chlorine is for Water Purification	0.217	0.187	0.206	0.030
	(0.012)	(0.015)	(0.010)	
Caretaker Asked about Chlorine Test	0.143	0.164	0.150	-0.021
	(0.010)	(0.015)	(0.008)	
Attrited (endline)	0.125	0.128	0.126	-0.003
	(0.010)	(0.013)	(0.008)	
N	1154	648	1802	
F-test of joint significance (F-stat)				1.110

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Standard errors are robust. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 35: Baseline Balance: Spillover Sample (Info-tool Only)

Variable	(1) Info-Tool (no IT neighbor) Mean/SE	(2) Info-Tool (with IT neighbor) Mean/SE	(3) Total Mean/SE	T-test Difference (1)-(2)
Any Child Had Motions	0.253	0.358	0.291	-0.105**
	(0.026)	(0.037)	(0.021)	
Reported Highest Diarrhea in Summer	0.793	0.867	0.820	-0.074
	(0.024)	(0.027)	(0.018)	
Number of Children <5	1.498	1.521	1.507	-0.023
	(0.038)	(0.076)	(0.037)	
Has Heard of Chlorine	0.235	0.248	0.240	-0.013
	(0.025)	(0.034)	(0.020)	
Would consider using chlorine	0.867	0.897	0.878	-0.030
	(0.020)	(0.024)	(0.015)	
Enumerator Observes Dirt in Water	0.172	0.127	0.156	0.045
	(0.022)	(0.026)	(0.017)	
Reports Dirt in Water	0.754	0.745	0.751	0.009
	(0.026)	(0.034)	(0.020)	
Boils, Bleaches, or Chlorinates Water	0.126	0.182	0.147	-0.056
	(0.020)	(0.030)	(0.017)	
Strains or Filters Water	0.618	0.661	0.633	-0.043
	(0.029)	(0.037)	(0.023)	
Believes Chlorine is for Water Purification	0.239	0.182	0.218	0.057
	(0.025)	(0.030)	(0.019)	
Caretaker Asked about Chlorine Test	0.123	0.182	0.144	-0.059
	(0.019)	(0.030)	(0.017)	
Attrited (endline)	0.105	0.127	0.113	-0.022
	(0.018)	(0.026)	(0.015)	
N	285	165	450	
F-test of joint significance (F-stat)				1.382

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Standard errors are robust. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 36: Baseline Balance: Spillover Sample (Incentives Only)

Variable	(1) Incentives (no IT neighbor) Mean/SE	(2) Incentives (with IT neighbor) Mean/SE	(3) Total Mean/SE	T-test Difference (1)-(2)
Any Child Had Motions	0.315	0.285	0.304	0.030
	(0.028)	(0.035)	(0.022)	
Reported Highest Diarrhea in Summer	0.836	0.806	0.825	0.030
	(0.022)	(0.031)	(0.018)	
Number of Children <5	1.486	1.467	1.479	0.019
	(0.040)	(0.050)	(0.031)	
Has Heard of Chlorine	0.220	0.145	0.193	0.075*
	(0.025)	(0.028)	(0.019)	
Would consider using chlorine	0.895	0.873	0.887	0.022
	(0.018)	(0.026)	(0.015)	
Enumerator Observes Dirt in Water	0.185	0.152	0.173	0.034
	(0.023)	(0.028)	(0.018)	
Reports Dirt in Water	0.773	0.691	0.743	0.082
	(0.025)	(0.036)	(0.021)	
Boils, Bleaches, or Chlorinates Water	0.140	0.115	0.131	0.025
	(0.021)	(0.025)	(0.016)	
Strains or Filters Water	0.612	0.630	0.619	-0.018
	(0.029)	(0.038)	(0.023)	
Believes Chlorine is for Water Purification	0.234	0.194	0.220	0.040
	(0.025)	(0.031)	(0.020)	
Caretaker Asked about Chlorine Test	0.133	0.152	0.140	-0.019
	(0.020)	(0.028)	(0.016)	
Attrited (endline)	0.154	0.139	0.149	0.014
	(0.021)	(0.027)	(0.017)	
N	286	165	451	
F-test of joint significance (F-stat)				0.776

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Standard errors are robust. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 37: Baseline Balance: Spillover Sample (Chlorine Only)

Variable	(1) Chlorine (no IT neighbor) Mean/SE	(2) Chlorine (with IT neighbor) Mean/SE	(3) Total Mean/SE	T-test Difference (1)-(2)
Any Child Had Motions	0.338	0.388	0.356	-0.050
	(0.028)	(0.038)	(0.023)	
Reported Highest Diarrhea in Summer	0.835	0.861	0.844	-0.026
	(0.022)	(0.027)	(0.017)	
Number of Children <5	1.475	1.588	1.517	-0.113
	(0.042)	(0.071)	(0.038)	
Has Heard of Chlorine	0.190	0.200	0.194	-0.010
	(0.023)	(0.031)	(0.019)	
Would consider using chlorine	0.870	0.897	0.880	-0.027
	(0.020)	(0.024)	(0.015)	
Enumerator Observes Dirt in Water	0.137	0.200	0.160	-0.063
	(0.020)	(0.031)	(0.017)	
Reports Dirt in Water	0.739	0.776	0.753	-0.036
	(0.026)	(0.033)	(0.020)	
Boils, Bleaches, or Chlorinates Water	0.144	0.109	0.131	0.035
	(0.021)	(0.024)	(0.016)	
Strains or Filters Water	0.641	0.685	0.657	-0.044
	(0.029)	(0.036)	(0.022)	
Believes Chlorine is for Water Purification	0.180	0.230	0.198	-0.051
	(0.023)	(0.033)	(0.019)	
Caretaker Asked about Chlorine Test	0.165	0.194	0.176	-0.028
	(0.022)	(0.031)	(0.018)	
Attrited (endline)	0.123	0.158	0.136	-0.034
	(0.020)	(0.028)	(0.016)	
N	284	165	449	
F-test of joint significance (F-stat)				0.636

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Standard errors are robust. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 38: Baseline Balance: Adult Outcomes (Endline Sample)

Variable	(1) Control Mean/SE	(2) Chlorine-tablets Mean/SE	(3) Incentives Mean/SE	(4) Info-tool Mean/SE	(5) Total Mean/SE	T-test Difference (1)-(2)	T-test Difference (1)-(3)	T-test Difference (1)-(4)
Any Child Had Motions	0.285	0.345	0.292	0.278	0.300	-0.061*	-0.007	0.006
	(0.022)	(0.024)	(0.023)	(0.022)	(0.012)			
Reported Highest Diarrhea in Summer	0.837	0.832	0.820	0.835	0.831	0.004	0.016	0.002
	(0.018)	(0.019)	(0.020)	(0.019)	(0.009)			
Number of Children <5	1.493	1.497	1.479	1.511	1.495	-0.005	0.013	-0.019
	(0.035)	(0.039)	(0.034)	(0.039)	(0.018)			
Has Heard of Chlorine	0.200	0.204	0.201	0.238	0.211	-0.003	-0.000	-0.038
	(0.020)	(0.020)	(0.020)	(0.021)	(0.010)			
Would consider using chlorine	0.866	0.879	0.880	0.872	0.874	-0.013	-0.014	-0.006
	(0.017)	(0.017)	(0.017)	(0.017)	(0.008)			
Enumerator Observes Dirt in Water	0.146	0.155	0.169	0.158	0.157	-0.009	-0.023	-0.012
	(0.018)	(0.018)	(0.019)	(0.018)	(0.009)			
Reports Dirt in Water	0.748	0.747	0.758	0.752	0.751	0.000	-0.010	-0.004
	(0.022)	(0.022)	(0.022)	(0.022)	(0.011)			
Boils, Bleaches, or Chlorinates Water	0.141	0.121	0.130	0.148	0.135	0.020	0.011	-0.007
	(0.017)	(0.017)	(0.017)	(0.018)	(0.009)			
Strains or Filters Water	0.651	0.660	0.622	0.639	0.643	-0.009	0.029	0.012
	(0.024)	(0.024)	(0.025)	(0.024)	(0.012)			
Believes Chlorine is for Water Purification	0.198	0.196	0.232	0.216	0.210	0.002	-0.034	-0.018
	(0.020)	(0.020)	(0.022)	(0.021)	(0.010)			
Caretaker Asked about Chlorine Test	0.136	0.173	0.143	0.140	0.148	-0.037	-0.007	-0.004
	(0.017)	(0.019)	(0.018)	(0.017)	(0.009)			
N	404	388	384	399	1575			
F-test of joint significance (F-stat)						0.640	0.606	0.256

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Standard errors are robust. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 39: Baseline Balance: Child Outcomes (Endline Sample)

	(1) Control		(2) Chlorine-tablets		(3) Incentives		(4) Info-tool		T-test Difference		
Variable	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	(1)-(2)	(1)-(3)	(1)-(4)
Female	572	0.479	541	0.457	541	0.481	565	0.485	0.022	-0.002	-0.006
	[386]	(0.022)	[370]	(0.024)	[370]	(0.022)	[383]	(0.022)			
Age (in months)	568	30.885	539	30.638	539	30.226	564	31.599	0.247	0.659	-0.715
	[384]	(0.584)	[370]	(0.563)	[368]	(0.600)	[383]	(0.605)			
Child Weight (kg)	568	10.562	539	10.581	539	10.674	564	10.691	-0.019	-0.112	-0.130
	[384]	(0.119)	[370]	(0.122)	[368]	(0.129)	[383]	(0.120)			
Child MUAC (cm)	572	14.345	540	14.274	542	14.317	565	14.395	0.072	0.028	-0.050
	[386]	(0.053)	[370]	(0.049)	[371]	(0.057)	[383]	(0.058)			
Number of Motion Days	572	0.769	541	0.852	542	0.769	565	0.796	-0.083	-0.001	-0.027
	[386]	(0.074)	[370]	(0.080)	[371]	(0.081)	[383]	(0.084)			
Child Had > 0 Motion Days	572	0.215	541	0.262	542	0.218	565	0.219	-0.047*	-0.003	-0.004
	[386]	(0.018)	[370]	(0.020)	[371]	(0.018)	[383]	(0.018)			

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are clustered at variable HHID. Fixed effects using variable block are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

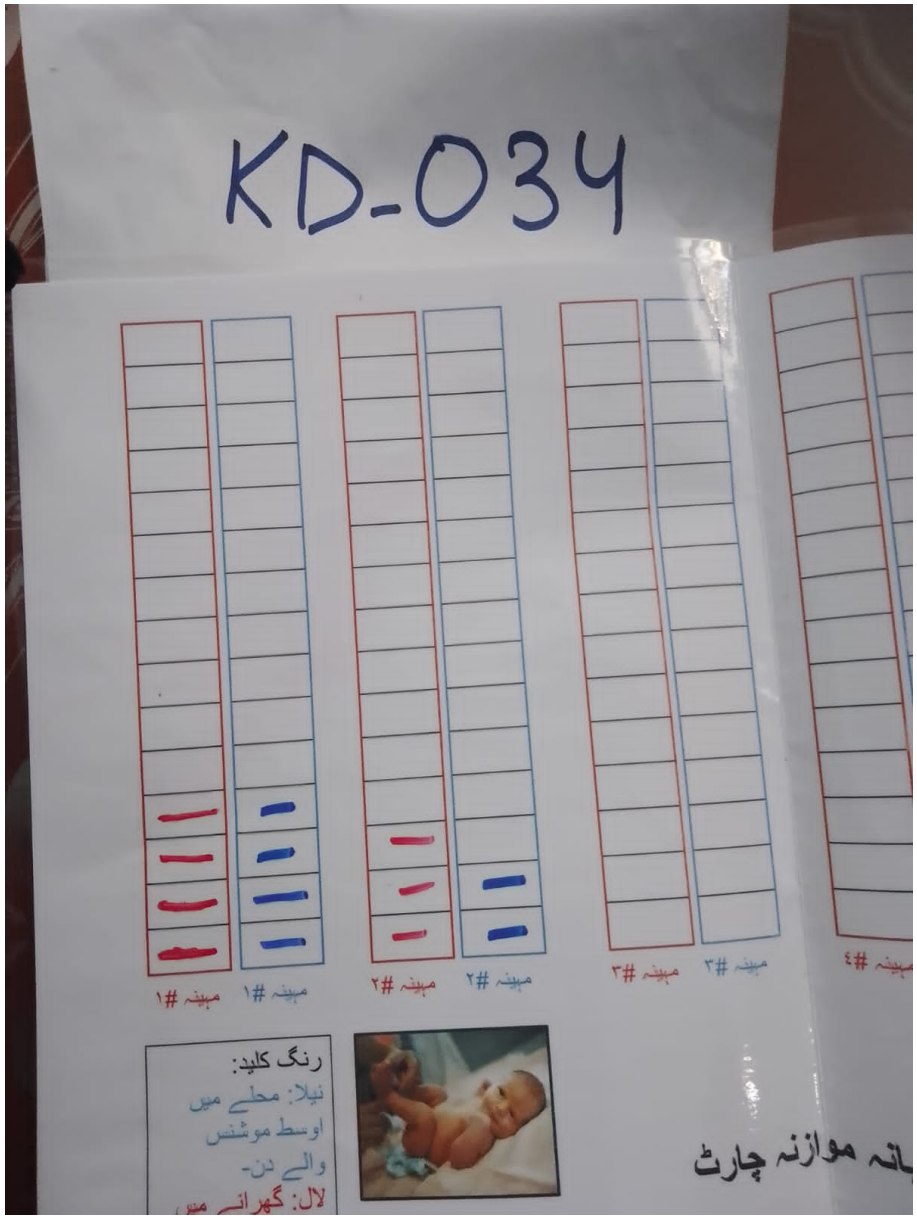


Figure 13: Experiential Learning Intervention: Info-Tool

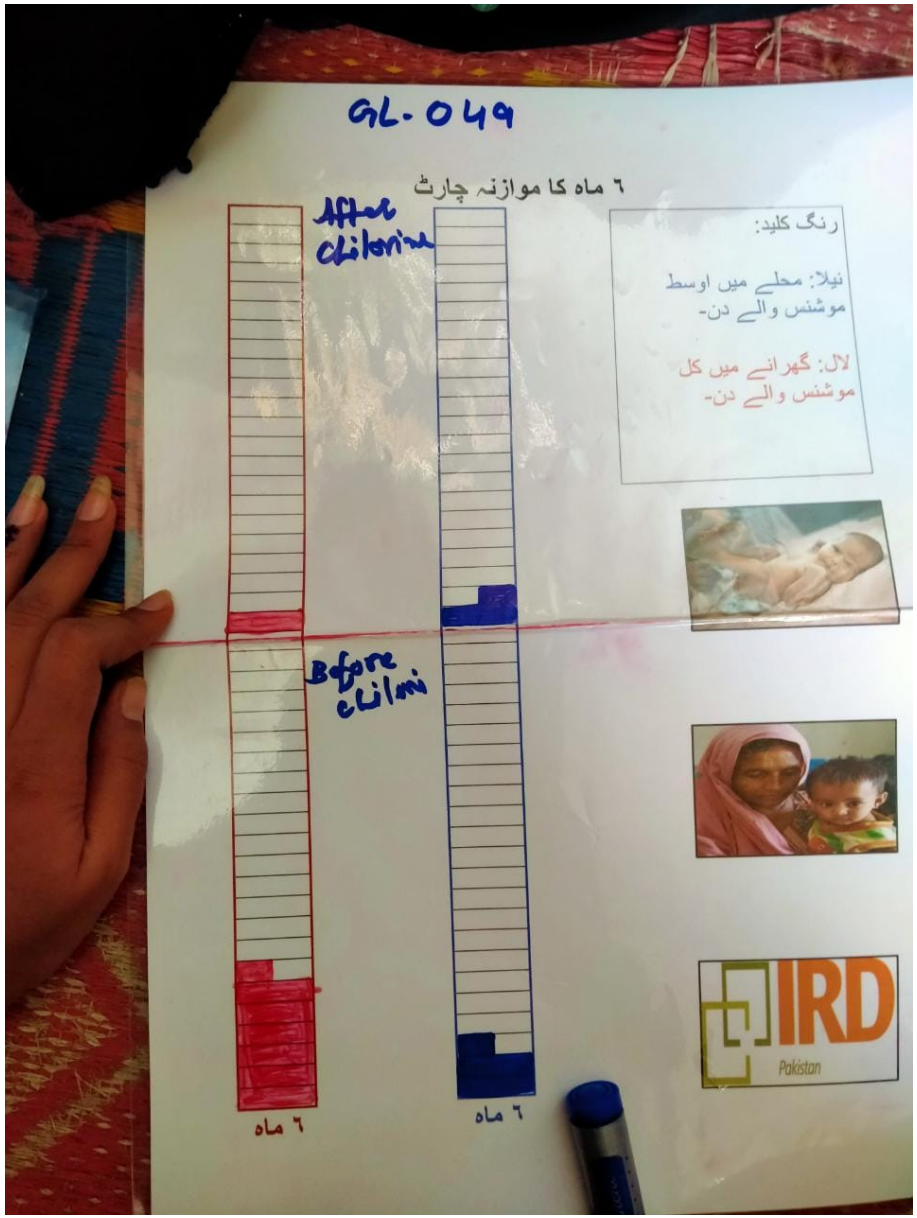


Figure 14: Experiential Learning Intervention: Info-Tool (Difference-in-Differences Comparison)



Figure 15: Sample Water Vessel

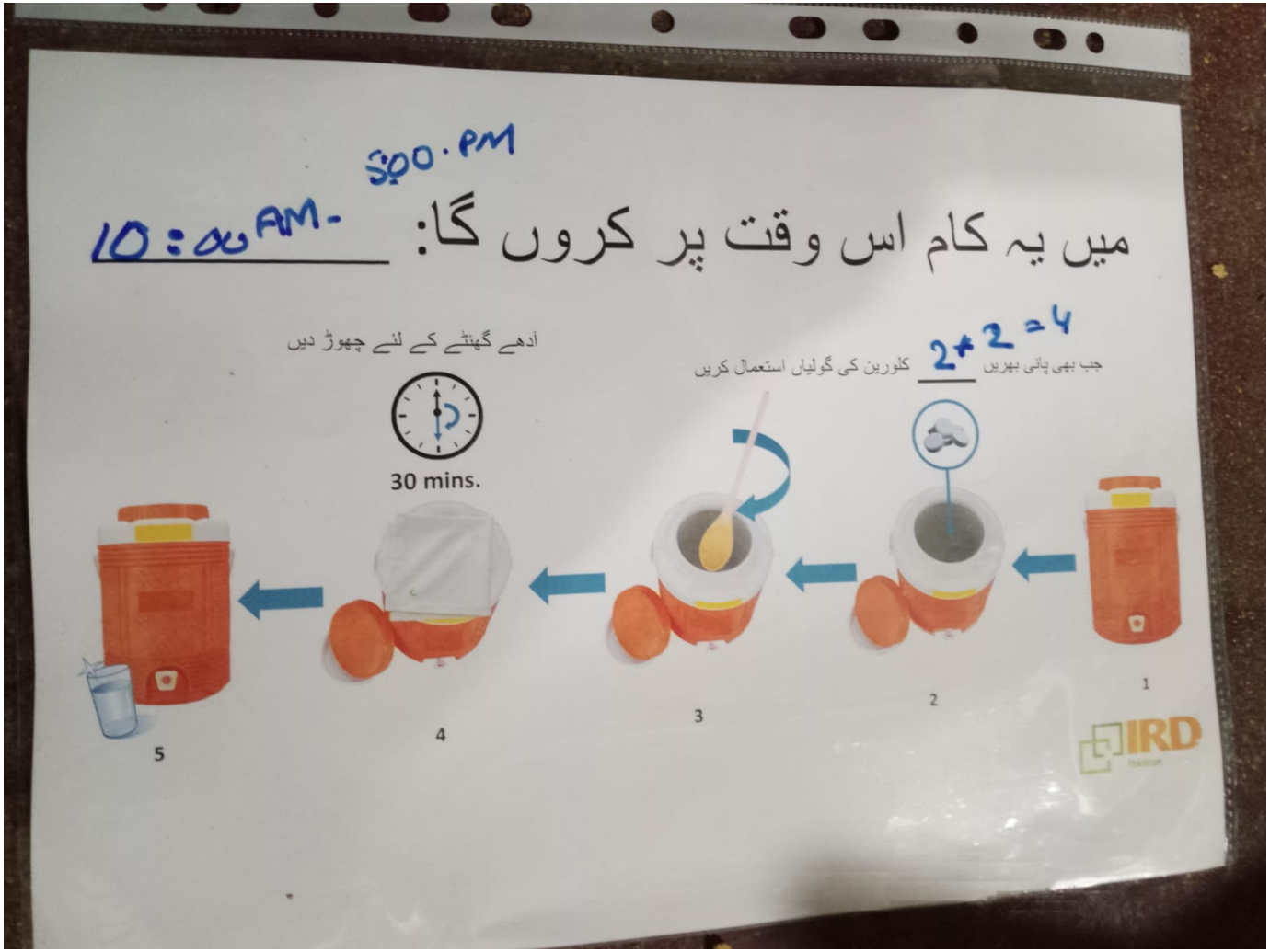


Figure 16: Leaflet with Instructions on Chlorination

8.6 Detailed Literature Review

This paper speaks to a small literature on the value of experiential and social learning in adopting health technology. There exists a large literature investigating how experts themselves learn about health from their own experiences, but this literature is more focused on physician skill development than on belief formation (Halm et al. (2002); Facchini (2022)). While health experts have the information that is most likely to be accurate on average, drawing from clinical trials and personal experiences, there are large inequities in access to experts (Dussault and Franceschini, 2006). Conditional on access to experts, there are still large inequities in how much trust people have in experts. Oftentimes, mistrust is a result of past wrongs committed by the medical community on groups with whom patients identify (Alsan and Wanamaker (2018); Lowes and Montero (2021); Martinez-Bravo and Stegmann (2022)), but there is evidence that knowing that doctors have financial incentives linked with the medical care they provide is enough to sow mistrust (Banerjee et al. (2023)).

Disseminating health information through laypeople or social networks may help solve the problems of trust and access associated with relying on experts to relay health information (Alsan and Eichmeyer (2024),Banerjee et al. (2019)). However, the most socially isolated people may be excluded from information transmission through the social network. Furthermore, the types of information that individuals have access to may differ in quality. Calónico et al. (2023) find that a clinically unsupported treatment for COVID-19 spread through Argentina in a pattern that follows rational learning from neighbors, suggesting that networks can facilitate the spread of medically dubious information. Similarly, Chen et al. (2022) find that access to well-informed networks can generate health inequities. Individuals in Sweden who have a doctor in the family invest more in preventive health, and consequently enjoy healthier and longer lives.

There is scant evidence of experiential learning as a method through which people form beliefs about the efficacy of health inputs and behaviors. Bennett et al. (2018) find that a hygiene course in Pakistan leads to more hygienic behavior when the course includes showing participants microbes under a microscope, suggesting that "seeing is believing." Corno (2014) finds that individuals in rural Tanzania are more likely to seek clinical care if they previously healed after utilizing clinical care, or if they previously *did not* heal after *foregoing* clinical care. Akram and Mendelsohn (2021), whose treatment arm we replicate in our learning arm, find that the Info-tool leads to an increase in water chlorination in the long run. We draw from a large literature in development economics showing that individuals learn about agricultural technologies through their own experiences and their neighbors' experiences (Foster and Rosenzweig, 1994; Hanna et al., 2014; Conley and Udry, 2010). Choosing an agricultural input is similar to choosing a health input in that it is a high-stakes, high-dimensional problem, subject to random shocks and with potential for misattribution.

Globally, 663 million people lack access to an improved source of drinking water while fecal contamination affects 1.8 billion people UNICEF and WHO (2015). Annually, contaminated

water contributes to 1.7 billion cases of diarrheal disease and 1.6 million deaths, including half a million under-five children, with most of the disease burden concentrated in developing countries (WHO, 2017). Safe drinking water results in better health, particularly among young children (Kremer et al., 2023), and yields long-term health and cognitive improvements (Scharf et al., 2014). Point-of-use decontamination technologies such as chlorine tablets can drastically reduce the burden of diarrheal disease (Reller et al. (2003), Quick et al. (2002), Quick et al. (1999)).

8.6.1 Habit Formation versus Learning

Recent literature on long-term behavior change recognizes habit formation, or the generation of complementarities in use across time, as playing a key role in sustained change (e.g., Becker and Murphy (1988), Wellsjo (2021), Celhay et al. (2015), Allcott and Rogers (2014), Royer et al. (2015), Aggarwal et al. (2020)). In both the case of habit formation and learning, higher initial use implies higher future use: in learning, due to initial exposure generating knowledge of the positive returns to a product and increasing the likelihood of future use; and in habit formation, due to intertemporal complementarities in consumption wherein greater initial consumption stock raises future desire to consume. With a few exceptions, however, these studies make no distinction between these two mechanisms, with persistent behavior change in the post-intervention periods often being explained by one mechanism without consideration of the other.

Recognition that these mechanisms may act in concert or be conflated with one another is rather recent in the literature. Caro-Burnett et al. (2021) examine sanitary latrine use in Kenyan slums and isolate the impact of habit formation interventions above and beyond that of learning by holding short-run use constant across treatment arms and examining only long-run behavior change. They find no detectable difference in long-run behavior across arms, though this is likely due to the study context (with sanitary latrine use less amenable to habit formation) and the nature of the interventions themselves (with a time-constrained subsidy intervention intended to embed the behavior in a habit loop, but challenging to do in practice given the potential unpredictability of defecation). Hussam et al. (2022) consider these mechanisms in the context of handwashing in rural West Bengal and attempt to distinguish habit formation from learning about returns by comparing households who experienced the same level of short-run financial incentives, but varied levels of health returns to their behavior. They find that those who experience larger improvements in health (whether between weeks or in aggregate) are no more likely to persist in their handwashing, suggesting that long-run behavior change is not driven by households independently engaging in learning about the value of handwashing from their health experiences. Alpízar et al. (2022) offer a conceptual framework for how these mechanisms may result in long-term adoption (theoretically distinguishing learning how to use a good and learning the returns to use from changes in taste via habit formation), but are unable to disentangle these channels empirically in an experiment which generates long run use of water-saving technologies through short run incentives to engage.

Yet the distinction between learning and habit formation is critical to policy, as each mechanism implies a substantively different behavior change process and therefore intervention design. Should learning about returns be more effective at generating sustained behavior change or technology adoption, policy design efforts should focus on developing information campaigns that make returns to a behavior explicit. Should habit formation be more effective in motivating long term change, resources may be better spent incentivizing high short-run engagement, yielding long term intertemporal complementarities, and embedding behaviors within habit loops (a la Duhigg (2012) and Neal et al. (2015)).

Preventive health behaviors are often mundane acts that require repetition in order to generate meaningful health impacts - a setting in which habit-formation may be especially relevant. Our experiment allows us to distinguish between the role of learning through salient signals and learning through early adoption (either through habit formation or through the accumulation of more signals) in the long-term adoption of chlorine tablets.

8.7 Audit Visits

Our main outcome, presence of chlorine residual in drinking water, is a good approximation for consistent use of chlorine tablets *unless* participants are more motivated to use chlorine prior to visits from enumerators, in anticipation of the enumerator's visit. There are several reasons we do not expect this to happen. T3 participants received monetary incentives based on the number of empty chlorine tablet wrappers they could present, rather than based on the presence of residual chlorine in the water. If T3 respondents did not like using chlorine tablets, they could throw away the tablets, present the wrappers, and still receive the prize. Thus, there is no pecuniary motivation for any group to make sure their water presents chlorine residual on the day of the enumerator's visit.

However, there may be social desirability motivations for participants to use chlorine in anticipation of a visit from an enumerator. It may still be difficult for participants to plan perfectly. Chlorine residual is only present for 24 hours, so participants would have to plan to the exact day. Furthermore, enumerators tested for chlorine residual once per month, rather than at every visit, and did not explain the results to participants except in the case of over-chlorination. When possible, enumerators were instructed to take a cup of water outside and test for the presence of chlorine residual without the caregiver observing, to further eliminate the possibility that caregivers expected a reward for water quality.

To be sure that caregivers did not plan chlorine use around enumerator visits, we conducted audit visits throughout the behavioral intervention period (Phase 1 Round 2). In these audit visits, an enumerator arrived unexpectedly to test for the presence of chlorine residual. Audits were conducted by a different enumerator than the enumerator who usually visited the household. During Phase 1, households were visited once every two weeks but chlorine was only tested once per month. We conducted audit visits during the visit period where water was not tested for chlorine residual, and visited selected households several days before or after their

assigned survey visit date.

In each month, we randomly selected sixty-five households with whom to conduct audit visits, or just under 5% of treatment group households per month. We were most concerned about anticipatory behavior among households who frequently showed presence of chlorine residual at regularly scheduled visits, but did not want to oversample so heavily from this group that there would be a treatment effect from extra visits. As such, when selecting households for audit visits, we stratified by the frequency of testing positive for the presence of chlorine residual and ensured that no household was audited more than twice.

8.8 Competing Mechanisms

8.8.1 More Interactions and Mimickry

It is possible that the novelty of the Info-tool leads participants to interact with one another more when they are both in this group, and that they simply mimic one another's behaviors rather than learn from their information. For example, perhaps they assist one another in filling out the chart. However, we do not have any evidence to suggest that participants interacted in this way, with only 12% of Info-tool respondents reporting that they ever discussed the Info-tool with anybody else. Info-tool participants' endline social networks were not larger than any other groups' (Appendix Table 20). Furthermore, we have no evidence of Info-tool participants mimicking one another on other behaviors, including take-up of other health programs that were present during our study trial, or choice of household savings technology (Appendix Table 23).

8.8.2 More Conversations

We asked participants if they ever discussed health or water purification with members of their social network. On average, participants discussed health with one person, and this did not differ across any treatment group. Control group participants were much less likely to discuss water purification than treated participants, but Info-tool participants were no more likely to discuss water purification with their social network connections than the other two groups. If anything, the Incentives group participants were the most likely to discuss water purification with others (Appendix Table 20).

8.8.3 Social Norms

We asked participants at endline if they believe a guest at their home would drink chlorinated water if they were to serve it. While the majority of the Control group believed a guest would accept chlorinated water (67%), there was a large effect of being in any treatment group on this measure (Appendix Table 24). However, the Info-tool group was no more likely to believe a guest would accept water than the other two groups. If anything, Incentives participants were the most likely to believe a guest would accept chlorinated water. We also asked participants

to report to the best of their ability how each of their network connections purifies their water. Control group participants were very unlikely to believe that their network connections purified their water with chlorine (6.4% believed one network connection purified their water with chlorine), but the rate was 92.9% higher for participants in any treated group. Again, Info-tool was no more likely to believe their network connections purified water than any of the other treated groups (Appendix Table 24).

8.9 Diarrhea

In each visit, we asked participants to recount the number of days in the last week that each child had diarrhea ("experienced motions"). Ultimately, we consider child anthropometrics as our key measure of child health because it is objective. However, since we have panel data on diarrhea, we can use our diarrhea data to understand how child health changed over time with the treatments. We analyze the aggregate of child-diarrhea-days across all children in the household, controlling for the number of children. Household total child-days of diarrhea is the same measure that Info-tool households recorded as a part of the Info-tool treatment.

Although we recorded data on child-days of diarrhea during the three months prior to the distribution of chlorine, we do not analyze data from this time period because we consider this a training period for the Info-tool group. Indeed, we see much higher rates of recorded diarrhea in the Info-tool group during this period, which we believe is likely over-reporting in response to the Info-tool treatment. To ensure uniformity across treatment groups in how respondents interpret and report loose stools, we used the Bristol stool chart to help caregivers identify loose stools at baseline. This chart provides illustrations of different potential stool consistencies. At endline, we showed households the chart again and asked them to identify which illustrations they would consider to be motions, to see if the treatments had changed participants' interpretation of what constitutes a loose stool. Although we expected the Info-tool treatment to be the most likely treatment to change participants' interpretation of loose stools (because the treatment would lead participants to more closely attend to children's stools), it was actually the Incentives group whose interpretation of loose stools appears to have changed. The Incentives group was more likely to consider illustrations of solid stools to be loose, and therefore might have overestimated their children's rate of diarrhea. There were no differences in the number of illustrations that Pure Control, Chlorine Only, and Info-tool considered to be loose stools. Consequently, we think that child-days of diarrhea is a good proxy for child health when comparing the Info-tool, Chlorine Only, and Control groups. Incentives households may have over-estimated their children's diarrhea rates, so we interpret comparisons between the diarrhea rates in the Incentives group and other groups cautiously.

Table 16 presents the results. Across all treatment groups, the impact of chlorine tablet dispensation is substantial: diarrhea rates drop by approximately 30% ($p < 0.005$) for Chlorine only, Incentives, and Info-tool households in the short run. These effects are largely sustained over the medium run, although Incentives households fall short, consistent with their chlorine

residual patterns. Treated households broadly continue to demonstrate large and significant reductions in child diarrhea into the long-run, again with the exception of Incentive households.

Average treatment effects suggest that chlorine provision effectively and substantially reduces child diarrhea rates, these effects persist over the course of a year, and those in the Info-tool arm experience the largest gains, followed by those in Chlorine Only, and lastly the Incentives arm. While these patterns are merely suggestive given the self-reported nature of the outcome, we do observe the same pattern of results in our objective anthropometric measures of child health that we collect at endline, which we turn to next. Our findings suggest that continuous chlorine use, even during the season where diarrhea poses the lowest risk (our medium-run period, during which Incentives households most drastically reduce their chlorine use), is important for building young children's stock of health. This pattern also speaks to the importance of tools to help participants attend to subtle health signals. Diarrhea is rare in this season, so differences in health with the introduction of chlorine will be more difficult to observe than during the season when diarrhea rates are higher.