STOCKHOLM SCHOOL OF ECONOMICS
HOUSE OF SUSTAINABLE SOCIETY

HOSS WORKING PAPER SERIES

NO. 2026-03, JUNE 2026

EXPERIENTIAL AND SOCIAL LEARNING

Agha Ali Akram
Bill and Melinda Gates Foundation

Gabriella Fleischman
University of Pittsburgh

Reshmaan Hussam
Harvard Business School

Akib Khan
House of Sustainable Society,
Stockholm School of Economics

House of Sustainable Society (HoSS) | Saltmätargatan 19B, 113 59 Stockholm

Experiential and Social Learning*

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

June 25, 2026

Abstract

How does a person's own learning experience affect their ability to learn from others? We conduct a field experiment on chlorination in Pakistan, where randomized "learning-arm" households use a tool to track their children's diarrhea before and after chlorine distribution. Learning-arm households with learning-arm neighbors chlorinate significantly more one year after the withdrawal of the tool, with children's health improving by 0.08 SD relative to all other households receiving chlorine. Neither learning households without learning-arm neighbors, nor non-learning households with learning-arm neighbors, exhibit sustained behavior change, results which have significant implications for intervention and evaluation design.

1 Introduction

Learning through shared experience is common, whether socialized through exploring new activities together or institutionalized through education in the classroom. In canonical models of learning about a new technology, however, what an agent learns is fully characterized by the signals she observes: their number, their sign, and their precision. Social learning enters these models as an additional source of signals, which may be noisier or harder to interpret than one's own (Foster and Rosenzweig, 1994), and which may need to arrive from multiple independent sources before an agent acts (Beaman et al., 2021). Common to these frameworks is the assumption that signals from different sources are

*We are thankful for comments and suggestions from Marcella Alsan, Lori Beaman, Emily Breza, Lucia Corno, Jon Denton-Schneider, Seth Gitter, Eliana La Ferrara, Mushfiq Mobarak, Martina Björkman Nyqvist, Frank Schilbach, Anna Tompsett, Zachary Wagner, and seminar participants at numerous conferences. Fatimeh Munawar, Syed Ali Rehan, and Ibadullah Channa provided invaluable research assistance. We gratefully acknowledge financial support from the United States Agency for International Development (USAID) and the World Bank. Special thanks to Jed Friedman, who played a critical role at the beginning of this project. This project would not have been possible without Interactive Research and Development (IRD) Pakistan. The contents are the authors' sole responsibility and do not necessarily reflect the views of USAID or the United States Government. The experiment was approved by IRD-IRB under IRD_IRB_2019_12_009 and was pre-registered under RCT ID AEARCTR-0003673.

†Akram: agha.ali.akram@gmail.com; Fleischman: University of Pittsburgh and corresponding author, gfleischman@pitt.edu; Hussam: Harvard Business School, rhussam@hbs.edu; Khan: HoSS, Stockholm School of Economics, akib.khan@hhs.se.

aggregated into a single posterior, so that a sufficient number of signals from others can substitute for one's own. We study a setting in which this assumption fails: when interpreting information requires familiarity with the process that generated it, learning from others is predicated on having first learned through the same experience oneself. We term this complementarity between learning from one's own experience and learning from others' "joint experiential learning," and ask: what is its role in the decision to adopt new technologies?

We study the role of joint experiential learning in the adoption of a health technology among 1800 caregivers of children in the peri-urban slums of Karachi, Pakistan. We conduct a randomized controlled trial where we randomize 900 households into two primary treatment groups, the "chlorine-only arm" and the "learning arm." Chlorine-only households receive free chlorine tablets for water purification for fifteen months. Learning-arm households receive the same free chlorine tablets and, for the first six months, a visual learning tool. This "Info-Tool," which builds on an intervention from Akram and Mendelsohn (2021) (hereafter, AM), is designed to help households learn about their household-specific health returns to chlorine use by tracking child diarrhea in the three months before and three months after free chlorine distribution.

We make three experimental design choices that allow us to identify the role of joint experiential learning as the mechanism underlying the tool's efficacy. First, while AM cluster-randomizes the treatment arms at the neighborhood block level, we randomize the treatment arms at the individual level. The variation in proximity to neighboring learning-arm households that we generate through individual randomization allows us to disentangle the effect of joint experiential learning from the effects of individual experiential learning and social learning in isolation. Second, we randomize 450 households to a pure control arm, which allows us to estimate the effects of free chlorine distribution alone and benchmark the magnitude of the effect of joint experiential learning against this policy-relevant estimate. Third, we randomize another 450 households to receive free chlorine tablets and small daily financial incentives for chlorine use during the first three months of chlorine distribution, which allows us to rule out habit formation as a competing mechanism to learning.

Throughout, our comparison group is chlorine-only households without learning-arm neighbors. To identify the individual experiential learning effect, we compare this group to learning-arm households *without learning-arm neighbors*. To identify social learning, we compare it to chlorine-only households *with learning-arm neighbors* — the cross-treatment spillover. To test for a complementarity between individual and social learning, we ask whether the difference between *learning-arm households with learning-arm neighbors* and the comparison group exceeds the sum of the individual-learning and social-learning effects.

We find that neither individual experiential learning nor social learning alone produces sustained behavioral change. After the withdrawal of the incentives and Info-Tool interventions, learning-arm households without learning-arm neighbors, who only have access to individual experiential learning, closely track chlorine-only households without learning-arm neighbors in their chlorine use. Chlorine-only households with a learning-arm neighbor, who only have access to social learning, behave similarly. In contrast, learning-arm households with learning-arm neighbors chlorinate at significantly higher rates over the course of the study (41% more; $p = 0.001$). The effect in this group is larger than the sum of the treatment effects of individual learning and social learning alone ($p = 0.031$), implying that there is a complementarity between experiential learning and social learning. In our context, experiential and social learning appear to *rely* on each other. Individual learning does not persist unless it is reinforced through neighbors' experiences, and caregivers cannot learn from their neighbors' experiences unless they have first gone through the same learning experience themselves.

The consequences of this complementarity between experiential and social learning are substantial. One year after the withdrawal of the behavioral interventions, children in learning-arm households with learning-arm neighbors exhibit the largest improvements in health. Households with chlorine alone, social learning alone (chlorine-only households with any number of learning-arm neighbors), or experiential learning alone (learning-arm households without learning-arm neighbors) each experience some improvement in an index of child anthropometrics relative to pure control households, who receive no chlorine: 0.10 SD, 0.04 SD, and 0.08 SD increases, respectively, that are not distinguishable from each other. In contrast, learning-arm households with learning-arm neighbors experience a 0.16 SD ($p = 0.005$) increase in child health relative to pure control after one year. When we pool all other households who receive free chlorine into one comparison group to maximize power and identify the policy-relevant effect, we find that these learning-arm households with learning-arm neighbors exhibit a 0.08 SD ($p = 0.071$) higher treatment effect in child health, nearly doubling the impact of free chlorine distribution on child health.

Beyond the direct public health consequences, this phenomenon of joint experiential learning has direct bearing upon how research design choices can shape treatment effect estimates: joint experiential learning implies that average treatment effects increase with treatment saturation. We assess how much joint experiential learning may impact treatment effect estimates by comparing our results with those from AM, who cluster-randomize their learning intervention at the neighborhood-block level with full treatment saturation in neighborhoods comparable to those we operate in. Fifteen months after withdrawing their Info-Tool treatment, AM finds a 197% increase in the rate of chlorine detection in Info-Tool households' water relative to households who only received free chlorine. In contrast, we find no effect of individual learning alone, suggesting that estimates of learning in fully saturated cluster-randomized trials risk conflating the effects of joint experiential learning with those of individual learning, and individually-randomized trials risk underestimating the impact of the policy if it were to be fully saturated in a community.

To discipline our exploration of the learning process that learning-arm households may be engaging in, we outline a model of joint experiential learning. Our model describes a mode of learning where interpreting information requires an understanding of the context in which the information is generated. Participants follow a Bayesian learning process when forming their beliefs about the efficacy of a health technology. We allow experience with the signal-generating technology to affect adoption through two channels: by sharpening the perceived precision of signals (and thereby shifting posterior beliefs) or by raising the weight participants place on those signals when deciding whether to act on their beliefs. Within our experiment, the learning intervention provides caregivers with firsthand experience in acquiring information through the Info-Tool. This is a key feature of our intervention: the Info-Tool not only communicates the impact of chlorine on health; it also engages users in the process through which this estimate is generated. Our theoretical framework posits that, by working with the Info-Tool themselves, these participants become more responsive to any signal generated through the Info-Tool, including those generated by their Info-Tool neighbors.

We first check if households are behaving in a manner consistent with Bayesian learning. If behavioral change is driven by the signals that inform posterior beliefs, then chlorine adoption should be heterogeneous by the sign of the health signals that participants observe. We use a machine learning algorithm on a host of baseline variables in our pure control group to identify which types of households experience health improvements during the time period when the behavioral interventions are administered. This enables us to predict which households in the learning-arm are ex-ante more likely to acquire positive signals about chlorine efficacy while using the Info-Tool and should therefore be more likely to alter their behavior. Indeed, during the three months when they receive chlorine tablets and use the Info-Tool, learning-arm households who are predicted to experience health improvements use chlorine significantly more ($p = 0.004$) than those not predicted to improve, suggesting that participants respond to the information that they observe in their environment. This alone is a necessary but not sufficient condition: households we identify through our ML algorithm may simply be the types of households who possess certain attributes correlated with behavioral change, a possibility we cannot rule out. We therefore go one step further and examine the relationship of Info-Tool households with their neighbors.

Specifically, we ask whether learning-arm participants respond to their learning-arm neighbors' health signals. This exercise exploits exogenous variation in Info-Tool treatment assignment: even if households may be endogenously exposed to certain types of neighbors with attributes correlated with higher chlorination, the likelihood that these neighbors are Info-Tool neighbors is itself random. Consequently, this exercise identifies the effect of random exposure to positive health signals generated by the Info-Tool. Indeed, we find that the within-treatment spillover effect is entirely explained by learning-arm neighbors whose diarrhea rates were ex-ante predicted to improve ($p = 0.002$). Reassuringly, we find that learning-arm households with *chlorine-only* neighbors whose health was predicted to improve do not measurably alter their behavior, implying that learning-arm households respond differentially to exogenous positive signals from neighbors when those signals are generated by the Info-Tool.

Why do learning-arm households respond more to signals generated by learning-arm neighbors? While we find no evidence of differential communication among learning-arm households nor differences in posterior beliefs about chlorine efficacy, we document that learning-arm households are differentially more aware of their learning-arm neighbors, that they trust other learning-arm households differentially more on their knowledge about child health, and that learning-arm households with learning-arm neighbors express differentially higher *motivation* to use chlorine for the purpose of improving their children's health.

Finally, we rule out a series of competing mechanisms. Using the incentives arm, we are able to rule out that behavior change is driven by habit formation through the accumulation of consumption (chlorination) stock, a competing mechanism to learning-through-adoption that is often conflated in the literature with the effects of learning (Caro-Burnett et al., 2021).[1] Participants in the incentives arm use chlorine tablets at higher rates in the short term, but converge to the chlorination rates of chlorine-only households immediately following the withdrawal of incentives, suggesting that water chlorination in our context is not a habit-forming activity. We rule out a variety of other competing mechanisms as explanations for the joint experiential learning we observe, including social network formation, differential comprehension of information about chlorine, mimicry, and evolving social norms.

Our work contributes to three bodies of literature. First, we build upon the empirical and theoretical literature on technology adoption and behavior change. We identify what, to our knowledge, is both the first theoretical formulation and the first empirical evidence that information acquired from individual experience and from social learning can interact multiplicatively to shape behavior.[2] A large body of literature identifies

[1]See Section H.2 for a detailed review of the existing literature.

[2]Our model differs from standard social learning models because both *who* information comes from and *the order* in which one receives information from different sources is critical. Standard models of

social learning—learning from others' information or experiences, which an individual did not herself learn from or partake in—as an important driver of technology adoption and behavioral change (Arrow, 1962; Beaman et al., 2021; Bikhchandani et al., 2024; Breza and Chandrasekhar, 2019; Conley and Udry, 2010; de Janvry et al., 2016; Dupas, 2014; Foster and Rosenzweig, 1994; Khandelwal, 2024; Kondylis et al., 2023). A growing body of work demonstrates that social learning is limited when agents are unwilling to share information, receive information from certain people, or internalize information acquired by others (BenYishay and Mobarak, 2019; BenYishay et al., 2020; Banerjee et al., 2024; Chandrasekhar et al., 2022; Conlon et al., 2022). Our study suggests that, when social learning is predicated on personal experience—for example, if social learning can effectively confirm beliefs established from personal observation, but is ineffective at conveying novel information—social learning from an information intervention will only arise as a complement to individual learning.

Second, we speak to the literature on the design of randomized controlled trials and information campaigns (Athey and Imbens, 2017; Duflo et al., 2007). Cluster-randomized designs with full treatment saturation are often motivated by the desire to eliminate cross-treatment spillovers and thereby identify individual treatment effects. However, in the absence of random variation in exposure to other treated units among treated units themselves, within-treatment spillovers are not identified, and the threat of a SUTVA violation remains (Hudgens and Halloran, 2008; List, 2025).[3] Notably, as we explore in detail in Appendix G, empirical economics research on technology adoption and communication utilizes experimental designs that are cluster-randomized with full treatment saturation more often than any other experimental design, concealing the potentially significant presence of within-treatment spillovers in their estimated treatment effects.[4] These lessons on research design have, in turn, important policy implications. When an intervention's success relies on within-treatment spillovers rather than individual treatment effects, achieving efficacy, even at the individual level, requires a sufficient degree of treatment saturation. In other words, the size of the within-treatment spillovers of an intervention has direct implications for how a social planner administers a policy. Existing social learning literature focuses on how information diffuses through a network,

simple and complex contagion are both consistent with our results, but *only among Info-Tool participants*: social learning becomes relevant to one's behavior only upon experiencing personal learning.

[3] Al-Ubaydli et al. (2019), in their model of scaling, formalize within-treatment spillovers as a distinct channel of the treatment effect and call for empirical measurement of their magnitude (see also List, 2025, ch. 16). Our study enables us to distinguish within-treatment spillover effects from individual treatment effects and compare these estimates with those of a cluster-randomized test of the same intervention. In AM's fully saturated cluster-randomized design, the estimated individual-learning effect bundles individual learning with joint experiential learning and is correspondingly large. Our individually randomized design lets us separate the two, and we find joint-learning effects that are, conservatively, six times larger than individual-learning effects alone.

[4] See Section G for an analysis of trials registered with the American Economics Association RCT Registry.

rather than the impact upon the seeds themselves or the value of experiential learning (Banerjee et al., 2019; Beaman et al., 2021; Chandrasekhar et al., 2022). *Experience cannot be diffused through seeding*, suggesting that there may be large efficacy costs to seeding, rather than saturating, certain types of information.

Lastly, the public health implications of this study are substantial. Clean water is a major determinant of both child and adult morbidity and mortality (Alsan and Goldin, 2019; Buchmann et al., 2019; Burlig et al., 2026; Kremer et al., 2023), and chlorination via tablets effectively cleans water. But sustainable, long-term behavior change required to generate meaningful health impacts from chlorination remains elusive (Ashraf et al., 2010; Dupas et al., 2016, 2023; Kremer et al., 2023; Luoto et al., 2011; Null et al., 2018). Our study suggests that, in the high-stakes context of slums in flood-prone Karachi, pairing chlorine tablet provision with a learning tool that is distributed to not just oneself but also to one's neighbor can double the health impact of free chlorine provision alone and sustain chlorine use at rates approximately 40% above those of other chlorine recipients one full year after the withdrawal of the learning tool.

The rest of the paper proceeds as follows. Section 2 describes the experimental design, sample, and data. Section 3 presents a model of joint experiential learning in technology adoption. Section 4 reports the effects of the intervention on chlorination, child health, and cost-effectiveness. Section 5 examines the mechanisms, and Section 6 concludes.

2 Experimental Design

We implement our randomized controlled trial in Ibrahim Hyderi, a peri-urban neighborhood of Karachi, Pakistan. We identify 1,800 eligible households through a census. Households are eligible if there is at least one child between the ages of six months and five years old at baseline and if the caregiver (most often the child's mother) is home during the day. We also screen on several dimensions of water usage to ensure that chlorination has the potential to affect children's health, detailed in Appendix Section A. Water supply comes from a variety of sources, with 79% of households accessing water from centrally delivered but untreated water piped into participants' own households or plots, or publicly supplied groundwater built by NGO projects. Households can also access water from central tanks and, in extreme circumstances such as low water supply and flooding, from tanker-trucks.

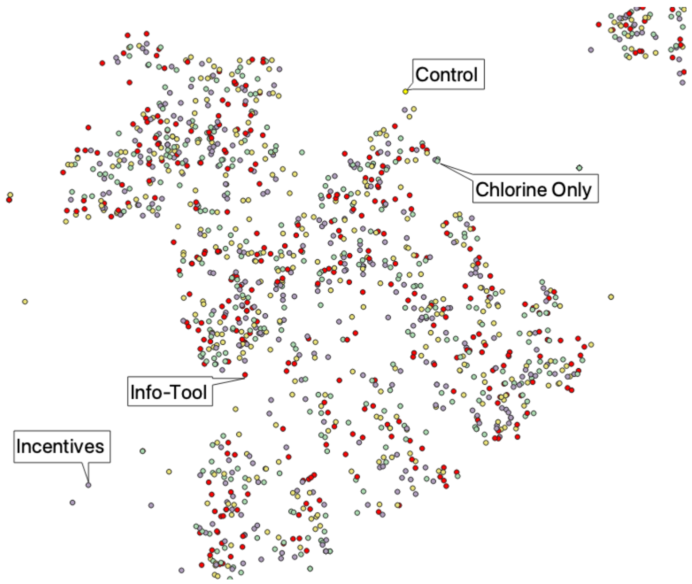


Figure 1: Distribution of Households Across Treatment Arms

Notes: Each point represents an eligible study household, and point color indicates randomized treatment assignment. The figure shows the spatial distribution of households across the four experimental arms within the study area.

2.1 Baseline Survey, Randomization, and Balance

Our baseline survey ran from May to June of 2022. We collect data on chlorine exposure, chlorine knowledge, diarrhea prevalence, and child anthropometrics.[5] At baseline, no household's water vessel is positive for chlorine residual, with 98% of households reporting that chlorine tablets are not available in local markets. Awareness of chlorine tablets in this community is likewise low at baseline: 20% of the sample report having heard of chlorine tablets, but 40% of this subsample cannot properly identify what chlorine tablets are used for. Despite low awareness, 87% report that they are open to using chlorine tablets after enumerators explain the purpose of chlorine for purifying water.

We electronically randomize households to treatment groups between the baseline visit and the first biweekly visit. Eligible households are individually randomized into the four 450-person experimental arms as follows (Figure 1):[6]

[5]See Appendix Section A for more details on how we measure self-reported diarrhea prevalence and child anthropometrics.

[6]Treatment groups are balanced on most, but not all, baseline measures. We control for unbalanced variables in all of our empirical specifications. In Table B.3, we compare our four main analytical samples: T1 households without T2 neighbors, T1 households with T2 neighbors, T2 households without T2 neighbors, and T2 households with T2 neighbors. At baseline, T1 households without T2 neighbors are more likely to have had any diarrhea in the past two weeks than T2 households without T2 neighbors, and are less likely to have heard about chlorine than T2 households with T2 neighbors. In Tables B.5 and B.7, we compare each treatment group with the pure control group. The same two variables are unbalanced in this comparison. 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 B.5), or when considering both a binary measure and the number of diarrhea-days at the child level (Table B.7). However, there are no significant differences

Control (C): No chlorine tablets (Pure control)

Treatment 1 (T1): Free chlorine tablets (Chlorine only)

Treatment 2 (T2): Free chlorine tablets + Info-Tool (Info-Tool)

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

<table>
<tr><th rowspan="3"></th><th rowspan="3">Month 1</th><th colspan="2">Phase 1</th><th rowspan="2">Phase 2</th><th rowspan="3">Month 19</th><th rowspan="3">Month 24</th></tr>
<tr><th>Round 1</th><th>Round 2</th></tr>
<tr><th>Pre-Treatment: Months 2-4</th><th>Treatment: Months 5-7</th><th>Post-Treatment: Months 8-18</th></tr>
<tr><td>Chlorine</td><td rowspan="4">Baseline Survey</td><td></td><td>Chlorine Distribution</td><td rowspan="3">Chlorine Distribution</td><td rowspan="4">Endline Survey 1</td><td rowspan="4">Endline Survey 2</td></tr>
<tr><td>Incentives</td><td></td><td>Chlorine Distribution + Incentives</td></tr>
<tr><td>Info-Tool</td><td>Info-tool Training</td><td>Chlorine Distribution + Info-tool</td></tr>
<tr><td>Control</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>Twice-monthly visits for child health data</td><td>Twice-monthly visits for child health and chlorine presence data</td><td>Once-monthly visits for child health and chlorine presence data</td><td></td><td></td></tr>
</table>

Figure 2: Study Timeline

Notes: This figure summarizes the sequence of study activities from baseline to endline. Phase 1 consists of biweekly visits during the behavioral-intervention period, including treatment-status revelation, Info-Tool training, chlorine distribution, and incentives delivery. Phase 2 consists of monthly follow-up visits after the behavioral interventions ended. The timeline is schematic and is intended to show the order and duration of major study activities.

2.2 Phase 1: Biweekly Visits

Phase 1, during which we introduce the behavioral interventions and visit households every two weeks, consists of two rounds. In Round 1, households are made aware of their treatment status and T2 households are trained on how to use the Info-Tool. In Round 2, we distribute chlorine tablets to all three treatment groups and deliver the incentives treatment to T3 households, while T2 households continue to use the Info-Tool. Figure 2 lays out the study timeline and key events. We describe in greater detail the activities of each round below.

in child anthropometrics, nor do we see lower levels of diarrhea among T1 households in the following visit (before households had received any treatment interventions), indicating these differences are likely spurious.

2.2.1 Round 1 of Phase 1

In the first Round 1 visit by enumerators, we reveal treatment status to caregivers. We explain each treatment group to caregivers and inform them that a lottery will determine their treatment status. We explain that the lottery will run through the enumerator's tablet, with the number that appears on the screen, which was pre-assigned to properly stratify by neighborhood block, determining their treatment status.

Throughout the remainder of Round 1 (June to August 2022), enumerators visit households every two weeks for brief rounds of data collection. Enumerators measure child diarrhea days over the previous two weeks, test for the presence of chlorine residual, and train caregivers in the T2 group on how to use the Info-Tool.

Info-Tool

The Info-Tool is a simple pen-and-paper chart that allows caregivers to track their children's diarrhea (Figure A.1). Each chart consists of two bars to represent each month: one bar in which caregivers fill 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 fills in a benchmark diarrhea rate. The benchmark diarrhea rate is 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).[7]

2.2.2 Round 2 of Phase 1

In Round 2 of Phase 1 (August to November 2022), we continue to visit households every two weeks and launch the distribution of chlorine tablets to all T1, T2, and T3 households.[8] We continue to collect information on diarrhea prevalence and test water for the presence of chlorine in all households. We test for chlorine residual every other visit (once per month). Incentives begin and the Info-Tool continues.

Incentives

Among households randomized into T3, caregivers are offered tokens redeemable for child and household goods based upon the number of chlorine tablet wrappers they

[7]When we use the diarrheal incidence in the Control group, the benchmark is updated daily to reflect the past fourteen days of data collection from the pure control households. We switched to the diarrhea rate in the epidemiological literature out of concern that the incidence of diarrhea in the Control group may be impacted by a reduced disease environment due to the presence of treated households in the community. Households were not informed of this change, and, regardless of how the benchmark was generated, enumerators explained to participants that the benchmark rate was the average rate of diarrhea from people in the community who do not use chlorine to purify their water.

[8]At distribution, all chlorine-recipient households also receive a pictorial leaflet with instructions on chlorination (Figure A.3).

show enumerators as proof of usage. The chlorine tablets come individually wrapped, and participants are instructed to save the empty wrappers in a pouch provided to all households in our experiment. Each daily reward for proper chlorine use is equal to approximately 0.05 USD (with 'proper use' calibrated to the household's pre-intervention water consumption). To hold income effects constant, we provide comparable products to participants in the remaining groups – *unconditioned* on chlorine tablet wrappers – framed as a token of appreciation for participating in our surveys. Since T3 households can redeem tokens for household goods that have varying values, we implement a lottery to determine the value of the unconditional gifts that non-T3 households receive.

Info-Tool

At the end of Round 2, we aggregate the monthly Info-Tool chart statistics across three-month intervals and present the aggregated data to T2 participants visually (Figure A.1). T2 participants are able to visualize their own children's diarrhea rate relative to the average diarrhea rate among households that do 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, they engage in a household-specific difference-in-differences exercise.

Throughout Round 2, we conduct random unscheduled audits where enumerators test water for the presence of chlorine residual. We conduct these audits to ensure that households are 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 Section A.2).

2.3 Phase 2: Monthly Visits

In Phase 2 (November 2022 to October 2023), we reduce our household visits to once monthly. We continue to distribute chlorine tablets and test water for the presence of chlorine residual, but we cease providing incentives or gifts, information to Info-Tool participants of the benchmark diarrhea rate, Info-Tool sheets, or assistance to Info-Tool participants to track their children's diarrhea rates. In other words, beyond free chlorine tablet provision, our interventions cease.

2.4 Endline Visits

In our first endline survey (December 2023 to January 2024), we collect all measures that we collected at baseline, including diarrhea rates, presence of chlorine residual, and child anthropometrics. We additionally elicit willingness-to-pay for chlorine tablets using a take-it-or-leave-it offer with a randomized price, and collect detailed social network data to understand potential mechanisms for spillovers.

We achieve an 87% follow-up rate in the first endline. Attrition is non-differential across Chlorine-only and Info-Tool households, with and without Info-Tool neighbors (Table B.3).[9] We conduct a second brief endline survey in June 2024 to provide participants with chlorine tablets for the summer months (the period in which diarrhea rates peak), monitor stockpiling of chlorine tablets, and ask additional survey questions to disentangle mechanisms. Taken together, we administer 26 surveys per household: a baseline survey, eleven rounds of bi-weekly visits in Phase 1, twelve monthly visits in Phase 2, one endline visit immediately after Phase 2, and a second endline visit four months later.

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. Enumerators test for chlorine residual in each household's drinking water vessel 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 detectable by the test strips is 0.5 parts per million). Chlorine residual only presents itself if the chlorine tablet was added to the water within twenty-four hours of testing, meaning we are likely to underestimate true chlorination given the preponderance of 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. However, this measure is subject to measurement and misclassification errors. The Info-Tool intervention is likely to change the *reporting* of diarrhea in children, since we explicitly encourage participants to track diarrhea within this arm. We may thereby be capturing a treatment effect on reporting, rather than a treatment effect on the actual incidence of diarrhea. Further, recent literature has suggested that the presence of non-infectious diarrhea and asymptomatic infection makes diarrhea presence a poor outcome for measuring the effectiveness of water and sanitation interventions (Watson et al., 2022).

Our preferred child health measure is instead child anthropometrics, which we measure directly, and which is not self-reported. Frequent diarrhea can result in under-absorption

[9] While attrition is non-differential across learning, chlorine, and pure control arms (our margins of comparison in our primary analyses), incentive households are 4.2 percentage points more likely to attrit at endline than pure control households ($p < 0.05$; Table B.5). Across a range of baseline observables, the endline incentives respondents do not statistically differ from their counterparts in the other arms (Tables B.4, B.6, and B.8).

of nutrients, leading to suboptimal physical development in children. We therefore collect height-for-age, weight-for-age, weight-for-height, and mid-upper-arm-circumference-for-age for all children under 80 months old.[10]

3 A Model of Joint Experiential Learning in Technology Adoption

Before turning to results, we offer a conceptual framework to motivate our empirical analysis. We present a Bayesian model of learning about a technology.[11] In contrast to standard Bayesian models, where adoption depends only on the sign, magnitude, and precision of the signals an individual receives, we allow experience with the signal-acquisition technology to enter the adoption decision, independent of the signal itself.

Let A be a distribution of prior beliefs about the efficacy of a health technology:

$$A \sim \mathcal{N}(\mu_0, \sigma_0^2)$$

In each period, individuals receive a health signal Y with precision given by σ_Y^2:

$$Y|A = a \sim \mathcal{N}(a, \sigma_Y^2)$$

Let $\gamma_Y \in \{\Gamma\}$ denote the source of the signal Y from a set of potential sources Γ (for example, the Info-Tool, community observation, or messages from an enumerator). Each individual accumulates experience with each technology over time, denoted $E(\gamma_Y)$. We assume that an individual's behavioral responsiveness to a signal depends on the technology that generates that signal and on the individual's cumulative experience with that technology.

Motivated by evidence from the psychology and behavioral economics literatures, we allow experience with signal-generating technologies to enter decisions through two distinct channels: it may affect how individuals update their beliefs about returns, and it may, separately, affect their willingness to act on those beliefs. Ex ante, the psychology and behavioral economics literatures offer evidence consistent with both channels, and we do not take a stance on the channel through which the Info-Tool may operate. Through the first channel, experience with a signal-generating technology reduces perceived uncertainty around signals produced by that technology, raising their weight in posterior beliefs (Hertwig et al., 2004). Through the second channel, experience raises an individual's willingness to act on a given set of beliefs without altering the beliefs themselves,

[10]These children are up to 80 months at endline because they were included in our baseline survey if they were up to 60 months old.

[11]We follow the notation of Kondylis et al. (2023).

consistent with models of behavioral change in psychology where motivation, not just knowledge, is a necessary prerequisite for action (Ryan and Deci, 2000; Ajzen, 1991; Sheeran and Webb, 2016).[12]

To model the first channel, we allow perceived signal precision $\hat{\sigma}_Y^2$ to be a function of *actual* signal precision σ_Y^2 and prior experience with the signal-generating technology:

$$\hat{\sigma}_Y^2 = \lambda(\sigma_Y^2, E(\gamma_Y))$$

with $\frac{\partial \lambda}{\partial E(\gamma_Y)} \leq 0$. Individuals update their posterior beliefs about the efficacy of the health technology according to Bayes' rule:

$$A|Y \sim \mathcal{N}(M, \hat{\sigma}^2)$$

where $M \equiv \ell Y + (1-\ell)\mu_0$ represents the updated expected returns to using the technology (in our case, M represents posterior beliefs about chlorine efficacy). The weight $\ell \equiv \sigma_0^2/(\sigma_0^2 + \hat{\sigma}_Y^2)$ represents the weight that agents assign to signal Y in their posterior beliefs, considering only the perceived precision of signals. Then, participants' updated uncertainty is $\sigma^2 \equiv (1-\ell)\sigma_0^2$.[13]

To model the second channel, we introduce "experience weights" $\alpha_Y \in [0, 1]$, which affect action independently of beliefs. As with perceived signal precision, experience weights depend on experience with the signal-acquisition technology. Formally, the experience weight associated with signal Y is:

$$\alpha_Y = \alpha\big(\gamma_Y, E(\gamma_Y)\big),$$

with $\frac{\partial \alpha}{\partial E(\gamma_Y)} \geq 0$. That is, the more familiar or practiced an individual is with a given signal-acquisition process, the more she relies on signals produced by that process when deciding whether to act.

M represents *stated* posterior beliefs about the returns to the technology (or explicit knowledge). We define $M_\alpha \equiv \ell\alpha_Y Y + (1-\ell\alpha_Y)\mu_0$ to be the experience-weighted posterior belief. An individual adopts the technology when $M_\alpha > C$, where C is the cost of adopting the technology.[14] When $\alpha_Y = 1$ and $\hat{\sigma}_Y^2 = \sigma_Y^2$, individuals behave as in the canonical

[12]Moreover, there are numerous phenomena from the behavioral economics literature that we permit to operate through 'willingness to act'. For example, people may develop psychological ownership over the information-generating process (Morewedge, 2021; Conlon et al., 2022), derive utility from acting consistently with an identity formed around using a technology (Akerlof and Kranton, 2000), or interpret actions taken on the basis of personally acquired information as signals about their own type or competence (Bénabou and Tirole, 2002; Bodner and Prelec, 2003).

[13]The only departure from a standard Bayesian model here is that we allow individuals' beliefs to depend on their perception of the signal precision, rather than the true signal precision.

[14]In a standard model, participants adopt the technology when $M > C$, with $\hat{\sigma}_Y^2 = \sigma_Y^2$.

Bayesian model, acting fully on their updated beliefs, which depend only on true features about signals Y. When $\hat{\sigma}_Y^2 \neq \sigma_Y^2$, experience with signal-generating technologies governs how readily individuals update their beliefs given a signal realization Y. When $\alpha < 1$, inexperience with signal-generating technologies inhibits individuals from acting on a given set of beliefs formed through signals generated by those technologies.

At its most extreme, when $\alpha_Y = 0$, individuals are sufficiently untrusting, inexperienced, or detached from the signal-acquisition process that they fail to act on the signal altogether, even if they superficially believe and recognize a signal to be positive and precise. When experience weights are operational—that is, when $\frac{\partial \alpha}{\partial E(\gamma_Y)} > 0$—people can intellectually understand and articulate the returns to a technology, but it is their experience-weighted beliefs that affect their decision to act on that knowledge. The function $\alpha(\cdot)$ can reflect a wide range of behavioral processes.

The objective of our model is to demonstrate how experience with a signal-generating technology can give rise to a complementarity between experiential and social learning through a shared learning process, not to estimate and compare ℓ with α_Y. Crucially, perceived signal precision $\hat{\sigma}_Y^2$ and experience weights α_Y are defined at the level of the *technology*, not the signal realization. Regardless of which channel operates, two individuals who receive the same signal Y may behave differently if one has prior experience with the technology that produced it. However, in Section 5, we offer empirical evidence for both channels via stated beliefs (representing ℓ) and stated motivations (or, willingness to act on beliefs, representing α_Y).

First period adoption

Initial priors are diffuse: at baseline, 12.7% of respondents report having ever heard of chlorine and knowing that its purpose is water purification. We therefore consider information sent by the enumerator as the first signal that participants receive about chlorine, Y_{ENUM}. It is likely that $\sigma^2_{Y_{ENUM}}$ is low, as the enumerators represent a known NGO and they have been visiting households for three months prior to chlorine distribution. It is also likely that C is close to zero in the first period, since the largest reported cost associated with using chlorine is an unpleasant taste if the chlorine dose is incorrectly titrated, which caregivers would not yet have experienced.

Interventions

<u>All chlorine groups:</u> We reduce C, the cost of using chlorine, in all time periods through free distribution and delivery.

<u>Incentives:</u> We further reduce the cost of using chlorine C in the Incentives group for the

first three months of chlorine distribution by providing monetary incentives for use.

Info-Tool: We change the signal-acquisition technology γ_Y in the first three months of chlorine distribution. This in turn raises cumulative experience with the Info-Tool, $E(\gamma_Y) \mid \gamma_Y = \textit{Info-Tool}$, in perpetuity. Then, any signal Y personally observed via the Info-Tool or sent by other Info-Tool participants (when the recipient is aware that the sender observed their signal through the Info-Tool) has a weakly lower $\hat{\sigma}_Y^2$ and weakly higher α_Y.

Comparative statics

Prediction 1 (Incentives: Shock to C): *Increased contemporaneous adoption relative to chlorine only households*

A negative shock to C increases the probability in any period that $M_\alpha > C$, leading to increased contemporaneous chlorine adoption. This effect is immediate from the first period of chlorine distribution but fades out as soon as the enumerator stops providing chlorination incentives.

Prediction 2 (Info-Tool: Shock to γ_Y): *Increased contemporaneous adoption relative to chlorine-only households among households with larger positive early-period health signals, with fade-out*

The Info-Tool leads to a change in the signal-acquisition technology in the first three months of chlorine adoption. Consequently, participants place a contemporaneously higher weight ℓ or α_Y on signals acquired through observation during the treatment period. After the enumerator stops assisting participants in the Info-Tool, costs are the same for all groups, but Info-Tool participants' posterior M or experience-weighted posterior M_α is higher. Thus, we should observe higher average use following the withdrawal of the Info-Tool.

Since individuals learn from short-term health signals, we should expect heterogeneity by the sign on individual contemporaneous health signals, and the Info-Tool group should be the most sensitive to differences in health signals. This effect should fade out as all participants acquire new signals—signals upon which the Info-Tool group, no longer using the tool, places no special weight.

Prediction 3 (Info-Tool: Shock to $E(\gamma_Y) \mid \gamma_Y = \textit{Info-Tool}$): *Increased adoption relative to chlorine-only households among Info-Tool households with connections to other Info-Tool households*

As soon as an Info-Tool participant has used the Info-Tool, any other Info-Tool-acquired signals, including those acquired by *other* Info-Tool participants, likewise carry additional weight. Info-Tool caregivers with an Info-Tool member in their network will be more responsive to those connections' health signals than caregivers in other treatment groups. This can have effects after the Info-Tool period has ended, if participants share cumulative or historical, rather than contemporaneous, information about their experiences.

Simulated Model

As a check on the plausibility of our theoretical framework, in Appendix F, we plot chlorine adoption patterns predicted by our model using simulated data. Under a set of reasonable parametric assumptions outlined in Appendix F, we document adoption patterns that closely follow our empirical results, which we present below.

4 Results: Chlorination

4.1 Spillover Definition

First, we define our spillover measure. Because we conduct randomization at the individual level, individuals are randomly exposed to Info-Tool neighbors. We consider an individual 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 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) over $r \in \{20(20)200\}$ meters for each outcome. 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 within the Chlorine Only and Info-Tool arms (Tables B.3 and B.4).

Households are, on average, 13.6 meters from the closest other participant; 16.0 meters from the closest participant in any chlorine treatment group; and 27.5 meters from the closest Info-Tool participant.[15] Forty-three percent of participants have a "nearby neighbor", or a participant within 20 meters of the respondent, who is in the Info-Tool

[15] The median distance to the closest household is 10.4 meters; the median distance to the closest participant in any chlorine treatment group is 12.1 meters; and the median distance to the closest Info-Tool households is 22.7 meters.

group. When we recenter this measure to account for endogenous neighborhood density (following Borusyak and Hull (2023)), 35% of participants are exposed to a nearby neighbor in the Info-Tool group by random variation in the treatment assignment (Table B.2 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 are close neighbors with other Info-Tool participants, while 2% of Chlorine Only participants are.[16] This means both that, in our trial, Chlorine Only 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 conceals Info-Tool treatment effects; and if there is a learning spillover from Info-Tool participants onto other Info-Tool neighbors, the treatment is less powerful in our setting.

4.2 Empirical Specification

We test for the presence of learning spillovers using the following specification:[17]

$$\begin{aligned} Y_i =& \theta_0 + \theta_{1,r} T1_i + \theta_{2,r} T2_i + \sum_{r=20}^{R} \theta_{3,r} AnyT2_i^r \times T1_i + \sum_{r=20}^{R} \theta_{4,r} AnyT2_i^r \times T2_i \\ &+ \sum_{r=20}^{R} \theta_{5,r} AnyT2_i^r + NV_i + X_{i0} + \gamma_b + \epsilon_i \end{aligned}$$

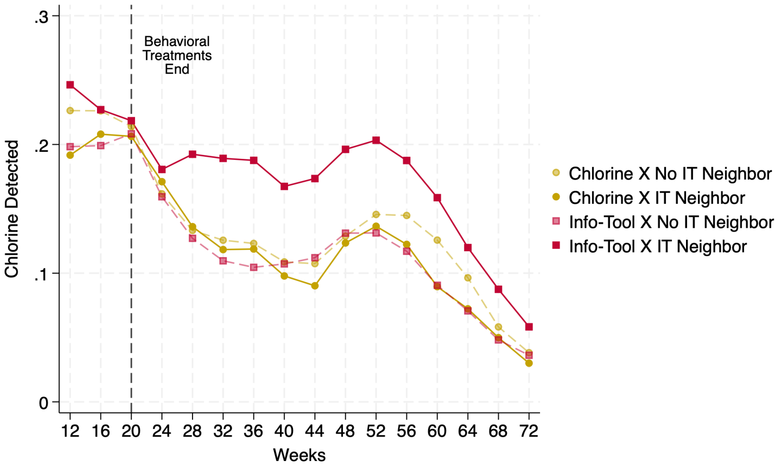
where Y_i is the total number of visits with chlorine presence detected in the water; $T1_i$ and $T2_i$ are binary variables representing Chlorine Only and Info-Tool participants, respectively; NV_i captures the number of visits with non-missing outcome measurements; X_{i0} is a vector of unbalanced baseline covariates, the density of study participants within twenty meters, and other covariates that we select using the double-lasso method proposed by Urminsky et al. (2016); and γ_b are block fixed effects (geographic units of stratification). Then $\theta_{2,r}$ is our estimate of individual learning; $\theta_{3,r}$ is our estimate of social learning; and $\theta_{4,r}$ is our estimate of the complementarity between individual and social learning.

To maximize power, we use a household-level specification where the outcome is the total number of times we detect chlorine across all visits in each time period, following

[16] There are no Incentives or Pure Control groups in AM.

[17] We pre-specified an investigation of learning spillovers by conducting a heterogeneity analysis by exposure to Info-Tool households. In our pre-specified empirical estimation, we specified analyzing the impact of the number of Info-Tool neighbors on the whole sample. We did not pre-specify interacting exposure to Info-Tool neighbors with individual treatment status, which is the main empirical specification we use. Estimating the spillover within each treatment group is a natural extension to the empirical estimation we pre-specified.

Figure 3: Seasonal Time Trends in Chlorine Detection (Moving Averages, Raw Data) by exposure to any Info-Tool neighbor



Notes: This figure plots raw moving averages of chlorine detection over time by treatment and spillover status. We code a participant as having an Info-Tool neighbor if they are more exposed to neighbors randomized to Info-Tool within 20 meters than they would be in expectation based on endogenous spatial factors. The vertical dashed line marks the end of the behavioral-intervention period.

McKenzie (2012).[18] This specification requires imputing zeros for survey visits where a household attrits, which would bias estimates if treatment groups participate in our monthly surveys at different rates. Survey participation rates do not differ significantly across treatment groups, but we cannot rule out that attrition is differentially correlated with true non-chlorine-use. We take a data-driven approach to address this concern and minimize attrition: using lasso, we find that endline survey participation is the strongest predictor of survey participation throughout the study, including the period over which we aggregate chlorine use.[19] The subsample of respondents who participate in the endline survey is also a natural sample for our analysis since our child anthropometric outcomes, willingness-to-pay measures, social network data, and stated-belief modules are all collected at endline. Within this subsample, baseline characteristics and survey participation rates are balanced across treatment groups. We report estimates for the endline sub-sample and the full sample, and find consistent results.

4.3 Results: Chlorine Use

We detect chlorine residual in the water of over 20% of treatment households while the behavioral interventions are ongoing (Figure 3). Chlorine use declines among participants in the Chlorine Only arm—whether or not they have an Info-Tool neighbor—and participants in the Info-Tool treatment group without any Info-Tool neighbors. Info-Tool

[18]For robustness, we also conduct our regressions using a household-survey panel specification, where the results hold (Appendix Section E).

[19]We report the full set of lasso coefficients in Table B.1 and survey participation patterns by endline-participation status in Figures ?? and ??.

participants *with* Info-Tool neighbors continue chlorinating at a stable high rate until week 52 of the experiment, after which their rate of use also slowly and steadily declines. While our trial is not designed to identify the causes of chlorine use patterns across time, the serial pattern follows diarrhea rates, with all groups chlorinating at higher rates during the high-diarrhea season.[20]

Table 1 presents the regression analog. Proximity to an Info-Tool household alters own chlorination behavior only among households who are themselves in the Info-Tool group. As is evident in Figure 3, this pattern emerges immediately after the behavioral treatments conclude and persists through the end of the study. Note that, in this period, Info-Tool households are no longer collecting any new data through the Info-Tool. Over this time period, Info-Tool participants with an Info-Tool neighbor chlorinate 30% more often ($p = 0.006$) than participants with no Info-Tool neighbor. While we use an indicator for *any* random exposure to Info-Tool neighbors in our main specifications, Figure D.2 documents that chlorine use is indeed increasing in the number of Info-Tool neighbors an Info-Tool household is proximate to. Reassuringly, in Appendix C, we reanalyze the data from Akram and Mendelsohn (2021), implementing our definition of proximity to an Info-Tool neighbor, and find comparable estimates for this within-treatment Info-Tool spillover effect.[21]

Three pieces of evidence underscore the presence of joint experiential learning among Info-Tool households. First, we observe no meaningful differences in chlorine use (statistically or in magnitude) by proximity to an Info-Tool household in the Chlorine Only group (Table 1 and Figure D.4).[22] Second, we replicate our analysis to test for spillovers from

[20]Overall trends in chlorine use might be explained by seasonality and by experimental changes:

Seasonality: Diarrheal prevalence peaks in the summer, during the hottest and rainiest months. We began our study at the peak season, which allows chlorine use to generate starker changes in diarrhea rates and gives the Info-Tool a greater chance of showing the efficacy of chlorine. Chlorine use tracks closely with self-reported diarrhea (Figure D.1). In the second summer of the experiment, diarrhea rates remain relatively low. This could be due to herd immunity generated via the intervention, or a milder monsoon season.

Experimental Changes: Chlorine rates fall sharply after the behavioral interventions end (week 20), including in the Chlorine Only group, which did not receive a behavioral intervention. There were two experimental changes in week 20: (1) we began visiting households every month, rather than every two weeks; and (2) we stopped providing participation gifts to all households. Towards the end of the study, it is possible that households stockpiled chlorine tablets, knowing that they would soon lose access to the free supply we offered and that the high diarrhea season was still several months away. We found some evidence of stockpiling in our second endline: 44% of respondents reported that they ran out of their chlorine tablets later than they should have if they were chlorinating daily, and for 24% of households the discrepancy was by more than one month.

[21]We are only able to replicate the within-treatment spillover because, as noted earlier, AM's cluster-randomized design leaves essentially no variation in Info-Tool exposure among Chlorine Only households. Only 2% of their Chlorine Only households have an Info-Tool neighbor within twenty meters, so the cross-treatment spillover is not identified in their data — a comparison our individual-level randomization is designed to enable.

[22]Nor do we observe differences by exposure to Info-Tool households in the Incentives group. We report Incentive group outcomes in Section J.0.1.

Table 1: Chlorine Detection Post-Behavioral Interventions
by exposure to any Info-Tool neighbor

	(1) Full Sample		(2) Endline Sample	
Chlorine × Spillover	0.071	(0.172)	0.055	(0.180)
Info-Tool × No Spillover	0.055	(0.142)	0.051	(0.149)
Info-Tool × Spillover	0.528***	(0.170)	0.646***	(0.178)
Observations	1282		1191	
Chlorine No Spillover Mean	1.543		1.558	
P-values:				
Info-Tool: Spillover = No Spillover	0.006		0.001	
Info-Tool × Spillover = Info-Tool × No Spillover + Chlorine × Spillover	0.092		0.031	

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 post-behavioral intervention time period. The Chlorine Only treatment group without spillover neighbors is the omitted group. An indicator for Pure Control participants, who are included in the sample, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor (someone within 20 meters) than they would be in expectation based on endogenous spatial factors. See Section 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

proximity to Chlorine Only or Incentives households, rather than Info-Tool households. As documented in Tables D.1 and D.2, there is no time period in which we find spillover treatment effects from Chlorine Only participants or Incentives participants onto any treatment group, including their own. Third, we show that, among non-Info-Tool participants, chlorine use remains unchanged by, not only the presence, but the number, of Info-Tool neighbors they are proximate to (Figure D.2). These tests rule out alternative mechanisms driven by changes in the diarrheal disease burden in the local environment or observation of others using chlorine, strengthening the interpretation that Info-Tool neighbors impact one another through shared experience with a health signal technology.

4.4 Results: Child Health

4.4.1 Diarrhea

Diarrhea rates over time offer a lens into how the various interventions impact children's health over the course of the experiment. Reductions in diarrhea rates mirror relative levels of chlorine use throughout the experiment (Figure D.1), with the Info-Tool group experiencing the largest reductions in diarrhea prevalence in the long-run.[23] More details on measurement and results on child diarrhea are reported in Section I. However, as a self-reported measure, children's diarrhea rates are vulnerable to respondent bias; that it is directly connected to the subject of the Info-Tool also leaves room for concerns of endogeneity in outcome reporting.

[23] During the behavioral interventions, treatment group participants report 36% (Chlorine Only), 26% (Info-Tool), and 38% (Incentives) reductions in diarrhea relative to the Pure Control group (Table I.1).

4.4.2 Child Anthropometrics: ITT Estimates

As an objective and thus preferred measure of health, we measure child anthropometrics at endline. To analyze the impact of our treatment on child anthropometrics, we use the following specification:

$$Y_{c,h} = \beta_0 + \beta_1 T1_h + \beta_2 T2_h + \beta_3 AnyT2_h^{r=20} \times T1_h + \beta_4 AnyT2_h^{r=20} \times T2_h + X_{h0} + \gamma_b + \epsilon_{c,h}$$

where $Y_{c,h}$ is health outcome Y for child c in household h, $T1_h$ indicates if household h is in the Chlorine Only group, $T2_h$ indicates if household h is in the Info-Tool group, $AnyT2_h^{r=20}$ indicates if a household is more exposed to an Info-Tool neighbor within 20 meters than in expectation, and X_{h0} are household-level baseline control variables. We cluster standard errors at the household level for all child-level specifications. Following Anderson (2008), we create a summary index combining height-for-age (HAZ), weight-for-height (WHZ), weight-for-age (WAZ), and MUAC-for-age z-scores (mid-upper arm circumference-for-age).

Table 2 presents the results, and Appendix Table I.4 disaggregates by the subcomponents of the index. As we show in Column 1, relative to chlorine-only households who have no Info-Tool neighbors, those in the Info-Tool treatment who have Info-Tool neighbors exhibit a 0.066 SD improvement in their child anthropometric index. No other group (chlorine households with Info-Tool neighbors or Info-Tool households without Info-Tool neighbors) exhibits health improvements. As health measurements are often noisy, these estimates remain imprecise. To increase power, Column 2 pools all chlorine-only households and Info-Tool households without an Info-Tool neighbor into the omitted category; under this specification, the child anthropometrics effect size of being an Info-Tool household with an Info-Tool neighbor is 0.079 SD ($p < 0.10$). If we further increase power by pooling Incentives households into the omitted category (Appendix Table I.5), the impact of being an Info-Tool household with an Info-Tool neighbor increases to 0.10 SD with high precision ($p < 0.05$).

4.4.3 Child Anthropometrics: TOT Estimates

Treatment compliance is imperfect: we detect chlorine in the water of treated participants 13% of the time across the study, yet 37% of treated participants self-report using chlorine to treat their water at endline. To recover treatment-on-the-treated (TOT) estimates, we instrument for self-reported use of an effective water purification method (boiling, bleaching, or chlorinating; 48% of the sample) with treatment assignment. We use self-reports rather than objective chlorine detection for two reasons. First, chlorine presents in water only if added within 24 hours of testing, so objective detection captures a narrow flow outcome rather than the cumulative practice relevant for child health. Second, pure con-

Table 2: Child Index of Anthropometry (ITT):
where "Spillover" is exposure to any Info-Tool neighbor

	(1)		(2)		(3)	
	Omitted: Chlorine × No Spillover		Omitted: Other Chlorine Only & IT Groups		Omitted: Pure Control	
Chlorine × No Spillover					0.099**	(0.045)
Chlorine × Spillover	-0.051	(0.060)			0.040	(0.056)
Info-Tool × No Spillover	-0.012	(0.049)			0.079*	(0.044)
Info-Tool × Spillover	0.066	(0.060)	0.079*	(0.044)	0.158***	(0.056)
Observations	2616		2616		2616	

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, the number of study participants within twenty minutes, and lasso-selected controls. The regression in column (1) controls for indicator variables for being in the following groups: Incentives × No Spillover, and Incentives × Spillover. The regression in columns (2) and (3) control for indicator variables for being in the following groups: Control, Incentives × No Spillover, and Incentives × Spillover. 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 20 meters) than they would be in expectation based on endogenous spatial factors. See Section 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

trol participants are more likely than treated participants to boil their water at endline, suggesting that the study may have prompted purification behavior *other* than chlorination. Defining compliance via an extensive-margin measure of effective purification treats these control boilers as always-takers and preserves the exclusion restriction.[24]

We estimate three specifications that correspond to three distinct complier samples: households induced to use an effective water purification strategy because they were randomized into (1) any treatment group, (2) Info-Tool treatment, and (3) Info-Tool treatment with random exposure to an Info-Tool neighbor. Because chlorine use is persistent, endline reports of purification in narrower complier samples correspond to higher cumulative chlorine use throughout the study: a complier in sample (3) has, on average, purified water on more days than a complier in sample (1). Because health is a stock, we therefore expect larger IV estimates in the more selective complier samples.

Table 3 reports the results. Among compliers in any treated arm (Column 1), water purification increases the child anthropometric index by 0.24 SD.[25] Among Info-Tool compliers (Column 2), the estimate rises to 0.33 SD — 37% larger than the pooled-treatment estimate — consistent with the higher cumulative chlorination we detect in this group. Among Info-Tool compliers with an Info-Tool neighbor (Column 3), water purification increases the anthropometric index by 0.51 SD, 54% larger than the average

[24] Among respondents who report chlorinating their water at endline, we detect chlorine in only 3.8%, confirming that our self-report measure captures broader purification behavior than a single-visit chlorine test would.

[25] In Table 3, regressions for columns (2) and (3) control for assignment to the Incentives and Chlorine Only arms; column (3) additionally controls for being in the spillover sample. Appendix Table I.6 reports the IV treatment effects on each index subcomponent using this same complier sub-sample.

Table 3: IV Impacts on Endline Child Health:
where "Spillover" is exposure to any Info-Tool neighbor

	(1)	(2)	(3)
	Instrument: Any Treatment	Instrument: Info-Tool	Instrument: Info-Tool × Spillover
Boils, Bleaches, or Chlorinates Water	0.240**	0.329***	0.506***
	(0.113)	(0.126)	(0.185)
Observations	2616	2616	2616
Endline Control Mean	-0.017	0.023	0.031
First-Stage Coefficient	0.312	0.329	0.312
Weak-IV robust F statistic	119.931	86.571	35.258
C-statistic p-value	0.027	0.005	0.002

Standard errors in parentheses
Child-level cross-section of the endline survey. Standard errors are clustered at the household level. Column 1 uses assignment to any treatment arm as an instrument for whether the household reports boiling, bleaching, or chlorinating water at endline. Column 2 uses assignment to the Info-Tool arm as the instrument and controls for assignment to the Incentives and Chlorine Only arms. Column 3 uses assignment to the Info-Tool arm interacted with spillover exposure as the instrument, controls for assignment to the Incentives and Chlorine Only arms, and additionally controls for being in the spillover sample. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, neighborhood block fixed effects, the number of study participants within twenty meters, 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 20 meters) than they would be in expectation based on endogenous spatial factors. See Section 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Info-Tool effect and more than twice the pooled-treatment effect.[26] The gradient across columns—larger effects in more selective complier samples—is consistent with joint experiential learning raising the intensity of chlorine use, and with that intensity translating into larger health gains for the children of complier households.

We note two threats to the exclusion restriction in this exercise. First, treatment may induce other sanitation behaviors that affect child health, in which case our estimates capture the effect of purification bundled with correlated changes. Second, Info-Tool households with Info-Tool neighbors benefit from both own chlorine use and neighbors' use, so part of the effect that we estimate in Column (3) may reflect herd protection rather than the direct effect of household-level purification. Both channels underscore why treatment saturation matters for policy—a topic we return to in Section 6—but they do caution against interpreting the estimate in Column (3) as simply the individual treatment effect of water purification.

4.4.4 Benchmarking against the literature

Comparing our results on child anthropometrics with those from other recent programs in South Asia, we find our estimates for weight-for-age to be significantly larger than those from handwashing, hygiene, nutrient supplements, or early childhood education programs (Table 4) (Hussam et al., 2022; Bennett et al., 2018; Sazawal et al., 2013; Soofi et al., 2022; Bos et al., 2024). In contrast to several of these alternative programs, we do

[26] Table I.7 reports the corresponding estimates for the four components of the index.

Table 4: Benchmarking Child Health Estimates

Intervention	Paper	HAZ	WHZ	WAZ	Muac-for-age
Chlorine: Average	Appendix Table I.6	0.050	-0.017	0.371*	0.198
Chlorine: Saturated Info-Tool	Appendix Table I.7	-0.083	0.584	0.789**	-0.428
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	–

Notes: Entries report point estimates for child anthropometric outcomes from this paper and selected studies from South Asia. The two rows from this paper correspond to the IV estimates in Appendix Tables I.6 and I.7. Reported stars follow the significance notation in the source estimates. Dashes indicate outcomes not reported in the cited study. Estimates are shown for comparability only and are drawn from studies with different samples, intervention designs, and empirical specifications.

not observe effects on children's height-for-age.

Are our effect sizes plausible? Our measure of chlorination is a flow, while anthropometrics are a stock: even modest increases in water purification, when accumulated over an extended period of time, may have large effects (Bekele et al., 2020). Furthermore, our measures of chlorine use are almost certainly underestimates, as participants must have used chlorine within the 24 hours prior to a visit in order for enumerators to detect chlorine presence in their visit.[27] Finally, there may exist a complementarity between herd protection and personal protection, leading to a further amplification of the health impacts of increased chlorine use (Deutschmann et al., 2024; Duflo et al., 2015; Fuller and Eisenberg, 2016). Info-Tool households with Info-Tool neighbors are the participants most likely to both use chlorine *and* have neighbors using chlorine, meaning their children are the most likely to be protected from infection transmitted through their drinking water as well as their environment.

4.4.5 Cost-Benefit Analysis

The Info-Tool is non-trivially more expensive than the other treatment arms, as it requires three months of healthworker time prior to chlorine distribution, implying three months of additional labor costs. Furthermore, the Info-Tool is only effective when offered to many people, implying higher total costs. Are the benefits of the within-treatment spillovers large enough to offset these costs?

We assume that every chlorine detection is equivalent to 30 days of water chlorination, and that 11,000 days of water chlorination equates to 1 DALY averted (International, 2017). Each treatment household costs 24.78 USD. Each Info-Tool household costs an additional

[27] We detect chlorine 13% of the time across the whole study, but 37% of treatment participants *report* that they were using chlorine to treat their water at endline. Furthermore, we detect higher rates of chlorine use in our unscheduled audit studies during the treatment period than we do during regularly scheduled visits.

7.84 USD (paper materials and salaries for enumerator time), and each Incentives household costs an average of an additional 0.88 USD (short-term incentives provision). We estimate that the cost per DALY averted for an Info-Tool participant with at least one other Info-Tool neighbor is 2778 USD, while the cost per DALY averted for all other participants is 38% to 131% larger (depending on the treatment group and density of Info-Tool neighbors). Indeed, we find that the cost per DALY averted is *decreasing* in the density of Info-Tool participants in an area (Figure 4). It is *only* cost-effective, relative to our other treatment groups, to implement the Info-Tool if the treatment is highly saturated in a geographic area.

4.4.6 Demand for Chlorine Tablets

While the Info-Tool meaningfully increases take-up of chlorine tablets, our results underscore that the cost and availability of chlorine tablets remains an important barrier to adoption. In our endline take-it-or-leave-it willingness-to-pay exercise, only 2.1% of the sample were willing and able to pay *anything* for a one-month supply of chlorine tablets, and only 1.7% were willing and able to pay the market price.[28] This is likely driven by an inability to pay rather than low valuation of chlorine, because participants demonstrate demand for chlorine in other ways. Seven percent of participants demonstrate willingness to give up their time for the chance to purchase chlorine by asking the enumerator to return at a later date to try selling again, hoping that they may have cash available at another time. In the same visit, 37% report that they currently use chlorine to purify their water,[29] and 77% accept free chlorine tablets.

5 Mechanisms

Having established the significant behavioral and health impacts of the Info-Tool intervention, particularly among those with Info-Tool neighbors, we now turn to a closer examination of the underlying channels and competing explanations therein. Why does a simple pen-and-paper chart lead to behavioral change?

5.1 Responsiveness to Health Signals

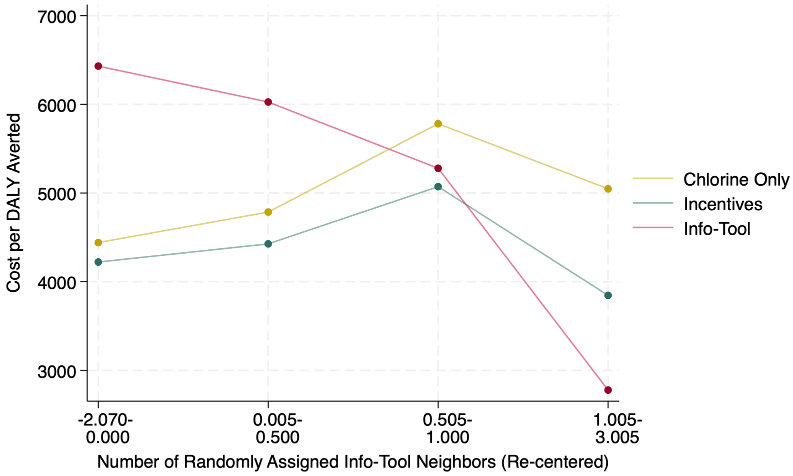
5.1.1 Average Effects

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

[28]Tables D.10 and D.11 report the full set of willingness-to-pay outcomes by treatment group, including a breakdown by spillover exposure; the patterns do not differ meaningfully across treatment arms.

[29]The correlation between self-reported chlorine use at endline and the probability that we ever detect chlorine residual in the prior three months is 0.22.

Figure 4: Cost-Effectiveness Analysis



Notes: This figure plots estimated cost per DALY averted by treatment status and Info-Tool saturation. Calculations assume that one chlorine detection corresponds to 30 days of chlorination and that 11,000 days of chlorination avert one DALY. Treatment-arm costs include the common intervention cost for all treated households, plus arm-specific costs for the Info-Tool and Incentives arms. Spillover exposure and treatment density are defined using the same 20-meter neighborhood measure used throughout the paper.

measure of health improvement. We define "health improvement" as the difference in diarrhea days between the three months prior to and following chlorine distribution. This is the relevant period of time to analyze health signals, as this is when Info-Tool participants gain experience acquiring information about their health via the Info-Tool. To predict which households are likely to experience positive health signals during this period, we employ lasso among the pure control sample of households using variables collected prior to the revelation of treatment status (e.g., all variables from the baseline survey and the first half of the first follow-up survey). To avoid bias that may arise from endogenous stratification, we implement the leave-one-out procedure proposed by Abadie et al. (2018).[30]

Note that testing for differences in chlorine adoption by one's own predicted health signals is a necessary but not sufficient condition, since the types of households that are likely to generate positive health signals may possess other unobserved attributes correlated with behavioral change. We therefore view this exercise as a sanity check of our model, but not a dispositive test.

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

$$Y_i = \theta_0 + \theta_1 T1_i + \theta_2 T2_i + \theta_3 \hat{I}_i + \theta_4 \hat{I}_i \times T1_i + \theta_5 \hat{I}_i \times T2_i + Y_{i0} + X_{i0} + \gamma_b + \epsilon_i$$

[30]The correlation between "actual health improvement" and "predicted health improvement" is 0.39 in the pure control group and 0.40 in the treatment groups.

Table 5: Chlorine Detection
by Own Positive Signals

	(1) During BI Full Sample		(2) Post-BI Full Sample		(3) During BI Endline Sample		(4) Post-BI Endline Sample	
Chlorine × Self Improved	0.104	(0.066)	0.271	(0.166)	0.097	(0.069)	0.195	(0.173)
Info-Tool × Self Not Improved	-0.046	(0.066)	0.190	(0.163)	-0.057	(0.068)	0.156	(0.169)
Info-Tool × Self Improved	0.107	(0.067)	0.468***	(0.169)	0.149**	(0.071)	0.508***	(0.177)
Observations	1192		1216		1108		1136	
Chlorine × Self Not Improved Mean	0.557		1.399		0.569		1.449	
P-values:								
Info-Tool × Self Improved = Info-Tool × Self Not Improved	0.024		0.097		0.004		0.044	

Standard errors in parentheses

Each observation is at the household level. The outcome is the total number of times that chlorine was detected across all visits in the specified time period. During-BI refers to the time period during the Behavioral Interventions, while Post-BI refers to the time period after the Behavioral Interventions. The Chlorine Only treatment group who were not predicted to improve in their diarrhea rate during the Behavioral Intervention period is the omitted group. An indicator for Pure Control participants, who are included in the sample to maximize power, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. Self 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 5.1 for a detailed explanation for how the predicted-health-improvement measure is defined.

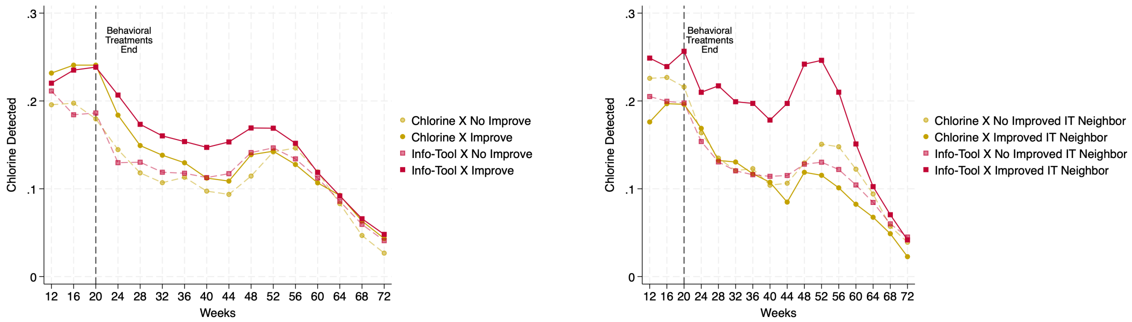
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

where $\hat{I}_i = 1$ if an individual is ex ante predicted to improve. θ_5 is our object of interest. We use θ_4 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 acquired through experience with the Info-Tool.

We find that Info-Tool participants whose health is ex-ante predicted to improve in the first three months of chlorine distribution use chlorine at a significantly higher rate than Info-Tool participants whose health is not predicted to improve. These differences are large: predicted-improved Info-Tool participants chlorinate 30% more often than not-predicted-improved Info-Tool participants during the behavioral intervention period (Columns 1 and 3 of Table 5, $p = 0.024$). After the behavioral interventions end, this effect declines over time. In the post-behavioral-intervention period (Columns 2 and 4), predicted-improved Info-Tool respondents continue using chlorine at a higher rate than not-predicted-improved Info-Tool participants, but the estimate is smaller and noisier (18% higher rate of use; $p = 0.097$), suggesting, consistent with our theory, that favorable own signals matter most when participants are actively generating and aggregating those signals through the Info-Tool.

The left panel of Figure 5 depicts the same pattern in the raw data. Unsurprisingly, we also observes differences by predicted-improved among the chlorine only group (Figure 5) and the incentives group (Figure D.3). The magnitude of this difference by predicted-improved is largest among Info-Tool households, though not statistically distinguishable from the difference within the other treatment arms. These patterns are consistent with a model where all households learn from signals in their environment, and Info-Tool

Figure 5: Chlorine Detection by Predicted Health Improvement (Moving Averages, Raw Data)



Notes: This figure plots raw moving averages of chlorine detection over time. The left panel plots data for Info-Tool and Chlorine-Only participants, separately by whether predicted health improvement is above or below the median. Predicted health improvement is constructed using pre-treatment variables only. The right panel plots raw moving averages of chlorine detection for Chlorine-Only and Info-Tool participants, separately by whether they are randomly more exposed to a predicted-improved Info-Tool neighbor than would be expected based on endogenous spatial factors. A predicted-improved Info-Tool neighbor is an Info-Tool participant within 20 meters whose diarrhea rate was ex-ante predicted to improve during the behavioral-intervention period; see Section 5.1 for details. The vertical dashed line marks the end of the behavioral-intervention period.

households place heavier weights on these signals.

We now turn to a causally identified exercise exploiting the random placement of Info-Tool neighbors to further examine underlying channels.

5.1.2 Spillovers

Tables 1 and 2 demonstrate that both behavioral and health improvements among Info-Tool households are driven by those who have at least one Info-Tool neighbor. If Info-Tool participants are indeed able to act on information from their Info-Tool neighbors to a greater degree than other participants are, then we should observe that chlorine use increases specifically when such neighbors communicate a positive health signal. 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 is any participant within radius r of participant i who was predicted to see their health improve and who was randomized to the Info-Tool group. By exploiting variation in Info-Tool treatment assignment of neighbors, this exercise therefore identifies the effect of random exposure to positive health signals generated by the Info-Tool.

We present results in Table 6. During the period of the behavioral intervention, differences in chlorine use by this measure of spillover exposure are present, but statistically indistinguishable (Columns 1 and 3). After the behavioral interventions end, however (Columns 2 and 4), the spillover effects remain concentrated among Info-Tool participants with a predicted-improved Info-tool neighbor. These patterns are temporally illustrated in

Table 6: Chlorine Detection by Info-Tool Neighbors' Positive Signals

	(1) During BI Full Sample		(2) Post-BI Full Sample		(3) During BI Endline Sample		(4) Post-BI Endline Sample	
Chlorine × IT Neighbor Improved	-0.084	(0.079)	0.083	(0.196)	-0.120	(0.083)	0.061	(0.204)
Info-Tool × No IT Neighbor Improved	-0.054	(0.053)	0.070	(0.132)	-0.056	(0.056)	0.084	(0.137)
Info-Tool × IT Neighbor Improved	-0.011	(0.079)	0.645***	(0.197)	-0.007	(0.084)	0.731***	(0.204)
Observations	1192		1216		1108		1136	
Chlorine × No IT Neighbor Improved Mean	0.643		1.538		0.660		1.554	
P-values:								
Info-Tool: No IT Neighbor Improved = IT Neighbor Improved	0.588		0.004		0.560		0.002	

Standard errors in parentheses
Each observation is at the household level. The outcome is the total number of times that chlorine was detected across all visits in the specified time period. During-BI refers to the time period during the Behavioral Interventions, while Post-BI refers to the time period after the Behavioral Interventions. The omitted group is households in the Chlorine Only treatment group who were not 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. An indicator for Pure Control participants, who are included in the sample to maximize power, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. IT Neighbor Improved is a binary indicator for if the participant's was randomly more exposed to an Info-Tool neighbor (someone within 20m) whose health was predicted to improve at the beginning of chlorine distribution more than the median. See Section 5.1 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

the right panel of Figure 5. Figure D.5 replicates this exercise but limits the sample to households who have at least one Info-Tool neighbor.

These dynamics align with the predictions of our model, with early adoption being linked to favorable personal signals as Info-Tool participants accumulate their own experience with the technology, and sustained adoption operating through the sharing and reinforcement of favorable cumulative experience generated through the same information-acquisition technology.

Because we exploit random variation in the Info-Tool assignment, we can rule out that participants who are predicted to improve simply have characteristics that make them more likely to influence their neighbors. However, it may be that neighbors predicted to improve become more influential when they are assigned to any treatment. To rule out this possibility, we conduct a placebo exercise where we test heterogeneity in chlorination behavior by the total number of predicted-improved Chlorine Only neighbors. Table D.3 demonstrates that there are no measurable differences in use by exposure to this spillover measure in the treatment groups.

5.2 Beliefs and Motivations

We argue that Info-Tool participants change their behavior because the treatment amplifies weights they apply to signals in a Bayesian learning process, but the previous exercise cannot rule out that Info-Tool participants simply acquire more or different signals, for example, by engaging in more conversations about chlorine. However, we find no evidence of differential information-sharing (discussion in Appendix Section J.0.4). While conversational content, frequency, and knowledge about others' water treatment choices do not differ measurably across treatment groups, we do find evidence that receipt of chlorine alone affects communication. Participants who receive any chlorine treatment are more likely to say that they discussed water purification with their social network members, with 34% reporting at least one person with whom they discuss water purification (Appendix Table J.5). Similarly, all participants are more likely to believe that their network links use chlorine when those friends are participants receiving chlorine in our study, suggesting that participants are knowledgeable about who receives chlorine.

5.2.1 Trust in information-acquisition technology

To explore if participants apply weights to signals differentially and if these weights depend on the signal-acquisition process itself, we first establish that neighbors' engagement with the Info-Tool is salient, and then ask participants directly: whose signal would you trust the most?[31]

To assess respondents' awareness of one another's participation in the study, we remind all participants in our final survey about the three treatment groups in our study, and ask participants to guess how many neighbors they have from each group. 71% of participants believe they have at least one treated neighbor. There are no measurable differences across treatment groups in beliefs about the presence of Chlorine-Only neighbors or Incentives neighbors, while Info-Tool participants are significantly more likely to believe that they have an Info-Tool neighbor than participants in any other treatment group (Appendix Table D.5). This suggests that Info-Tool treatment status is a characteristic of neighbors that other Info-Tool participants are differentially more likely to know or care about.

We then ask participants to imagine a neighbor from each group and indicate which one they believe has the most knowledge about children's health. We observe two striking facts. First, participants in every treatment group place greatest trust (i.e. belief in someone's knowledge about health) in members of their *own* treatment group, suggesting that one's own shared experience matters (Table 7 and Appendix Table D.6). Second, differencing out the propensity to choose one's own group using Chlorine Only households, participants in the Info-Tool group are 10 percentage points ($p = 0.046$) more likely to list the Info-Tool respondent as most trustworthy in terms of knowledge about children's health, suggesting that shared experience with the specific signal-generating technology of the Info-Tool amplifies this phenomenon.[32][33]

[31] We piloted questions that directly measured participants' ability to interpret data from the Info-Tool, but found that participants who were not in the Info-Tool group were unable to even guess at an interpretation. It is possible that participants outside the Info-Tool group are still able to understand that information generated by the Info-Tool is valuable. Thus, we view trust in information gathered from the Info-Tool as a lower bound on deeply understanding the Info-Tool data-generating process.

[32] The corresponding estimate is 13.3 percentage points ($p = 0.004$) if we use the Incentives households to net out the propensity to choose one's own group (Table D.6).

[33] Interestingly, Info-Tool participants *without* access to an Info-Tool spillover are more likely to rate

Table 7: Endline Stated Trust
Omitted group: Pure Control

	(1) Info-Tool Knows Most	(2) Chlorine Knows Most
Info-Tool	0.222***	-0.138***
	(0.027)	(0.034)
Chlorine Only	-0.043	0.122***
	(0.027)	(0.034)
Observations	1516	1516
Control Mean	0.196	0.527
DID Estimate:		
Info-Tool picks Info-Tool		
– Chlorine Only picks Chlorine Only		0.100 [p= 0.046]

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). The outcome for column (2) is an indicator for if the respondent chose a hypothetical Chlorine Only respondent as the person most likely to be knowledgeable about child health (rather than a Incentives or Info-Tool participant). All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

In our final survey, 53% of Info-Tool households still possess the Info-Tool, a mere sheet of paper that has not been touched in eighteen months. Cumulatively, these observations suggest that Info-Tool households come to value and trust the technology as a source of information, uniquely priming them to apply heavier weights to signals from other Info-Tool participants.

5.2.2 Motivation to translate beliefs into actions

In our model, experience generating health signals with the Info-Tool induces caregivers to place higher weights on any signal acquired from the Info-Tool. These behavioral changes can arise from changes in beliefs about the expected returns to chlorine tablet adoption, or from a greater willingness to act on a given set of beliefs. Our model remains agnostic on which channel may drive the behavior change, but we now explore this question empirically to the best of our ability.

We report estimates for our main model parameters in Table 8: beliefs about chlorine efficacy M, experience-weighted beliefs M_α, and chlorine use decision as determined by $\mathbb{1}(M_\alpha > C)$ (where Column 3 replicates Table 1, Column 1). Beliefs about chlorine efficacy, which are elicited halfway through Phase 2 in week 44, reflect respondents' beliefs that their children's diarrhea rates decreased due to chlorine, and are incentivized

other Info-Tool participants as the most knowledgeable, indicating that personal experience with the Info-Tool itself is enough to change participants' trust in the Info-Tool as an information source (Table D.7).

for accuracy (we discuss this measure and its validity in detail in Appendix Section J.0.5). As evident in Column 1, we find no measurable differences in stated beliefs across treatment groups or proximity to an Info-Tool neighbor.[34] However, we do find differences in the degree to which participants report that they are *motivated* to use chlorine to attain health for their family (Appendix Table D.4 reports treatment effects on all stated motivations).[35] .

This variation in M_α parallels the observed differences in chlorine adoption, $\mathbb{1}(M_\alpha > C)$: the treatment effect on motivation is larger among Info-Tool households who have an Info-Tool neighbor than among Info-Tool households without an Info-Tool neighbor (Column 2), while Chlorine Only participants do not rate health as a motivating factor differentially by access to Info-Tool neighbors. These estimates are not statistically distinguishable and therefore not conclusive on their own, but the pattern aligns with our model and is consistent with our earlier finding that Info-Tool households differentially trust their Info-Tool neighbors on knowledge related to child health, the very grounds upon which Info-Tool households differentially report higher motivation to use chlorine.

We leverage a final piece of data to distinguish between the channels of increased signal precision vs. greater weight on existing signals. While enumerators trained all households in how to use the Info-Tool, we recorded whether, in a given visit, a household filled in the Info-Tool themselves (prior to the enumerator's visit) or waited for the enumerator to fill it in for them. Both groups have access to equally precise information, but they differ in the extent to which they participate in the experience of data collection. As we show in Appendix Table D.9 and Figure D.7, Info-Tool households who filled the Info-Tool out themselves chlorinate at higher rates than Info-Tool participants for whom the enumerator always filled out the chart on their behalf. Similarly, as we show in Appendix Table D.8 and Figure D.6, the spillover effect is driven almost entirely by individuals with an Info-Tool neighbor who independently fill out the Info-Tool chart for at least one two-week period without the enumerator's assistance. These final exercises are exploratory, as proactive households may be different from their counterparts in numerous ways, but the patterns we observe are consistent with a mechanism of greater weight placed on existing signals, or a greater willingness to act, translating into higher chlorine use, especially by those with shared experience with the Info-Tool.

[34]We examine a number of questions on stated beliefs about chlorine efficacy and the disease environment in Appendix Section J.0.5. As we outline in the model, experience with the Info-Tool may lead participants to view all Info-Tool signals with higher precision, generating the same complementarity in chlorine use that we observe in our data via raising their posterior beliefs on chlorine efficacy. However, we find no differences across groups in participants' incentivized beliefs about chlorine's efficacy or beliefs about the riskiness of the disease environment. Beliefs *do* differ by education, with more educated participants exhibiting more accurate beliefs about chlorine efficacy, suggesting that the questions we ask to elicit beliefs are sensitive to relevant characteristics that affect belief formation.

[35]See Appendix Section A for details on how we measure motivation.

Table 8: Model Parameters

	(1) Stated M: Diarrhea Went Down		(2) Stated M_α: Motivation: Health		(3) Observed $\mathbb{1}(M_\alpha > C)$: Times Detected Chlorine	
Chlorine × Spillover	0.022	(0.048)	-0.094	(0.092)	0.071	(0.172)
Info-Tool × No Spillover	0.021	(0.040)	0.122	(0.075)	0.055	(0.142)
Info-Tool × Spillover	-0.003	(0.048)	0.222**	(0.090)	0.528***	(0.170)
Observations	1123		703		1282	
Chlorine × No Spillover Mean	0.711		6.081		1.543	
P-values:						
Info-Tool × Spillover = Info-Tool × No Spillover	0.621		0.278		0.006	
Info-Tool × Spillover = Info-Tool × No Spillover + Chlorine × Spillover	0.491		0.127		0.092	

Standard errors in parentheses

Each observation is at the household level. The omitted group is Chlorine × No Spillover. The outcome for column (1) is an indicator for participants who said they believe that their childrens' diarrhea rate dropped after using chlorine in an incentivized elicitation about the efficacy of chlorine. 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). The outcome for column (3) is the number of times that we detected chlorine residual in the participant's water during the post-treatment period. This regression controls for the number of times we tested their water during the post-treatment period, and an indicator variables for belonging to the following groups: Control × No Spillover (except column (2) – the outcome variables was not collected for Control participants), and Control × Spillover (except column (2) – the outcome variables was not collected for Control participants). All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, 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 4.1 for a detailed explanation of how the measure is defined.

* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

5.3 Competing Mechanisms

Early Adoption: Long-term adoption could be explained by early adoption leading to habit formation, or intertemporal complementarities in chlorine use. We use the Incentives arm to disentangle this mechanism, since it generates short-term use through financial incentives but does not provide a specific learning opportunity. During the behavioral intervention period, Incentives households chlorinate significantly more than those in Info-Tool or Chlorine Only. However, immediately following incentive withdrawal (week 20), chlorination rates for incentive households plummet and remain low, implying that chlorination is not a habit-forming activity (discussion in Appendix Section J.0.1), and sustained use is more likely attributed to learning.

More Interactions and Mimicry: Info-Tool participants may interact with one another more when they are in the same group, and mimic rather than learn from one another. However, Info-Tool participants' endline social networks are not larger than any other groups', and Info-Tool participants do not mimic one another in the take-up of other health programs (discussion in Appendix Section J.0.2).

Social Norms: Our results may be interpreted through the lens of social norms, wherein the Info-Tool sensitizes individuals to a certain type of norm which they then adopt from their neighbors. This remains consistent with our model. Our results are inconsistent with a model, however, wherein Info-Tool participants act on what they think *others* believe, via a preference to conform, rather than their own health preferences (Bernheim, 1994) (see discussion in Appendix Section J.0.3). Furthermore, participants do not respond to having Incentives neighbors, who also chlorinate at higher rates during the behavioral intervention, nor to Info-Tool neighbors who do not get positive signals. Finally, participants do not report different beliefs about others' chlorine use nor others' willingness to accept chlorinated water.[36]

6 Conclusion

We study the process of joint experiential learning about health by leveraging a learning intervention whose effectiveness in changing behavior hinges on the interaction between individual experiential learning and reinforcement through social learning. While social learning alone proves insufficient to disseminate novel information or alter behavior, its interaction with individual experiential learning yields significant behavioral change, leading to important downstream improvements in child health.

We document consistent differences across two objectively measured primary outcomes: water chlorination and child anthropometrics. Relative to all other comparison arms, Info-Tool households with Info-Tool neighbors exhibit higher chlorination rates and larger improvements in child health. All other groups chlorinate at lower, statistically comparable levels, experiencing correspondingly smaller health gains relative to our pure control of no chlorine.

We offer two 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. Indeed, we find that the cost per DALY averted is *decreasing* in the density of Info-Tool participants in an area (Figure 4). 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, our results speak to public health programs across a wide geographic space. In Pakistan, only 0.3% of the population report usage of chlorine tablets, with 7.1% adopting any purification technology (Pakistan DHS 2017-18). Many features of our 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

[36]Our results also provide little evidence for threshold-based coordination mechanisms. Info-Tool households appear to be linearly affected by the number of Info-Tool neighbors, suggesting there is no critical mass of neighbors that generates coordination (Granovetter, 1978).

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 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 (Perry and Hodgins, 2021), 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 scalable.

Beyond point-of-use chlorination for water purification, what other technologies may feature returns to joint experiential learning? We argue 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 through other trusted sources. Technologies and behaviors that fit these criteria are common in the health domain, from hygiene (e.g., hand-washing and proper treatment of food), to mental health care practices (e.g., therapy and meditation) and lifestyle changes (e.g., diet and exercise). Non-health technologies where shared experiential learning may be critical include agricultural practices (Vasilaky and Islam, 2018; Vasilaky and Leonard, 2018)[37], childcare practices (Bunting, 2004)[38], and learning or study habits (List and Uchida, 2024).[39] Exploring these spaces as potential sites of complementarity between individual and social learning remains a promising avenue for future research.

[37] Vasilaky and Islam (2018) find that female Ugandan farmers are more likely to learn new information disseminated between intervention rounds when they face team-based incentives to learn rather than tournament-based (individual) incentives. Consistent with our findings, it is only information that is novel that specifically benefits from joint experiential learning. Vasilaky and Leonard (2018) find that, among farmers in Uganda who go to an agricultural training, being paired with another farmer and encouraged to discuss the techniques they learned throughout the farming season leads to higher productivity relative to only attending the training.

[38] An illustrative example is the parental education groups organized in many countries under the coordination of trained health workers. These groups provide a centralized source of reliable information, communicated by the worker, while also fostering peer-to-peer learning. Parents share specific details and insights about various practices based on their personal experiences, thereby complementing the centralized guidance with practical, experiential knowledge (Bunting, 2004).

[39] 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 behaviors drive these cognitive gains, it is conceivable that students learn behaviors that promote cognitive development in preschool, which the students reinforce in one another if they are in classrooms together in kindergarten and grade school.

References

Aea rct registry dataverse. https://www.socialscienceregistry.org/site/csv, 2024. Accessed: 2024-11-12.

A. Abadie, M. M. Chingos, and M. R. West. Endogenous stratification in randomized experiments. *Review of Economics and Statistics*, C(4):567–580, 2018.

S. Aggarwal, R. Dizon-Ross, and A. D. Zucker. Incentivizing behavioral change: The role of time preferences. *NBER Working Paper No. w27079*, 2020.

I. Ajzen. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2):179–211, 1991. doi: 10.1016/0749-5978(91)90020-T.

G. A. Akerlof and R. E. Kranton. Economics and identity. *Quarterly Journal of Economics*, 115(3):715–753, 2000. doi: 10.1162/003355300554881.

A. A. Akram and R. Mendelsohn. Diaries to increase the adoption of chlorine tablets for water purification by poor households. *Water Economics and Policy (WEP)*, 7(2): 1–34, 2021.

O. Al-Ubaydli, J. A. List, and D. L. Suskind. The science of using science: Towards an understanding of the threats to scaling experiments. Working Paper 25848, National Bureau of Economic Research, 2019.

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

F. Alpízar, M. Bernedo Del Carpio, P. J. Ferraro, and B. S. Meiselman. Exposure-enhanced goods and technology disadoption. *Working Paper.*, 2022.

M. Alsan and S. Eichmeyer. Experimental evidence on the effectiveness of nonexperts for improving vaccine demand. *American Economic Journal: Economic Policy*, 16(1): 394–414, 2024.

M. Alsan and C. Goldin. Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, 127:586–638, 2019.

M. Alsan and M. Wanamaker. Tuskegee and the health of black men. *The Quarterly Journal of Economics*, 133(1):407–455, 2018.

M. Alsan, O. Garrick, and G. Graziani. Does diversity matter for health? experimental evidence from oakland. *American Economic Review*, 109:4071–4111, 2019.

M. L. Anderson. 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*, 103(484):1481–1495, 2008.

K. J. Arrow. The economic implications of learning by doing. *The Review of Economic Studies*, 29(3):155–173, 1962.

N. Ashraf, J. Berry, and J. M. Shapiro. Can higher prices stimulate product use? evidence from a field experiment in zambia. *American Economic Review*, 100(5):2383–2413, 2010.

N. Ashraf, A. Banerjee, and V. Nourani. Learning to teach by learning to learn. *Working Paper*, 2021.

S. Athey and G. W. Imbens. The econometrics of randomized experiments. *Handbook of Economic Field Experiments*, 1:73–140, 2017.

A. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson. Using gossips to spread information: Theory and evidence from two randomized controlled trials. *The Review of Economic Studies*, 86(6):2453–2490, 2019.

A. Banerjee, A. Chowdhury, J. Das, J. Hammer, R. Hussam, and A. Mohpal. The market for healthcare in low income countries. *Working Paper*, 2023.

A. Banerjee, E. Breza, A. Chandrasekhar, and B. Golub. When less is more: Experimental evidence on information delivery during india's demonization. *Review of Economic Studies*, 91:1884–1992, 2024.

L. Beaman, A. BenYishay, J. Magruder, and A. M. Mobarak. Can network theory-based targeting increase technology adoption? *American Economic Review*, 111(6), 2021.

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

T. Bekele, A. Rahman, and P. Rawstorne. Effect of water, sanitation and hygiene interventions alone and combined with nutrition on child growth in low and middle income countries: a systematic review and meta-analysis. *BMJ Open*, 10(7):e034812, 2020. doi: 10.1136/bmjopen-2019-034812.

R. Bénabou and J. Tirole. Self-confidence and personal motivation. *Quarterly Journal of Economics*, 117(3):871–915, 2002.

D. Bennett, A. Naqvi, and W.-P. Schmidt. Learning, hygiene and traditional medicine. *The Economic Journal*, 128(612):F545–74, 2018. doi: https://doi.org/10.1111/ecoj.12549.

A. BenYishay and A. M. Mobarak. Social learning and incentives for experimentation and communication. *Review of Economic Studies*, 86:976–1009, 2019.

A. BenYishay, M. Jones, F. Kondylis, and A. M. Mobarak. Gender gaps in technology diffusion. *Journal of Development Economics*, 143:1–27, 2020.

B. D. Bernheim. A theory of conformity. *Journal of Political Economy*, 102(5):841–877, 1994.

S. Bikhchandani, D. Hirshleifer, O. Tamuz, and I. Welch. Information cascades and social learning. *Journal of Economic Literature*, 62:1040–1093, 2024.

R. Bodner and D. Prelec. Self-signaling and diagnostic utility in everyday decision making. In I. Brocas and J. D. Carrillo, editors, *The Psychology of Economic Decisions, Volume 1: Rationality and Well-Being*, pages 105–126. Oxford University Press, Oxford, 2003.

K. Borusyak and P. Hull. Nonrandom exposure to exogenous shocks. *Econometrica*, 91: 2155–2185, 2023.

J. M. Bos, A. S. Shonchoy, S. Ravindran, and A. Khan. Early childhood human capital formation at scale. *Journal of Public Economics*, 231, 2024.

E. Breza and A. Chandrasekhar. Social networks, reputation, and commitment: Evidence from a savings monitors experiment. *Econometrica*, 87(1):175–216, 2019.

N. Buchmann, E. M. Field, R. Glennerster, and R. N. Hussam. Throwing the baby out with the drinking water: Unintended consequences of arsenic mitigation efforts in bangladesh. Technical report, National Bureau of Economic Research, 2019.

L. Bunting. Parenting programmes: The best available evidence. *Child Care in Practice*, 10(4):327–343, 2004.

F. Burlig, A. Jina, and A. Sudarshan. The value of clean water: Experimental evidence from rural india. *American Economic Review*, 116:1148–1187, 2026.

S. Calónico, R. Di Tella, and J. C. 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.

J. Caro-Burnett, J. A. Chevalier, and A. M. Mobarak. "is habit a powerful policy instrument to induce prosocial behavioral change? *Cowles Foundation Discussion Papers*, 2600, 2021.

P. A. Celhay, P. J. Gertler, P. Giovagnoli, and C. Vermeersch. Long-term effects of temporary incentives on productivity in medical care clinics. *American Economic Journal: Applied Economics*, 11(3):92–127, 2015.

A. G. Chandrasekhar, E. Dufo, M. Kremer, J. F. Pugliese, J. Robinson, and F. Schilbach. Blue spoons: Sparking communication about appropriate technology use. *NBER Working Paper 30423*, 2022.

Y. Chen, P. Persson, and M. Polyakova. The roots of health inequality and the value of intrafamily expertise. *American Economic Journal: Applied Economics*, 14(3):185–223, 2022.

T. G. Conley and C. Udry. Learning about a new technology: Pineapple in ghana. *The American Economic Review*, 100(1):35–69, 2010.

J. Conlon, M. Mani, G. Rao, M. Ridley, and F. Schilbach. Not learning from others. *Revise Resubmit, Econometrica*, 2022.

L. Corno. Learning (or not) in health-seeking behavior: Evidence from rural tanzania. *Economic Development and Cultural Change*, 63(1):27–72, 2014.

F. D'Acunto, U. Malmendier, J. Ospina-Tejeiro, and M. Weber. Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, 129:1615–1639, 2021.

M. E. Darden and M. Macis. Trust and health care-seeking behavior. *NBER Working Paper 32028*, 2024.

A. de Janvry, K. M. Macours, and E. Sadoulet. Learning for adopting: Technology adoption in developing country agriculture. *FERDI*, 2016.

J. W. Deutschmann, M. Lipscomb, L. Schechter, and J. Zhu. Spillovers without social interactions in urban sanitation. *American Economic Journal: Applied Economics*, 16 (3):482–515, 2024.

E. Duflo, R. Glennerster, and M. Kremer. Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4:3895–3962, 2007.

E. Duflo, M. Greenstone, R. Guiteras, and T. Clasen. Toilets can work: Short and medium run health impacts of addressing complementarities and externalities in water and sanitation. Technical report, National Bureau of Economic Research, 2015.

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

P. Dupas. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica*, 82(1):197–228, 2014.

P. Dupas, V. Hoffman, M. Kremer, and A. Peterson Zwane. Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in kenya. *Science*, 353 (6302):889–895, 2016.

P. Dupas, B. Nhlema, Z. Wagner, A. W. Wolf, and E. Wroe. Expanding access to clean water for the rural poor: experimental evidence from malawi. *American Economic Journal: Economic Policy*, 15(1):272–305, 2023.

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

D. Egger, J. Haushofer, E. Miguel, P. Niehaus, and M. Walker. General equilibrium effects of cash transfers: experimental evidence from kenya. *Econometrica*, 90(6):2603–2643, 2022.

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

M. Fafchamps, A. Islam, D. Pakrashi, and D. Tommasi. Diffusion in social networks: Experimental evidence on information sharing vs persuasion. *Working Paper*, 2024.

A. D. Foster and M. R. Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *The Journal of Political Economy*, 103(6): 1176–1209, 1994.

J. A. Fuller and J. N. Eisenberg. Herd protection from drinking water, sanitation, and hygiene interventions. *The American journal of tropical medicine and hygiene*, 95(5): 1201, 2016.

M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.

E. A. Halm, C. Lee, and M. R. Chassin. Is volume related to outcome in health care? a systematic review and methodologic critique of the literature. *Annals of Internal Medicine*, 137(6), 2002.

R. Hanna, S. Mullainathan, and J. Schwartzstein. Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3): 1311–1353, 2014.

R. Hertwig, G. Barron, E. U. Weber, and I. Erev. Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8):534–539, 2004.

M. G. Hudgens and M. E. Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008. doi: 10.1198/016214508000000292.

R. Hussam, A. Rabbani, G. Reggiani, and N. Rigol. Rational habit formation: Experimental evidence from handwashing in india. *American Economic Journal: Applied Economics*, 14(1):1–41, 2022.

P. S. International. Psi impact calculator. http://impactcalculator.psi.org/., 2017.

V. Khandelwal. Learning in networks with idiosyncratic agents. *Games and Economic Behavior*, 144:225–249, 2024.

F. Kondylis, J. A. Loeser, M. Mobarak, M. R. Jones, and D. K. Stein. Learning from self and learning from others : Experimental evidence from bangladesh (english). *Policy Research working paper ; no. WPS 10545; Impact Evaluation series Washington, D.C. : World Bank Group.*, 2023.

M. Kremer, S. P. Luby, R. Maertens, B. Tan, and W. Wiecek. Water treatment and child mortality: A meta-analysis and cost-effectiveness analysis. Technical report, National Bureau of Economic Research, 2023.

J. A. List. *Experimental Economics: Theory and Practice.* University of Chicago Press, Chicago, 2025.

J. A. List and H. Uchida. Here today, gone tomorrow? toward an understanding of fade-out in early childhood education programs. *NBER Working Paper 33027*, 2024.

S. Lowes and E. Montero. The legacy of colonial medicine in central africa. *American Economic Review*, 111(4):1284–1314, 2021.

S. P. Luby, M. Agboatwalla, J. Painter, A. Altaf, W. Billhimer, B. Keswick, and R. 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*, 11(4):479–489, 2006.

J. Luoto, N. Najnin, M. Mahmud, J. Albert, M. S. Islam, S. Luby, L. Unicomb, Levine, and D. I. What point-of-use water treatment products do consumers use? evidence from a randomized controlled trial among the urban poor in bangladesh. *PloS one*, 6 (10), 2011.

U. Malmendier and S. Nagel. Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87, 2015.

M. Martinez-Bravo and A. Stegmann. In vaccines we trust? the effects of the cia's vaccine ruse on immunization in pakistan. *Journal of the European Economic Association*, 20 (1):150–186, 2022.

D. McKenzie. Beyond baseline and follow-up: The case for more t in experiments. *Journal of development Economics*, 99(2):210–221, 2012.

E. Miguel and M. Kremer. Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72:159–217, 2004.

C. K. Morewedge. Psychological ownership: implicit and explicit. *Current Opinion in Pyschology*, 39:125–132, 2021.

K. Muralidharan and V. Sundararaman. Teacher performance pay: Experimental evidence from india. *Journal of Political Economy*, 119(1):39–77, 2011.

D. Neal, J. Vujcic, O. Hernandez, W. 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.

C. Null, C. P. Stewart, A. J. Pickering, H. N. Dentz, B. F. Arnold, C. D. Arnold, and e. a. Benjamin-Chung, Jade Benjamin-Chung. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural kenya: A cluster-ran- domised trial. *Lancet Global Health*, 6:e316–29, 2018.

H. B. Perry and S. Hodgins. Health for the people: Past, current, and future contributions of national community health worker programs to achieving global health goals. *Global Health: Science and Practice*, 9:1–9, 2021.

H. Royer, M. Stehr, and J. 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*, 7(3):51–84, 2015.

R. M. Ryan and E. L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78, 2000. doi: 10.1037/0003-066X.55.1.68.

S. Sazawal, A. A. Habib, U. Dhingra, A. Dutta, P. Dhingra, A. Sarkar, S. Deb, J. Alam, A. Husna, and R. 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*, 13(514), 2013.

P. Sheeran and T. L. Webb. The intention–behavior gap. *Social and Personality Psychology Compass*, 10(9):503–518, 2016. doi: 10.1111/spc3.12265.

U. Simonsohn, N. Karlsson, G. Loewenstein, and D. Ariely. The tree of experience in the forest of information: Overweighing experienced relative to observed information. *Games and Economic Behavior*, 62:263–286, 2008.

S. B. Soofi, G. N. Khan, S. Ariff, Y. Ihtesham, M. Tanimoune, A. Rizvi, M. Sajid, C. Garzon, S. de Pee, and Z. 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*, 4:1–11, 2022.

W. J. W. Supply and S. M. Programme. *Progress on drinking water and sanitation: 2014 Update*. World Health Organization, 2014.

M. A. Tremblay, C. M. Blanchard, S. Taylor, L. G. Pelletier, and M. Villeneuve. Work extrinsic and intrinsic motivation scale: Its value for organizational psychology research. *Canadian Journal of Behavioural Science*, 41(4):213–226, 2009.

O. Urminsky, C. Hansen, and V. Chernozhukov. Using double-lasso regression for principled variable selection. *Working Paper*, 2016.

K. N. Vasilaky and A. M. Islam. Competition or cooperation? using team and tournament incentives for learning among female farmers in rural uganda. *World Development*, 103: 216–225, 2018.

K. N. Vasilaky and K. L. Leonard. As good as the networks they keep? improving outcomes through weak ties in rural uganda. *Economic Development and Cultural Change*, 66:755–792, 2018.

S. I. Watson, R. T. Rego, T. Hofer, and R. J. Lilford. Evaluations of water, sanitation and hygiene interventions should not use diarrhoea as (primary) outcome. *BMJ Global Health*, 7(5):e008521, 2022.

A. S. Wellsjo. Simple actions, complex habits: Lessons from hospital hand hygiene. Technical report, Working paper, 2021.

Z. Xu, Y. Liu, Z. Ma, G. S. Toloo, W. Hu, and S. Tong. Assessment of the temperature effect on childhood diarrhea using satellite imagery. *Scientific Reports*, 4(5389), 2014.

Online Supplementary Material

Appendix A Experimental Design Details

A.1 Study Procedures

Exclusion Criteria: Households are excluded if they do not store drinking water in a separate vessel from the water for other uses, if children drink from a separate vessel from adults, if the vessel's capacity is less than 10 liters (to ensure there would be enough water to avoid over-chlorination), or if children frequently drink bottled water (see Figure A.2 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.

Blinding: The study was non-blind with respect to study subjects and treatment administrators (enumerators). 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-blind. Enumerators used electronic tablets and smartphones to collect data.

Random Chlorine Testing: Before chlorine distribution began, chlorine tablets were not widely available in local markets, so we did not test for chlorine broadly. Instead, we randomized ten percent of households for chlorine testing during every other visit of Round 1 of Phase 1. We test chlorine in this pre-distribution period to ensure that knowledge of treatment status does not affect people's ability to procure chlorine tablets on their own.

A.2 Audit Visits

Our main outcome, the 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.

A.3 Measurement

Diarrhea Prevalence: 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 child-days 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.

Child Anthropometric Measurements: 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.

First Bi-Weekly Survey: In the first biweekly survey, we also collect 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 are 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 collect anthropometric data for 338 households for whom we were not able to collect anthropometric data in the baseline visit.

Endline Take-it-or-Leave-it: After providing respondents with a one-month supply of chlorine tablets, free of charge, we offer them a second month's supply of chlorine at a randomized price (market price, a 29% subsidy, or a 53% subsidy).

Endline Motivation for Using Chlorine: 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").

Endline 2 Stockpiling: 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, by the time of this survey, only 3% of respondents reported that they still had chlorine tablets remaining, 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 more than one month.

A.4 Photos of Intervention Materials

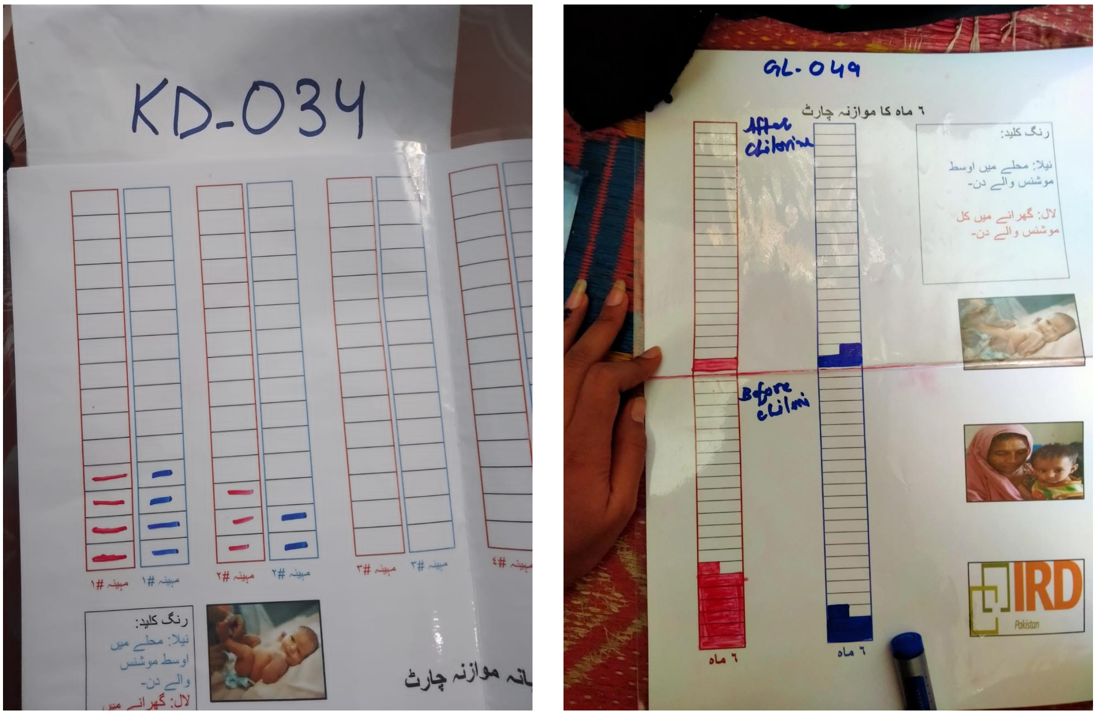


Figure A.1: Experiential Learning Intervention

Red represents a household's own diarrhea rate, that they fill in on their own in the two weeks between visits. If they did not fill in the data during the previous two weeks, the enumerator helps them fill it in retrospectively. During these bi-weekly visits, the enumerator also fills in the blue bar with the average rate among people who do not use chlorine (from Luby et al. (2006)).

In Panel B, data aggregated in the first three months before chlorine was distributed are in the bottom panel, and aggregated in the three months after chlorine was distributed are in the top panel. The enumerator aggregates this data for the respondent. She also fills in the aggregated rate among people who do not use chlorine in blue (from Luby et al. (2006)).

Figure A.2: Sample Water Vessel

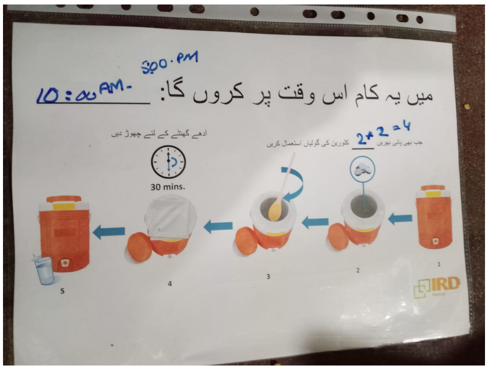


Figure A.3: Leaflet with Instructions on Chlorination

Appendix B Balance

Table B.1: Variables Predicting Number of Non-Attrited Visits Post-Behavioral Intervention

Variable	Lasso Coefficient
Endline Participant	5.07
Survey 10 Participant	2.35
Endline 2 Participant	1.78
Child 4 diarrhea: Missing	1.71
Child 1 weight: Missing	1.4
Survey 9 Participant	1.4
Survey 4 Participant	0.91
Survey 5 Participant	0.86
Survey 8 Participant	0.72
Block 3	0.46
Block 28	0.45
Block 16	0.3
Number of times visited before surveyed in follow-up 1	0.18
Closest water source is a pipeline tap	0.09
Child 1 weight	0.04
Mother's weight	0.04
Closest water source is a filtered pump	0
Any child had diarrhea in past 2 weeks	-0.24
Reports dirt in water	-0.27
Uses bottled water when children are sick	-0.31
Block 18	-0.4
Block 23	-0.44
Block 17	-0.63
Uses a water filter	-0.68
Believes diarrhea is highest in November	-0.98
Believes diarrhea is highest in October	-1

Figure B.1: Participation, Endline Participants (Panel A) and Attriters (Panel B)

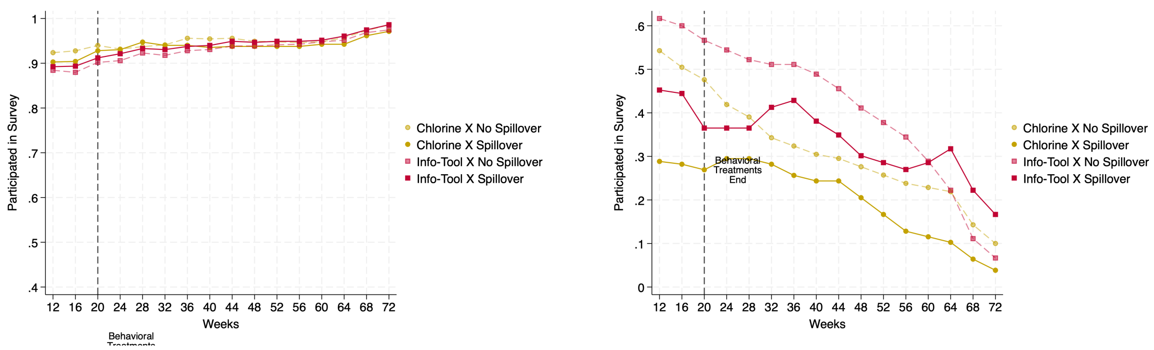


Table B.2: Treatment Balance: Spillover

	(1) Spillover Sample		(2) Spillover Sample	
Any Treatment Group	0.019	(0.025)		
Chlorine Only			0.018	(0.031)
Incentives			0.019	(0.031)
Info-Tool			0.019	(0.031)
Observations	1690		1690	
P-values:				
Incentives = Chlorine Only			0.975	
Info-Tool = Chlorine Only			0.981	
Info-Tool = Incentives			0.995	

Standard errors in parentheses. * $p < .1$, ** $p < 0.05$, *** $p < 0.01$
Each observation is at the household level. This regression includes neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.

Table B.3: Baseline Balance: Adult Outcomes

Variable	(1) Chlorine (no IT neighbor) Mean/SE	(2) Chlorine (with IT neighbor) Mean/SE	(3) Info-Tool (no IT neighbor) Mean/SE	(4) Info-Tool (with IT neighbor) Mean/SE	(5) Total Mean/SE	T-test Difference (1)-(2)	(1)-(3)	(1)-(4)
Any Child Had Motions	0.338	0.388	0.253	0.358	0.324	-0.050	0.085**	-0.020
	(0.028)	(0.038)	(0.026)	(0.037)	(0.016)			
Reported Highest Diarrhea in Summer	0.835	0.861	0.793	0.867	0.832	-0.026	0.042	-0.032
	(0.022)	(0.027)	(0.024)	(0.027)	(0.012)			
Number of Children <5	1.475	1.588	1.498	1.521	1.512	-0.113	-0.023	-0.046
	(0.042)	(0.071)	(0.038)	(0.076)	(0.026)			
Has Heard of Chlorine	0.190	0.200	0.235	0.248	0.217	-0.010	-0.045	-0.058**
	(0.023)	(0.031)	(0.025)	(0.034)	(0.014)			
Would consider using chlorine	0.870	0.897	0.867	0.897	0.879	-0.027	0.003	-0.027
	(0.020)	(0.024)	(0.020)	(0.024)	(0.011)			
Enumerator Observes Dirt in Water	0.137	0.200	0.172	0.127	0.158	-0.063	-0.035	0.010
	(0.020)	(0.031)	(0.022)	(0.026)	(0.012)			
Reports Dirt in Water	0.739	0.776	0.754	0.745	0.752	-0.036	-0.015	-0.006
	(0.026)	(0.033)	(0.026)	(0.034)	(0.014)			
Boils, Bleaches, or Chlorinates Water	0.144	0.109	0.126	0.182	0.139	0.035	0.018	-0.037
	(0.021)	(0.024)	(0.020)	(0.030)	(0.012)			
Strains or Filters Water	0.641	0.685	0.618	0.661	0.645	-0.044	0.023	-0.020
	(0.029)	(0.036)	(0.029)	(0.037)	(0.016)			
Believes Chlorine is for Water Purification	0.180	0.230	0.239	0.182	0.208	-0.051	-0.059	-0.002
	(0.023)	(0.033)	(0.025)	(0.030)	(0.014)			
Caretaker Asked about Chlorine Test	0.165	0.194	0.123	0.182	0.160	-0.028	0.043	-0.016
	(0.022)	(0.031)	(0.019)	(0.030)	(0.012)			
Attrited (endline)	0.123	0.158	0.105	0.127	0.125	-0.034	0.018	-0.004
	(0.020)	(0.028)	(0.018)	(0.026)	(0.011)			
N	284	165	285	165	899			
F-test of joint significance (F-stat)						0.636	1.363	0.778

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 B.4: Baseline Balance: Adult Outcomes (Endline Sample)

	(1)	(2)	(3)	(4)	(5)		T-test	
	Chlorine (no IT neighbor)	Chlorine (with IT neighbor)	Info-Tool (no IT neighbor)	Info-Tool (with IT neighbor)	Total		Difference	
Variable	Mean/SE	Mean/SE	Mean/SE	Mean/SE	Mean/SE	(1)-(2)	(1)-(3)	(1)-(4)
Any Child Had Motions	0.333	0.367	0.235	0.354	0.311	-0.034	0.098**	-0.021
	(0.030)	(0.041)	(0.027)	(0.040)	(0.017)			
Reported Highest Diarrhea in Summer	0.819	0.856	0.804	0.889	0.834	-0.037	0.015	-0.070
	(0.024)	(0.030)	(0.025)	(0.026)	(0.013)			
Number of Children <5	1.454	1.576	1.502	1.528	1.504	-0.122	-0.048	-0.074
	(0.042)	(0.080)	(0.039)	(0.085)	(0.028)			
Has Heard of Chlorine	0.197	0.216	0.239	0.236	0.221	-0.019	-0.042	-0.039*
	(0.025)	(0.035)	(0.027)	(0.036)	(0.015)			
Would consider using chlorine	0.867	0.899	0.863	0.889	0.875	-0.032	0.005	-0.021
	(0.022)	(0.026)	(0.022)	(0.026)	(0.012)			
Enumerator Observes Dirt in Water	0.137	0.187	0.176	0.125	0.156	-0.051	-0.040	0.012
	(0.022)	(0.033)	(0.024)	(0.028)	(0.013)			
Reports Dirt in Water	0.731	0.777	0.761	0.736	0.750	-0.046	-0.030	-0.005
	(0.028)	(0.035)	(0.027)	(0.037)	(0.015)			
Boils, Bleaches, or Chlorinates Water	0.133	0.101	0.125	0.188	0.135	0.032	0.007	-0.055
	(0.022)	(0.026)	(0.021)	(0.033)	(0.012)			
Strains or Filters Water	0.643	0.691	0.631	0.653	0.649	-0.048	0.011	-0.010
	(0.030)	(0.039)	(0.030)	(0.040)	(0.017)			
Believes Chlorine is for Water Purification	0.177	0.230	0.231	0.188	0.206	-0.054	-0.055	-0.011
	(0.024)	(0.036)	(0.026)	(0.033)	(0.014)			
Caretaker Asked about Chlorine Test	0.157	0.201	0.122	0.174	0.156	-0.045	0.035	-0.017
	(0.023)	(0.034)	(0.021)	(0.032)	(0.013)			
N	249	139	255	144	787			
F-test of joint significance (F-stat)						0.721	1.302	0.838

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 B.5: Baseline Balance (Full Sample): 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 B.6: 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)	(1)-(3)	(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 B.7: 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)	676	10.465	680	10.451	667	10.625	675	10.630	0.014	-0.160	-0.165
	[452]	(0.110)	[449]	(0.104)	[451]	(0.131)	[450]	(0.110)			
Child MUAC (cm)	676	14.546	680	14.463	667	14.501	675	14.613	0.083	0.045	-0.067
	[452]	(0.046)	[449]	(0.041)	[451]	(0.048)	[450]	(0.048)			
Number of Motion Days	676	0.760	680	0.929	667	0.887	675	0.778	-0.169*	-0.127	-0.018
	[452]	(0.067)	[449]	(0.077)	[451]	(0.080)	[450]	(0.073)			
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)			
F-test of joint significance (F-stat)									2.364*	1.368	0.530

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 clustered at variable HHID. Fixed effects using variable block are included in all estimation regressions. All missing values in balance variables are treated as zero.***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table B.8: 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	542	0.481	565	0.485	0.022	-0.002	-0.006
	[386]	(0.022)	[370]	(0.024)	[371]	(0.022)	[383]	(0.022)			
Age (in months)	572	30.885	541	30.638	542	30.226	565	31.599	0.247	0.659	-0.715
	[386]	(0.580)	[370]	(0.561)	[371]	(0.597)	[383]	(0.604)			
Child Weight (kg)	572	10.562	541	10.581	542	10.674	565	10.691	-0.019	-0.112	-0.130
	[386]	(0.118)	[370]	(0.122)	[371]	(0.128)	[383]	(0.120)			
Child MUAC (cm)	572	14.345	541	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)			
F-test of joint significance (F-stat)									0.960	0.423	0.264

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 clustered at variable HHID. Fixed effects using variable block are included in all estimation regressions. All missing values in balance variables are treated as zero.***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Appendix C 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 (they 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 (Figure C.1). Also similar to our study, the spillover and non-spillover samples converge towards the end of the trial. Figure C.1 reproduces the corresponding plot from our study using the same outcome scaling.

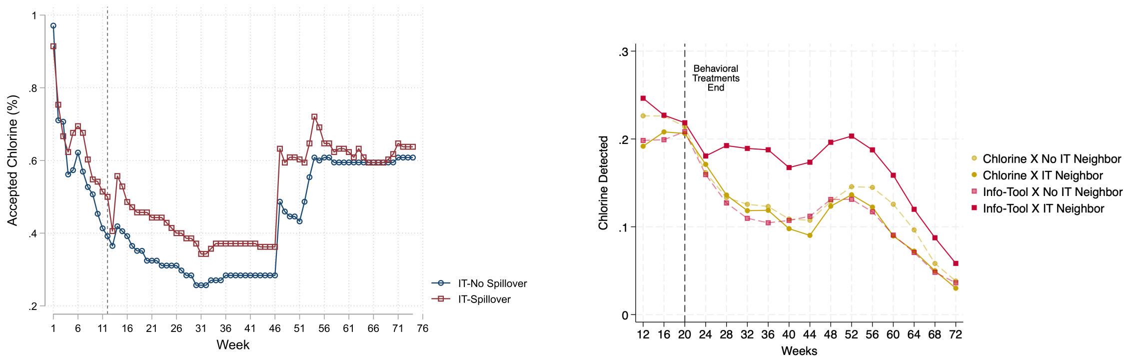


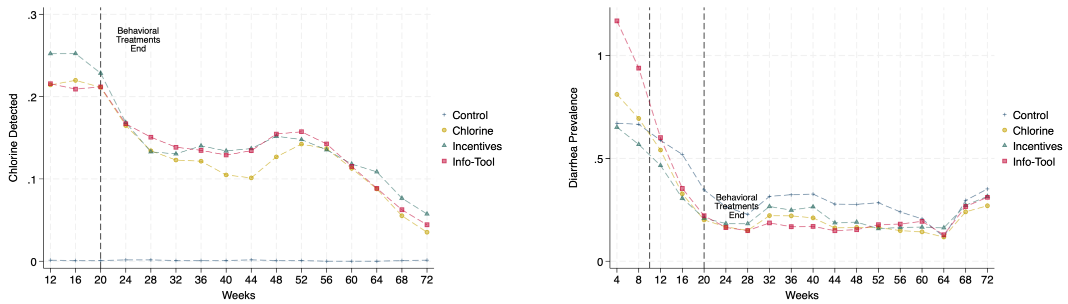
Figure C.1: Rates of Chlorine Acceptance
Akram and Mendelsohn (2021) Reanalysis (Panel A) versus Our Study (Panel B)
where "Spillover" is exposure to any Info-Tool neighbor

Panel A Notes: This figure replicates the analysis of Figure 3 using data from Akram and Mendelsohn (2021). It plots rates of chlorine tablet acceptance over time separately for Info-Tool participants with and without a nearby Info-Tool neighbor, defined using the 20-meter recentered proximity measure described in Section 4.1. The vertical dashed line marks the end of the Info-Tool intervention period in that study.

Panel B Notes: This figure reproduces Figure 3 to facilitate direct comparison with the AM reanalysis. It plots raw moving averages of chlorine detection for Chlorine-only and Info-Tool participants with and without a nearby Info-Tool neighbor, as defined in Section 4.1. The vertical dashed line marks the end of the behavioral-intervention period.

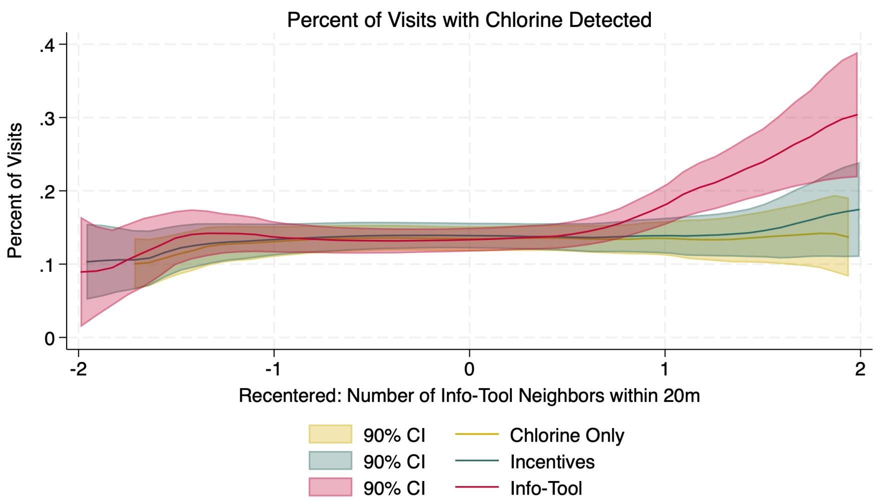
Appendix D Additional Figures and Tables

Figure D.1: Raw Moving Averages: Chlorine Detection and Diarrhea Prevalence



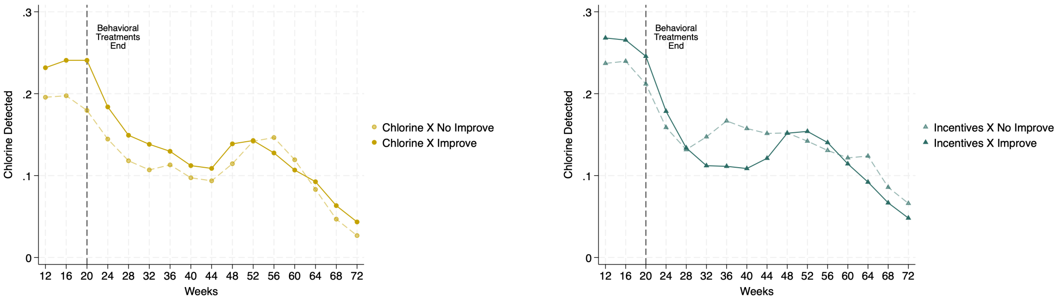
Notes: The left panel plots raw moving averages of chlorine detection across all four study arms over time. The right panel plots raw moving averages of self-reported child diarrhea prevalence across the same arms. The vertical dashed line marks the end of the behavioral-intervention period.

Figure D.2: Raw Rates of Chlorine Detection by Spillover Exposure (Continuous)



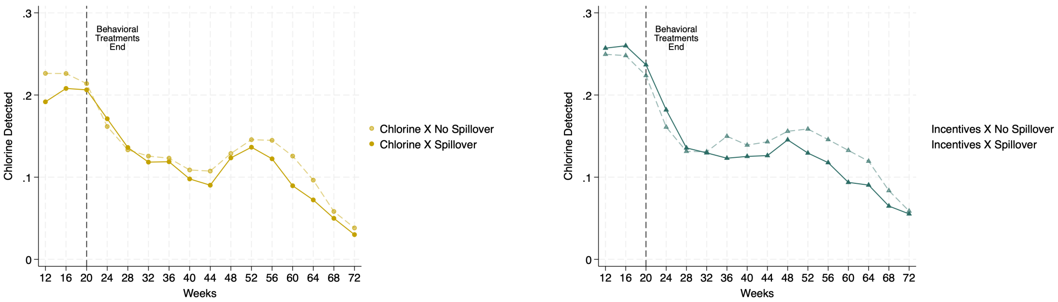
Notes: This figure plots the percent of visits where we detect chlorine by the recentered number of Info-Tool neighbors within 20m for each treatment group. The recentered measure is the actual number of Info-Tool neighbors within 20 meters, minus the average across 200 simulated treatment assignments. Negative values mean that an individual is less exposed than in expectation.

Figure D.3: Heterogeneity: Predicted Health Improvement



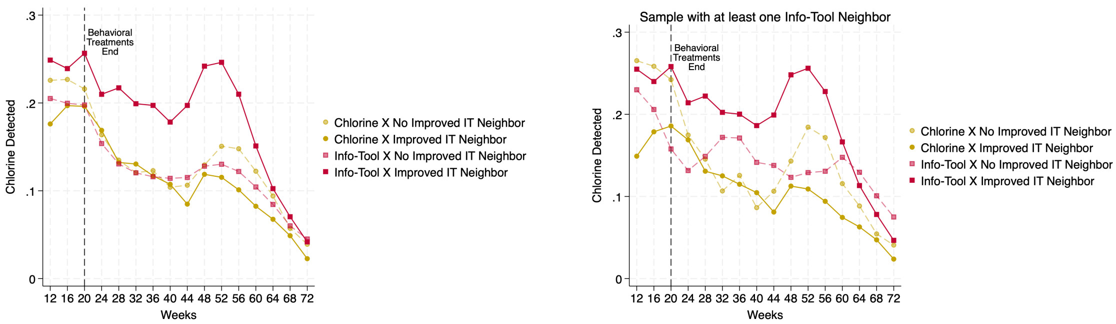
Notes: This figure plots raw moving averages of chlorine detection separately for households predicted to experience health improvements (above-median predicted improvement) and those not predicted to improve, for Chlorine Only (left panel) and Incentives participants (right panel). Predicted health improvement is constructed from pre-treatment variables using lasso among pure control households; see Section 5.1 for details. The vertical dashed line marks the end of the behavioral-intervention period.

Figure D.4: Heterogeneity: Info-Tool Neighbors (Re-centered)



Notes: This figure plots raw moving averages of chlorine detection by spillover exposure, separately for Chlorine Only participants (left panel) and Incentives participants (right panel). Within each panel, series are separated by whether the household was randomly more exposed to an Info-Tool neighbor within 20 meters than would be expected based on endogenous spatial factors. The spillover measure is defined as in Section 4.1. The vertical dashed line marks the end of the behavioral-intervention period.

Figure D.5: Heterogeneity: Predicted-Improved Info-Tool Neighbors



Notes: This figure plots raw moving averages of chlorine detection for Chlorine-Only and Info-Tool participants, separately by whether they are randomly more exposed to a predicted-improved Info-Tool neighbor than would be expected based on endogenous spatial factors. The left panel includes all participants; the right panel restricts to participants with at least one Info-Tool neighbor. A predicted-improved Info-Tool neighbor is an Info-Tool participant within 20 meters whose diarrhea rate was ex-ante predicted to improve during the behavioral-intervention period; see Section 5.1 for details. The vertical dashed line marks the end of the behavioral-intervention period.

Table D.1: Placebo Test – Chlorine Detection by exposure to any Chlorine Only neighbor

	(1) During-BI Full Sample		(2) Post-BI Full Sample		(3) During-BI Endline Sample		(4) Post-BI Endline Sample	
Chlorine × Chlorine-Only Neighbor	0.114	(0.097)	0.175	(0.241)	0.097	(0.102)	0.126	(0.253)
Info-Tool × No Chlorine-Only Neighbor	0.024	(0.058)	0.267*	(0.145)	0.030	(0.061)	0.286*	(0.151)
Info-Tool × Chlorine-Only Neighbor	0.000	(0.095)	0.210	(0.236)	0.016	(0.100)	0.237	(0.248)
Observations	1192		1216		1108		1136	
Chlorine-Only without Chlorine-Only Neighbors Mean	0.563		0.563		0.563		1.477	
P-values:								
Info-Tool: No Chlorine-Only Neighbors = Chlorine-Only Neighbors	0.806		0.810		0.888		0.845	

Standard errors in parentheses

Each observation is at the household level. The outcome is the total number of times that chlorine was detected across all visits in the specified time period. During-BI refers to the first three months of chlorine distribution during the Behavioral Interentions. Post-BI refers to all time periods after the Behavioral Interventions. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. The omitted groups is Chlorine-Only participants who were less exposed to a Chlorine Only neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors.

* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.2: Placebo Test – Chlorine Detection by exposure to any Incentives neighbor

	(1) During-BI Full Sample		(2) Post-BI Full Sample		(3) During-BI Endline Sample		(4) Post-BI Endline Sample	
Chlorine × Incentives Neighbor	0.004	(0.106)	0.276	(0.255)	-0.015	(0.111)	0.153	(0.267)
Info-Tool × No Incentives Neighbor	-0.040	(0.064)	0.287*	(0.152)	-0.041	(0.066)	0.275*	(0.159)
Info-Tool × Incentives Neighbor	0.013	(0.105)	0.293	(0.252)	0.037	(0.110)	0.283	(0.264)
Incentives × No Incentives Neighbor	0.150**	(0.062)	0.308**	(0.149)	0.129**	(0.065)	0.313**	(0.158)
Incentives × Incentives Neighbor	0.021	(0.106)	0.015	(0.254)	0.007	(0.111)	-0.052	(0.268)
Observations	1599		1625		1467		1503	
Chlorine-Only without Incentives Neighbors Mean	0.614		1.414		0.636		1.484	
P-values:								
Info-Tool: No Incentives Neighbors = Incentives Neighbors	0.616		0.980		0.479		0.976	
Incentives: No Incentives Neighbors = Incentives Neighbors	0.222		0.249		0.277		0.174	

Standard errors in parentheses

Each observation is at the household level. The outcome is the total number of times that chlorine was detected across all visits in the specified time period. During-BI refers to the three months during chlorine distribution and the Behavioral Interventions. Post-BI refers to all visits after the Behavioral Interventions. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. The omitted group are Chlorine-Only households who were randomly less exposed to an Incentives neighbor (someone within 20m) than they would be in expectation based on endogenous spatial factors.

* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.3: Placebo Test – Chlorine Detection by exposure to any Chlorine-Only neighbor predicted to improve

	(1) Full Sample		(2) Endline Sample	
Chlorine × Improved Chlorine Neighbors	-0.181	(0.196)	-0.173	(0.205)
Info-Tool × No Improved Chlorine Neighbors	0.165	(0.130)	0.218	(0.136)
Info-Tool × Improved Chlorine Neighborss	0.134	(0.195)	0.160	(0.205)
Observations	1282		1191	
Chlorine No Improved Chlorine Neighbors Mean	1.532		1.556	
P-values:				
Info-Tool: Spillover = No Spillover	0.876		0.774	
Info-Tool × Spillover = Info-Tool × No Spillover + Chlorine × Spillover	0.574		0.685	

Standard errors in parentheses
Each observation is at the household level. The outcome is the aggregate of the total number of child-days of diarrhea that a household self-reports across all visits in the post-behavioral intervention time period. The Chlorine Only treatment group without spillover neighbors is the ommitted group. An indicator for Pure Control participants, who are included in the sample, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Long-run is defined as the remainder of the trial (twelve months). The Spillover Sample is defined as participants who were randomly more exposed to a Chlorine Only neighbor (someone within 20m) whose diarrhea rate 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.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.4: Endline Stated Motivations
Omitted group: Chlorine Only

	(1) Motivation: Health		(2) Motivation: Habit		(3) Motivation: Money		(4) Motivation: Intrinsic	
Incentives	0.088	(0.062)	-0.031	(0.064)	-0.020	(0.063)	-0.023	(0.032)
Info-Tool	0.191***	(0.062)	-0.049	(0.064)	-0.068	(0.062)	-0.027	(0.033)
Observations	1049		1057		1051		1058	
P-values:								
Incentives = Info-Tool	0.094		0.769		0.444		0.880	

Standard errors in parentheses
Each observation is at the household level. The outcome in column (1) is a rating from 1 to 7 for how true the following statement felt: "Because using treatments like chlorine tablets help me to attain a certain level of health for my family." The outcome in column (2) is a rating from 1 to 7 for how true the following statement felt: "Because it has become a fundamental part of my routine." The outcome in column (3) is a rating from 1 to 7 for how true the following statement felt: "Because it could allow me to earn money." The outcome in column (4) is the average of ratings from 1 to 7 for the following statements: "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." All regressions control for the average motivation score the respondent gave across all motivation questions, the order of questions (randomized), neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.5: Endline Stated Perception of Neighbors in Treatment Groups
Omitted group: Pure Control

	(1) Any IT Neighbor		(2) Any Chlorine-Only Neighbor		(3) Any Incentives Neighbor	
Chlorine-Only	0.006	(0.028)	0.050*	(0.029)	-0.021	(0.030)
Incentives	0.039	(0.028)	-0.024	(0.029)	0.024	(0.030)
Info-Tool	0.189***	(0.028)	0.005	(0.029)	0.057*	(0.030)
Observations	1455		1455		1455	
Control Mean	0.284		0.563		0.496	
P-values:						
Info-Tool = Chlorine-Only	0.000		0.128		0.010	
Info-Tool = Incentives	0.000		0.337		0.280	

Standard errors in parentheses
Each observation is at the household level. The outcome takes the value of 1 if the participant guesses that they have at least one neighbor from the treatment group in the Info-Tool group (column 1), the Chlorine-Only group (column 2), or the Incentives group (column 3), and 0 otherwise. The omitted group is the Pure Control group. 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 total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.6: Endline Stated Trust
Omitted group: Pure Control

	(1) Info-Tool Knows Most	(2) Chlorine Knows Most	(3) **Robustness:** Incentives Knows Most
Chlorine Only	-0.043	0.122***	-0.079**
	(0.027)	(0.034)	(0.031)
Incentives	-0.004	-0.085**	0.088***
	(0.027)	(0.034)	(0.031)
Info-Tool	0.222***	-0.138***	-0.083***
	(0.027)	(0.034)	(0.031)
Observations	1516	1516	1516
Control Mean	0.196	0.527	0.276
DID Estimate:			
Info-Tool picks Info-Tool			
– Other Group picks Own Group		0.100 [p= 0.046]	0.133 [p= 0.004]

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). The outcome for column (2) is an indicator for if the respondent chose a hypothetical Chlorine Only respondent as the person most likely to be knowledgeable about child health (rather than a Incentives or Info-Tool participant). The outcome for column (3) is an indicator for if the respondent chose a hypothetical Incentives respondent as the person most likely to be knowledgeable about child health (rather than an Chlorine Only or Info-Tool participant). All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.7: Endline Stated Trust
Omitted group: Pure Control

	(1) Info-Tool Knows Most	(2) Incentives Knows Most	(3) Chlorine Knows Most
Chlorine × No Spillover	-0.023	-0.105***	0.128***
	(0.034)	(0.038)	(0.042)
Chlorine × Spillover	-0.079*	-0.030	0.109*
	(0.046)	(0.053)	(0.058)
Incentives × No Spillover	0.001	0.069*	-0.070*
	(0.034)	(0.038)	(0.042)
Incentives × Spillover	-0.013	0.125**	-0.112*
	(0.046)	(0.052)	(0.058)
Info-Tool × No Spillover	0.256***	-0.115***	-0.141***
	(0.034)	(0.039)	(0.043)
Info-Tool × Spillover	0.159***	-0.025	-0.134**
	(0.046)	(0.052)	(0.058)
Observations	1516	1516	1516
P-values:			
Outcome Group: Spillover = No Spillover	0.091	0.394	0.792
DID Estimate:			
Info-Tool Spillover Effect			
– Other Group Spillover Effect		-0.152 [p= 0.110]	-0.078 [p= 0.468]

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). The outcome for column (2) is an indicator for if the respondent chose a hypothetical Incentives respondent as the person most likely to be knowledgeable about child health (rather than a Chlorine Only or Info-Tool participant). The outcome for column (3) is an indicator for if the respondent chose a hypothetical Chlorine Only respondent as the person most likely to be knowledgeable about child health (rather than an Incentives or Info-Tool participant). All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, 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 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.8: Chlorine Detection
by exposure to any 'Proactive' Info-Tool Neighbor

	(1) Full Sample		(2) Endline Sample	
Chlorine × Proactive IT Neighbor	0.071	(0.179)	0.027	(0.186)
Info-Tool × No Proactive IT Neighbor	0.068	(0.143)	0.062	(0.149)
Info-Tool × Proactive IT Neighbor	0.493***	(0.181)	0.564***	(0.187)
Observations	1216		1136	
Chlorine-Only without Proactive IT Neighbor Mean	1.520		1.539	
P-values:				
Info-Tool: No Proactive IT Neighbor = Proactive IT Neighbor	0.020		0.008	
Info-Tool × Proactive IT Neighbor = Chlorine-Only × Proactive IT Neighbor + Info-Tool × No Proactive IT Neighbor	0.150		0.064	

Standard errors in parentheses
Each observation is at the household level. The outcome is the total number of times that chlorine was detected across all visits in the specified time period. During-BI refers to the time period during the Behavioral Interventions, while Post-BI refers to the time period after the Behavioral Interventions. The omitted group is households in the Chlorine Only treatment group who were not 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. An indicator for Pure Control participants, who are included in the sample to maximize power, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. The omitted group is Chlorine-Only households who were randomly more exposed to an Info-Tool neighbor (someone within 20m) who ever filled out the Info-Tool themself without enumerator assistance over a two-week period, than they would be in expectation based on endogenous spatial factors..
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

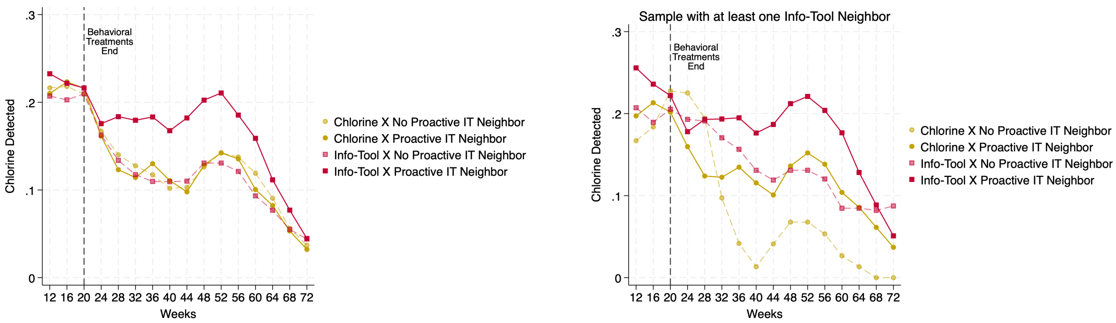


Figure D.6: Heterogeneity: Proactive Info-Tool Neighbors

Notes: This figure plots raw moving averages of chlorine detection for Chlorine-Only and Info-Tool participants, separately by whether they are exposed to a proactive Info-Tool neighbor. The left panel includes all participants; the right panel restricts to participants with at least one Info-Tool neighbor. A proactive Info-Tool neighbor is an Info-Tool participant within 20 meters who independently filled in the Info-Tool chart for at least one two-week period without enumerator assistance. The vertical dashed line marks the end of the behavioral-intervention period.

Table D.9: Chlorine Detection
by 'Proactivity' of Info-Tool Participants

	(1) Full Sample		(2) Endline Sample	
Info-Tool × Not Proactive	-0.107	(0.176)	-0.060	(0.181)
Info-Tool × Proactive	0.302**	(0.126)	0.340**	(0.132)
Observations	1216		1136	
Chlorine-Only Mean	1.494		1.518	
P-values:				
Info-Tool: Self Proactive = Not Proactive	0.027		0.036	

Standard errors in parentheses
Each observation is at the household level. The outcome is the total number of times that chlorine was detected across all visits in the time period after the Behavioral Interventions. The omitted group is households in the Chlorine Only treatment group. An indicator for Pure Control participants, who are included in the sample to maximize power, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. Info-Tool neighbors are considered proactive if they ever filled out the Info-Tool themselves without enumerator assistance over a two-week period.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

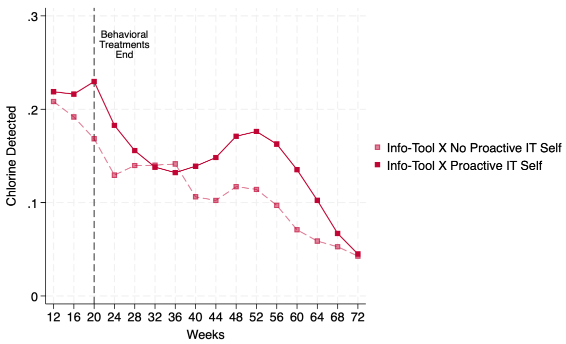


Figure D.7: Heterogeneity: Proactive Info-Tool Participants

Notes: This figure plots raw moving averages of chlorine detection separately for Info-Tool participants who independently filled in the Info-Tool chart for at least one two-week period without enumerator assistance ("proactive") and those who always waited for the enumerator to complete the chart ("non-proactive"). The vertical dashed line marks the end of the behavioral-intervention period.

Table D.10: Endline Willingness-to-Pay

	(1) Purchased Tablets	(2) Price Accepted	(3) Hypothetical WTP	(4) Sold or Revisited
Chlorine Only	-0.020*	-2.832*	0.124	-0.044**
	(0.010)	(1.487)	(1.472)	(0.019)
Incentives	-0.012	-2.220	0.215	-0.037*
	(0.010)	(1.491)	(1.482)	(0.019)
Info-Tool	0.004	0.026	0.598	-0.020
	(0.010)	(1.485)	(1.494)	(0.019)
Observations	1503	1503	1099	1503
Control Mean	0.027	4.059	5.217	0.106
P-values:				
Incentives = Chlorine Only	0.487	0.685	0.949	0.685
Info-Tool = Chlorine Only	0.023	0.057	0.742	0.204
Info-Tool = Incentives	0.115	0.136	0.791	0.389

Standard errors in parentheses
Each observation is at the household level. This regression includes neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls. The outcome for column (1) is an indicator for if the household ultimately purchased chlorine tablets. The outcome for column (2) is the price at which the household purchased chlorine tablets (the price is 0 for households who did not purchase tablets). The outcome for column (3) is a hypothetical price at which the household said they would be willing to purchase chlorine tablets, if they refused the offered price. The outcome for column (4) is an indicator for if the household purchased the chlorine tablets *or* asked the enumerator to return at a later date when they expected to have cash on hand.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table D.11: Endline Willingness-to-Pay

	(1) Purchased Tablets	(2) Price Accepted	(3) Hypothetical WTP	(4) Sold or Revisited
Chlorine × No Spillover	-0.020*	-2.779	-0.735	-0.047**
	(0.012)	(1.714)	(1.689)	(0.022)
Chlorine × Spillover	-0.019	-2.975	1.786	-0.037
	(0.016)	(2.261)	(2.147)	(0.029)
Incentives × No Spillover	-0.009	-1.671	0.267	-0.056***
	(0.012)	(1.692)	(1.691)	(0.022)
Incentives × Spillover	-0.019	-3.463	0.125	0.006
	(0.016)	(2.343)	(2.226)	(0.030)
Info-Tool × No Spillover	0.005	0.224	-0.453	-0.018
	(0.012)	(1.711)	(1.717)	(0.022)
Info-Tool × Spillover	0.002	-0.404	2.572	-0.022
	(0.016)	(2.263)	(2.183)	(0.029)
Observations	1503	1503	1099	1503
Control Mean	0.027	4.059	5.217	0.106
P-values:				
Chlorine: No Spillover = Spillover	0.918	0.939	0.293	0.760
Incentives: No Spillover = Spillover	0.589	0.492	0.955	0.062
Info-Tool: No Spillover = Spillover	0.856	0.806	0.216	0.908

Standard errors in parentheses
Each observation is at the household level. This regression includes neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls. The outcome for column (1) is an indicator for if the household ultimately purchased chlorine tablets. The outcome for column (2) is the price at which the household purchased chlorine tablets (the price is 0 for households who did not purchase tablets). The outcome for column (3) is a hypothetical price at which the household said they would be willing to purchase chlorine tablets, if they refused the offered price. The outcome for column (4) is an indicator for if the household purchased the chlorine tablets *or* asked the enumerator to return at a later date when they expected to have cash on hand.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Appendix E Main Tables: Robustness to Alternate Specification

For robustness, we also estimate the following specification using household-survey-visit panel specification:

$$
\begin{aligned}
Y_i =& \theta_0 + \theta_{1,r} T1_i + \theta_{2,r} T2_i + \sum_{r=20}^{R} \theta_{3,r} AnyT2_i^r \\
&+ \sum_{r=20}^{R} \theta_{4,r} AnyT2_i^r \times T1_i + \sum_{r=20}^{R} \theta_{5,r} AnyT2_i^r \times T2_i \\
&+ Y_{i0} + X_{i0} + \gamma_b + \epsilon_i
\end{aligned}
$$

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, the density of study participants within twenty meters[40], 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.[41] Table E.1 reports panel-specification estimates of the spillover specification for the post-behavioral-intervention period; Table E.2 reports the corresponding aggregate estimates for the behavioral-intervention period; and Tables E.3 and E.4 report panel-specification analogues of the heterogeneity results in Section 5.1. Across all four, the patterns from the aggregate specification hold.

[40] Section 4.1 details the selection of this bandwidth following Egger et al. (2022).

[41] 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.

Table E.1: Chlorine Detection: Panel Specification
where "Spillover" is exposure to any Info-Tool neighbor

	(1) Full Sample		(2) Endline Sample	
Chlorine × Spillover	0.007	(0.014)	0.007	(0.014)
Info-Tool × No Spillover	0.005	(0.013)	0.007	(0.013)
Info-Tool × Spillover	0.041**	(0.017)	0.050***	(0.018)
Observations	15341		14647	
Chlorine No Spillover Mean	0.127		0.125	
P-values:				
Info-Tool: No Spillover = Spillover	0.037		0.014	
Info-Tool × Spillover = Info-Tool × No Spillover + Chlorine × Spillover	0.191		0.105	

Standard errors in parentheses
Each observation is a household-level survey visit in the post-Behavioral Intervention period. The outcome is a binary indicator for whether or not chlorine residual was detected in the respondent's water during the survey visit. The Chlorine Only treatment group without spillover neighbors is the omitted group. An indicator for Pure Control participants, who are included in the sample, is in the estimating equation but not reported. All specifications include survey-round fixed effects, neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, 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.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table E.2: Chlorine Detection: Aggregate Specification, Behavioral Intervention Period
where "Spillover" is exposure to any Info-Tool neighbor

	(1) Full Sample		(2) Endline Sample	
Chlorine × Spillover	-0.008	(0.070)	-0.037	(0.073)
Info-Tool × No Spillover	-0.034	(0.058)	-0.028	(0.061)
Info-Tool × Spillover	-0.033	(0.069)	-0.023	(0.073)
Observations	1234		1143	
Chlorine No Spillover Mean	0.645		0.665	
P-values:				
Info-Tool: Spillover = No Spillover	0.993		0.938	
Info-Tool × Spillover = Info-Tool × No Spillover + Chlorine × Spillover	0.928		0.678	

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 during the behavioral interventions. The Chlorine Only treatment group without spillover neighbors is the ommitted group. An indicator for Pure Control participants, who are included in the sample, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. Short-run is defined as the first three months of chlorine distribution (while the behavioral interventions were ongoing); Long-run is defined as the remainder of the trial (twelve 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 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table E.3: Chlorine Detection: Panel Specification
by predicted health improvement

	(1) During-BI		(2) Post-BI	
Chlorine × Self Improved	0.031	(0.029)	0.021	(0.014)
Info-Tool × Self Not Improved	-0.028	(0.026)	0.018	(0.014)
Info-Tool × Self Improved	0.028	(0.028)	0.036**	(0.014)
Observations	2286		10082	
Chlorine × Self Not Improved Mean				
P-values:				
Info-Tool × Self Improved				
= Info-Tool × Self Not Improved	0.043		0.229	

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. During-BI refers to the time period during the Behavioral Interventions, while Post-BI refers to the time period after the Behavioral Interventions. The Chlorine Only treatment group who were not predicted to improve in their diarrhea rate during the Behavioral Intervention period is the omitted group. An indicator for Pure Control participants, who are included in the sample to maximize power, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. Self 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 5.1 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table E.4: Chlorine Detection: Panel Specification
by exposure to any Info-Tool neighbor predicted to improve

	(1) During-BI		(2) Post-BI	
Chlorine × IT Neighbor Improved	-0.029	(0.033)	0.017	(0.015)
Info-Tool × No IT Neighbor Improved	-0.024	(0.022)	0.012	(0.011)
Info-Tool × IT Neighbor Improved	-0.007	(0.031)	0.043**	(0.020)
Observations	3534		10082	
Chlorine × No Spillover Mean	0.226		0.226	
P-values:				
Info-Tool: No Spillover = Spillover	0.581		0.111	

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. During-BI refers to the time period during the Behavioral Interventions, while Post-BI refers to the time period after the Behavioral Interventions. The omitted group is households in the Chlorine Only treatment group who were not 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. An indicator for Pure Control participants, who are included in the sample to maximize power, is in the estimating equation but not reported. 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 total number of study participants within twenty meters, and lasso-selected baseline controls. IT Neighbor Improved is a binary indicator for if the participant's was randomly more exposed to an Info-Tool neighbor (someone within 20m) whose health was predicted to improve at the beginning of chlorine distribution more than the median. See Section 5.1 for a detailed explanation for how the predicted-health-improvement measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Appendix F Model Appendix: Simulation

Under a reasonable set of parameters, our model generates dynamic predictions that match our empirical data. We simulate individuals' learning and technology adoption over sixteen periods, where the first four periods represent the "treatment period" (the behavioral interventions), and the following twelve periods represent the "post-treatment period" (free chlorine distribution in all groups). First, we simulate a dataset of 1350 individuals and divide them into three groups (450 per group, as in our experimental design). All three groups undergo the same learning and adoption process, with the following exceptions: one group, modeled on the Incentives group, faces lower costs of technology adoption in the first four periods; and another group, modeled on the Info-Tool group, places heavier experience-weights on the signals observed through the Info-Tool in the first four periods (whether those signals are their own or their Info-Tool friends').

F.1 Simulated Data Set-Up

Individuals have a prior belief about chlorine efficacy, μ_0, which is distributed in the population as: $\mu_0 \sim \mathcal{N}(.3, .3)$. We let variance $\sigma_0^2 = 0.3$ be constant in the population. We assume that prior beliefs are largely positive and certain because most participants' only information about chlorine is from the enumerator, who sends a positive signal about chlorine and represents a trusted information source.

In every period where participants adopt chlorine, they receive a signal about chlorine efficacy from nature $Y \sim \mathcal{N}(0.58, 0.49)$ in the population. This is modeled on the experimental sample mean and standard deviation of the probability that participants' diarrhea rate decreased in the three months after chlorine distribution relative to the three months before. We assume that every participant tries chlorine at least once in period 0 (which represents "day 1" of the treatment period), so that every participant begins period 1 with a signal Y. We model participants' uncertainty about this belief with $\sigma_Y^2 = 1$. Participants apply an *experience-weight* to this belief, $\alpha_Y = 0.3$.

Social Learning

We randomly assign friends using $Pr(F = 1) = 0.0015$, sampling with replacement. On average, participants have 2.7 friends each (maximum = 9, minimum = 0), and 0.6 Info-Tool friends each (maximum = 4, minimum = 0). In each period, participants talk with their friends about chlorine with probability $Pr(T = 1) = 0.2$. In period t, the friend sends a signal, which is the average of the signals about chlorine that they received in the previous three periods $Y_{\bar{t}}$:

$$Y_{\bar{t}} = \frac{\sum_{p=t-2}^{p=t} Y_p \cdot \mathbb{1}(M_\alpha^p > C^p)}{\sum_{p=t-2}^{p=t} \mathbb{1}(M_\alpha^p > C^p)}$$

This means that, if a friend has not adopted chlorine in the preceding three periods, they do not share information about chlorine, even if they talk. When participants learn from themselves (signals from nature), they only update with the information they gain in this particular period, Y_t. Different from friends, who we assume only learn from each others' signals if they talk (which happens with some random probability), individuals will *always* incorporate information that they generate if they adopt chlorine. In other words, there is no analog for randomly talking in the way that people incorporate their own signals.

Technology Adoption

In each period, i incorporates signals into her prior (one at a time) from herself and her neighbors. After incorporating all the new signals in the time period, she decides to adopt chlorine if $M_\alpha > C$, where we assume that $C = 0.6$. Recall that, if she does *not* adopt chlorine, she will not gain any new information about chlorine in the next time period. However, she can still learn from her friends about chlorine (if they talk in that time period, and if her friend has adopted chlorine in the past three time periods).

Interventions

For 450 individuals, we let $C = 0.5$ in the first four time periods (the Incentives group). For another 450 individuals, we let $\alpha_Y = 1$ for signals that are acquired with the Info-Tool (signals they receive from themselves in the first four periods, or signals they receive from others in the same group in the first 7 periods, since individuals share prior information up to three periods later). The remaining 450 participants undergo "standard" learning, and represent the Chlorine Only group.

F.2 Dynamic Simulations

Summary of base parametric assumptions:

$$\begin{aligned} \mu_0 &\sim \mathcal{N}(0.3, 0.3) \qquad & \sigma_0 &= 0.3 \\ Y &\sim \mathcal{N}(0.58, 0.49) \qquad & \sigma_Y^2 &= 1 \\ \alpha_Y &= 0.3 \\ C &= 0.6 \end{aligned}$$

Summary of interventions:

$$\begin{aligned} \text{Incentives: } & C = 0.5, \forall t <= 4 \\ \text{Info-Tool: } & \alpha_Y = 1, \forall Y_{t<=4 \text{ \& } i=j} \\ \text{Info-Tool: } & \alpha_Y = 1, \forall Y_{t<=7 \text{ \& } j\in\{IT\} \text{ \& } i\neq j} \end{aligned}$$

Our model generates a pattern of adoption that resembles our raw data (Figure F.1, left-hand side).

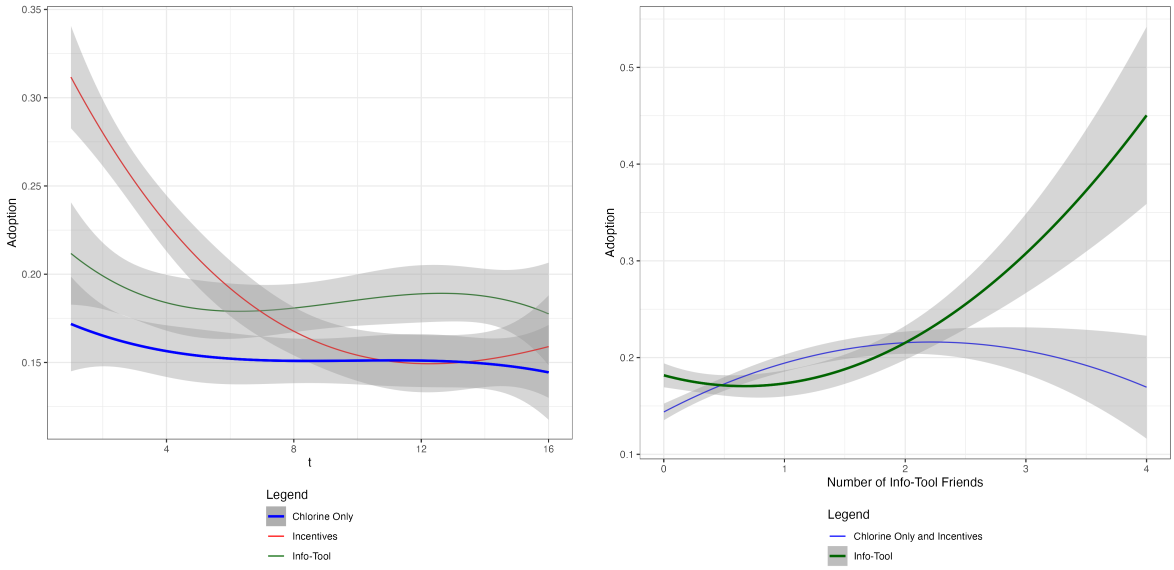


Figure F.1: Model Simulation: Chlorine Adoption

We observe a positive and *increasing* relationship between the number of Info-Tool friends and chlorine adoption within the Info-Tool group. However, this relationship is not present in the other groups. (Figure F.1, right-hand side). While adoption dynamics differ by intervention groups, and according to the number of Info-Tool neighbors, *stated* posterior beliefs M (i.e. explicit knowledge) do not (Figure F.2).

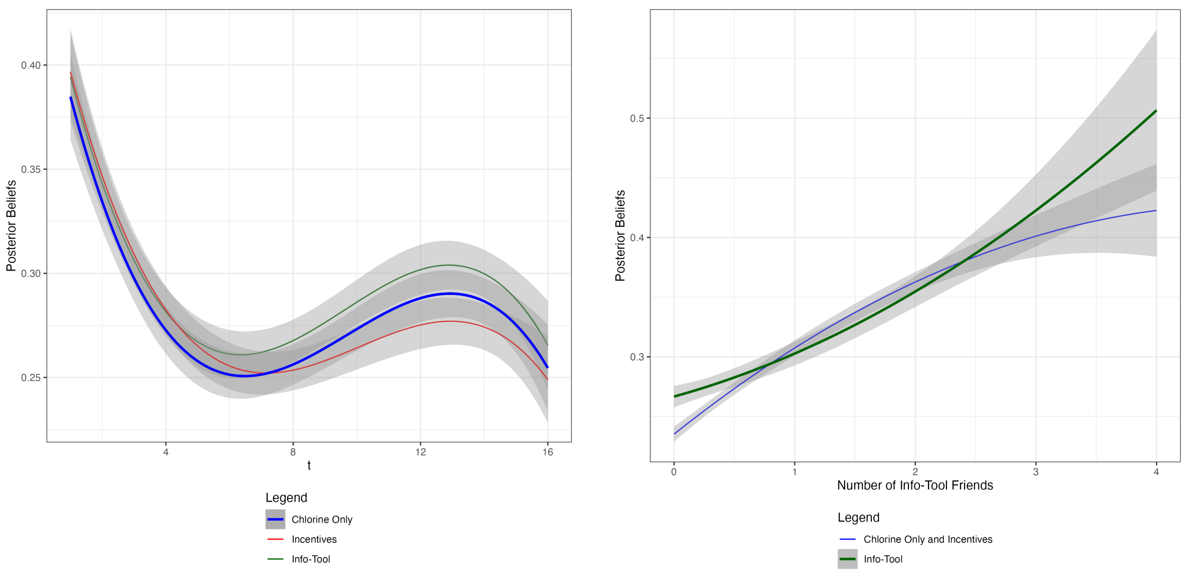


Figure F.2: Model Simulation: Stated Posterior Beliefs

Appendix G AEA Registry Analysis

G.1 Background

Randomized controlled trials (RCTs) are a cornerstone of modern empirical research, but their validity often rests on the assumption that individual treatment effects are independent of other individuals' treatment status, commonly referred to as the Stable Unit Treatment Value Assumption (SUTVA). Economists have made substantial progress in analyzing cross-treatment spillovers, where treated units influence control units, to understand general equilibrium effects. However, much less attention has been given to *within*-treatment spillovers, where one treated unit affects *another* treated unit, despite their relevance for economic theory and policy. For individually randomized experiments where there are cross-treatment spillovers, comparing average outcomes between treated and control units may conceal individual treatment effects in the treatment group. A common way to address this is through clustered randomization. In cluster-randomized experiments, researchers randomize clusters of individuals – groups that ostensibly do not interact with each other – to treatment or control, reducing the possibility that treated units' treatment status will influence outcomes among control units. However, cluster-randomized designs with full treatment saturation, where entire clusters are assigned to treatment or control, can *not* address the possibility of complementarity between receiving treatment and being exposed to other people who receive treatment.

Two-stage cluster-randomized designs with varying treatment saturation *can* identify within-treatment spillovers by randomizing clusters to different treated-to-control ratios. This creates random variation in exposure to treated units for both treated and control units, making it possible to measure and understand spillovers comprehensively. However, these designs are rarely used in practice due to logistical challenges and reduced statistical power for detecting individual treatment effects. Moreover, when implemented, they often focus on forces generating cross-treatment spillovers, such as general equilibrium effects in markets, and not on the forces that we argue are similarly likely to generate within-treatment spillovers: the behavioral and social forces behind learning and technology adoption.

Researchers identify or negate cross-treatment spillovers through two primary methods: analyzing the treatment effect of distance (geographic or social) to treated units among control units, and through cluster-randomized designs. In designs with random variation in distance to treated units among control units, cross-treatment spillovers are identified but not neutralized (designs such as Miguel and Kremer (2004)). In cluster-randomized designs where every unit is treated in some clusters and no units are treated in other clusters, cross-treatment spillovers are neutralized but not identified (designs that randomize at a group level such as school-level randomization in Muralidharan and Sundararaman

(2011)). Designs that randomize treatment saturation across clusters, and then randomize individuals to treatment or control within clusters according to the predetermined ratio, are cluster-randomized designs that effectively randomize individuals' distance to treated units (designs such as Egger et al. (2022)). Concerns about the inability to detect true treatment effects due to cross-treatment spillovers have led cluster-randomized designs to be highly prevalent in economics randomized controlled trials, especially within development economics.

G.2 Methods

We downloaded the AEA RCT Registry data on November 11, 2024 (aea, 2024). Since there is no single variable indicating the randomization procedure, we utilized ChatGPT 4o to help conduct a text analysis of text-response variables.

Cluster-randomized versus individually-randomized trials: We identify individually-randomized trials from cluster-randomized trials using the "randomization unit", "sample size number of clusters", and "sample size number of observations" variables. The assignment rules were as follows.

ClusteredTrial = 0 if any of the following conditions are true:

1. "Sample size number clusters" is equal to any of the following (ignoring cases and punctuation): "no cluster", "no clusters", "NA", "none", "same as observations", "no", "individual", or is blank
2. "Sample size number clusters" includes any of the following phrases (ignoring cases and punctuation): "treatment is not clustered", "no clustering", "there are no clusters", "equals the total number of participants", "not clustered", "not applicable", "treatment will not be clustered", "there are no planned clusters", "this is not cluster-randomized", "this is not a cluster-randomized", "NA ", "N/A ", "does not have clustered randomization", "does not cluster randomize", "there is no need for clustering", or "no cluster"
3. "Sample size number of clusters" = "Sample size number observations"
4. "Randomization unit" is equal to any of the following (ignoring cases and punctuation): "individual", "we will randomize individuals", "individual is the unit of randomization", "individual level", "individual randomization for treatments", "treatments are randomized within-session at the individual level", "randomization takes part on the participant level", "individual survey participant", "individual worker", "individual participant", "the randomization was done using stratified random sampling at the individual level", or "randomization will be done at the individual level"

5. "Randomization unit" includes any of the following (ignoring cases and punctuation): "without clustering"
6. The only difference between the values in "Randomization Unit" and "Sample size number of observations" is that "Sample size number of observations" has a number that is not present in "Randomization Unit"
7. The number in "Sample size number of clusters" is the same as the number in "Sample size number of observations" and "Randomization unit" includes the word "individual" and "Randomization Unit" does not include the word "cluster", and "Sample size number of clusters" does not include any words (only numbers)
8. The number in "Sample size number of clusters" is the same as the number in "Sample size number of observations" and "Randomization unit" does not include the word "cluster", and "Sample size number of observations" does not include any words

Among cases where ClusteredTrial $\neq 0$, ClusteredTrial $= 1$ if any of the following conditions are true:

1. "Sample size number of clusters" includes both a number and a unit (for example, "300 schools", "100 households" or "50 villages")
2. "Experimental design" or "Experimental design details" include the words "medium treatment saturation", "high treatment saturation", or "low treatment saturation"
3. "Experimental design" or "Experimental design details" include the words "% saturation" or "% treatment saturation"
4. There is a number included in both the "Sample size number of observations" variable and the "Sample size number of clusters" variable, and these numbers are not the same
5. "Sample size number of clusters" and "Sample size number of observations" are both just numbers, and "Sample size number of clusters" is a smaller number than "Sample size number of observations"

After implementing these rules, we are able to classify 96.8% of trials (the treatment design of 304 out of 9494 trials remain unidentified).

Full- vs. partial-treatment saturation: We then classify a trial as cluster-randomized with varying treatment saturation across clusters if the the trial is classified as cluster-randomized, and if the registration ever uses terms related to "treatment saturation" or "two-stage randomization" across any variable.

KeywordSat $= 1$ for observations that mention any word stemming with "saturat" across any variable. Otherwise, KeywordSat = 0.

KeywordTwoStage = 1 if any variable includes the phrase "two stage randomization", "two-stage randomization", "2-stage randomization" or "2 stage randomization". Otherwise, KeywordTwoStage = 0.

We consider a trial to be cluster-randomized with varying treatment saturation if ClusteredTrial = 1 and either KeywordSat =1 or KeywordTwoStage = 1.

Research Topics: For each trial, researchers select "keywords" associated with their trial. First, we make minor cleaning adjustments for cohesion across variations in specific keywords that people use:

1. "behavior" includes: "behavior", "behavioral", "behavioral economics"
2. "environment" includes: "environment", "climate change", "environment and energy"
3. "health" includes: "health", "covid-19", "mental health", "nutrition"
4. "crime violence and conflict" includes: "crime violence and conflict", "post-conflict"
5. "experiment" includes: "experiment", "rct", "online experiment", "survey experiment", "field experiment", "lab"
6. "firms" includes: "firms", "firms and productivity"
7. "productivity" includes: "productivity", "firms and productivity"

Next, we identify all of the keywords that are associated with at least fifty trials (we do not analyze keywords that are excessively niche). This leaves us with thirty-nine keywords to analyze.

G.3 Results: Prevalence of Clustered Randomization

Among all projects registered on the American Economics Association RCT Registry, 35% of trials in high-income countries (HICs) featured cluster-randomized designs, whereas 62% of trials in low- and middle-income countries (LMICs) featured cluster-randomized designs. Of trials that did not specify or publish the country, 48% were cluster-randomized. We used the World Bank classification of country income.

Of all the cluster-randomized trials registered to the American Economics Association RCT Registry, only 2.3% were classified as cluster-randomized with varying treatment saturation (98 trials total). An additional 26 trials met the criteria based on the keywords "treatment saturation" or "two-stage randomized", but were individually randomized trials (a case of type 1 error). Of the cluster-randomized trials with varying treatment saturation, 38% were within LMICs, 9% were within HICs, and 53% did not specify or publish the country.

G.4 Results: By Topic

For which questions do researchers implement cluster-randomized designs, especially those with varying treatment saturation? To analyze the prevalence of research topics across experimental designs, we use the following regression for each keyword K:

$$K_t = \beta_0 + \beta_1 C_t + \beta_2 S_t + \beta_3 C_t \times S_t + \gamma_t + \epsilon_t \tag{1}$$

where $K_t = 1$ if trial t was registered with keyword K, and 0 otherwise; $C_t = 1$ if the trial is cluster-randomized; $S_t = 1$ if the trial included the terms "treatment saturation" or "two-stage randomization" anywhere in the registration; and γ_t is a world-region fixed effect (HIC, LMIC, or unspecified). Then β_1 is the probability that a fully-saturated cluster-randomized trial is associated with keyword K relative to individually-randomized trials; and $\beta_2+\beta_3$ is the probability that a cluster-randomized trial with varying treatment saturation is associated with keyword K relative to fully-saturated cluster-randomized trials.

Keywords that are *more* likely to be associated with fully-saturated cluster-randomized trials, relative to individually-randomized trials, include "agriculture", "cash transfers", "communication", "early childhood development", "education", "gender", "health", "incentives", "poverty", and "technology adoption". Keywords that are *less* likely to be associated with fully-saturated cluster-randomized trials, relative to individually-randomized trials, include "altruism", "behavior", "beliefs", "electoral", "fairness", "inequality", "labor", "migration", "redistribution", "social media", "trust", and "other". There are no statistically significant differences in the probability of individual- or clustered-randomization among the remaining keywords.[42]

Keywords that are more likely to be associated with cluster-randomized trials with varying treatment saturation, relative to fully saturated cluster-randomized trials, are "agriculture" and "migration" (there are no statistically distinguishable differences across any other keywords).

<u>Keywords Associated with Our Trial</u>

We analyze the prevalence of treatment designs across trials that use top keywords that we group together as "behavioral", "health", "technology adoption", or "communication".

[42]Keywords with no differences in treatment design: "cooperation", "crime violence and conflict", "discrimination", "entrepreneurship", "environment", "experiment", "finance", "financial literacy", "firms", "governance", "information", "productivity", "savings", "social norms", "taxation", "training", and "welfare".

Behavioral trials are most likely individually randomized, and, when cluster-randomized, almost always with full treatment saturation. Health, technology adoption, and communication trials are very likely to be cluster-randomized, but usually with full treatment saturation.

Cluster-randomized trials with full saturation are 11pp ($24\%, p < 0.001$) *less* likely to be associated with a top behavioral economics keyword ("behavior", "beliefs", "fairness", or "altruism") than individually-randomized trials. Cluster-randomized trials with varying treatment saturation are another 8pp ($25\%, p = 0.099$) less likely to include a top behavioral economics keyword, relative to fully saturated cluster-randomized trials.

Cluster-randomization with full saturation – the design that allows for the strongest within-treatment spillovers *but conceals them* – is the most common experimental design researchers implement for economics trials focused on technology adoption, health, or communication. Technology adoption trials are 0.7pp ($163\%, p < 0.001$) more likely to be cluster-randomized than individually randomized; health trials are 1.9pp ($12\%, p = 0.020$) more likely to be cluster-randomized than individually randomized; and communication trials are 0.5pp ($55\%, p = 0.024$) more likely to be cluster-randomized than individually randomized.

Appendix H Detailed Literature Review

H.1 Experiential Learning in Health

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.

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; Fafchamps et al., 2024); rather, in our context, learning appears to be most effective when it is action-based and similarly experienced by others in one's network. This has relevance for 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 find that a part of the psychological mechanism underlying an Info-Tool participant's behavioral change is increased trust in other Info-Tool participants' knowledge about child health. 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).[43]

Our findings also speak to the role of the identity of the transmitter – determined here by how one acquires signals – on the perceived value of the transmitted information. Actors are sensitive to *how* information is acquired and more readily respond to information whose acquisition process they are personally familiar with. This may contribute to why people are hesitant to trust health information relayed by experts (Alsan and Wanamaker, 2018; Banerjee et al., 2023; Darden and Macis, 2024; Lowes and Montero, 2021; Martinez-Bravo and Stegmann, 2022), yet update their beliefs and actions strongly in response to personal experiences and information from in-group members (Alsan et al., 2019; Alsan and Eichmeyer, 2024; Bennett et al., 2018; Conlon et al., 2022; Corno, 2014; D'Acunto et al., 2021; Malmendier and Nagel, 2015; Simonsohn et al., 2008).[44] Our findings suggest that, for certain activities, embedding learning experiences within a community rather than importing information from outside may maximize long-term behavior change. 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 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.

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

[44] Reliance on experts may also inadvertently deepen social and economic inequities given disparities in trust of experts, 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).

H.2 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.

Appendix I Child Health

I.1 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 I.1 presents the results. Across all treatment groups, the impact of chlorine tablet dispensation is substantial: diarrhea rates drop by approximately 30% ($p < 0.05$) 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 report and discuss in Section 4.4. 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.

I.2 Anthropometrics

ITT Results: Any Treatment

Table I.2 reports the ITT effects of *any* treatment on the various measures of child anthropometry, whereas Table I.3 reports the treatment-specific estimates. Table I.3 shows that the improvements in anthropometric outcomes are consistently higher for households in the Info-Tool group. Table I.6 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. We find that treatment compliance leads to a 0.24 SD increase in the index of child anthropometrics ($p < 0.05$).

Table I.1: Child-Days of Diarrhea (Household-Survey Panel)

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine Only	-0.186***	(0.062)	-0.070	(0.045)	-0.091***	(0.032)
Incentives	-0.198***	(0.057)	-0.027	(0.045)	-0.050	(0.033)
Info-Tool	-0.136**	(0.062)	-0.092**	(0.042)	-0.091***	(0.031)
Observations	7872		6354		14066	
Control Mean	0.515		0.240		0.277	
P-values:						
Chlorine = Incentives	0.821		0.271		0.187	
Chlorine = Info-Tool	0.378		0.528		0.990	
Incentives = Info-Tool	0.228		0.070		0.167	

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, the number of study participants within twenty meters, 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 I.2: 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**	0.016	-0.005	0.112*	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, the number of study participants within twenty meters, and lasso-selected controls. Indeces are constructed using following Anderson (2008).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.3: 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**	0.020	0.007	0.080	0.025
	(0.039)	(0.084)	(0.099)	(0.081)	(0.064)
Incentives	0.030	0.022	-0.098	0.065	0.061
	(0.042)	(0.078)	(0.101)	(0.079)	(0.066)
Info-Tool	0.108***	0.006	0.072	0.187**	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	0.186	0.979	0.310	0.853	0.572
Chlorine = Info-Tool	0.455	0.867	0.523	0.193	0.386
Incentives = Info-Tool	0.044	0.836	0.104	0.130	0.761

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, the number of study participants within twenty meters, and lasso-selected controls. Indeces are constructed using following Anderson (2008).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.4: ITT Impacts on Endline Child Health (Separated by Treatment and Spillover)

	Omitted: Chlorine × No Spillover				Omitted: Other Chlorine Only & IT Groups				Omitted: Pure Control			
	(1) HAZ	(2) WHZ	(3) WAZ	(4) MUAC-for-age	(5) HAZ	(6) WHZ	(7) WAZ	(8) MUAC-for-age	(9) HAZ	(10) WHZ	(11) WAZ	(12) MUAC-for-age
Chlorine × No Spillover									-0.022	0.038	0.103	0.002
									(0.097)	(0.114)	(0.094)	(0.074)
Chlorine × Spillover	0.153	-0.086	-0.056	0.081					0.129	-0.048	0.050	0.081
	(0.128)	(0.147)	(0.121)	(0.088)					(0.113)	(0.134)	(0.109)	(0.084)
Info-Tool × No Spillover	-0.017	0.018	0.066	0.154*					-0.041	0.055	0.167*	0.153**
	(0.104)	(0.125)	(0.103)	(0.079)					(0.088)	(0.111)	(0.090)	(0.074)
Info-Tool × Spillover	0.088	0.065	0.152	-0.044	0.056	0.080	0.141	-0.120	0.063	0.104	0.258**	-0.045
	(0.128)	(0.168)	(0.131)	(0.102)	(0.111)	(0.152)	(0.116)	(0.093)	(0.115)	(0.157)	(0.120)	(0.100)
Observations	2371	2439	2492	1954	2371	2439	2492	1954	2371	2439	2492	1954

Standard errors in parentheses

Child-level cross-section of the endline survey. Standard errors are clustered at the household level. Columns (1)–(4) omit Chlorine × No Spillover, columns (5)–(8) omit all other Chlorine Only and Info-Tool groups, and columns (9)–(12) omit Pure Control. Within each specification, columns report the four anthropometric components used to construct the Anderson (2008) index: HAZ, WHZ, WAZ, and MUAC-for-age. All regressions include baseline unbalanced household-level controls, baseline child anthropometrics and diarrhea rates, child gender, child age, neighborhood block fixed effects, the number of study participants within twenty meters, and lasso-selected controls. The Spillover Sample is defined as participants who were randomly more exposed to an Info-Tool neighbor within twenty meters than they would be in expectation based on endogenous spatial factors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.5: Child Index of Anthropometry (ITT):
where "Spillover" is exposure to any Info-Tool neighbor

	(1) Omitted: Other Chlorine Only & IT Groups		(2) Omitted: All Other Treated Groups		(3) Omitted: Pure Control		(4) Omitted: Chlorine × No Spillover	
Chlorine × No Spillover					0.099**	(0.045)		
Chlorine × Spillover					0.040	(0.056)	-0.051	(0.060)
Info-Tool × No Spillover					0.079*	(0.044)	-0.012	(0.049)
Info-Tool × Spillover	0.079*	(0.044)	0.100**	(0.042)	0.158***	(0.056)	0.066	(0.060)
Observations	2616		2616		2616		2616	
P-values:								
Chlorine × Spillover = Chlorine × No Spillover					0.331			
Chlorine × Spillover = Info-Tool × No Spillover					0.519		0.523	
Chlorine × Spillover = Info-Tool × Spillover					0.081		0.087	
Info-Tool × No Spillover = Info-Tool × Spillover					0.186		0.196	
Info-Tool × Spillover = Chlorine × Spillover + Info-Tool × No Spillover							0.126	

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, the number of study participants within twenty meters, and lasso-selected controls. The regression in column (1) controls for indicator variables for being in the following groups: Incentives × No Spillover, and Incentives × Spillover. The regression in columns (2) and (3) control for indicator variables for being in the following groups: Control, Incentives × No Spillover, and Incentives × Spillover. 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 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table I.6: 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	0.240**	0.050	-0.017	0.371*	0.198
	(0.112)	(0.207)	(0.260)	(0.212)	(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	119.93	127.97	121.36	122.46	95.19
C-statistic p-value	0.027	0.491	0.392	0.139	0.056

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, the number of study participants within twenty meters, and neighborhood block fixed effects. Indeces are constructed using following Anderson (2008).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.7: 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	0.506***	-0.083	0.584	0.789**	-0.428
	(0.183)	(0.376)	(0.516)	(0.398)	(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	35.26	34.36	33.85	36.26	27.02
C-statistic p-value	0.002	0.986	0.471	0.062	0.379

Standard errors in parentheses

The endogenous variable is whether the household reported at endline that it boils, bleaches, or chlorinates its water. The excluded instrument is assignment to the Info-Tool group interacted with spillover exposure, where "Spillover" is exposure to at least one Info-Tool neighbor within twenty meters. 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, the number of study participants within twenty meters, and neighborhood block fixed effects. Indeces are constructed using following Anderson (2008).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix J Competing Mechanisms Discussion

J.0.1 Early Adoption

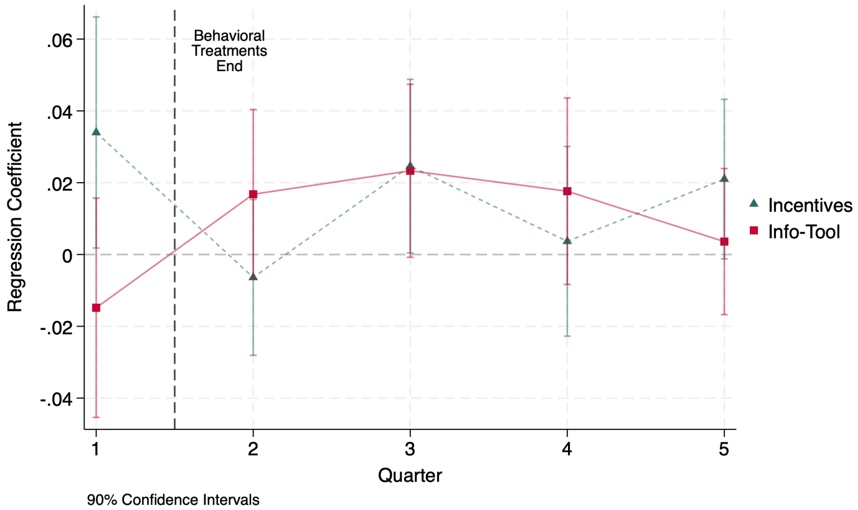
Alternatively, the Info-Tool might lead participants to adopt chlorine early on at a higher rate than they would without the Info-Tool. Long-term adoption could then be explained by early adoption leading to habit formation, or intertemporal complementarities in chlorine use. We use the Incentives arm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

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.[45] 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-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.

During the behavioral intervention period, Incentives households chlorinate significantly more than those in Info-Tool or Chlorine Only. However, immediately following incentive withdrawal (week 20), chlorination rates for incentive households plummet, while the decay among Info-Tool households is gradual. Figure J.1 plots the regression coefficients for the impact of Incentives and Info-Tool, relative to Chlorine Only, on chlorination using the panel dataset from household visits (also reported in Tables J.2 and J.1). Consistent with the observations from the raw chlorination plots, during the behavioral interventions period (Quarter 1), households assigned to Incentives outperform those in Info-Tool ($p = 0.014$) and Chlorine Only ($p = 0.090$). In the immediate post-behavioral-intervention period (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

[45] 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.

Figure J.1: Regression Coefficients of Chlorine Detection (Omitted: Chlorine Only)



Notes: This figure plots quarterly regression coefficients for the impact of the Info-Tool and Incentives arms on chlorine detection, relative to the Chlorine Only arm, estimated using the household-survey panel specification reported in Table J.2. Quarter 1 covers the behavioral-intervention period; Quarters 2–4 cover successive quarters of the post-behavioral-intervention period. Error bars indicate 95% confidence intervals. Standard errors are clustered at the household level.

(p=0.142). Effects converge thereafter.

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 detect chlorine at a higher rate in Incentives than in Info-Tool. Furthermore, we detect 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 are not more likely to use chlorine when they expect that we will come to test their water than at other times.

Since the Incentives arm builds 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 borne out in the long term in the Incentives arm. The Incentives group immediately reverts to the same rate of chlorine use as the Chlorine Only group when the behavioral treatments end, so the mechanism that helps the Info-Tool sustain chlorine adoption in the months following the behavioral interventions cannot be explained by habit formation or any story related to early adoption.

Table J.1: 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)	0.545***	(0.046)	1.085***	(0.089)
Info-Tool	0.622***	(0.050)	0.647***	(0.046)	1.059***	(0.089)
Observations	1599		1661		1653	
P-values:						
Chlorine = Incentives	0.039		0.490		0.091	
Chlorine = Info-Tool	0.706		0.139		0.162	
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, the number of study participants within twenty meters, 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 J.2: Chlorine Detection (Household-Survey Panel)

	(1) Short-Run		(2) Medium-Run		(3) Long-Run	
Chlorine Only	0.221***	(0.015)	0.147***	(0.011)	0.108***	(0.009)
Incentives	0.255***	(0.015)	0.142***	(0.010)	0.126***	(0.010)
Info-Tool	0.208***	(0.014)	0.168***	(0.011)	0.122***	(0.009)
Observations	4711		6354		14066	
P-values:						
Chlorine = Incentives	0.090		0.689		0.117	
Chlorine = Info-Tool	0.484		0.142		0.201	
Incentives = Info-Tool	0.014		0.055		0.749	

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, the number of study participants within twenty meters, 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).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

J.0.2 More Interactions and Mimicry

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 are not larger than any other groups' (Table J.3), and this holds within the Info-Tool spillover sample (Table J.4). Furthermore, we have no evidence of Info-Tool participants mimicking one another on other behaviors, including take-up of other health programs that are present during our study trial, or choice of household savings technology (Table J.6).

Table J.3: Endline Social Networks

	(1) Number in Social Network	(2) Number Discuss Health	(3) Number Discussed Water Purification
Incentives	-0.051	-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	0.377	0.730	0.015

Standard errors in parentheses
Each observation is at the household level. All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table J.4: 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	0.157	-0.065	0.133**
	(0.115)	(0.081)	(0.067)
Incentives × No Spillover	-0.019	-0.050	0.148***
	(0.096)	(0.068)	(0.056)
Incentives × Spillover	0.046	-0.037	0.186***
	(0.114)	(0.080)	(0.067)
Info-Tool × No Spillover	-0.048	0.004	0.037
	(0.097)	(0.068)	(0.057)
Info-Tool × Spillover	-0.085	-0.081	0.083
	(0.115)	(0.081)	(0.067)
Observations	1116	1116	1116
P-values:			
Incentives × No Spillover = Incentives × Spillover	0.574	0.877	0.576
Info-Tool × No Spillover = Info-Tool × Spillover	0.754	0.297	0.502

Standard errors in parentheses
Each observation is at the household level. All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, 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 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

J.0.3 Social Norms

It is possible that participants act on what they think *others* believe, via a preference to conform to social norms, rather than their own beliefs. We ask 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 believe a guest would accept chlorinated water

Table J.5: Endline Social Networks

	(1) Number in Social Network	(2) Number Discuss Health	(3) Number Discussed Water Purification
Any Treatment Group	-0.030	-0.000	0.117***
	(0.062)	(0.042)	(0.035)
Observations	1503	1503	1503
Control Mean	1.715	1.007	0.307

Standard errors in parentheses
Each observation is at the household level. All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table J.6: 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	0.041
	(0.038)	(0.026)	(0.011)	(0.065)
Chlorine × Spillover	-0.041	0.011	0.008	-0.079
	(0.053)	(0.032)	(0.014)	(0.080)
Incentives × No Spillover	0.046	0.017	0.002	0.033
	(0.038)	(0.025)	(0.011)	(0.065)
Incentives × Spillover	0.102*	-0.028	0.025*	-0.045
	(0.052)	(0.031)	(0.013)	(0.080)
Info-Tool × No Spillover	0.102***	0.016	0.012	-0.086
	(0.038)	(0.026)	(0.011)	(0.065)
Info-Tool × Spillover	0.050	-0.018	0.021	-0.166**
	(0.053)	(0.032)	(0.014)	(0.080)
Observations	1512	1512	1512	1503
P-values:				
Chlorine × No Spillover = Chlorine × Spillover	0.092	0.722	0.689	0.165
Incentives × No Spillover = Incentives × Spillover	0.388	0.182	0.104	0.373
Info-Tool × No Spillover = Info-Tool × Spillover	0.424	0.314	0.510	0.353

Standard errors in parentheses
Each observation is at the household level. All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, 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 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

(67%), there is a large effect of being in any treatment group on this measure (Table J.7). However, the Info-Tool group is no more likely to believe a guest would accept water than the other two groups. If anything, Incentives participants are the most likely to believe a guest would accept chlorinated water. We also ask participants to report to the best of their ability how each of their network connections purifies their water. Control group participants are very unlikely to believe that their network connections purify their water with chlorine (6.4% believe one network connection purifies their water with chlorine), but the rate is 92.9% higher for participants in any treated group. Again, Info-Tool is no more likely to believe their network connections purify water than any of the other treated groups (Table J.7). The same holds when we further break the analysis down by Info-Tool spillover exposure (Table J.8).

Table J.7: Endline Social Norms

	(1) Believes Neighbor Would Accept Chlorinated Water	(2) Number of Friends Believes Chlorinates
Chlorine Only	0.082***	0.041*
	(0.030)	(0.022)
Incentives	0.139***	0.079***
	(0.030)	(0.022)
Info-Tool	0.098***	0.061***
	(0.030)	(0.022)
Observations	1503	1690
P-values:		
Chlorine = Incentives	0.061	0.073
Chlorine = Info-Tool	0.601	0.347
Incentives = Info-Tool	0.176	0.403

Standard errors in parentheses
Each observation is at the household level. All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table J.8: 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**	0.041*
	(0.037)	(0.025)
Chlorine × Spillover	0.088*	0.040
	(0.051)	(0.030)
Incentives × No Spillover	0.157***	0.072***
	(0.037)	(0.025)
Incentives × Spillover	0.109**	0.092***
	(0.050)	(0.030)
Info-Tool × No Spillover	0.060	0.049**
	(0.037)	(0.025)
Info-Tool × Spillover	0.164***	0.082***
	(0.051)	(0.030)
Observations	1503	1690
P-values:		
Chlorine × No Spillover = Chlorine × Spillover	0.892	0.969
Incentives × No Spillover = Incentives × Spillover	0.445	0.531
Info-Tool × No Spillover = Info-Tool × Spillover	0.100	0.316

Standard errors in parentheses
Each observation is at the household level. All specifications include neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, 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 4.1 for a detailed explanation of how the measure is defined.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

J.0.4 More Conversations

We ask participants if they have 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 are much less likely to discuss water purification than treated participants, but Info-Tool participants are no more likely to discuss water purification with their social network connections than the other two treated groups. If anything, the Incentives group participants are the most likely to discuss water purification with others (Table J.3).

J.0.5 Testing the Perceived Precision Channel

Recall from Section 3 that our model permits experience with the Info-Tool to affect adoption through two channels: by reducing perceived signal uncertainty $\hat{\sigma}_Y^2$ (and thus shifting posterior beliefs M through the updating weight ℓ), or by raising the weight α_Y that participants place on Info-Tool signals when deciding whether to act. In this section, we test the first channel directly. If the Info-Tool operates by sharpening participants' comprehension of signals about chlorine — and thereby shifting beliefs — two conditions should hold: (1) the Info-Tool leads participants to better comprehend the signals about chlorine that they observe; and (2) Info-Tool participants share different information about chlorine than participants in the other treatment groups, since the comprehension channel implies a downstream effect on the content of social transmission. We find evidence of neither, which suggests that the perceived-precision channel is empirically inactive in our setting. Behavioral change must therefore operate through the experience-weight channel α_Y.

No Evidence of Differential Comprehension

Halfway through Phase 2 (the week 44 survey), we conduct an incentivized belief elicitation activity. We simply asked participants to tell us whether their children's diarrhea rates increased, decreased, or remained the same during the three months following the start of chlorine distribution, compared to the three months prior. If participants answer correctly, they receive a gift (children's goods such as pencils or a notepad). If the Info-Tool improves signal comprehension, then participants should be more likely to answer the question correctly. However, there are no differences in correctly answering the question across any groups, with or without access to Info-Tool neighbors (Table J.9). Access to Info-Tool neighbors also does not change which answer Info-Tool participants give, indicating non-differential knowledge or optimism about chlorine efficacy.[46]

[46] Control participants were less likely to say that their diarrhea rate had improved than the treatment groups (accurately so), but there were no differences between any groups that received chlorine on average.

Table J.9: Midline Memory/Beliefs about Chlorine Efficacy

	(1) Correct Answer	(2) Went Down	(3) Didn't Change	(4) Went Up	(5) Don't Know
Chlorine × Spillover	0.014	0.107**	-0.036	-0.042	-0.028*
	(0.050)	(0.045)	(0.028)	(0.039)	(0.014)
Info-Tool × No Spillover	0.062	0.094***	-0.027	-0.040	-0.032***
	(0.039)	(0.036)	(0.022)	(0.031)	(0.011)
Info-Tool × Spillover	0.008	0.029	-0.003	-0.010	-0.018
	(0.051)	(0.046)	(0.029)	(0.040)	(0.015)
Observations	1123	1123	1123	1123	1123
Control Mean	0.484	0.705	0.071	0.205	0.017
P-values:					
Info-Tool × No Spillover = Info-Tool × Spillover	0.332	0.198	0.435	0.493	0.404

Standard errors in parentheses

Each observation is at the household level. This regression includes neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls. In the week 44 survey, we ask participants to tell us whether their children's diarrhea rates increased, decreased, or remained the same during the three months following the start of chlorine distribution, compared to the three months prior. If participants answer correctly, they receive a gift (children's goods such as pencils or a notepad). The outcome of column (1) takes the value of 1 if respondents report the correct answer based on our own data, and 0 otherwise. The outcome of column (2) takes the value of 1 if respondents report that their child's diarrhea rate decreased, and 0 otherwise. The outcome of column (3) takes the value of 1 if respondents report that their child's diarrhea rate didn't change, and 0 otherwise. The outcome of column (4) takes the value of 1 if respondents report that their child's diarrhea rate increased, and 0 otherwise. The outcome of column (5) takes the value of 1 if respondents report that they do not know what happened to their children's diarrhea rate, and 0 otherwise.

* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

To ascertain that answering the question correctly is sensitive to comprehension skills, we test whether response accuracy is differential by education. Reassuringly, participants in any treatment group who had ever been to school (sixty-four percent of the sample had received zero years of education) are 6.6 percentage points (13%, $p = 0.035$) more likely to answer the question correctly (Table J.10).[47] Interestingly, Control participants with any education are *less* likely to answer the question correctly, partially explained because educated Control participants are more likely to answer "I don't know" than any other group. This suggests that access to technological experimentation alone *does* lead participants to increase the attention they pay to signals from their observed environment, especially among people with education. Jointly, this set of results demonstrates that the Info-Tool does not actually increase the accuracy of participants' knowledge or the optimism of their beliefs about chlorine efficacy. At endline, we ask several more non-incentivized questions about beliefs in chlorine efficacy and the riskiness of the disease environment, none of which yield differential responses across treatment groups.[48]

[47] The education treatment effect was non-differential across treated arms among participants who received any chlorine treatment.

[48] 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 before 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). It is possible that beliefs updated im-

No Evidence of Differential Information Sharing

We can rule out that Info-Tool participants talk to more people, or discuss water purification or child health with more people (see Section J.0.2 and Section J.0.4). Since we do not observe participants' conversations, we cannot determine exactly *what* they say about water purification or child health in these conversations. However, recall that Info-Tool participants are no more likely to answer positively or accurately about chlorine efficacy when correct answers are incentivized, and there is no reason to believe that only the Info-Tool group would convey a different set of beliefs to each other within their private conversations. Furthermore, we ask a few questions at endline to understand the nature of participants' conversations about water purification, and there are no differences across treatment groups.

We ask participants to guess the water purification method of each member of their social network. Using each participant's name, nickname, husband's name, and neighborhood block, we are able to link social network nodes with participants in our sample. If Info-Tool participants are sharing information about chlorine tablets that is more favorable towards chlorine than other groups, we might expect participants to believe that *their* Info-Tool friends are chlorine users. We see no differences in participants' beliefs about their treated friends' chlorine tablet use (Table J.11). Interestingly, participants *are* more likely to believe that their friends in *any* treatment group are more likely to use chlorine than friends in the Control group or outside the study sample. This implies that participants have some awareness of who uses chlorine tablets, and that this question captures relevant information about participants' propensity to share information about the water treatment methods they use.

It is possible that some Info-Tool participants share especially optimistic information about chlorine tablets, and some share especially pessimistic information. Then, these two forces might counteract each other so that participants believe their Info-Tool and other-treatment-group friends use chlorine at the same rates on average. If this is the case, then participants should have more *accurate* information about Info-Tool participants' chlorine use, even if not more optimistic. Using our participants' self-reported water purification methods at endline, we can determine if participants accurately guess the water purification method of their within-sample network connections. We find no evidence that participants, regardless of their treatment group, have a better understanding of their Info-Tool friends' water purification methods than the purification methods that any of their other treatment-group friends use (Table J.12).[49] Furthermore, Info-Tool

mediately 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, so we take these null results as consistent with our results but not conclusive.

[49]The same holds true when we use other measures of "true treatment method" using our objective measures of water chlorine detection.

Table J.10: Midline Memory/Beliefs about Chlorine Efficacy: By Education

	(1) Correct Answer	(2) Went Down	(3) Didn't Change	(4) Went Up	(5) Don't Know
Any Education	-0.127**	0.017	0.008	-0.049	0.032**
	(0.053)	(0.048)	(0.029)	(0.042)	(0.014)
Any Treatment Group	-0.050	0.069**	-0.023	-0.033	-0.011
	(0.036)	(0.033)	(0.020)	(0.029)	(0.009)
Any Treatment Group × Any Education	0.193***	0.020	-0.036	0.051	-0.040**
	(0.061)	(0.055)	(0.033)	(0.048)	(0.016)
Observations	1427	1427	1427	1427	1427
Control Mean (No Education)	0.551	0.650	0.115	0.214	0.019
P-values:					
Any Education + Any Treatment Group × Any Education = 0	0.035	0.197	0.103	0.955	0.371

Standard errors in parentheses
Each observation is at the household level. This regression includes neighborhood block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, and lasso-selected baseline controls.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table J.11: Endline Beliefs About Friends' Chlorine Use

	(1) Believes Uses Chlorine: Full Sample		(2) Believes Uses Chlorine: Info-Tool Only	
Control Friend	-0.010	(0.038)	0.065	(0.067)
Chlorine Only Friend	0.086**	(0.041)	0.137	(0.087)
Incentives Friend	0.013	(0.032)	0.016	(0.068)
Info-Tool Friend	0.035	(0.046)	0.093	(0.089)
Observations	2614		617	
Mean Belief about Non-Participant Friends	0.085		0.052	
P-values:				
Info-Tool Friend = Chlorine Friend	0.281		0.713	

Standard errors in parentheses
Standard errors are clustered at the household level. Each observation is at the network-link level, for observations where the node is within our sample. The omitted group is links between respondents and their friends who are not participants in the study sample. The outcome in column (1) is an indicator for whether the respondent believes that the friend uses chlorine. Column (1) includes the full sample of reported social-network links. Column (2) uses the same outcome but restricts the sample to respondents in the Info-Tool group. All specifications include neighborhood-block fixed effects, unbalanced baseline controls, the total number of study participants within twenty meters, lasso-selected baseline controls, and respondent fixed effects.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

participants do not do a better job of guessing than any other group.

Finally, we also ask participants how long ago they first and last discussed water purification with each social network connection. Info-Tool participants do not begin engaging in conversations about water purification earlier than any of the other groups (Chlorine begins engaging in these conversations the earliest), nor have they had a conversation more recently (Incentives had the most recent water purification discussions) (Table J.13). This evidence supports our model, where treated participants are all equally aware of the benefits of chlorine tablets for water purification, and share this information non-differentially. Experience with the Info-Tool makes the belief in chlorine's efficacy actionable.

How come the Info-Tool intervention, which assists participants in generating conclusions based on data-driven evidence, *does not* act as an educational tool and increase data

Table J.12: Endline Accurate Beliefs About Friends' Water Purification Methods

	(1) Correct Guess (User or Not: SR)		(2) Correct Guess (User or Not: OB)		(3) Correct Guess (Is a Super User)	
Info-Tool Participant	-0.046***	(0.016)	0.011	(0.023)	-0.008	(0.008)
Info-Tool Friend	-0.024	(0.020)	-0.056	(0.036)	-0.007	(0.012)
Info-Tool Participant × Info-Tool Friend	-0.005	(0.038)	0.044	(0.057)	0.015	(0.020)
Observations	2040		2824		2824	
No Info-Tool × No Info-Tool Friend Mean	0.953		0.440		0.035	
P-values:						
Info-Tool Friend + Info-Tool × Info-Tool Friend = 0	0.378		0.805		0.672	

Standard errors in parentheses
Standard errors are clustered at the household level. Each observation is at the network-link level, for observations where the node is within our sample. This regression include neighborhood block fixed effects, unbalanced baseline controls, the number of study participants within twenty meters, lasso-selected baseline controls, participant treatment group fixed effects, and network-link treatment group fixed effects. The outcome in column (1) is an indicator for if the participant's guess about if her friend uses chlorine aligns with what that friend reported using in the endline survey. The outcome in column (2) is an indicator for if the participant's guess about if her friend uses chlorine aligns with what we objectively observed (did we ever detect chlorine in the friend's water). The outcome in column (3) is an indicator for if the participant guessed that her friend uses chlorine tablets for water purification, and we detected chlorine in that participant's water at least three times.
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table J.13: Months Since Water Purification Discussions

	(1) First Water Discussion		(2) Last Water Discussion	
Chlorine	0.326	(0.200)	-1.236*	(0.703)
Incentives	0.099	(0.157)	-2.404***	(0.715)
Info-Tool	-0.003	(0.165)	-0.987	(0.704)
Observations	1355		1355	
Control Mean	1.005		1.005	
P-values:				
Incentives = Chlorine Only	0.202		0.114	
Info-Tool = Chlorine Only	0.072		0.732	
Info-Tool = Incentives	0.471		0.055	

Standard errors in parentheses
Each observation is at the household level. This regression include neighborhood block fixed effects, unbalanced baseline controls, the number of study participants within twenty meters, and lasso-selected baseline controls. For links where participants responded that they had never discussed water, we impute with the minimum (0), and with the maximum (24 months).
* $p < .1$, ** $p < 0.05$, *** $p < 0.01$

memory or comprehension? Different from other successful interventions in the learning-to-learn literature, our intervention is significantly lighter-touch and is conducted with a sample who are majority-uneducated. Ashraf et al. (2021), for example, successfully improves teacher effectiveness in the classroom by implementing a pedagogy that encourages students to approach learning like scientists, taking into account "evidence and data gathered from everyday life".