TRANSLATING INFORMATION INTO ACTION

Reshmaan Hussam[*†]

Kailash Pandey[‡]

Abu Shonchoy[§]

Chikako Yamauchi[¶]

February 22, 2023

Revision Requested, *American Economic Review*

Abstract

While models of technology adoption posit learning as the basis of behavior change, information campaigns in public health frequently fail to change behavior. We design an information campaign embedding hand-hygiene edutainment within popular dramas using mobile phones, randomly distributed to households in Bangladesh. We document substantial improvements in handwashing and health, but no change in hygiene knowledge. Employing machine learning techniques with temporal data on media exposure and handwashing, we find that both cumulative and immediate exposure is correlated with washing, consistent with cue-based habituation. Results highlight how behavior change may be induced by tacit, rather than explicit, knowledge acquisition.

[*]We are indebted to research assistant Mohammed Helal Uddin and Nayamat Tasnim Ullah, as well as the survey firm MOMODa in Gaibandha, Bangladesh. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 25101002. This study is registered in the AEA RCT Registry and the digital object identifier (DOI) is: "10.1257/rct.8074-1.0." IRB approval for endline data collection at IRB18-1380. All errors are our own.

[†]Harvard Business School; corresponding author: rhussam@hbs.edu. (617) 495-6378. Address: Morgan Hall, 15 Harvard Way, Boston, MA 02163.

[‡]Harvard Business School; kpandey@hbs.edu

[§]Department of Economics, Florida International University, shonchoy@fiu.edu

[¶]GRIPS, c-yamauchi@grips.edu

1 Introduction

Models of technology adoption frame behavior change as precipitated by information acquisition: agents possess priors over the returns to a behavior, receive new information via information provision or experimentation, update their priors, and engage if returns outweigh costs (Arrow (1962); Janvry, Macours, and Sadoulet (2017); Foster and Rosenzweig (1995); Conley and Udry (2010)). A shift in knowledge is thus a prerequisite to behavior change. It is this theory of change that motivates the profusion of information campaigns in public health.

Against this backdrop, we administer an information campaign that fails to alter explicit knowledge, yet meaningfully improves behavior and health. How does such change transpire? Our high-frequency temporal data on informational inputs and behavioral outcomes points to the role of salience and tacit knowledge, or know-how via accumulated exposure and association that cannot be measured directly (Hadjimichael and Tsoukas, 2019), in transforming informational content into action.

We explore this process in the context of handwashing with soap in Bangladesh. To focus on the translation of information into action, we choose a setting in which neither the raw materials required for the act nor the social norms surrounding it are a constraint: 100% of our households own soap and 99% rinse their hands with water before eating. Families possess some, but not complete, knowledge of proper hygiene behavior: 83% of mothers believe soap removes germs from hands, but 64% do not think that soap will make hands clean if they already appear clean. 54% of mothers volunteer handwashing as a method of preventing diarrhea, but only 2% believe handwashing can prevent colds. We therefore operate in a space where an information campaign may alter behavior either by shifting explicit priors about returns, or through the 'translation function' of priors into action.

Such behavior change is critical to health across the developing world. Diarrheal and respiratory illnesses from unsafe sanitation and hygiene practices kill nearly one million people each year and stunt the growth of millions more (WWAP, 2019). Relative to expensive infrastructural investments, improvements in individual hygiene involve small changes with potentially substantial returns: handwashing with soap, for example, can drastically reduce illness by interrupting the transmission of pathogens into the body (Freeman et al., 2014; McGuinness et al., 2018). While successful in intensive and costly programs, practitioners face the challenge of identifying low cost, scalable interventions that yield sustained improvements in behavior and health. This study proposes one such program.

We design an edutainment campaign using existing public service announcements on hand hygiene. To focus on information transmission, the content is composed of simple but engaging depictions of why, when, and how to properly wash, with no celebrity appearances nor gamification. Our medium is the mobile phone, whose penetration in rural Bangladesh has grown rapidly (GSMA, 2014). Network reliability and internet accessibility remain poor, so households forego streaming and typically rely on SD cards preloaded with content (see Figure A1) for their entertainment needs.[1] Our intervention embeds edutainment within popular dramas and movies and distributes call-disabled smartphones with preloaded SD cards to randomly selected households in rural Bangladesh. We provide an equal platform for all treated households to view content by issuing them a simple smartphone (valued at 50 USD). We disable network capabilities to focus the intervention on the preloaded edutainment.

To measure handwashing behavior, we distribute handsoap dispensers embedded with time-stamped sensors to all households. This technological innovation addresses the serious challenges posed by standard participant observation measures of hand hygiene, making data collection unobtrusive, objective, and precise (Hussam et al., 2021; Ram et al., 2010; Biran et al., 2008) (see Figures A2 and A3 for sensor diagram and installation).

Our analysis proceeds in three steps. We first examine the relationship between edutainment consumption and behavior change. We find that daily handwashing rates, as measured by the sensor technology, increase significantly due to the intervention: treated households wash 22% ($p = 0.000$) more than their control counterparts.

Contrary to the purported intent of the campaign however, this effect is not driven by a shift in knowledge. We collect detailed information on household knowledge of hand hygiene and germ theory via open-ended questions with answers coded in order of the relevance given to them by respondents, allowing us to construct both an absolute knowledge index (a measure of whether households acquire any knowledge about hand hygiene or the returns to handwashing) and a relative knowledge index (a measure of how much importance households place on hand hygiene). We find that treated households exhibit no knowledge improvements along either margin: treatment effects are statistically indistinguishable from zero, and we can reject knowledge gains greater than 4% with 95% confidence for both indices.

If not a change in beliefs, what, then, is the nature of information exposure that leads to behavior change? To probe this, we exploit minute-level time series data from an application within the mobile phones that tracks media consumption. Paired with our minute-level data on handsoap dispenser use, we assemble a unique panel dataset of behavioral inputs (media exposure) and outputs (handwashing behavior) and employ machine learning techniques to uncover which patterns of entertainment exposure best predict future handwashing behavior. The machine learning approach is well suited to this objective, as *ex ante* existing literature

[1] Also seen in India (Tenhunen, 2018).

offers no priors around what temporal patterns of exposure may generate shifts in behavior. We find that cumulative exposure to edutainment during the month prior to a washing episode is most predictive of soap use, followed by immediate exposure to any entertainment 30 minutes prior to a washing episode. All other temporal features, including exposure to content beyond one month or 31 to 60 minutes prior, have near zero predictive value. This is consistent with handwashing becoming a cue-based habit, in which cumulative exposure to informative content familiarizes one with the behavior and associates it with phone exposure, and immediate exposure then serves as the necessary trigger to wash.

Because only treated households can produce data on watching habits, the patterns we uncover are correlational. However, the edutainment is randomly placed within the entertainment, limiting the role of endogenous viewer choice in information exposure. Our ML models also include day fixed effects to address concerns of parallel time trends or reverse causality from washing to watching, as well as household fixed effects to eliminate selection on household-type, or different household watching and washing preferences. We uncover a sensible set of patterns around immediate and cumulative exposure that can inform the parameters of future mechanism experiments and shape the implementation of public health information campaigns.

Having documented impacts of the intervention on hand-hygiene behavior, we next estimate the effect of the edutainment package on the health of treated children. We find substantial reductions among children in treated households relative to their control counterparts on the incidence of loose stool ($-54\%, p = 0.011$) and symptoms of acute respiratory infection (ARI) ($-29\%, p = 0.005$). We find no effect of the campaign on other sanitation or hygiene behaviors, suggesting that the estimated health impacts are indeed driven by changes in hand hygiene. These health improvements persist over the course of the twelve months of data collection despite an edutainment intervention which lasts only eight months.

To benchmark our health results against the status quo (households with soap but no edutainment campaigns *and* no dispensers), we supplement our randomized sample with a sample of households drawn from our initial census. While this "pure control" group is not randomized, their only observable distinction from our experimental sample is that mobile phones were already employed as a primary source of entertainment in these homes, making them ineligible for our experimental sample. We collect child health data from this sample for the final six months of the experiment. We find that, relative to this pure control sample, children who received the dispenser (but no edutainment treatment) exhibit 68% lower incidence of loose stool ($p = 0.015$) and 52% lower incidence of symptoms of ARI ($p = 0.029$). This suggests that the dispenser alone had substantial effects on hand hygiene behavior and subsequent health, echoing a finding of Hussam et al. (2021) in West Bengal. Reframing the results of the edutainment campaign in light of the effect of the dispenser alone, we interpret the dispenser as a potentially critical complement to edutainment: households require not only the material resources of soap and water (which all households in our setting possess), but also a convenient and user-friendly medium of use, in order to act upon the content delivered in a public health campaign.

We view the contributions of this study as threefold. First, our results on exposure recency build upon literature around the value of reminders, often delivered via text messages, for building healthy behaviors (e.g. Patrick et al. (2009); Koshy, Car, and Majeed (2008); Karlan et al. (2016).[2][3] Beyond the binary presence or absence of a reminder, our mobile phone and dispenser technology allow us to examine a broad set of potential temporal relationships between inputs (edutainment exposure) and outputs (handwashing behavior). We identify the features of edutainment exposure that are most predictive of handwashing, clarifying the windows of time during which exposure may be most impactful. This high-frequency time-series data on both information stimuli and subsequent behavior has not been collected or utilized, to our knowledge, in existing studies of behavior change or technology adoption, and offers a path forward to empirically constructing the translation function of information exposure into action even in the absence of explicit changes in knowledge.

Second, we document that prolonged information campaigns can generate meaningful changes in behavior despite no measurable change in knowledge. This suggests that designers of information campaigns may be well served to consider the tacit means by which their campaigns can engender behavior change, rather than focusing on the dispensation of facts *per se*. While consumers may be unable to consciously recognize factual information about the returns to a technology or behavior, the act of repeatedly conveying informative content to a captive audience can serve to familiarize, and associate a cue, with the promoted behavior (cumulative exposure), then make the behavior salient by triggering the cue (immediate exposure), resulting in increased adoption with no commensurate change in explicit knowledge. We denote this combination of effects under the umbrella of 'tacit knowledge,' distinct from explicit knowledge in a manner similar to Romer (2000)'s theoretical distinction between 'feeling' and 'thinking.' Our findings are consistent with economic models of cue-based habit formation (Laibson, 2001), psychological literature around the predominance of System 1 over System 2 thinking (Kahneman, 2012) in routine behaviors, and recent work

[2]Related is recent work by Bettinger et al. (2021), which finds that 'content-less' text message reminders are as effective as texts bearing informative content about a child's inputs to education. The primary channel in this context, however, remains information acquisition: content-less messages encourage parents to secure the relevant information to encourage their childrens' educational performance.

[3]While the underlying mechanism of increasing salience may be the same, text message interventions require the decision of timing to be made by the experimenter, precluding an exploration of which temporal patterns of exposure are most predictive of behavior change.

in neuroscience that identifies the brain's 'default mode network' to serve the role of autopilot: through repeated exposure and contextual triggers, we engage in behavior with no conscious awareness of why or that we are doing so (Raichle, 2015; Vatansever, Menon, and Stamatakis, 2017). Our results, however, also offer insight into why impacts of information campaigns may be short-lived: in the absence of the intervention, both associations and cues to trigger behavior disappear. As underlying priors about the returns have not shifted, neither will subsequent behavior.

Third, we devise a simple and scalable intervention that manages to not only shift hygiene behavior but also generate meaningful and sustained improvements in child health. The greater part of information campaigns in the hygiene and sanitation space have been unable to produce health improvements (see, for example, Biran et al. (2009); Chase and Do (2012); Galiani et al. (2016); Null et al. (2018); Lewis et al. (2018); Bennett, Naqvi, and Schmidt (2018), with its innovative use of microscopes, is a notable exception in health outcomes and affordability), and the few that record changes in behavior without subsequent health effects employ self-reports or observational data (such as Tidwell et al. (2019), which also documents improvements in handwashing from a media campaign) with their concomitant challenges (Ram et al., 2010; Biran et al., 2008). Health effects of the magnitudes we document are uncommon in the literature: a lower bound of \$6.50 USD per household for the cost of the SD card (\$2 USD), dispenser (\$3.50 USD), and ten months of soap (\$1 USD), and an upper bound of \$65 USD for the cost of the SD card, dispenser, soap, card delivery, and phone - both estimates that are likely to drop as phone and internet penetration grow, and dispensers are produced domestically at scale, across the developing world.

Beyond hand hygiene, a behavior of increasing importance in the wake of the global COVID-19 pandemic, this work may speak to the design and dissemination of public health information campaigns for other low cost, high return, and repetitive behaviors, with particular relevance to behaviors such as water treatment and mask-wearing.

The remainder of the paper proceeds as follows. Section 2 describes the design and implementation of the experiment, Section 3 presents the analysis and results, and Section 4 concludes.

2 Experimental Design

Our study was conducted in Gaibandha District, Bangladesh, among 333 households across 34 villages. All households had at least one child of primary school age, access to a latrine, and a female head of household. All households received a handsoap dispenser with a sensor embedded inside. Randomization was executed via computer, with 50% of households allocated to treatment.

Once per month, all households were visited, and their dispensers were refilled. Given our limited supply of sensors, a randomly selected third of dispensers included sensors in any given month; in the subsequent month, these sensors were extracted, data downloaded, and sensors then inserted into the next third of households, and so on over the course of eight months. As such, we have approximately two months of sensor data per household, but representative data of a balanced sample of control and treatment households.

The intervention lasted from April 2017 to November 2017. During this period, enumerators collected sensor data as well as data on child health and [for treatment households] self-reported entertainment exposure during their monthly visits. Using an application preloaded onto the smartphones, enumerators also extracted data for treated households on mobile phone watching patterns. An endline survey was conducted in April 2018. Follow-up rates vary by data type: the endline survey was completed for 86% of the sample, interim health surveys for 97% of the sample, the sensor data for 85% of the sample, and the mobile phone data for 54% of the [treated] sample. Lower followup rates for sensor and mobile phone data come not from household attrition but rather technical failures. Enumerators faced difficulties transferring data to laptops in the field, and many files were corrupted during extraction.

Table A1 demonstrates balance across treatment and control households at baseline along a host of sociodemographic, hygiene behavior, and hygiene knowledge characteristics. Table A2 presents balance along these features for the subsample of households in each data source. We see no evidence of differential attrition: baseline characteristics are balanced between treated and control households for whom we were able to secure followup data in each source, and the subsample of households for whom we have mobile phone data are comparable to the full sample of households.

2.1 Edutainment campaign

The edutainment campaign was delivered via a smartphone for which the phone technology had been disabled, leaving only a screen. The device was provided to treatment households after the baseline with an SD card preloaded with three hours of dramas and cartoons.

Between each preloaded [non-informative] drama or cartoon, we embedded an ad campaign around proper hand hygiene. These ads ranged from thirty seconds to seven minutes and were drawn from a set of publicly available material (for example, see links to the following: Meena Cartoon, Bangladesh campaign, and Sesame Street). Enumerators delivered SD cards with new dramas and cartoons to all treatment households monthly.

3 Analysis and Results

We present the results in three stages. First, we examine whether the intervention reached its intended audience. We then estimate its impact on hand hygiene behavior and explore the underlying mechanisms. Third, we consider whether these behavioral changes were consequential in terms of child health.

3.1 Impact of edutainment campaign on media consumption

To document that our intervention reached its intended audience, we run the following regression:

$$Media_h = \alpha + \beta Edutainment_h + X_h + \epsilon_h \tag{1}$$

Where $Media_h$ represents a series of outcomes around media engagement for household h drawn from the endline data, namely: whether the phone is used for entertainment by the mother, how many minutes per day is spent watching entertainment on the phone, whether the child edutainment content ('cartoons') are watched on the phone, whether the child uses the phone for entertainment, whether the child watches daily, how many minutes per day the child watches, and whether the child watches the edutainment cartoons on the phone. X_h are baseline covariates (including baseline value of the outcome and sociodemographic controls) selected via double-selection LASSO (Belloni, Chernozhukov, and Hansen, 2014).

Results are presented in Table 1. Treated mothers report using a mobile device for entertainment 73 pp (340%, Column 1) more than control households. Treated children are 38 pp (132%, Column 4) more likely to employ their phone as a source of entertainment than control children, and three times more likely to watch the device daily (Column 6). All other measures exhibit similar magnitude effects, are significant at the one percent level, and are robust to the inclusion of a rich set of sociodemographic controls.

3.2 Impact of edutainment campaign on handwashing behavior

We now examine whether edutainment exposure resulted in behavior change, as documented from the dispenser sensor data. To do so, we run the following regression:

$$Handwashing_{ht} = \alpha + \beta Edutainment_{ht} + \gamma_t + \delta_v + X_h + \epsilon_{ht} \tag{2}$$

Where $Handwashing_{ht}$ represents daily dispenser use as measured either in binary form (one if the dispenser was pressed at all in the day and zero otherwise) or continuously (the total number of presses that day), γ_t is day level fixed effects, δ_v is village level fixed effects, and X_h are baseline sociodemographic covariates selected via double-selection LASSO (Belloni, Chernozhukov, and Hansen, 2014). Standard errors are clustered at the household level.

Results are presented in Table 2. While treated and control households are equally likely to use the dispenser at all on a given day, treated households use the soap dispenser 22% ($p = 0.000$) *more* per day than their control counterparts. Figure 1 depicts the evolution of handwashing behavior over the course of the eight months during which we collected sensor data. Both treated and control households exhibit enthusiasm with the dispenser in the initial weeks, with treated households particularly engaged, and engagement declining over time.

Is this change in hygiene behavior consequential? Section 3.5 considers whether this increase in hand soap use yields health improvements. The short answer is yes: this impact, though temporary, has significant and lasting consequences for child health. We therefore turn to the mechanisms of information internalization: what dimensions of exposure to the edutainment campaign generate the behavioral - and consequently health - improvements we observe?

3.3 Mechanisms

3.3.1 Impact of edutainment campaign on knowledge

We first estimate whether the edutainment program shifted the knowledge of treated households, ostensibly the central and intended channel through which the intervention should alter behavior. At baseline and endline, we ask respondents a series of questions regarding their knowledge of hand hygiene, described in detail in Appendix A.2: if and why soap is useful, if and how it differs from washing only with water, and what actions can prevent colds and diarrhea. We designed this knowledge module with two features in mind. First, respondents are asked open-ended questions, rather than being equipped with answer choices, in an effort to eliminate anchoring or leading effects and elicit only the knowledge content that the respondent believes to be relevant. Second, we allow respondents to rank up to four answers per question. We do this in order to gauge not only whether the respondent has possession of the edutainment information, but additionally how pertinent or important she believes it to be.

For example, consider the following question: *"What are some ways in which you can keep from getting a cough or cold?"* A respondent may answer by first reflecting that dressing warmly is important (as 63% of edutainment treated respondents give as their first answer), then suggesting that one regularly change their clothes (as 33% say as their second answer), then mentioning that washing one's hands can help too (as 10% offer as their third). We score each question not by what is technically correct (as indeed, staying warm may reduce vulnerability to a cold), but by whether or not the information that is imparted in the edutainment programs appears in the respondent's answers.

We employ two methods for scoring each question: First, an absolute knowledge metric, in which we consider whether the information exists at all in the respondent's answers. In the case of the question above, the respondent would receive a score of 1, given that she mentioned that washing hands can prevent colds in one of her answer slots. Second, a relative knowledge metric, in which we estimate how much weight the respondent places on this answer, with the first slot receiving a weight of 1, the second a weight of 0.75, and so on, resulting in a score of 0.5 for the respondent in the question above. This method allows us to gauge relative magnitudes of, in this case, the expected returns to handwashing: while most individuals may know that washing hands can reduce vulnerability to colds and coughs, the treated respondent should have learned from the edutainment content that this is among the most effective prophylactics available.

We aggregate these scores into a Knowledge Index, an inverse covariance weighted index of the eight knowledge questions described in Appendix A.2 (Anderson, 2008). We regress this index on treatment status, selecting baseline covariates (including baseline value of the outcome and sociodemographic controls) via double-selection LASSO (Belloni, Chernozhukov, and Hansen, 2014). Results are presented in Table 3. Panel A estimates the impact of the edutainment treatment on absolute knowledge (Knowledge Index I) and its subcomponents. Treated respondents exhibit a 1.5% higher knowledge index than their control counterparts, statistically indistinguishable from zero. This null effect is exhibited across all of the subcomponents of the index and is relatively precise: we can reject any gains in the absolute knowledge index greater than 4% with 95% confidence.

Perhaps most respondents possess the knowledge in question nominally, but edutainment treated respondents better understand the importance of germ theory and handwashing practices as imparted by the treatment. Panel B thus examines relative knowledge, or the relevance individuals give to the key content of the edutainment program within their answers. We observe here a similarly precise zero impact: treated respondents exhibit a 0.4% higher score than their control counterparts, statistically indistinguishable from zero, and we can reject any effect of edutainment on the relative knowledge index greater than 3.5% with 95% confidence.

Our results on knowledge therefore suggest that the edutainment treatment failed to alter treated households' explicit beliefs about the returns to handwashing. We offer a final piece of evidence on this margin. As part of their consent process at the onset of the experiment, all respondents (both control and treatment) were verbally informed of the following: *"If you use the handsoap dispenser regularly to wash your hands after you defecate, before you prepare food, and before you eat, then your own health and especially your childs health should improve drastically. Your child will experience less diarrhea, fewer colds, and grow stronger, healthier, and taller by avoiding these sicknesses."* This information was delivered just prior to the administration of the baseline survey. Moments later in the hygiene knowledge module of the survey, however, only 2.4% of individuals articulate that handwashing with soap is an effective way of preventing colds. Indeed, this may have been because few respondents paid sufficient attention to process the information provided to them. But this is precisely why information campaigns may not be effective at altering priors, as our evidence indicates: new information requires attention and energy, or System 2 thinking (Kahneman, 2012), to internalize and transform into beliefs.

What, then, generated a change in behavior? The content delivered through the information campaigns may have had tacit, rather than explicitly measurable, impacts on knowledge, affecting behavior by developing associations through repetitive visual stimuli to engage in improved hand hygiene. We turn to this possibility next.

3.3.2 Temporal dynamics of information acquisition and behavior change

To probe how the dynamics of exposure may have altered behavior, we examine our time-series data on media consumption and handwashing and consider how the temporality of information exposure translates to behavior change. Our high-frequency data on the input of edutainment exposure and the output of handwashing, paired with machine learning techniques, presents a unique opportunity to shed light on the temporal nature of information-driven behavior change.

Our dataset for this exercise is composed of the handwashing outcome, which takes the form of a binary variable in which a one indicates that the dispenser was pressed at least once during the breakfast (dinner)-time range on a given day and zero otherwise, and a series of temporal feature variables around exposure to the hygiene information campaign: namely, binary and continuous measures of exposure to edutainment and entertainment in the previous thirty minutes, hour, week, twelve weeks, and interim periods. All observations are collapsed to the household-day-mealtime level to ensure a balanced panel across washing and non-washing household-time cells. Details on all temporal features and data construction decisions are described in Appendix A.3 (Chawla et al. (2002), Norberg (2016), Bergstra and Bengio (2012)).

Given that we observe media exposure data for treated households alone, and watching media is a potentially endogenous choice, this exercise cannot establish a definitive causal link. One may have opted for a randomized experiment to determine causation, randomizing the timing or frequency of messaging across households. Bettinger et al. (2021), for example, randomize the frequency of text messages to parents on children's educational performance in Brazil, and finds that multiple weekly nudges are significantly more effective at raising child attendance than a single nudge. While valuable in itself, we view this as distinct from the objective of our exercise: we aim here to put shape, with minimal constraints, to the process by which exposure to entertainment may translate into behavior change. *Ex ante*, existing literature provides no sense of what patterns or frequencies of messaging may be most impactful in altering hygiene behavior, so a determination of treatments along these two margins would be both arbitrary and limiting (barring an enormous sample size, infeasible in the context of dispenser, soap, and SD card provision in rural geographies). Does exposure over the course of weeks matter? Does immediate exposure matter while earlier exposure is forgotten? How should we define "early" or "immediate"?

To explore this broad space of possible hypotheses, we opt for a machine-learning analysis of a non-experimental data generation process with features embedded to limit the role of endogeneity and address concerns of selection. Edutainment videos are interspersed at random within the entertainment, reducing 'choice' in edutainment conditional on entertainment exposure. We include household fixed effects in all models, eliminating the selection channel of certain types of households choosing to both watch and wash. Neither can this be a story of reverse causality of washing leading to watching, as we include day fixed effects and consider minute- and hour-level *lags* in watching behavior for every mealtime-washing episode.

We train and test our data using the lasso, elastic net, and random forest algorithms. A comparison of the predictive performance of these algorithms is presented in Table 4. The elastic net exhibits the highest testing accuracy at 62%. We then employ this algorithm to rate features by 'importance,' a means of classifying the contribution of each variable to the model which, in the case of the elastic net, is the absolute value of the coefficient for each variable in the tuned model (Kuhn, 2020). Figure 2 presents the top four features selected by the elastic net algorithm. The algorithm identifies the total number of minutes of exposure to the edutainment campaign over the previous three and four weeks as most important (with importance scores of 0.12 and 0.09 respectively), followed by whether the entertainment portions were watched in the past half hour (importance of 0.06). All other features, including binary or cumulative exposure of edutainment or entertainment at, for example, five or more weeks, two weeks, and two hours, exhibit importance scores of 0.001

or below. These three exposure features of three weeks, four weeks, and thirty minutes thus appear distinctively predictive of handwashing among the 46 temporal features considered.

The selected temporal features imply that a combination of cumulative and immediate exposure is predictive of whether a household will wash in a given mealtime. Interestingly, while the cumulative features [of three to four weeks prior] rely on exposure to the edutainment campaign, the immediate feature [of half an hour prior] relies on exposure to the *non-informative* entertainment components. This pattern is consistent with a story of cue-based habit formation: agents exposed to the edutainment content develop an association between the media content and the act of handwashing, such that watching the entertainment alone is eventually a sufficient trigger to catalyze a handwashing episode.

Importantly, our exercise also implies that the influence of the content, or edutainment 'memory,' does not exceed one month: despite including all temporal watching patterns up to three months prior to each observation in the selection process, no features beyond one month have predictive power on the likelihood of handwashing in a given mealtime-day.

3.4 Potential confounds

We find significant impacts of the edutainment campaign on handwashing behavior and child health, with no commensurate change in knowledge. We consider here several alternative channels beyond the edutainment itself through which this may have transpired.

1. Time away from peers: Perhaps the time children spend watching media substitutes away from time spent playing with children, an activity located further away from handsoap dispensers (potentially less washing) and more prone to germ transmission (poorer health). While plausible, treated children watch 37 minutes of phone media daily, relative to 20 minutes among control children. While nearly a doubling of the control mean, the magnitude of the difference is small relative to the total time children are likely to be exposed to their peers each day: 87% of our sample attends pre-school or school for at least four hours per day, after which they return home and are likely to play outside until sundown. The remainder are likely to spend their entire day playing outside in the dirt and local ponds with neighborhood children, as is typical in this environment. While we cannot rule this channel out definitively, we suspect that a 17 minute reduction in exposure to peers during a full day of engagement is unlikely to drive the large health effects we document (as discussed in detail in Section 3.5).

2. TV as an incentive to wash: Perhaps parents use the phone entertainment as a means of incentivizing, or bribing, their children to wash their hands: "You can only watch TV if you go wash your hands after."[4] In order to use the entertainment as an incentive, however, this channel requires that parents first recognize the value of proper hand hygiene. Our null effect on hygiene knowledge suggests that such conscious knowledge acquisition is unlikely. Alternatively, parents may already possess sufficient knowledge of the importance of hand hygiene and simply need a proper bribing instrument, which arrives with the experiment in the form of the phone entertainment. Two pieces of evidence suggest this is not the case. First, only immediate exposure to entertainment, and not cumulative exposure to *edutainment*, should then be predictive of handwashing in our machine learning exercise. Second, effect sizes should be smaller in households where the 'carrot' of phone entertainment already existed at baseline. Panel B of Table 2 and Panel C of Table 5 present the impacts of the intervention on behavior and health, respectively, for the subsample of households who report at baseline that their children already use mobile phones for entertainment (33% of the sample, balanced between treatment and control). We find that the magnitude of the treatment effect in both behavior and in health persists (with a 35% increase in daily handwashing, 60% decline in loose stool incidence, and 22% decline in ARI symptoms), although estimates lose some precision given the substantially reduced sample size. These effect sizes among those families who already possess a bribing instrument suggests that such a strategy is unlikely to be a primary mechanism in the effects we estimate.

Notably, because we examine households who already utilize a phone for entertainment purposes, this subsample exercise further underscores that the *edutainment* content of the intervention, rather than the phone or entertainment provision, is the plausible driver behind the intervention's impacts on handwashing and health.

3. Experimenter demand: Given the variety of data we collect, there may exist several potential spaces for experimenter demand effects to arise. We take each in turn.

> **Knowledge** Did edutainment treated respondents alter their responses to the knowledge questions in order to please the enumerators? This would require that they mention hand hygiene or germ theory, the key dimensions along which the experiment intervened, differentially more than control households. We document a precise null effect of the intervention on such responses in the knowledge module.
>
> **Behavior** Did edutainment treated respondents utilize the handsoap dispenser more

[4] Being granted television time as a reward for washing is precluded by our machine-learning exercise, in which we find that *lags* in watching are predictive of washing.

to reciprocate enumerators for their generosity? Recall that all participants, control and treatment alike, received handsoap dispensers and soap and were informed of their value. It is plausible that such a gift would generate reciprocal behavior in the form of using the gift, but this should be equally true across treated and control households. Increased handwashing as reciprocal behavior for the phone and media content (our treatment) is less plausible: this would require treated households to explicitly recognize that the purpose of the media intervention was to improve handwashing *and* wish to please enumerators by acting on this awareness, two channels that the null impact on hygiene knowledge suggests did not transpire. The patterns we uncover around recency from the machine-learning exercise, in which immediate exposure to the media is predictive of washing behavior, is likewise inconsistent with experimenter demand.

Health Did edutainment treated respondents wish to appear healthier to impress enumerators? While we cannot rule this out, the seasonality we document in both health levels and treatment effects suggests this is not the case (described further in Section 3.5): households report health statuses in a manner consistent with expected variation in the incidence of respiratory and water-borne illnesses over the course of the year. Alternatively, perhaps the edutainment intervention made health more salient to treated households. Salience, however, should lead to increased parental attention to child health, resulting in a heightened awareness and reporting of children's coughs, colds, and loose stool, the opposite of the impacts we document.

4. Features of edutainment content: Finally, we consider whether specific features of the media content may have precipitated behavior change.

Visual role models Perhaps viewers of the edutainment encountered role models to emulate (Chong and La Ferrara, 2009; Bernard et al., 2019). The edutainment content we provided included no celebrity actors nor high-status roles (with adult characters playing village housewives and children's characters in the form of cartoons and puppets in schools and villages), so it is unlikely that viewers internalized prestige-related returns to handwashing. However, it is certainly possible that viewers found these characters fun and appealing and thus wished to mimic their behavior.

Engaging content Relatedly, while simple, the content was likely to be engaging in other ways: the children's edutainment cartoons had songs and bright colors, and 56% of parents reported that these cartoons were their children's favorite piece of media content; similarly, 46% of adults reported one of the edutainment pieces to be their own favorite piece of media content provided.

We cannot rule out the possible role of these features in catalyzing behavior change, nor do we seek to. Interesting content with appealing characters and storylines is likely essential to the behavior change we document and underpins the central argument of this paper: the provision of information, and subsequent explicit knowledge acquisition, is not the driving mechanism for behavior change in our context of preventive health behavior adoption, despite being the central intent of the educational intervention. Rather, an association between watching the media and the hand hygiene behaviors enacted, likely strengthened through engaging content, compels behavior change. We leave a dissection of precisely what features of the content maximize engagement to future work.

3.5 Impact of edutainment campaign on child health

We document statistically significant changes in handwashing behavior and evidence that such changes are generated not by updating beliefs, but rather by altering implicit associations with repeated exposure. Is this change in hand hygiene meaningful enough to impact health? We run the following regression using the health data obtained from our monthly surveys:

$$Health_{ht} = \alpha + \beta * Edutainment_{ht} + \gamma_t + \delta_v + \epsilon_{ht} \tag{3}$$

Where $Health_{ht}$ represents child health as measured by (1) the presence of any symptoms of acute respiratory infection (ARI) such as coughs, colds, or runny noses in the previous two weeks and (2) the presence of loose stool, a proxy for diarrhea.[5] γ_t is survey round fixed effects, δ_v is village level fixed effects, and X_h are baseline covariates (including baseline value of the outcome and sociodemographic controls) selected via double-selection LASSO (Belloni, Chernozhukov, and Hansen, 2014). Standard errors are clustered at the household level. Our health sample is composed of children ages twelve years and below at baseline, though results are robust to expanding and reducing the age cutoff.

Results are presented in Panel A of Table 5. The edutainment campaign leads to a 54.4% reduction ($p = 0.036$) in incidence of loose stool and a 28.8% reduction ($p = 0.016$) in symptoms of ARI over the course of the campaign.

Notably, the average incidence of reported illness is low, at 1.5 percent of households reporting loose stool in any given two-week period and 6.6 percent of households reporting any symptoms of ARI. This masks heterogeneity over the course of the year: diarrhea and

[5]Diarrhea is defined as three or more loose motions in a day. Because mothers often do not observe every child-motion episode, we elicit any observations of loose stool. The presence of loose stool does not necessitate diarrhea, but it is the key symptom.

ARI are seasonal, with diarrhea most likely during summer monsoon months and ARI during the transition into winter months.

Panel A of Figure 3 plots the temporal evolution of illness. Consistent with the seasonality of water-borne diseases, rates of loose stool peak during monsoon season (June to October), during which the impact of the edutainment campaign is most apparent; rates fall rapidly thereafter, with nearly zero loose stool incidence reported for both groups in the winter. This seasonality in treatment effect is not apparent for ARI: both treated and control households exhibit a decline in symptoms over the first two months, a low incidence thereafter (with treated households, notably, hovering near zero during the winter months), and a stable gap between treatment and control.

Our results point to the direct impact of an edutainment program on child health. Can these health improvements be attributed to better hand hygiene alone, or might the intervention have precipitated other hygiene and sanitation improvements among exposed households? Table A4 estimates the impact of the campaign on water treatment practices, open defecation, and construction of a sanitary latrine by endline, and finds no effect on any other margin.

3.6 Impact of hand soap dispensers on child health

Panel A of Figure 3 exhibits a steep decline in ARI and diarrhea incidence in the early months of the intervention *regardless* of treatment assignment. Because all households received dispensers at the outset of the experiment, we suspected this decline was due to the dispenser alone. We subsequently added a group of "pure control households to the sample to measure illness incidence among households who received no dispenser. Having been added *ex-post*, these households were not randomly selected; they were rather the subset of households who had been excluded from the experiment because the female head of household owned a mobile phone that was already used as a primary source of entertainment. We returned to these households and collected an abridged baseline survey and monthly surveys of child health six months after the experiment launched.

Table A3 presents balance between these pure control households and the dispenser control households. There exist no significant differences between the two groups along any surveyed margins except phone use. While phone use may be correlated with unobservables such as wealth, these are likely associated with better health status among the pure control, making estimates of the health impacts of the dispensers lower bounds.

We run the identical regression to Equation 3, with our treated sample now defined as the dispenser control group (who received a dispenser but no edutainment), and the control defined as the pure control. Results are presented in Panel B of Table 5. We find that the impact of the handsoap dispenser alone is substantial. Reported incidence of loose stool is 67.6% lower and symptoms of acute respiratory infection 51.8% lower among households who received a dispenser. These effects appear on a larger base: 3.8% of pure control households report that their children experienced loose stool in each two week period, and 9.2% of households report symptoms of ARI.

Panel B of Figure 3 plots the evolution of illness incidence for the pure control, the control (with dispenser), and the treated (with dispenser and edutainment) groups during the last six months of the experiment, the period during which pure control health data was collected. The seasonality of ARI now emerges: incidence is highest among pure control households during winter and steadily declines thereafter. Loose stool rates are relatively stable over winter months and rise as summer approaches. The dispenser alone entirely eliminates loose stool incidence and drastically reduces ARI incidence over the time period observed.

4 Conclusion

Using sensors in handsoap dispensers, this study finds that a simple hand hygiene edutainment campaign, viewed amidst popular dramas via SD cards in mobile phones, yields significant improvements in both handwashing behavior and health with no commensurate improvement in knowledge. We offer two points of consideration here.

First, the substantial effect of the handsoap dispenser alone is striking. It begs the question: is the behavior change generated by the edutainment campaign meaningful? We posit that the impact of the edutainment campaign must be understood within the context of households who have access to appropriate infrastructure. The results suggest that human-centered design plays an important role in the provision of resources: the dispenser was enjoyable to use and situated near common sites of use. We stress the relevance of the product itself beyond simple resource provision: the edutainment may have had little effect on household behavior and health absent the dispensers, despite the ready availability of soap and water in all households. As such, the edutainment campaign may best be seen as a valuable complement to a necessary infrastructural investment. The impacts of this complementary campaign are alone substantial: the campaign is able to halve loose stool incidence during peak monsoon months and effectively eliminate symptoms of ARI during the peak early winter months relative to those households who received a dispenser but no edutainment campaign. As mobile smartphone penetration continues to grow, a program such as this becomes rapidly scalable at low marginal cost.

Our results also problematize a common framing of knowledge acquisition. In classical models of technology and behavior adoption, consumers are Bayesian updaters who learn about the returns to a behavior, update beliefs, and alter their behavior accordingly. Related policy recommendations of subsidizing experimentation or information provision assume a conscious acquisition of knowledge. However, knowledge change does not guarantee behavior change, a fact that comes to bear in study after study of information campaigns which document improvements in self-reported hygiene awareness with no corresponding change in behavior or health. The results of this study suggest that the reverse may also be true: behavior change does not require a change in explicit knowledge. The value of an edutainment campaign, when embedded into an everyday activity such as watching television, may be not to educate, but rather to familiarize one with, then serve as a visual reminder of, an activity. In other words, campaigns may be more impactful as tools to subconsciously habituate individuals - rather than consciously shift priors around returns - to an activity.

As such, behavior change programs must consider not merely the provision of information, but also the means by which such information is delivered, to be effective. And importantly, impact evaluations that estimate improvements in knowledge as well as behavior may be misattributing the latter to the former. As we find here, behavior may change regardless of the state of one's explicit beliefs, so evaluations that ignore the timing, frequency, and context by which information interventions are presented may be missing the central mechanism behind the behavior change they document.

References

Anderson, Michael L. 2008. "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. URL `https://doi.org/10.1198/016214508000000841`. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/016214508000000841.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning by Doing." *The Review of Economic Studies* 29 (3):155–173. URL `https://www.jstor.org/stable/2295952`. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies* 81 (2):608–650. URL `https://doi.org/10.1093/restud/rdt044`.

Bennett, Daniel, Asjad Naqvi, and Wolf-Peter Schmidt. 2018. "Learning, Hygiene and Traditional Medicine." *The Economic Journal* 128 (612):F545–F574. URL `https://www.onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12549`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12549.

Bergstra, James and Yoshua Bengio. 2012. "Random Search for Hyper-Parameter Optimization." *Journal of Machine Learning Research* 13:281–305.

Bernard, Tanguy, Stefan Dercon, Kate Orkin, and Alemayehu Seyoum Taffesse. 2019. "Parental Aspirations for Children's Education: Is There a "Girl Effect"? Experimental Evidence from Rural Ethiopia." *AEA Papers and Proceedings* 109:127–132. URL `https://www.aeaweb.org/articles?id=10.1257/pandp.20191072`.

Bettinger, Eric, Nina Cunha, Guilherme Lichand, and Ricardo Madeira. 2021. "Are the Effects of Informational Interventions Driven by Salience?" SSRN Scholarly Paper ID 3633821, Social Science Research Network, Rochester, NY. URL `https://papers.ssrn.com/abstract=3633821`.

Biran, A., T. Rabie, W. Schmidt, S. Juvekar, S. Hirve, and V. Curtis. 2008. "Comparing the performance of indicators of hand-washing practices in rural Indian households." *Tropical Medicine & International Health* 13 (2):278–285. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-3156.2007.02001.x`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-3156.2007.02001.x.

Biran, Adam, Wolf-Peter Schmidt, Richard Wright, Therese Jones, M. Seshadri, Pradeep Isaac, N. A. Nathan, Peter Hall, Joeleen McKenna, Stewart Granger, Pat Bidinger, and Val Curtis. 2009. "The effect of a soap promotion and hygiene education campaign on handwashing behaviour in rural India: A cluster randomised trial." *Tropical Medicine & International Health* 14 (10):1303–1314. Tex.copyright: 2009 Blackwell Publishing Ltd.

Chase, Claire and Quy-Toan Do. 2012. "Handwashing Behavior Change at Scale: Evidence from a Randomized Evaluation in Vietnam." Tech. rep., The World Bank.

Chawla, Nitesh, Kevin Bowyer, Lawrence Hall, and W. Philip Kegelmeyer. 2002. "SMOTE: Synthetic Minority Over-sampling Technique." *Journal of Artificial Intelligence Research* 16:321–357.

Chong, Alberto and Eliana La Ferrara. 2009. "Television and Divorce: Evidence from Brazilian "Novelas"." *Journal of the European Economic Association* 7 (2/3):458–468. Publisher: Oxford University Press.

Conley, Timothy G. and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *The American Economic Review* 100 (1):35–69. Publisher: American Economic Association Section: Articles tex.address: Nashville, United States, Nashville tex.copyright: Copyright American Economic Association Mar 2010.

Foster, Andrew D. and Mark R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *The Journal of Political Economy* 103 (6):1176–1209. Publisher: University of Chicago Press, The University of Chicago Press, University of Chicago, acting through its Press tex.address: Chicago.

Freeman, Matthew C., Meredith E. Stocks, Oliver Cumming, Aurelie Jeandron, Julian P. T. Higgins, Jennyfer Wolf, Annette Prss-Ustn, Sophie Bonjour, Paul R. Hunter, Lorna Fewtrell, and Valerie Curtis. 2014. "Systematic review: Hygiene and health: systematic review of handwashing practices worldwide and update of health effects." *Tropical Medicine & International Health* 19 (8):906–916. Tex.copyright: 2014 John Wiley & Sons Ltd.

Galiani, Sebastian, Paul Gertler, Nicolas Ajzenman, and Alexandra Orsola-Vidal. 2016. "Promoting Handwashing Behavior: The Effects of Large-Scale Community and School-Level Interventions." *Health Economics* 25 (12):1545–1559.

GSMA. 2014. "Country overview: Bangladesh." Tech. rep., GSMA Intelliegence.

Hadjimichael, Demetris and Haridimos Tsoukas. 2019. "Toward a Better Understanding of Tacit Knowledge in Organizations: Taking Stock and Moving Forward." *Academy of Management Annals* 13 (2):672–703. URL https://journals.aom.org/doi/10.5465/annals.2017.0084. Publisher: Academy of Management.

Hussam, Reshmaan, Giovanni Reggiani, Atonu Rabbani, and Natalia Rigol. 2021. "Rational Habit Formation: Experimental Evidence from Handwashing in India." *American Economic Journal: Applied Economics* Forthcoming.

Janvry, Alain De, Karen Macours, and Elisabeth Sadoulet. 2017. *Learning for Adopting: Technology Adoption in Developing Country Agriculture*. FERDI.

Kahneman, Daniel. 2012. "Of 2 Minds: How Fast and Slow Thinking Shape Perception and Choice [Excerpt]." *Scientific American* URL https://www.scientificamerican.com/article/kahneman-excerpt-thinking-fast-and-slow/.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62 (12):3393–3411. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2015.2296. Publisher: INFORMS.

Koshy, Elizabeth, Josip Car, and Azeem Majeed. 2008. "Effectiveness of mobile-phone short message service (SMS) reminders for ophthalmology outpatient appointments: observational study." *BMC ophthalmology* 8:9.

Kuhn, Max. 2020. "varImp function | R Documentation." URL https://www.rdocumentation.org/packages/caret/versions/6.0-86/topics/varImp.

Laibson, David. 2001. "A Cue-Theory of Consumption." *The Quarterly Journal of Economics* 116 (1):81–119. URL https://www.jstor.org/stable/2696444. Publisher: Oxford University Press.

Lewis, Henrietta E., Katie Greenland, Val Curtis, and Wolf-Peter Schmidt. 2018. "Effect of a School-Based Hygiene Behavior Change Campaign on Handwashing with Soap in Bihar, India: Cluster-Randomized Trial." *The American Journal of Tropical Medicine and Hygiene* 99 (4):924–933. Publisher: The American Society of Tropical Medicine and Hygiene.

McGuinness, Sarah L., S. Fiona Barker, Joanne O'Toole, Allen C. Cheng, Andrew B. Forbes, Martha Sinclair, and Karin Leder. 2018. "Effect of hygiene interventions on acute respiratory infections in childcare, school and domestic settings in low- and middle-income countries: A systematic review." *Tropical Medicine & International Health* 23 (8):816–833. Tex.copyright: 2018 John Wiley & Sons Ltd.

Norberg, Robert. 2016. "Data Splitting: Time Slices With an Irregular Time Series." URL `https://r-norberg.blogspot.com/2016/08/data-splitting-time-slices-with.html`.

Null, Clair, Christine P Stewart, Amy J Pickering, Holly N Dentz, Benjamin F Arnold, Charles D Arnold, Jade Benjamin-Chung, Thomas Clasen, Kathryn G Dewey, Lia C H Fernald, Alan E Hubbard, Patricia Kariger, Audrie Lin, Stephen P Luby, Andrew Mertens, Sammy M Njenga, Geoffrey Nyambane, Pavani K Ram, and John M Colford. 2018. "Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: A cluster-randomised controlled trial." *The Lancet Global Health* 6 (3):e316–e329.

Patrick, Kevin, Fred Raab, Marc A. Adams, Lindsay Dillon, Marian Zabinski, Cheryl L. Rock, William G. Griswold, and Gregory J. Norman. 2009. "A text message-based intervention for weight loss: randomized controlled trial." *Journal of Medical Internet Research* 11 (1):e1.

Raichle, Marcus. 2015. "The Brain's Default Mode Network." *Annual Review of Neuroscience* 38:433–447.

Ram, Pavani K., Amal K. Halder, Stewart P. Granger, Therese Jones, Peter Hall, David Hitchcock, Richard Wright, Benjamin Nygren, M. Sirajul Islam, John W. Molyneaux, and Stephen P. Luby. 2010. "Is Structured Observation a Valid Technique to Measure Handwashing Behavior? Use of Acceleration Sensors Embedded in Soap to Assess Reactivity to Structured Observation." *The American Journal of Tropical Medicine and Hygiene* 83 (5):1070–1076. URL `https://www.ajtmh.org/view/journals/tpmd/83/5/article-p1070.xml`. Publisher: The American Society of Tropical Medicine and Hygiene Section: The American Journal of Tropical Medicine and Hygiene.

Romer, Paul M. 2000. "Thinking and Feeling." *The American Economic Review* 90 (2):439–443. URL `https://www.jstor.org/stable/117265`. Publisher: American Economic Association.

Tenhunen, Sirpa. 2018. *A Village Goes Mobile: Telephony, Mediation, and Social Change in Rural India*. Studies in Mobile Communication. New York: Oxford University Press, Oxford University Press USA - OSO.

Tidwell, James B., Anila Gopalakrishnan, Stephen Lovelady, Esha Sheth, Arathi Unni, Richard Wright, Shonali Ghosh, and Myriam Sidibe. 2019. "Effect of Two Complementary Mass-Scale Media Interventions on Handwashing with Soap among Mothers." *Journal of Health Communication* 24 (2):203–215. URL https://doi.org/10.1080/10810730.2019.1593554. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10810730.2019.1593554.

Vatansever, Deniz, David K. Menon, and Emmanuel A. Stamatakis. 2017. "Default mode contributions to automated information processing." *Proceedings of the National Academy of Sciences* 114 (48):12821–12826. URL https://www.pnas.org/content/114/48/12821. Publisher: National Academy of Sciences Section: Biological Sciences.

WWAP, (UNESCO World Water Assessment Programme). 2019. "The United Nations World Water Development Report 2019: Leaving No One Behind." Tech. rep., Paris, UNESCO.

Tables

Table 1: Mobile phone and edutainment use

	(1) Entertainment	(2) Minutes	(3) Cartoons	(4) Child use	(5) Daily	(6) Child min	(7) Child cartoons
Edutainment Treatment	0.73	33.65	0.27	0.38	0.64	17.19	0.22
	(0.04)	(4.45)	(0.04)	(0.04)	(0.04)	(1.95)	(0.06)
Mean of control	0.12	4.59	0.04	0.59	0.27	20.01	0.52
Observations	287	287	287	287	287	287	287

Notes: Outcomes come from the endline survey with a two-week lookback period. Standard errors in parentheses. Baseline covariates (including baseline value of the outcome and sociodemographic controls) selected via double-selection LASSO (Belloni, Chernozhukov, and Hansen, 2014). Outcome measures are as follows: "Entertainment": Is the phone used for entertainment? "Minutes": How many minutes is the phone watched by an adult per day? "Cartoons": Does the adult watch any children's cartoons? (Note that the children's edutainment content is cartoons as well; this is a proxy for watching this content). "Child use": Does the child use the phone for entertainment? "Daily": Does the child use the phone daily? "Child min": How many minutes daily does the child watch the phone? "Child cartoons": Does the child watch any children's cartoons?

Table 2: Handsoap dispenser use

Panel A: Full sample

	(1) Used at all	(2) Total daily use
Edutainment Treatment	0.029 (0.024)	2.035 (0.725)
Mean of control	0.59	9.27
Observations	12846	12846

Panel B: Subsample of households with children who use phone for entertainment at baseline

	(1) Used at all	(2) Total daily use
Edutainment Treatment	0.011 (0.061)	3.086 (1.875)
Mean of control	0.59	8.76
Observations	3798	3798

Notes: Outcomes come from dispenser sensor data. Observations are at the household-day level. "Used at all" is a binary variable equal to one if the dispenser was active at all during the given day. "Total daily use" is the total number of dispenser presses (with presses occurring within two seconds of each other collapsed) within a given day. All regressions include village and day fixed effects. Controls are mother's age, age at marriage, literacy level, whether she completed primary education, whether she owns the home, the number of rooms in the home, whether the house has electricity, and respondent religion. Standard errors in parentheses are clustered at the household level. Sensors observed from April 19, 2017 to November 9, 2017.

Table 3: Hand hygiene knowledge

Panel A: Absolute Knowledge

		Individual Components of Knowledge Index I							
	(1) **Knowledge I**	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Edutainment Treatment	0.014 (0.012)	-0.000 (0.058)	0.006 (0.018)	-0.081 (0.052)	-0.035 (0.046)	0.034 (0.020)	0.053 (0.037)	0.012 (0.034)	0.019 (0.033)
Mean of control	0.896	0.589	0.973	0.315	0.822	0.952	0.863	0.904	0.904
Observations	287	287	287	287	287	287	287	287	287

Panel B: Relative Knowledge

		Individual Components of Knowledge Index II							
	(1) **Knowledge II**	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Edutainment Treatment	0.003 (0.012)	-0.016 (0.049)	0.000 (0.019)	-0.046 (0.037)	0.011 (0.035)	0.026 (0.029)	0.025 (0.034)	-0.010 (0.034)	0.017 (0.033)
Mean of control	0.746	0.449	0.961	0.204	0.409	0.675	0.592	0.848	0.904
Observations	287	287	287	287	287	287	287	287	287

Notes: Outcomes from the endline survey. Baseline covariates selected through a double-selection LASSO procedure Belloni, Chernozhukov, and Hansen (2014). Standard errors are clustered at household level and in parentheses. 'Knowledge' represents the Knowledge Index, an inverse-covariance-weighted index of the eight binary knowledge questions in Columns 2-9 (Anderson, 2008). In Panel A ('Knowledge I'), respondents receive points for mentioning the answer specified below in *any* of their recorded responses (we include four answer slots for each question). In Panel B ('Knowledge II'), respondents receive points weighted by the importance which they gave the answer specified below.

(2): What do you think causes your child to have a cough or cold? Respondent scores a 1 if they say that one can catch a cough or cold from other children or by touching germs, and a 0 otherwise (eg. from cold weather, playing in water).

(3): How do colds or coughs spread to other people? Respondent scores a 1 if they say that sneezing or coughing can cause colds to spread, and a 0 otherwise (eg. from food allergies, thrashing rice).

(4): What are some ways in which you can keep a cough or cold from happening in the first place? Respondent scores a 1 if they say that such illness can be prevented by washing ones hands, and a 0 otherwise (eg. dress warmly, put oil on body, eat healthy food).

(5): What do you think causes diarrhea? Respondent scores a 1 if they say that dirty hands can cause diarrhea, and a 0 otherwise (ex. something in the water, something in the food).

(6): What are some ways in which you can keep you or your child from getting diarrhea in the first place? Respondent scores a 1 if they say that diarrhea can be prevented by washing ones hands, and a 0 otherwise (eg. dont eat too much).

(7): What do you think is the difference between washing your hands with water only and washing your hands with soap and water? Perhaps there is no difference? Respondent scores a 1 if they say that hands are cleaner when washed with soap, and a 0 otherwise (eg. No difference, hands smell different or look clean).

(8): In what way does soap make your hands cleaner? Respondent scores a 1 if they say that it removes germs or 'worms (another word for germs) from hands, and a 0 otherwise (eg. removes dust).

(9): If your hands look clean is there any need to wash them with soap? Why? Respondent scores a 1 if they say yes, you should wash with soap in order to get rid of the germs or invisible worms, and 0 otherwise (eg. to continue the habit, dont know).

Table 4: Machine learning algorithms comparison

Algorithm and Sampling		Hyperparameters		Training Accuracy	Testing Accuracy
Regression		α	λ		
	Lasso	1.00	0.00	71.61%	59.00%
	Elastic Net	0.85	0.01	74.01%	61.90%
Random Forest		Mtry			
		7.68		72.76%	53.40%

Notes: The λ is the penalty coefficient, or the degree of bias introduced into the ordinary least square regression to counter overfitting; α=1 signifies a LASSO regression. M is number of variables randomly sampled at each split. Both training and testing accuracy are highest for the elastic net algorithm.

Table 5: Child health

Panel A: Effect of edutainment campaign

	(1) Loose stool	(2) ARI symptoms
Edutainment Treatment	-0.007 (0.003)	-0.023 (0.009)
Mean of control	0.01	0.07
Observations	3284	3284

Panel B: Effect of dispenser alone

	(1) Loose stool	(2) ARI symptoms
Dispenser Treatment	-0.038 (0.011)	-0.071 (0.015)
Mean of control	0.04	0.09
Observations	4053	4052

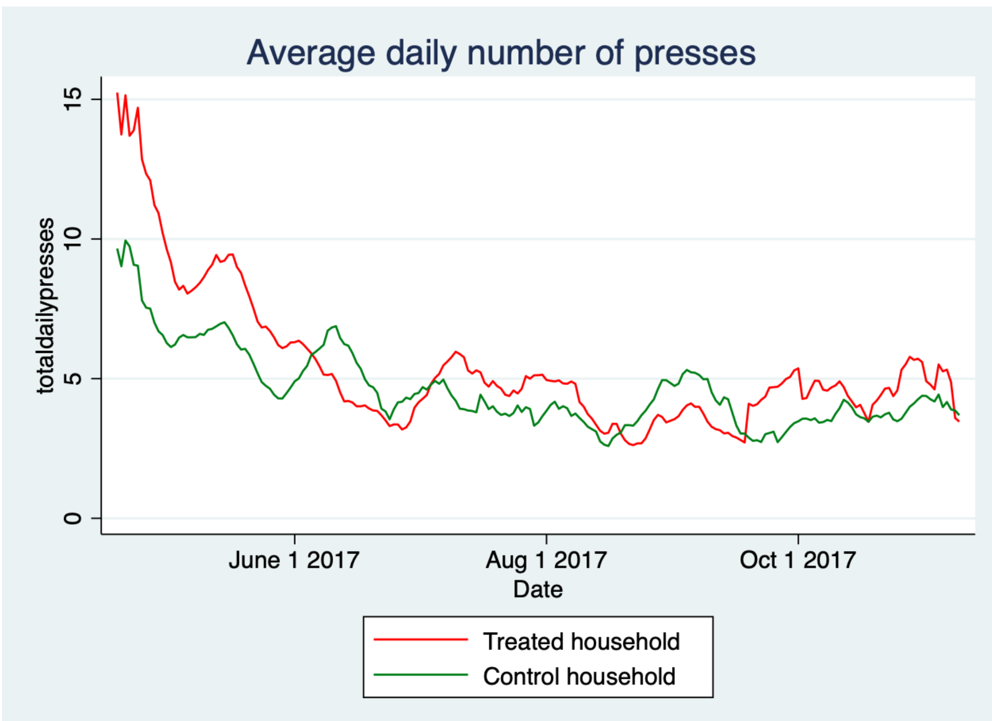
Panel C: Effect of edutainment on subsample with phone entertainment

	(1) Loose stool	(2) ARI symptoms
Edutainment Treatment	-0.012 (0.007)	-0.013 (0.016)
Mean of control	0.02	0.06
Observations	997	997

Notes: Health outcomes obtained from monthly health surveys. “Loose stool” is a binary variable equal to one if child had any loose stool in the previous two weeks. “ARI symptoms” is a binary variable equal to one if child had a cough, cold, or runny nose in the previous two weeks. All regressions include survey round fixed effects and standard errors clustered by household. “Edutainment treatment” are households who received the dispenser and the mobile phone edutainment campaign. “Dispenser treatment” are households who received the dispenser (but no edutainment campaign). “Pure control” are households who did not receive a dispenser (or edutainment campaign). Note that pure control households were not chosen randomly; these households were recruited in the initial sample (prior to randomization) but excluded because they owned a mobile phone which was already utilized for video entertainment by the female household head. “Subsample with phone entertainment” are those households in which the household head reported that their children use a mobile phone for entertainment at baseline.

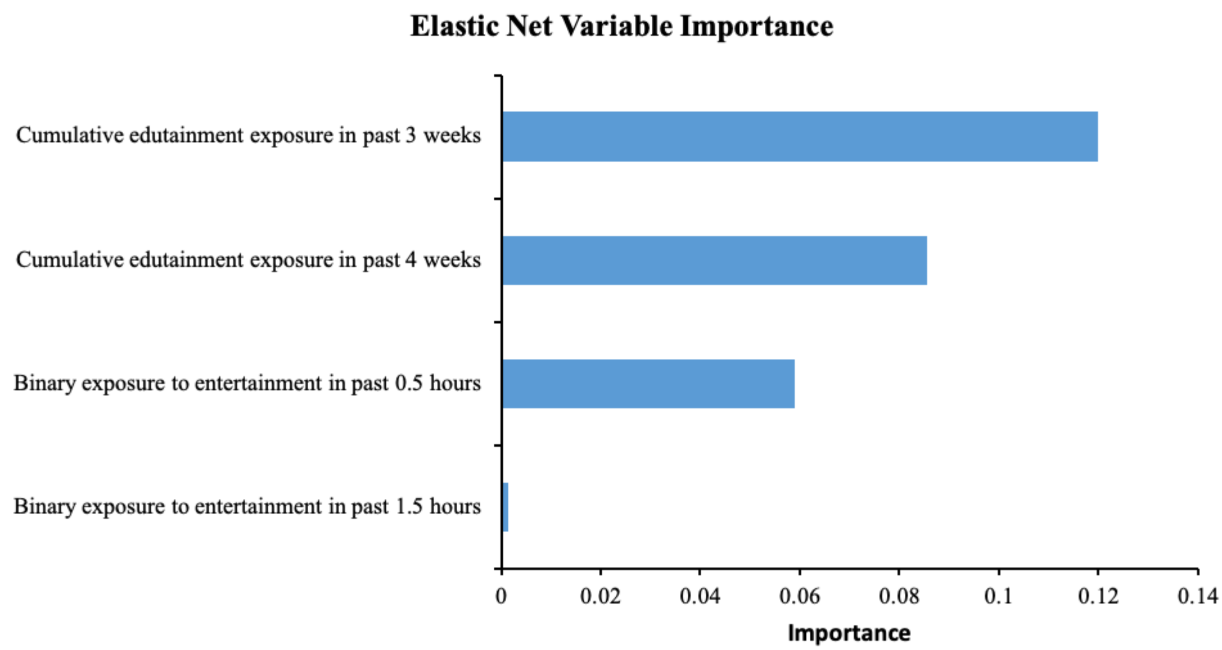
Figures

Figure 1: Dispenser use



Notes: Figure shows the average number of individual presses per day over the course of the eight months that sensor data was collected. Green line represents control households and red line represents households who received the edutainment intervention.

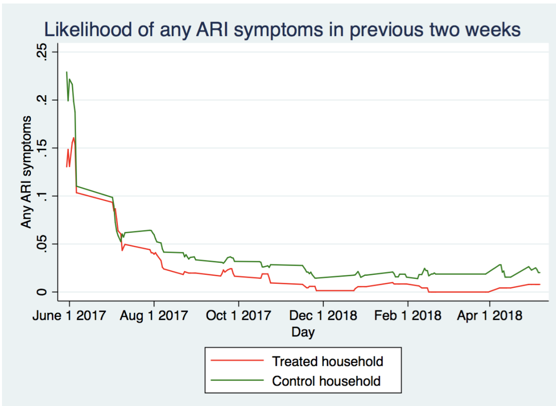
Figure 2: Elastic Net Feature Selection

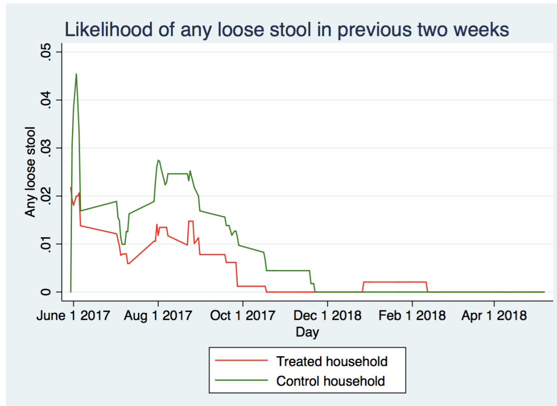


Notes: Figure shows the top four temporal features in order of importance selected by the elastic net algorithm. All 46 remaining features exhibited importance levels below that of 'binary exposure to entertainment in past 1.5 hours.'

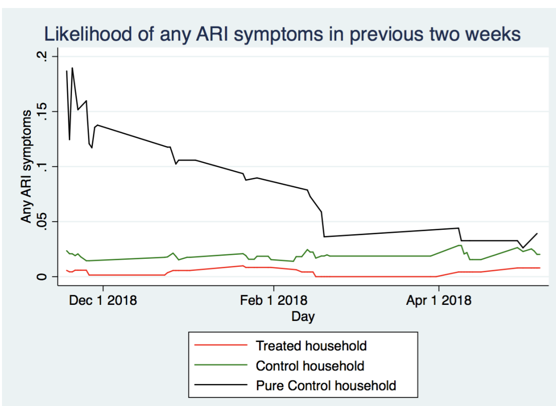
Figure 3: Child health over time

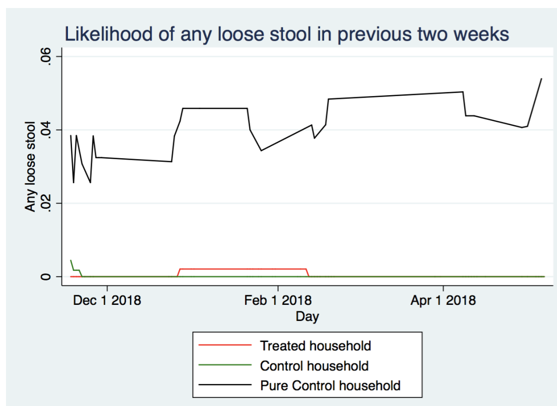
Panel A: Effect of edutainment campaign





Panel B: Effect of dispenser alone





Notes: Figures show the two-week moving average of reported incidence of loose stool and symptoms of ARI over the course of the experiment in Panel A, and during the last six months of the experiment (during which pure control data was collected) in Panel B. Green line represents dispenser control households, red line represents households who received the edutainment intervention in addition to the dispenser, and black line represents 'pure' control households, who received neither an edutainment program nor a dispenser.

A Appendix [online only]

A.1 Results

A.1.1 Tables

Table A1: Descriptives and balance

		Dispenser control mean	Edutainment treatment mean	p-value	N
Household and mother	Number of rooms	1.699	1.74	0.579	330
	Age at marriage	16.11	16.13	0.928	330
	Education	9.765	8.63	0.576	330
	Eat fish or meat every day	0.578	0.62	0.488	330
Hygiene practice	Drinking water is filtered	0.0241	0.01	0.419	330
	Open defecates	0.0120	0.02	0.404	330
	Owns a latrine	0.970	0.98	0.750	330
	Own soap	0.991	1.00	0.180	330
	Number of times washes hands with soap	4.494	4.40	0.749	330
	Whether hands washed with soap before eating	0.515	0.52	0.893	330
	Whether hands washed with soap before cooking	0.455	0.46	0.937	330
	Whether child washes hands with soap before eating	0.467	0.45	0.676	330
	Whether hands washed with soap after defecation	0.157	0.16	0.950	330
	Whether hands washed with soap after urination	0.506	0.52	0.715	330
Hygiene knowledge	Can get cold from germs	0.0904	0.10	0.823	330
	Handwashing with soap can prevent cold	0.0120	0.01	0.570	330
	Handwashing with soap can prevent diarrhea	0.542	0.53	0.832	330
	Soap makes hands clean even when they look clean	0.361	0.37	0.934	330
	Soap removes germs	0.542	0.52	0.665	206
Entertainment practice	Watches mobile phone for entertainment	0.217	0.16	0.176	330
	Minutes mobile phone watched for entertainment	6.988	5.73	0.413	330
	Child watches mobile phone for entertainment	0.289	0.34	0.308	330
	Minutes child watched mobile phone for entertainment	10.66	13.50	0.211	330
Child (60 months and below)	Any loose stool in last two weeks	0.00268	0.01	0.495	165
	Any ARI symptoms in last two weeks	0.0509	0.05	0.917	165
	Child height (cm)	80.97	79.01	0.613	165
	Weight (kg)	12.52	13.07	0.446	165
	Age (months)	37.09	38.07	0.728	165
	Male	0.550	0.56	0.853	165

Notes: Table reports the p-value and number of observations in a comparison of means between treated and control groups using data from the baseline survey.

Table A2: Test for differential attrition in followup data

		Interim survey		**Endline survey**		**Sensor data**	
		t-stat	**N**	**t-stat**	**N**	**t-stat**	**N**
Household and mother	Number of rooms	0.396	6,729	0.771	287	0.956	282
	Age at marriage	-0.750	6,729	0.301	287	-0.389	282
	Education	0.516	6,729	-0.466	287	-0.0208	282
	Eat fish or meat every day	0.116	6,729	0.730	287	0.671	282
Hygiene practice	Drinking water is filtered	-0.396	6,729	-0.785	287	-0.695	282
	Open defecates	0.890	6,729	1.378	287	0.570	282
	Owns a latrine	0.348	6,729	-0.0498	287	0.560	282
	Own soap	1.312	6,729	1.345	287	1.345	282
	Number of times washes hands with soap	-0.518	6,729	0.0450	287	-0.430	282
	Whether hands washed with soap before eating	0.0167	6,729	0.420	287	-0.627	282
	Whether hands washed with soap before cooking	-0.193	6,729	0.363	287	-0.584	282
	Whether child washes hands with soap before eating	-0.196	6,729	-0.276	287	-0.462	282
	Whether hands washed with soap after defecation	-0.620	6,729	0.273	287	0.00539	282
	Whether hands washed with soap after urination	0.640	6,729	-0	287	0.212	282
Hygiene knowledge	Can get cold from germs	1.116	6,729	-0.300	287	-0.108	282
	Handwashing with soap can prevent cold	-0.113	6,729	-0.552	287	0.0801	282
	Handwashing with soap can prevent diarrhea	-0.640	6,729	-0.151	287	-0.680	282
	Soap makes hands clean even when they look clean	-0.470	6,729	0.474	287	0.181	282
	Soap removes germs	-0.301	4,114	0.707	183	-0.680	282
Entertainment	Watches mobile phone for entertainment	-1.807	6,729	-0.605	287	-0.983	282
	Minutes mobile phone watched for entertainment	-1.529	6,729	-0.204	287	-0.630	282
	Child watches mobile phone for entertainment	0.558	6,729	1.089	287	0.996	282
	Minutes child watched mobile phone for entertainment	0.883	6,729	1.175	287	1.215	282
Child (60 months and below)	Any loose stool in last two weeks	0.599	6,788	na	287	na	282
	Any ARI symptoms in last two weeks	0.0977	6,788	0.0350	287	0.570	282
	Child height (cm)	-0.807	1,597	0.749	21	1.216	20
	Weight (kg)	0.641	1,597	0.275	21	0.340	20
	Age (months)	-0.0621	1,597	1.177	21	0.990	20
	Male	-0.331	1,597	-1.768	21	-1.502	20

Notes: Table reports the t-statistic and number of observations in a comparison of means between treated and control groups for the subsamples followed up in each specified data source, using data from the baseline survey.

Table A3: Descriptives and balance for pure control group

		Pure control mean	Dispenser treatment mean	t-statistic	N
Household	Education of household head	4.500	4.643	0.266	186
	Electricity	0.688	0.6369	-0.548	186
	Sanitary latrine	0.250	0.136	-1.615	186
Hygiene practice	Wash only with water	0.688	0.799	1.380	186
	Wash with ash	0.303	0.186	-1.511	194
	Wash with soap	0	0.00621	0.452	194
Entertainment	Videos on mobile primary source of entertainment	0.970	0.224	-10.06	194

Notes: Table reports the t-statistic and number of observations in a comparison of means between treated and control groups for the subsamples followed up in each specified data source, using data from the baseline survey. Selection of variables is smaller than previous balance tables as we conducted a significantly shorter survey among 'pure control' groups, who were more vulnerable to survey fatigue given that they did not receive any intervention from the research team.

Table A4: Other sanitation and hygiene actions

	(1)	(2)	(3)	(4)	(5)	(6)
	Filters water		Open defecates		Has latrine	
Treated	0	0.000436	0	0	-0.00710	-0.00809
	(0.00811)	(0.00911)	(0)	(0)	(0.00711)	(0.00820)
Control mean	0.024		0.012		0.970	
	(0.154)		(0.109)		(0.171)	
With controls		X		X		X
Observations	287	287	286	286	287	287

Notes: Standard errors are in parentheses. All regressions include the baseline value of the outcome as a control. Additional controls are mother's age, age at marriage, literacy level, whether she completed primary education, whether she owns the home, the number of rooms in the home, whether the house has electricity, and respondent religion. Corrected q-values using Anderson (2008). *** p<0.01, ** p<0.05, * p<0.1

A.1.2 Figures

Figure A1: SD card and mobile phone entertainment

Notes: Top two figures depict a typical street stall from which SD cards with preloaded entertainment are rented or purchased. Bottom figure depicts a family watching the entertainment through the SD card on the distributed mobile phone together.

Figure A2: Soap dispenser anatomy

Notes: The dispenser is a standard wall mounted handsoap dispenser with a foaming pump. It is opened with a special key available only to the surveyors. The sensor module is secured inside between the pump and the liter container.

Figure A3: Child using dispenser

Notes: A child uses the dispenser by pushing the black button once or twice. The foaming soap can be rubbed on the hands without water. He then goes to the nearby water pail or tubewell in the courtyard and rinses the soap off with the help of the mother, who pours the water.

A.2 Hygiene Knowledge Questions

Hygiene Knowledge: Now I would like to ask you some questions about hygiene.

SURVEYOR: Do not read out the codes. Wait for the respondent to answer, and fill in as many of the codes as they mentioned in order of importance.

What do you think causes your child to have a caugh or cold? (0=don't know; 1=chance; 2=it is cold outside; 3=they eat something bad; 4=they catch it from other children; 5=they touch germs that get into their body; 6=playing in water; 7=they don't wear proper clothing; 8=seasonal changes (beginning of winter, beginning of summer); 9=playing in dirt; 10=other (please specify))	**1**						Multi-select
How can you make a caugh or cold go away? (0=don't know; 1=nothing you can do, just wait; 2=give them medicine; 3=give them fluids; 4=feed them a specific food; 5=dress warmly; 6=take to doctor; 7=nothing I can do since I can't supervise them, I'm working; 8=other (please specify))	**2**						Multi-select
Can the cold or cough spread from one person to another? (0=don't know; 1=yes; 2=no)	**3**						Pick one
[IF YES] How does it spread? (0=don't know; 1=sneezing, coughing; 2=dust allergy, dhan thrashing; 3=food allergy; 4=other (please specify))	**4**						Multi-select
What are some ways in which you can keep a caugh or cold from happening in the first place? (0=don't know; 1=dress warmly; 2=eat healthy food; 3=drink clean water; 4=wash your hands; 5=stay clean (keep your whole body clean);6=changing clothes regularly, wearing clean clothes; 7=put oil on body; 8=nothing I can do since I can't supervise them, I'm working; 9=not playing in water; 10=other (please specify))	**5**						Multi-select
What do you think causes diarrhea? (0=don't know; 1=chance; 2=something bad or dirty in the food you eat/spoiled food; 3=something in the water you drink; 4=something on your hands; 5=other (please specify))	**6**						Multi-select
How can you make diarrhea go away? (0=don't know; 1=medicine; 2=drink extra fluids; 3=treat your water (boil, etc.); 4=eat a specific food; 5=saline-sugar solution; 6=take to doctor; 7=nothing I can do since I can't supervise them, I'm working; 8=other (please specify))	**7**						Multi-select

What are some ways in which you can keep you or your child from getting diarrhea in the first place? (0=don't know; 1=drink clean water; 2=eat healthy foods; 3=wash your hands with soap; 4=stay clean (keep your whole body clean); 5=moderate eating (don't eat too much); 6=nothing I can do since I can't supervise them, I'm working; 7=other	**8**					Multi-select
What do you think is the difference between washing your hands with water only and washing your hands with soap AND water? Perhaps there is no difference? (0=Don't Know; 1=No difference, 2=Hands smell different, 3=Hands look cleaner, 4=Hands are cleaner; 5=other (please specify))	**9**					Multi-select
[If 9 > 1] In what way does it make your hands cleaner? (0 = Don't know, 1 = Makes hands worm free, 2 = Removes dust/dirt, 3 = Removes germs, 4 = Other (please specify)	**10**					Multi-select
When do you think it's most important to wash your hands with soap? (0 = Don't know, 1 = Never, 2 = before cooking, 3 = before eating, 4 = after eating, 5 = after using the restroom, 6 = after picana, 7 = after returning from outside, 8 = others (please specify)	**11**					Pick one
Do you think that using soap on your hands can help prevent a sickness? (0=Don't know, 1=Yes, 2=No)	**12**					Pick one
If your hands look clean, is there any need to wash them with soap? (0=Don't know, 1=Yes, 2=No)	**13**					Pick one
IF YES, why? (0=Don't know, 1=to clean germs, 2 = to clear of invisible worms, 3 = to clear off dirt you can't see; 4=to continue the habit; 5=other (please specify))	**14**					Multi-select

Notes: Displayed above is the hygiene knowledge module of the baseline and endline surveys. To build the hygiene knowledge index, we exclude questions 2, 7, and 11, since no answer to these questions is more or less indicative of exposure to the edutainment. We also exclude question 12, for which 100% of control respondents answer correctly at endline (unsurprising given that all sample households received the soap dispenser and soap regularly over the course of the experiment). Knowledge results are robust to the inclusion of any or all of these questions.

A.3 Variable Construction for Machine Learning Exercise

We employ a binary rather than continuous measure of dispenser presses to define handwashing behavior. Our exercise is therefore transformed into one of classification of household-mealtimes into 'washing or 'non-washing.

We then collapse our data into two mealtime ranges, which generates a relatively balanced panel. We choose to collapse rather than preserve the original minute-level data as the latter would yield an unbalanced panel in which the vast majority of observations (household-minutes) are 'non-washing observations, making any machine learning algorithm we employ appear highly predictive yet uninformative by classifying all observations as 'non-washing. [6]

While handwashing during mealtimes was not the only focus of the edutainment intervention, our sensor data demonstrate that households are most likely to use the handsoap dispenser during the morning breakfast hours (6-11 am) and the evening dinner hours (5-11 pm). As such, we identify the peak handwashing time within each range for each household per day and define the household-day-specific mealtime range as the peak half hour plus and minus an hour (eg. an 8 pm peak implies a dinnertime range of 7 pm to 9 pm). For household-days with no presses (and therefore no peak times), we assign default mealtimes of 7 am and 8 pm, the peak washing times across all households and all days.

Finally, we define a broad set of temporal variables related to information campaign exposure that we generate from second-level data around when and for how long households were exposed to both the edutainment and the entertainment programs on their phones each day. The complete list of temporal features can be provided upon request: fifty features were included. These temporal features range from the cumulative number of minutes the household was exposed to the media in the twelve, eleven, ten, etc. weeks prior to the given mealtime observation, to a binary measure of whether the household was exposed in the half hour prior to the mealtime observation, and defines these measures of exposure separately for the edutainment campaign and the entertainment shows and dramas.

Feature selection proceeds as follows. We randomly subset 80% of treated households into a training dataset and the remaining 20% into the test dataset. A cross-validation exercise then reduces the likelihood of over- or under-fitting the data during the learning process; for this, we choose to employ a holdout cross-validation technique with time slices rather than the more common K-fold cross-validation technique because of the time-series nature of our data: a random splitting of data into K-folds would lead to situations in which future data was used to predict the past (Norberg (2016)). The holdout method instead allows us to divide the training dataset into several month-long folds; we then iteratively train the data on month N-3, N-2, and N-1 and test the resulting model on month N (results are robust to a two month training window as well). We follow this feature selection process using three algorithms: LASSO, elastic net, and random forest.[7]

[6] Because the sample is still not perfectly balanced, we also do a robustness check using the Synthetic Minority Oversampling technique (SMOTE) (Chawla et al. (2002)); we find that this technique produces no change in model accuracy nor the resulting selected features, suggesting that our model is not biased towards the majority class (of non-washers).

[7] While the hyperparameters are [by definition] fixed in the LASSO model, we identify the hyperparameters with a random search for the optimal model for the elastic net and random forest algorithms (Bergstra and Bengio (2012)).