

## WSH Evaluation Convening Report

On March 31<sup>st</sup>, 2011, researchers working in Water, Sanitation, and Hygiene came together with their colleagues for a two-day, interdisciplinary meeting to share key lessons from evaluations in this sector. The goal was to stimulate cross-project learning in order to make research stronger and more relevant to practitioners, policy-makers, and funders such as the Gates Foundation. This report highlights important takeaways throughout the evaluation process, from intervention design to the measurement of take-up and outcomes. We also offer recommendations on a framework for future convenings.

### SECTION I. Lessons Shared and Learned

#### Intervention Design

- ***Appropriate design is context specific, and can be reached through an iterative evaluation process.*** Chlorine dispensers in Kenya are a prime example of how learning from evaluations can continuously inform product design. Out of a series of randomized interventions under the Rural Water Project, these community dispensers (coupled with promoters) were found to be most effective in increasing take-up of chlorine. When the WashBenefits project began, the chlorine dispenser was adopted as part of the intervention in Kenya, but researchers quickly learned that this product needed to be adapted for use in the highly-concentrated and urban Bangladesh. Evidence from this project demonstrates that the mix of hardware and software promoted should be contextually appropriate and informed by formative work.
- ***Convenience and salience may be more important for take-up than price.*** As mentioned previously, the shift towards community dispensers significantly increased take-up of chlorine in western Kenya. Located at the point of water collection, these dispensers were easy to use and served as a constant reminder to chlorinate water. In addition, because they were installed at public springs, social pressures may have encouraged villagers to adopt the chlorination process. As a follow-up, future work may want to consider separating out the relative importance of salience and convenience, as some sanitation technologies involve a trade-off. Researchers may also want to consider using both demand and supply side promotion strategies to improve take-up, and test the efficacy of these strategies jointly.
- ***Ascertain take-up potential.*** Under encouragement designs, correct assumptions on uptake of the interventions (including latrine use and hand-washing behavior) are crucial to design evaluations that allow for adequate and meaningful minimum detectable size effects.
- ***Balance between product effectiveness and uptake potential.*** A criterion for any intervention is that it is an effective solution to the problem at hand. For example, women and children spend a lot of time collecting water, so giving households easier access to water may have large welfare consequences. Since water will be used for other reasons in addition to drinking, the capacity of any product (such as rain-harvesting water tanks) needs to be designed appropriately in order to substantially alleviate the problem of time allocation. However, sometimes what is most effective at targeting the problem may not be the most cost-effective solution – nor popular. The cholera trial in Bangladesh, conducted at scale, piloted interventions more upon cost-effectiveness and popularity than on their demonstrated effectiveness in preventing disease. This uncommon approach has benefits in that it avoids evaluation of a technically sound but unaffordable and unpopular intervention.
- ***Use the product to collect data.*** Electronic data capture can potentially be built into the equipment. For example, the passive latrine use monitor (PLUM) allows for the collection of objective data, which contributes to a better understanding of sanitation and latrine use behavior.

#### Evaluation Design

- ***Capturing long-term impacts.*** Long-term impacts of WSH interventions can be impressive and convincing evidence for policy level changes. For example, the Kenya Life Panel Survey project, which produces a unique longitudinal dataset of Kenyan youth, demonstrates that child health gains from treating worm infections translate into meaningful improvements in adult earnings (by over 20%) and

living standards. The project underscores the value of putting in place plans to rigorously follow up populations in intervention trials that would be expected to have long-term effects.

- ***High frequency surveying can change behavior.*** The frequency of surveyor interaction may be tantamount to an intervention itself. Researchers may want to consider integrating this factor as an additional measurement for future assessments. For example, subpopulations of different treatment arms can be surveyed with different frequencies to determine the impact of surveying alone. This also points back to the role of using products to collect data, so that enumerator contact is minimized.

### **Study Implementation**

- ***Invest in training to ensure better measurement of primary exposure and outcome variables.*** Evidence from the WashBenefits project in Kenya and Bangladesh suggest up to a 3-fold impact on required sample size attributable to variability in the measurement of just one outcome measure, height for age z-score. Training (and re-training during prolonged studies) for field staff might seem to be a "routine" task, but it deserves a significant investment of attention and resources. Failure to do so could result in an underpowered study unable to answer its primary research questions.
- ***Closely monitor and document field activities.*** There is a need to accompany effectiveness real world studies with good monitoring and process evaluations. When we analyze results from evaluations and don't find evidence supporting impact, this will help disentangle whether the intervention as designed is insufficient to generate the desired impacts, or if impacts are faltering because interventions were not implemented as anticipated.

### **Measuring Exposure to Contaminants, Take-Up, and Outcomes**

- ***Use a 7-day recall window to optimize the trade-off between bias and variance, especially to power evaluations.*** Unlike most recent recommendations to limit the recall window to 2 days, WSP recommends that investigators collect diarrhea symptoms with daily recall that extends up to 7 days if the measurement will be used to estimate differences between groups. Estimates will be minimally biased and the studies will have greater statistical power.
- ***Consider Environmental Enteropathy as the primary causal pathway towards child stunting.*** Environmental Enteropathy is a subclinical (i.e., no symptoms), highly prevalent condition of people living in resource constrained settings. It is a primary cause of stunting and under-nutrition, which in turns leads to increased child mortality, cognitive deficits, and reduced adult economic productivity. The Zvitambo project in Zimbabwe hypothesizes that EE, not diarrhea, is the primary causal pathway between WASH interventions and child stunting – that is, EE can be prevented or reduced by improved WASH. If this is true, WASH interventions have been underestimated because they have been primarily appraised for their effect on diarrhea. However, this also highlights the uncertainty about which biomarkers to use if we want to define and measure EE.

### **Other Considerations**

- ***Gender differences.*** WSH interventions, especially in schools, should consider the different sanitation needs of girls versus boys.
- ***Social networks.*** It is important to look at social effects of community interventions in sanitation, especially if it relies on collective action. Researchers should also be cautious of unintended consequences of interventions to community structures. For example, by paying promoters to deliver a message, there is a risk that their credibility may be undermined. This provokes serious consideration of how we might think about social networks and their contribution to behavior change.
- ***Path dependency in preferences.*** Individuals with access to a technology that is moderately better than the traditional technology may not be willing to pay the start-up costs of trying the superior technology. As an example, women in developing countries are increasingly using purchased products for menstrual management. The order of introduction of new or unfamiliar products may affect the formation of women's preferences for a particular category of menstrual management product over substitutes.

## SECTION II. Investment Recommendation for the Foundation

The Gates Foundation may want to consider dedicating resources for methodological research because it can make important contributions to the sector and improve learning about what works. For example, it is clear that we need to develop new markers for diarrhea and other measures of exposure to water contaminants. Allowing researchers to incorporate a methodological study within their evaluations can help produce measurement strategies that both accurately quantify outcomes and can be easily applied when going to scale, especially once routine monitoring of outcomes is necessary.

## SECTION III. Recommended Structure for Future Convenings

- ***Ideas for future, “start-up” convenings for other sectors:*** Extend the convening to 2 full days, especially given the long distance travel of several participants. An additional half-day gives participants more time to hear about the specific designs and difficulties of each project, and give detailed feedback. It would allow for concentrated discussion time to narrow in on certain topics and hash out the issues in an interdisciplinary and cross-project environment, which is hard to come by. In addition, incorporating topic-specific break-out groups may allow participants to discuss the issues that are most relevant to their projects. The convening could also extend to an optional third day of topic-specific meetings.
- ***Convening structure for future meetings with current WSH group:*** It is essential to cover the entire process, from design to implementation to results. However, it might be useful to allocate presentation and discussion time disproportionately to focus on priority topics depending on the primary objective(s) of the meeting. Now that participants have more background on colleagues’ projects, topic-specific discussion sessions could dominate the agenda.
- ***Content of the convening.*** For future meetings with the same group, inquire with participants about where their needs are during the agenda development process. In addition, have both the presenters and discussants prepare discussion points to focus the audience the issues they would like to address. This would allow for more focused and productive talks. We should also offer more guidance to discussants on their role – they should bring out the main arguments of the presentation and offer points of discussion, but avoid summarizing what has just been presented.
- ***Materials for Participants.*** In addition to the materials prepared for this event, prepare a participant list with contact information for continuing dialog.
- ***Invitees.*** In addition to CEGA guest invites, ask participants to invite affiliated graduate students to attend as they can greatly benefit from the exposure. Also, as these evaluations demonstrate impact, convenings should include more practitioners and policy-makers so that researchers can understand how their work influences development policies and implementation practices. In turn, these individuals can offer “on-the-ground” insight from their experiences. Bridging communication between researchers and practitioners could lead to fruitful collaboration and further evolution in the WSH sector.