EVALUATION: DECEMBER 2019 PUBLICATION: JANUARY 2022

SUSTAINABLE WATER AND SANITATION IN SIERRA LEONE

Impact evaluation of the 'Improved WASH Services in WAU and WAR Districts' project

Effectiveness Review Series 2019/20

Community members gather water at a newly constructed borehole. Credit: Muhammad Naveed/Oxfam.

JAYNIE VONK

Data gathering led by Robert Sam-Kpakra

OXFAM GB

www.oxfam.org.uk/effectiveness

OXFAM

ACKNOWLEDGEMENTS

The author would like to recognize and thank everyone who had a role in carrying out this Effectiveness Review. Thank you to everyone at Oxfam in Sierra Leone who contributed to the evaluation process, especially Muhammad Naveed and Innocent Mutabaruka.

Data gathering was led by consultant Robert Sam-Kpakra and his team – including Aisha Jalloh, Albert Williams, Benetta Konie Kargbo, David Momoh, Ibrahim K. Sheriff, Mohamed Daramy, Mohamed Feika, Zainab Jenneh Lavalie, and others (who chose to remain anonymous) – thank you, together you were a great team to work with on this evaluation. Appreciation is also due to all the interviewees, representing their households and communities, who took the time to share their information and experiences to make this evaluation possible.

Special thanks also to Simone Lombardini and Alexia Pretari for their support throughout the evaluation, to the Oxfam GB Sustainable Water practice group for input and feedback on measurement tools, to Franziska Mager for guidance on SenseMaker®, and to Sarah Barakat for support on gender integrated programming. Thank you to the HWISE team, especially Sera Young, Josh Miller and Chad Staddon, for your ongoing collaboration. Also, thank you to Sarah Dicken and Elijah Bisung for your support during our initial use of the EWI.

CONTENTS

Executive Summary 4

1 Introduction 9

2 Project description 11

3 Evaluation design 14

4 Data 16

5 Measuring Sustainable Water and Sanitation 22

6 Results 26

7 Conclusions 48

References 77

EXECUTIVE SUMMARY

Oxfam GB's Strategic Evidence Framework is part of the organization's effort to better understand and communicate its effectiveness, as well as enhance learning across the organization. Under this Framework, a small number of completed or mature projects are selected each year for an evaluation of their impact, known as an 'Effectiveness Review' (ER). During the 2019/20 financial year, one project selected for an ER was the 'Improved Water, Sanitation and Hygiene (WASH) Services in Western Area Urban (WAU) and Western Area Rural (WAR) Districts' project – also known as the Freetown WASH Consortium (FWC) programme's third phase (referred to hereafter as FWC Phase III). Funded by UK Aid, the consortium was led by Oxfam together with other members, Against Hunger, Concern Worldwide and Save the Children. Collaborating stakeholders included the Government of Sierra Leone (GoSL), WASH-NET, the Guma Valley Water Company (GVWC) and others.

FWC Phase III was implemented from October 2016 to March 2019, following Phase II (April 2013 to September 2016) and Phase I (February 2010 to March 2013). Project activities during Phase II were reprioritized to respond to the Ebola outbreak from late 2014 to early 2016 (WHO, 2020). Phase III was aligned with the GoSL's national agenda on Ebola recovery and increased preparedness against possible future outbreaks. Broadly, the project aimed to improve the availability, accessibility, affordability and sustainability of integrated WASH services.

EVALUATION APPROACH

The ER aimed to answer the following questions:

1. What is the impact of FWC Phase III for individuals and their households in terms of Sustainable Water and Sanitation?

2. What community-level factors are necessary and sufficient for sustainability of FWC Phase III?

3. How do community-level factors for Sustainable Water and Sanitation relate to the household-level outcomes observed?

To answer these questions, we used quasi-experimental methods using propensity score matching (PSM), qualitative coding of personal stories (inspired by SenseMaker®), and configurational analysis, using EvalC3. We assessed impact in intervention and comparison communities at the individual and household level, allowing us to make causal statements about project impact. We also used community-level information to understand which factors contribute to better individual- and household-level outcomes. Any impacts of broader activities conducted across the entire local area (e.g., campaigns and advocacy, support to Freetown City Council [FCC]) were not within in the scope of this evaluation.

Data were gathered in December 2019 and January 2020 in areas of Freetown where Oxfam had implemented the project, limiting the scope to WAU District only (i.e., we did not visit areas where other consortium members had implemented, except for two communities where Concern Worldwide had worked). The intervention group included 11 communities across the district – five where new boreholes were constructed by Oxfam, four where water kiosks were installed by Oxfam, and two where new boreholes were constructed and public toilets were rehabilitated by Concern Worldwide. The comparison group consisted of 11 communities identified by Robert Sam-Kpakra, the consultant leading data gathering, by visiting nearby areas to observe characteristics and check if there had been any similar WASH projects since 2016.

Households were randomly selected for interview among the entire population within each community using a random walk protocol. The gender of the interviewee representing each

household was randomly determined before the interview using SurveyCTO. A total of 1,079 household surveys were completed – 444 in the intervention group and 635 in the comparison group. During analysis, propensity score matching (PSM) and multivariate regression were used to control for apparent baseline differences (using recalled baseline data) between the groups, with a final sample size of 981 (410 intervention, 571 comparison). Of those interviewees, 35% chose to share a personal story about something important that had happened recently, where water and sanitation played a role; they also interpreted their stories through follow-up questions (an approach inspired by SenseMaker®). Fifteen Key Informant Interviews (KIIs) were also conducted to gather community-level information for configurational analysis.

RESULTS

The results of this evaluation are discussed below. Table 1 summarizes impact findings, with highlights in green if positive, red if negative and yellow if mixed (i.e., both positive and negative impacts). 'Limited' refers to cases where impact is observed among subgroups, but not overall. Subgroup analysis was done by implementation type, gender of interviewee and disability of household members.

Table 1: Summary of Effectiveness Review results for project impact.

Outcome	Evidence of impact?	Notes
 Improved water access	Yes	20% increase in during the dry season; no significant impact during the wet season
 Improved sanitation access	Mixed (limited)	No significant impact overall; 4% increase among those with a disability and 2% decrease among those with no disability
 Reduced water insecurity	Yes	54% reduction in water insecurity (HWISE score); 65% reduction in proportion of water insecure households; larger impact with borehole compared to kiosk
 Sustainable Water and Sanitation: Overall	Yes	24% increase in index score
 Sustainable Water and Sanitation: Water Security	Yes	119% increase in dimension score; larger impact with borehole compared to kiosk
 Sustainable Water and Sanitation: Equity	Yes	24% increase in dimension score; larger impact with kiosk compared to borehole; larger impact with public toilet rehabilitation
 Sustainable Water and Sanitation: Institutions	Yes	21% increase in dimension score
 Sustainable Water and Sanitation: Operations	Yes	15% increase in dimension score
 Sustainable Water and Sanitation: Well-Being	Yes (limited)	6% increase for men but no significant increase for women; 10% increase in Quality of Life indicator overall; overall larger impact with public toilet rehabilitation
 Sustainable Water and Sanitation: The Environment	Yes	23% increase in dimension score

	Empowerment in WASH	Mixed (limited)	No significant impact overall; 11% increase in proportion empowered those with a disability and 17% decrease among those with no disability

Overall, we find the project had large positive impacts, including increasing access to improved water and decreasing water insecurity. Impacts on improved sanitation were smaller and mixed. The project also had a large positive impact on the Sustainable Water and Sanitation Index, including in five of the six dimensions and for 12 of the 14 indicators; details shown in Figure 1. Impact on Empowerment in WASH was not found overall and was mixed among subgroups.

Figure 1: Sustainable Water and Sanitation Index dimensions (top row) and indicators showing project impact (PSM estimates; * p <0.1, ** p <0.05, *** p <0.01)

Personal stories (inspired by SenseMaker®) allowed additional information to emerge in relation to Sustainable Water and Sanitation. Most stories were more about water than sanitation and more negative than positive. Most stories were about Water Security, Well-Being, and The Environment. The main themes that emerged beyond these dimensions were gender-based violence (GBV) – sexual harassment and rape – and teenage pregnancy, illustrating the link between experience of violence and problems with water and sanitation.

Configurational analysis highlights the importance of the borehole intervention for Sustainable Water and Sanitation. We also see that is it necessary to collect fees for water and important to have women leaders in the WASH Management Team (WMT). Other community-level factors also contribute to more successful outcomes, including the public toilet intervention and having a WMT that was already established before the project and has a bank account. Furthermore, if we look at the household-level outcomes of water insecurity and empowerment in WASH, we find that it is also important to have women members in the WMT and to have a well-established WMT (in this case before the project started, meaning success is more likely if it is possible to work with an existing team rather than establishing a new one).

Finally, we consider Oxfam's accountability mechanisms and discuss which results would be of most interest for sharing back with communities. Based on information from an accountability module included within the household surveys, among those familiar with the project who had experienced a problem at some point, only 35% said they know how to provide feedback or make a complaint (30% of women and 39% of men). The top three ways people said they prefer to share feedback were face-to-face, hotline and in a group. People would most like to hear back about plans for future water and sanitation improvements in their communities and plans for using the results for influencing. Many raised issues of water accessibility, especially during the dry season in March and April each year, but also the opening hours each day and long queues. People requested upgrades to their water systems to make it safe for drinking and to address issues of water pricing

affordability, with differential pricing reported in some cases. Finally, a few people wanted to know more about any positive changes found by the survey, such as school attendance of children, well-being and any other good news. Note that among the people who had taken part in a similar survey in the past, only 15% had heard back about the results, representing an overall lack of accountability in survey research (Kashurha, 2019). Within the scope of this evaluation, we were able to share summary information back with community leaders and water point attendants addressing these requests but were unable to reach all community members who had participated in the survey because of various resource limitations. We acknowledge this is a critical limitation of this work.

PROGRAMME LEARNING CONSIDERATIONS

ENSURE GENDER MAINSTREAMING IN ALL FUTURE WASH PROGRAMMING, INCLUDING ADDRESSING WOMEN'S LEADERSHIP AND GENDER-BASED VIOLENCE (GBV)

Findings from the analysis of stories about personal experiences with water and sanitation reflected the emergence of key themes around gender-based violence and teenage pregnancy. Future programming should ensure robust gender mainstreaming using Oxfam's body of knowledge and building on Oxfam in Sierra Leone's learning from gender justice programmes. In addition to following guidance and best practice on gender mainstreaming, the country team can support this process by partnering with women's rights organizations to establish referral mechanisms while also following Oxfam's safe programming approach. Programmes could also work more closely with WMT women leaders; we do see that stories about GBV were shared less often in communities where the WMT has women leaders (6% of stories, compared to 14% in communities without WMT women leaders; the proportion of stories about teenage pregnancy did not vary substantially for these two groups), although there could be other reasons for this difference that we have not considered. In any case, WMT women leaders and members could benefit from additional training on how they can contribute to gender justice in relation to their role within the WMT. Broader awareness-raising activities could also help, as well as consulting women when designing and implementing infrastructure to ensure their needs are being met, particularly for safety (e.g., better lighting around water points and toilets).

CONSIDER ASSESSING AND IMPROVING WATER INFRASTRUCTURE TO MAINTAIN ADEQUATE SERVICE LEVELS DURING THE DRY SEASON

In their comments, people in both intervention and comparison areas raised ongoing issues with water access during the dry season, especially from March to May. For those in the intervention group, some reported that the boreholes provide less water during this time, and the queues also increase as others travel from further away to access that water when their own sources dry up. An infrastructure assessment is needed, which would have to be tailored to the design and context of each system. Then, based on the findings from that assessment, necessary improvements would need to be made to make it through the dry season in all communities. Any assessments and subsequent improvements would likely need to be done within the existing governance structures.

WASH MANAGEMENT TEAMS NEED FURTHER TRAINING AND SUPPORT, PARTICULARLY ON IMPLEMENTATION OF WATER PAYMENT SYSTEMS

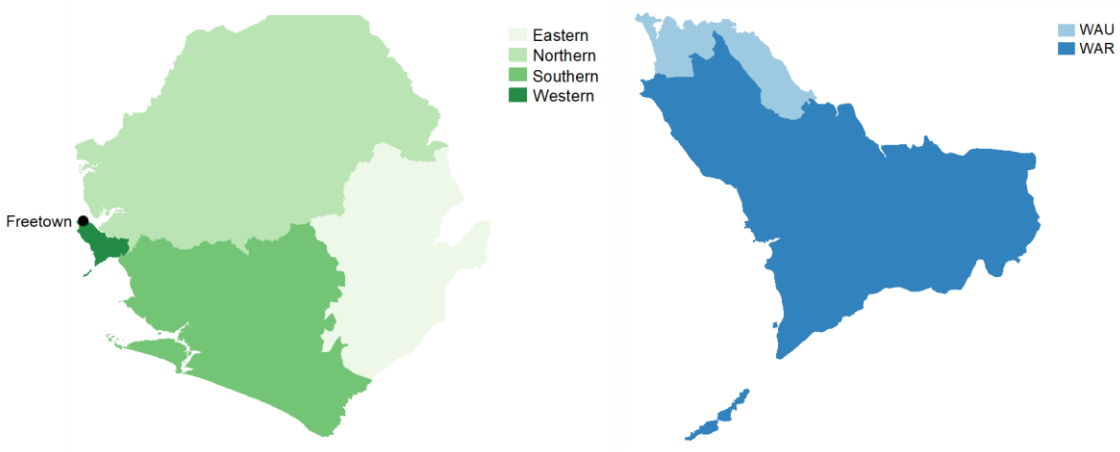
Through configurational analysis of community-level factors, this evaluation confirmed that collecting fees is necessary for increased sustainability of water systems, and it is also critical for achieving lower water insecurity. However, when asked what they would like to hear back from the survey, several people said that those collecting fees had raised the price and they wanted to find ways to reduce the price, or to keep the pricing fair (e.g., clearer or more transparent rules for any payment waivers or fee reductions given to specific households). Future programming should work to ensure payments happen regularly, to streamline the payment process, and to increase transparency and accountability related to the payment systems. For example, what else could be learned from existing invoice-based payment systems and/or mobile money payment systems in other countries? How can Guma Valley Water Company increase transparency for WMTs to know what happens to the funds they send? And how can WMTs increase transparency for water users to understand what happens with the fees they pay? Finally, some WMTs were not very active any more at the time of this evaluation. It is important to better understand the reasons for disengagement in order to identify what steps could be taken to support them in continuing to carry out their mandate.

1 INTRODUCTION

Every year since 2011, Oxfam Great Britain (OGB) has conducted rigorous impact evaluations known as Effectiveness Reviews (ERs). As part of our Strategic Evidence Framework, these reviews evaluate randomly select projects that have been active for at least two years and have a minimum budget of £200,000. We look for evidence of a cause–effect relationship between the project activities and any observed outcomes and impacts to understand whether our work leads to positive changes in the lives of the people with whom and for whom we work.

For the financial year 2019/20, projects were selected under five thematic areas – Livelihoods, Women's Empowerment, Resilience, Good Governance, and Sustainable Water and Sanitation. The 'Improved Water, Sanitation and Hygiene (WASH) Services in Western Area Urban (WAU) and Western Area Rural (WAR) Districts' project was selected under Sustainable Water and Sanitation. Figure 1.1. shows the location of the districts within the Western Area of Sierra Leone. Also known as the Freetown WASH Consortium (FWC) programme's third phase, this project is referred to as 'FWC Phase III' throughout this report. The consortium, funded by UK Aid, was led by Oxfam together with other members, Action Against Hunger, Concern Worldwide and Save the Children. Collaborating stakeholders included the Government of Sierra Leone (GoSL), WASH-NET, the Guma Valley Water Company (GVWC) and others.

Figure 1.1: Location of Western Area in Sierra Leone (left) and Western Area Urban (WAU) and Western Area Rural (WAR) in Western Area (right) (OCHA, 2020)



FWC Phase III was implemented from October 2016 to March 2019, following Phase II from April 2013 to September 2016 and Phase I from February 2010 to March 2013. Project activities during Phase II were reprioritized to respond to the Ebola outbreak from late 2014 to early 2016 (WHO, 2020). Phase III was aligned with the GoSL's national agenda on Ebola recovery and increased preparedness against possible future outbreaks. The project was implemented in 44 urban and peri-urban wards in Freetown, targeting the most vulnerable communities that were water deprived, without access to the GVWC water network, lacking access to functional public toilets, low-income and in areas of high population density. Potential hazards in each ward were also tracked (e.g., floods, cholera and Ebola) but not used as selection criteria.

Broadly, the project aimed to improve the availability, accessibility, affordability and sustainability of integrated WASH services. According to the project's final report, 127,479 people were reached through water supply improvements, 35,546 were reached through direct sanitation activities, 478,786 were reached through behaviour change communication and hygiene promotion and 52,134 who were disaster-affected were reached with integrated WASH services. Additionally, the project developed new models for decentralized water supply and public toilet management, developed a disaster risk reduction (DRR) strategy and trained 450 disaster management

volunteers, trained 102 youth groups in solid waste management and carried out additional campaigning and training activities.

The evaluation focused on water and sanitation services – namely new boreholes and water kiosks and rehabilitated public toilets – in areas of Freetown where Oxfam was the implementer, which limited the scope to WAU District only (i.e., we did not visit areas where other consortium members were the implementing organization, except for two communities where Concern Worldwide worked). The data gathering phase took place in December 2019, nine months after the end of the project in March 2019. The questions guiding the evaluation were:

1. What is the impact of FWC Phase III for individuals and their households in terms of Sustainable Water and Sanitation?
    a. How does impact vary by implementation type?
    b. How does impact vary by interviewee gender?
    c. How does impact vary by disability of household members?
2. What community-level factors contribute to the sustainability of FWC Phase III?
3. How do community-level factors contribute to individual and household outcomes of interest?

2 PROJECT DESCRIPTION

2.1 PROJECT ACTIVITIES

The project activities fit into five categories – (1) water, (2) sanitation, (3) behaviour change communication (BCC), (4) disaster risk reduction (DRR) and (5) advocacy. Different activities were carried out in different areas of Freetown, and certain project activities were done at the district or national level (e.g., strategy development, advocacy).

In detail, project activities for water included constructing 25 new boreholes and rehabilitating 16 existing boreholes, together with the installation of solar pumping systems. Also, 50 water kiosks were installed, which are large water tanks that are refilled by GVWC water trucks. While the boreholes were seen by the project as a longer-term solution to decentralized water supply, the kiosks are meant to be an interim solution. WASH Management Teams (WMTs) were established and trained for each new water point. The project also developed a new management model for these decentralized water supply systems – a GVWC-regulated Public Private Partnership (PPP) – and piloted it in selected areas.

The activities for sanitation included rehabilitating 10 public toilets and conducting sanitation marketing activities for household toilets. Additionally, the project trained 102 youth groups in solid waste management, as part of a tangential project called Operation Clean Freetown (OCF). The project also developed a new management model for public toilets and trained 65 public toilet attendants following the newly developed model.

For BCC, project activities included developing simplified messaging, undertaking hygiene promotion activities and campaigning via mass media (radio), door-to-door, public sign boards, and road shows/skits.

For DRR, project activities included developing a new DRR strategy, training 450 disaster management volunteers and supporting local government and the Office of National Security (ONS) on key aspects of disaster response.

The activities for advocacy included developing an advocacy strategy with WASH-Net, supporting local development councils in their planning processes, engagement with the GoSL and influencing through events, conferences, civil society organizations (CSOs) and documentation and sharing of good WASH practices.

2.2 THEORY OF CHANGE

The theory of change (ToC) for this project is illustrated in Figure 2.2.1. This version was recreated as part of the Effectiveness Review process during a two-day working session held on 28–28 November 2019 in Freetown, to reflect on what actually happened during the life of the project. The working session included representatives from Oxfam and the consultant.

Based on these activities, expected outcomes included increased access to sustainable, improved water services, local management and replication of decentralized, low-cost water systems, implementation of the new governance model that was recommended by the project and increased capacity for WASH service delivery overall. In terms of sanitation specifically, the expected outcomes were increased access to sustainable improved sanitation services, improved management of public toilets and solid waste as well as increased employment opportunities and better flood prevention. Further outcomes expected included improved hygiene practices and environmental sanitation, increased preparedness to mitigate against disasters and disease

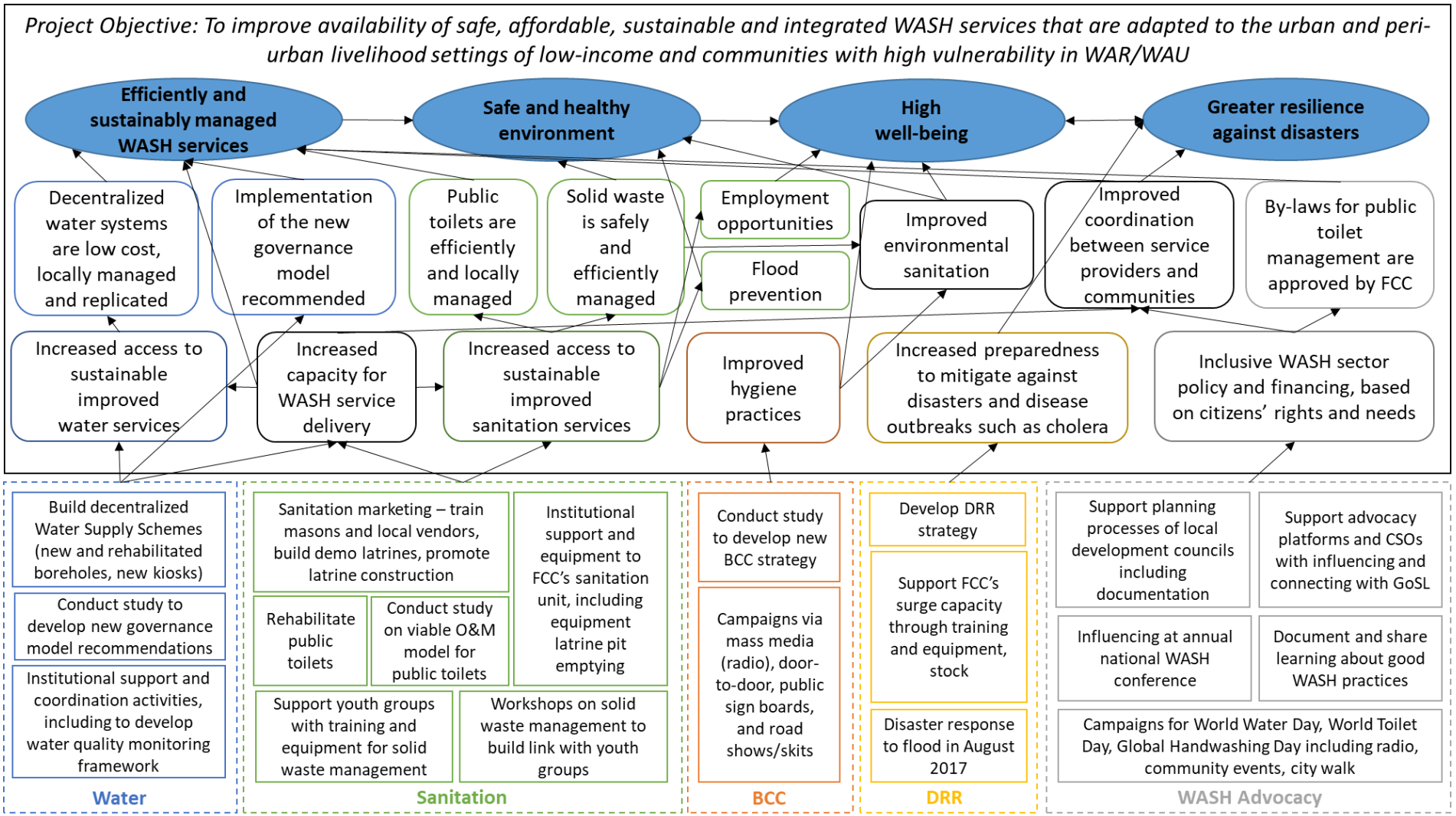
outbreaks, improved coordination between service providers and communities and inclusive laws and policies to meet the rights and needs of citizens.

The project outcomes aimed for impact – positive change in the lives of individuals and their communities. The expected impacts were: more efficiently and sustainably managed WASH services, a safer and healthier environment, increased well-being and greater resilience against disasters.

The ToC made assumptions about how change would happen. Any invalid assumptions may reduce the impact of the project. The key assumptions raised during the working session, which were made at the beginning of the project, were categorized into three groups: (1) valid, (2) partially valid, (3) invalid.

1. Valid assumptions did not present any issues for project implementation.
    a. Deep-well-drilling companies will be available for hire.
    b. Supplies and materials needed will be available on the local market.
    c. The groundwater resources in Freetown have the capacity (i.e., sustainable yield) to provide the required number of boreholes.
2. Partially valid assumptions presented some issues along the way, which may have reduced or reallocated impacts, but were mostly mitigated and did not require a complete change in the project design. Here we also include valid assumptions that posed challenges (f and g).
    a. The project will be able to balance its original plan while additionally supporting the GoSL's master plan (i.e., Dry Season Plan).
    b. Land will be made available for implementing infrastructure and kept public.
    c. GVWC will take over operation and maintenance (O&M) of water supply schemes.
    d. Currency fluctuation (both SLL and GBP) and inflation will not affect implementation.
    e. Fraud and corruption will be controlled and not affect implementation.
    f. Future disease outbreaks (e.g., cholera) will occur.
    g. Civil society space is narrowing.
3. Invalid assumptions required major changes in the project design, planning and implementation.
    a. The needed staff will be hired and retained for the duration of the project.
    b. It will be possible to secure partnerships with top research institutions to complete the planned study on BCC within the timeframe of the project.
    c. The full allocated budget will be available until the end of the project.
    d. Certain project activities (e.g., based on study findings) will continue as a Phase IV Urban WASH project.

Figure 2.2.1: The project's theory of change (recreated with Oxfam staff during a workshop held 27–28 November 2019 in Freetown, Sierra Leone)



3 EVALUATION DESIGN

This section presents the choices made for the evaluation design. The project worked with a wide range of actors on various types of activities at different levels. To rigorously assess its impact, we had to decide which aspects of the project could fall within the scope of the evaluation. Our goal was to carefully balance the need for both accountability and learning. The central problem in evaluating the impact of any project is understanding what changes are attributable to it versus what would have happened otherwise. To the extent possible, we also want to know which activities led to change – how did it happen? In this evaluation, we decided to examine the situation through household surveys and Key Informant Interviews (KIIs), focusing on individual, household and community-level change. Clearly, we could not directly observe what the situation would have been without the project. This 'counterfactual' situation could only be estimated.

Given a large number of direct project participants, we followed the common practice of estimating the counterfactual by comparing those who were part of the project (intervention group) to those who were not (comparison group). Assuming these two groups are the same, except for the project, observing the situation in both groups provides a good estimate of the counterfactual. In the case of such interventions, an ideal approach from a statistical perspective is to randomly assign individuals (or households, communities, etc.) to the intervention and comparison groups. Random assignment minimizes the chance of systematic differences between the groups and maximizes the confidence that any observed impacts were caused by the project. However, this approach is often not ideal for large-scale implementation, and not possible ex-post.

Thus, to answer the first evaluation question – what is the impact of the project for individuals and their households? – we adopted a 'quasi-experimental' evaluation design using propensity score matching (PSM) with data gathered through household surveys carried out with one representative of each household. We matched those in the intervention group with others with similar baseline characteristics in the comparison group during analysis. One key choice for this evaluation was to purposefully select intervention areas based on specific project activities, resulting in three intervention arms – (1) new borehole, (2) new water kiosk, (3) new borehole and rehabilitated public toilet. Having these three arms enabled us to assess the relative impact of the different types of interventions to inform future programming.

The matching process was done with a pre-defined set of baseline characteristics including information about the interviewee, household demographics, income sources, wealth, group participation, and water and sanitation access. The baseline data needed for PSM were not available, so survey interviewees were asked simple questions about their situation from 2016 (the year the project began, with implementation starting at the end of the year), thereby creating recall data (Nicola & Giné, 2013; Godlonton et al., 2018). While recall data may not be completely accurate, in this evaluation we assume these data will not bias the evaluation results because systematic variation between how interviewees remember information in the intervention and comparison groups is unlikely. Using recall data to recreate a baseline is not the ideal approach (methodologically); we opted to use it as a second-best option (pragmatically) because the needed baseline data were not available.

To answer the second and third evaluation questions – what community-level factors contribute to sustainability of FWC Phase III? and how do community-level factors contribute to individual and household outcomes of interest? – we analysed information from each community gathered through semi-structured Key Informant Interviews (KIIs). We did configurational analysis using EvalC3 (Davies, 2020) to further understand the main community-level factors leading to better outcomes at the household-level.

The overall evaluation design is summarized in Table 3.1. It allowed us to see project impacts at the household level and to understand how these impacts interact with factors at the community level. We were also able to hear individual stories from both household and community representatives. The main limitation of this design is that any impacts of broader activities conducted across the entire local area (e.g., campaigns and advocacy, support to the Freetown City Council [FCC]) were not explicitly evaluated. To do so would have required another evaluation methodology, which was not possible to integrate into this evaluation at the desired level of confidence, given the resources available.

Table 3.1: Summary of the evaluation design

Evaluation Question(s)	Data Source	Data Type	Analysis Description
1	Household survey closed questions asked to one individual per household	Quantitative	Quantitative analysis using propensity score matching (PSM) to estimate individual- and household-level project impacts, also assessing differential impacts for different interventions and social groups
1	Household survey open-ended questions asked to one individual per household	Qualitative	Qualitative analysis using thematic coding and quantitative analysis based on interviewee-coding of personal stories (inspired by SenseMaker®) using PSM to estimate individual-level project impacts
2 & 3	Key Informant Interviews (KIIs)	Qualitative	Quantitative configurational analysis using EvalC3 to identify community-level attributes that contribute to successful outcomes at the community level, on average

4 DATA

4.1 INTERVENTION AND COMPARISON

Oxfam and its consortium partners implemented FWC Phase III in 44 wards across WAU and WAR. Based on the resources available for this evaluation, it was not feasible to visit all areas. In total, we selected 22 communities, 11 each from the intervention and comparison groups.

Intervention Group: Given the focus of this ER on Sustainable Water and Sanitation, and areas where Oxfam was the implementing consortium member, we selected 11 communities for the intervention group based on participation in the project's water and sanitation activities, in three arms. First, we selected all five communities where new boreholes were constructed by Oxfam (out of 25 new boreholes constructed by the consortium). Oxfam did not do any of the borehole rehabilitations. Second, we randomly selected four communities where water kiosks were installed by Oxfam (out of 50 water kiosks installed by the consortium, 12 of which were installed by Oxfam). Third, we selected two communities where new boreholes were constructed AND public toilets were rehabilitated by consortium member Concern Worldwide (Oxfam did not rehabilitate any community-based public toilets).

Comparison Group: The evaluation consultant identified 11 comparison communities through a process of visiting neighbouring areas to observe whether various important characteristics were similar (e.g., housing characteristics, proximity to a main road, terrain, etc.) and verify that they had not had any similar WASH projects since 2016 (i.e., no new or rehabilitated boreholes, kiosks or public toilets).

4.2 HOUSEHOLD SURVEYS

The sampling strategy for selecting households is the same in the intervention group and the comparison group. In intervention areas, we have a stratified random sample of approximately 400 households from 11 communities. In comparison areas, we have a stratified random sample of approximately 600 households from 11 communities. The total sample size was expected to be at least 1000, stratified by community. For reference, Table 4.2.1 lists the number of communities and target household survey sample size by group and intervention type.

Table 4.2.1: Details of the sampled intervention and comparison communities

Group	Intervention Arm	# Communities	Sample Size (Target)
Intervention	New borehole	5	185
	New water kiosk	4	148
	New borehole and rehabilitated public toilet	2	74
Comparison	N/A	11	605
Total		22	1,012

Following a three-day training course, a team of three supervisors and 12 interviewers (contracted and managed by evaluation consultant Robert Sam-Kpakra) conducted the household surveys from 9 December 2019 to 16 January 2020 (more than 90% of surveys were completed before a holiday break from 24 December to 10 January, after which the remaining 10% of surveys were completed). The questionnaire was reviewed in English and Krio languages during the training for translation quality and to develop a common understanding of all questions among the team. All surveys were conducted digitally with SurveyCTO on mobile devices (with daily uploading); paper questionnaires were available as a backup. Throughout data gathering, privacy and data protection were

prioritized, following principles and protocols based on Oxfam's Responsible Data Policy and GDPR (Vonk, 2019; Vonk & Lombardini, 2020).

In each community, households were randomly selected using the following procedure:

1. For sparsely populated communities (e.g., a few main roads where all households are along the roads), community boundaries were defined in advance with guidance from a community leader who knew the boundaries. For densely populated communities (e.g., main roads and many small footpaths, where households are beyond the main roads), the community was divided into four approximately equal quadrants with guidance from a community leader who knew the boundaries.
2. From the centre of the community (for sparsely populated communities) or the centre of each quadrant (for densely populated communities), each interviewer spun a pen to indicate the direction in which to go.
3. The interviewers selected houses according to a pre-determined interval set by their supervisor (based on the sample size and the number of households in the community).
4. For each household, SurveyCTO randomly indicated whether the interviewee should be a woman or a man. This information was shown on the device.
5. The interviewer asked for an interviewee in the household who is knowledgeable about water and sanitation, specifying whether they needed to interview a man or woman.
    a. If a suitable interviewee was available and willing to participate in the survey, the interviewer proceeded with the interview.
    b. If a suitable interviewee was not available, the interviewer asked if they could come back another time when that person would be available.
    c. If it was not possible to interview a suitable interviewee because one did not exist in the household or was away for an extended period of time, the interviewer proceeded to interview an interviewee of the opposite gender.
6. In case of refusal or if a suitable interviewee could not be located after two attempts, the interviewer recorded the reason for no interview and proceeded from Step 3.

We aimed for a sample size of 1,012 (605 comparison, 407 intervention). Interviewee gender was randomly indicated for each household, so we expect a near to 50:50 ratio. This was done to ensure the lived experience and perspectives of women and men are represented in the data gathered for this evaluation (particularly important for questions at the individual level) (Pretari, 2018). In total, the team visited 1,088 households, of which six had no interviewee available (e.g., refused; 2 intervention, 4 comparison) and two did not consent to be interviewed (both comparison), thereby completing 1,080 surveys (636 comparison, 444 intervention). Of those surveys, one was dropped from analysis due to irreconcilable errors in the household roster. The final sample of interviewees who consented to and completed the household survey is shown in Table 4.2.2.

Table 4.2.2: Final sample of interviewees by gender

Group	Total Sample (N)	Men Interviewees	Women Interviewees
New borehole	192	97 (51%)	95 (49%)
New water kiosk	153	76 (50%)	77 (50%)
New borehole and rehabilitated public toilet	99	47 (47%)	52 (53%)
Intervention (total)	444	220 (50%)	224 (50%)
Comparison	635	343 (54%)	292 (46%)
Total	**1,079**	**563 (52%)**	**516 (48%)**

4.3 KEY INFORMANT INTERVIEWS

We sought to identify community representatives, including community leaders and tap (water point) attendants, from each intervention and comparison community to invite for interviews. In January 2020, 15 Key Informant Interviews (KIIs) were completed with five community leaders and 10 tap attendants. In many cases, the community leaders interviewed were the chief for more than one community, so we asked them about all relevant communities (to enable configurational analysis at community level). Table 4.3.1 shows the final combination of interviewees who provided information for each group and intervention arm.

Table 4.3.1: Final set of KIIs by group

Group	Intervention Arm	Community Leaders	Tap Attendants
Intervention	New borehole	X	X
	New water kiosk	X	X
	New borehole and rehabilitated public toilet	X	X
Comparison	N/A	X	N/A
Total		**5**	**10**

The KIIs were conducted in a semi-structured manner, with a pre-defined list of questions to guide the interviews. For all interviewees, we asked them to share information about specific themes to allow us to better link community and household-level information during analysis. However, the interviews were also allowed to flow naturally, rather than sticking to a specific structure. All interviews were conducted independently in Krio by the evaluation consultant, Robert Sam-Kpakra, and his team following principles and protocols based on Oxfam's Responsible Data Policy and GDPR (Vonk, 2019). Audio recordings were then transcribed and subsequently translated into English, organized by theme.

4.4 MATCHING PROCESS OVERVIEW

An overview of the most pertinent information from PSM process and other descriptive information is provided below. In short, we first filtered the data, dropping 96 observations (33 intervention, 63 comparison) from those who did not live in the same community during the first year of the project. Before matching we found several significant differences between the intervention and comparison groups. Summary statistics for the intervention and comparison groups before matching are available in Appendix 2.

Using PSM, we adjusted for the differences when estimating impacts; no significant differences remained after matching. In the matching process one intervention observation was dropped because no adequate matches could be found, and one comparison observation was dropped because the propensity score could not be calculated. Further details on how we did PSM and full specifications for this evaluation are available in Appendix 3. Table 4.4.1 shows the final sample sizes and gender distributions for each group.

Table 4.4.1: Final sample of interviewees after filtering and PSM by gender

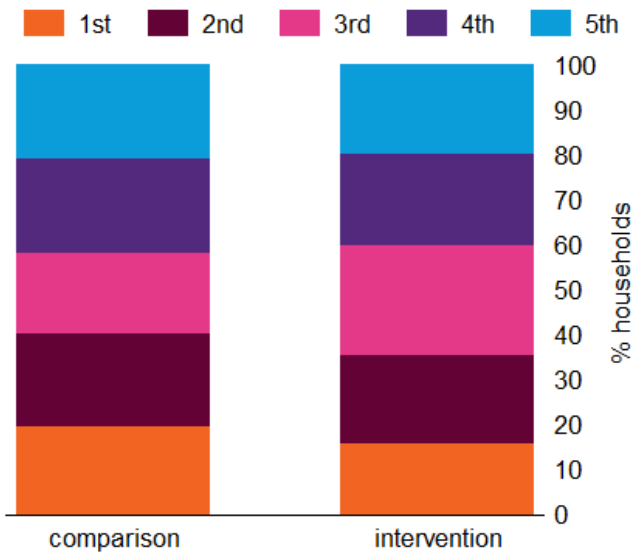
Group	Total Sample (N)	Men	Women
New borehole	176	88 (50%)	88 (50%)
New water kiosk	139	67 (48%)	72 (52%)
New borehole and rehabilitated public toilet	95	45 (47%)	50 (53%)
Intervention (total)	410	200 (49%)	210 (51%)

Comparison	571	309 (54%)	262 (46%)
Total	**981**	**509 (52%)**	**472 (48%)**

One key matching variable we use is a wealth index, which is based on household ownership of various assets (e.g., furniture, livestock, equipment) and housing conditions in 2013. When generating the index, we first verify internal consistency using Cronbach's alpha, following the guidance of Bland and Altman (1997), and then use a data reduction technique called principal component analysis (PCA) to assign appropriate weights to each variable in the index, following the approach of Filmer and Pritchett (2001). We ensure comparability of the wealth indexes from 2013 (based on recall data) and 2019 (based on the household's situation at the time of the survey), by pooling data by time period before undertaking PCA. We use wealth index quintiles for PSM, to avoid over constraining the matching process (i.e., households are matched with others based on similar wealth in 2013).

The significant differences before matching tell us a few important things. Households in the intervention group were more likely to be using an improved drinking water source in 2016. They also had slightly more household members, were less likely to have a household composition of two spouses and children and more likely to have a household composed of other family (e.g., siblings), and household members were participating in more community groups. The 2016 wealth quintiles show that the intervention group had more residents around the median (in the 3rd quintile) than the comparison group; also see Figure 4.4.1.

Figure 4.4.1: Wealth quintiles in 2016 (descriptive only)



4.5 CONFIGURATIONAL ANALYSIS OVERVIEW

Here we provide a brief overview of how we carried out configurational analysis using EvalC3. By combining community-level average outcome scores based on the household surveys (see Section 4.2) with qualitative information gathered from the KIIs (see Section 4.3), we compiled 22 cases for analysis. Each case represents one community.

The following 14 attributes were assessed as 'yes' (1) or 'no' (0) for each community:

1. Intervention type(s)
    a. New borehole
    b. New water kiosk
    c. Rehabilitated public toilet
2. Community characteristics
    a. Is along a paved road
    b. Terrain is mostly flat (i.e., not hilly)
3. Water and sanitation management team
    a. Was established in 2016 or earlier
    b. Is currently active
    c. Has women members
    d. Has women leaders
    e. Is collecting fees
    f. Has a bank account
    g. Has water safety plans in place
    h. Has plans in place that consider equity

The following outcomes were assessed as 'at or above the threshold' (1) or 'below the threshold' (0):

1. Sustainable Water and Sanitation Index (threshold set at the mean)
2. Empowerment in WASH Index (EWI) (threshold set at the mean)
3. Household Water Insecurity Experiences (HWISE) Scale (threshold set at the recommended value of 12).

More information on the definition of these outcomes is provided in Section 5. The resulting 'truth table' input into EvalC3 is provided in Appendix 5.

Each outcome had to be assessed separately. We used the configurational analysis to see which attributes, as well as combinations of attributes (also known as configurations), were good predictors for each outcome using the 'Design & Evaluate' and 'Decision Tree' features in EvalC3. We first assessed single attributes (i.e., one attribute at a time) and recorded whether each attributed was Necessary and/or Sufficient for each outcome, alongside the model accuracy for each attribute-outcome pairing. We also built a decision tree for each outcome of interest using an iterative process of assessing different configurations to maximize Accuracy.

An attribute is Sufficient if the outcome is always present when that attribute is present. An attribute is Necessary if the outcome is only present when that attribute is present. Accuracy is an overall performance measure that represents the proportion of True Positive cases (i.e., attribute and outcome both present) and True Negative cases (i.e., attribute and outcome both absent) (Davies, 2020). In other words, Accuracy presents how well an attribute or configuration predicts the outcome of interest.

5 MEASURING SUSTAINABLE WATER AND SANITATION

The FWC Phase III project under review was specifically aimed at sustainably improving water and sanitation systems and services. Oxfam GB's Sustainable Water and Sanitation Strategy (Mizniak et al., 2017) states that, to achieve sustainability, we must 'work with essential partners in a unified approach to connect poor and vulnerable communities to water and sanitation services, ensuring they are working equitably' to 'keep water and sanitation systems operational, accessible, and affordable'. This section outlines the measurement tools used to measure sustainable water and sanitation in line with these intentions.

SUSTAINABLE WATER AND SANITATION INDEX

Building on our approach for Effectiveness Reviews when measuring other complex themes, such as Women's Empowerment (Bishop & Bowman, 2014; Lombardini et al., 2017; Torre & Lombardini, 2019) and Resilience (Hughes & Bushell, 2013; Fuller & Lain, 2015; Pretari & Gaboune, 2019), we gathered a set of 'dimensions' to create an individual-level Sustainable Water and Sanitation Index (Miziniak et al., 2020). The dimensions shown in Table 5.1 were developed by listing out and then consolidating factors identified in the following sources:

1. Oxfam GB's Sustainable Water and Sanitation Strategy (Mizniak et al., 2017) and Outcomes-based Monitoring Framework (Medland, 2018)

2. Sustainable Development Goal 6 (UN, 2018)

3. A literature review of existing relevant frameworks and indexes (Banerjee & Morella, 2011; Bartram et al., 2014; Bratton & Gyimah-Boadi, 2016; Giné-Garriga & Pérez-Foguet, 2018; Kayser et al., 2013; Porteous, 2016; Shilling et al., 2013; Thomas et al., 2018; WHO/UNICEF, 2017; Wilbur & Danquah, 2015).

Table 5.1: Six dimensions of Oxfam's Sustainable Water and Sanitation Index

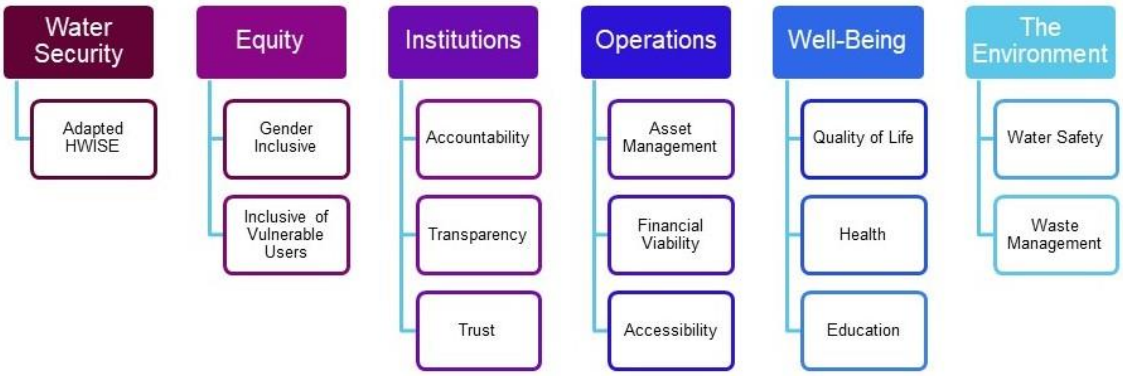
Dimension	Definition
Water Security*	People have consistent and sufficient access to preferable and acceptable water; water is not a source of worry, shame or anger.
Equity	WatSan facilities, services, and related management systems and decision-making processes are inclusive of all genders and vulnerable users.
Institutions	WatSan duty bearers are accountable and transparent, levels of trust are high, and there is an enabling environment for service provision.
Operations	WatSan systems and services have an asset management scheme in place to enable easy and continuous access that is affordable and financially viable.
Well-Being	People are satisfied with life, they are healthy and unburdened by waterborne diseases, and their children are attending school.
The Environment	Communities have plans in place for water safety, waste management, water conservation, etc., and people are aware of and participate in these systems.

*Adapted from the Household Water Insecurity Experiences (HWISE) Scale by applying a threshold and averaging 12 items, as indicated in Appendix 1 (HWISE, 2020; Young et al., 2019).

The goal is to capture all key aspects that are known or suspected to contribute to the sustainability of water and sanitation systems and services. Having a single aggregate number (the index score) gives us a general sense of the water and sanitation situation of a population. At the same time, we can look at the different dimensions and indicators to understand possible stronger and weaker points. In terms of impact, comparing the scores for intervention and comparison groups gives us a more holistic measurement of our successes and failures.

The index is designed to be adaptable to the given context. This is a key aspect of Oxfam GB's approach to hard-to-measure concepts (e.g., Women's Empowerment, Resilience, Sustainable Water and Sanitation), considering these as multi-dimensional and context specific. The six dimensions remain constant, but the set of indicators used to measure each dimension can vary, although indicators should not be removed simply because the project did not work on that aspect. The adapted version of the index for this evaluation shown below was developed together with Oxfam staff and the consultant. It is constructed from 14 indicators that are assessed with 85 questions. In this evaluation, the questions were asked through a survey to one individual per household; most indicators measure individual-level information, although some measure household-level information (e.g., under Water Security and Operations).

Figure 5.1: Sustainable Water and Sanitation Index dimensions (top row) and indicators



Each indicator is measured using one or more questions, where each question has a pre-defined threshold of acceptability, adapted from the approach of Alkire and Foster (2011). A positive result (above the threshold) receives a score of '1', while a negative result (below the threshold) receives a score of '0'. Each indicator is calculated as the average of the scores for the individual questions and reported as a percentage on a scale from 0 to 100. Similarly, each dimension is calculated as the average of its indicators, and the overall index is the average of the six dimensions, again on a scale of 0 to 100. A detailed table of questions and thresholds by indicator, and whether each indicator is directly related to the project's theory of change, is provided in Appendix 1.

We give equal weight to each dimension in the index; we have no theoretical justification to claim one is more important than another. Within each dimension, we also give equal weight to its indicators for the same reason. Therefore, questions are not weighted equally within the index; some indicators are measured with one question while others are measured by averaging 10 or more. Take the Equity dimension as an example with two indicators. Gender Inclusive is the average value across five variables (equity1, equity4, equity5, equity6, equity7) and Inclusive of Vulnerable Users is the average value across four variables (equity2, equity3, feeoptynwat, feeoptynsan). Equity equals Gender Inclusive plus Inclusive of Vulnerable Users, divided by two.

HOUSEHOLD WATER AND SANITATION ACCESS

Following the definitions used by the Joint Monitoring Programme (JMP), access levels for water reflect on whether households report that their main source is unimproved, improved, or piped water on premises. Access levels for sanitation consider whether households report that their main sanitation facility is unimproved or improved. Definitions for both water and sanitation are below for reference (WHO/UNICEF, 2017, p. 9).

Water Source Categories

Piped on premises – water tap in the house or on the plot/yard

Improved – public tap, standpipe or kiosk, tube well or borehole, protected dug well, protected spring, rainwater collection

Unimproved – unprotected dug well or spring, cart with a small tank/drum, tanker truck, surface water (river, dam, lake, pond, stream, canal, irrigation channel), bottled/sachet water

Sanitation Facility Categories

Improved – pit toilet/latrine that is closed/covered, flush toilet/latrine

Unimproved – no facility (open defecation), pit toilet/latrine that is open/uncovered

HOUSEHOLD WATER INSECURITY (HWISE SCALE)

The Household Water Insecurity Experiences (HWISE) Scale measures project impact on water insecurity. The scale is a standardized set of 12 questions that has been tested for cross-cultural reliability and validity in low- and middle-income countries. For each question in the scale, a household receives a score based on their response (i.e., from 0 for 'Never [0 times]' to 3 for 'Often [more than 10 times]'). The overall HWISE score, ranging from 0 to 36, is calculated as the sum of the values for all 12 questions. Higher scores indicate greater insecurity (HWISE, 2020; Young et al., 2019).

EMPOWERMENT IN WASH (EWI AND EMPOWERMENT SCORES)

The Empowerment in WASH Index (EWI) is survey-based tool used to understand agency, participation and empowerment in relation to WASH for both women and men (Dickin & Bisung, 2019). Like Oxfam's Women's Empowerment Index (Lombardi et al., 2017), the EWI considers three levels – personal, relational, environmental (adapted from 'individual, household, community to avoid confusion with levels of analysis in other sections of this evaluation). What is different is that the EWI assesses empowerment for both women and men, specifically in relation to WASH.

The EWI indicators for each level are shown in Table 5.2 below. Each indicator is calculated using information from one or more survey questions and relevant thresholds. Further details, including the questions, thresholds and modifications we made to the index for this evaluation, are provided in Table A1.7 at the end of Appendix 1.

Table 5.2: Empowerment in WASH Index (EWI) levels and indicators

Level	Indicator
Personal	Intrinsic attitudes about WASH roles and responsibilities
Relational	Input into decisions about WASH roles and responsibilities
	Input into decisions about WASH expenditures
	Input into decisions about involvement in community WASH activities
	Control over household assets
	Work balance
	Time for water collection
	Access to and sharing of WASH practices information
	Access to and sharing of information on WASH rights and responsibilities
Environmental	Group membership
	Leadership in WASH implementation
	Leadership in WASH accountability

PERSONAL STORIES (INSPIRED BY SENSEMAKER®)

SenseMaker® is a narrative method of enquiry aimed at capturing many short narratives and qualifying those narratives through interactive follow-up questions in an integrated fashion (Cognitive Edge, 2020; Mager et al., 2018). In this evaluation, household survey interviewees were invited to share a short personal story, based on a brief prompt, about something that happened recently in their life (in the last few years) in which water and/or sanitation played a significant role. Sharing a story was optional.

Those who did opt to share a story were also asked a series of follow-up questions to self-signify what it is about. These questions included rating how positive or negative the story was (sliding scale), whether it was more about water or sanitation (sliding scale), which Sustainable Water and Sanitation Index dimension related most to the story's topic (based on descriptions provided by the interviewer), how often it happened, and whether they think it was an experience common among others or unique only to them. A random selection of example stories is also shared under each Sustainable Water and Sanitation Index dimension. Following the approach of Lombardini & Mager (2019), some self-signification questions are also used for quantitative impact analysis of the qualitative data using PSM, as applicable.

CONFIGURATIONAL ANALYSIS (USING EVALC3)

EvalC3 is an Excel application that uses simple machine-learning algorithms to find good performing predictive models and identify the consequences of small changes (Davies, 2020). In this evaluation, configurational analysis is done using community-level information gathered through the KIIs. The input for this analysis is built from a list of intervention, community and water and sanitation management team attributes that were suspected to contribute to success for outcomes, including the Sustainable Water and Sanitation Index, HWISE Scale and EWI.

The analysis shows which attributes and combinations of attributes (also known as configurations) are good predictors of the outcomes of interest. An attribute is Sufficient if the outcome is always present when that attribute is present. An attribute is Necessary if the outcome is only present when that attribute is present. Accuracy is an overall performance measure for each attribute or combination of attributes that represents the proportion of True Positive cases (i.e., attribute and outcome both present) and True Negative cases (i.e., attribute and outcome both absent), of all cases (Davies, 2020).

6 RESULTS

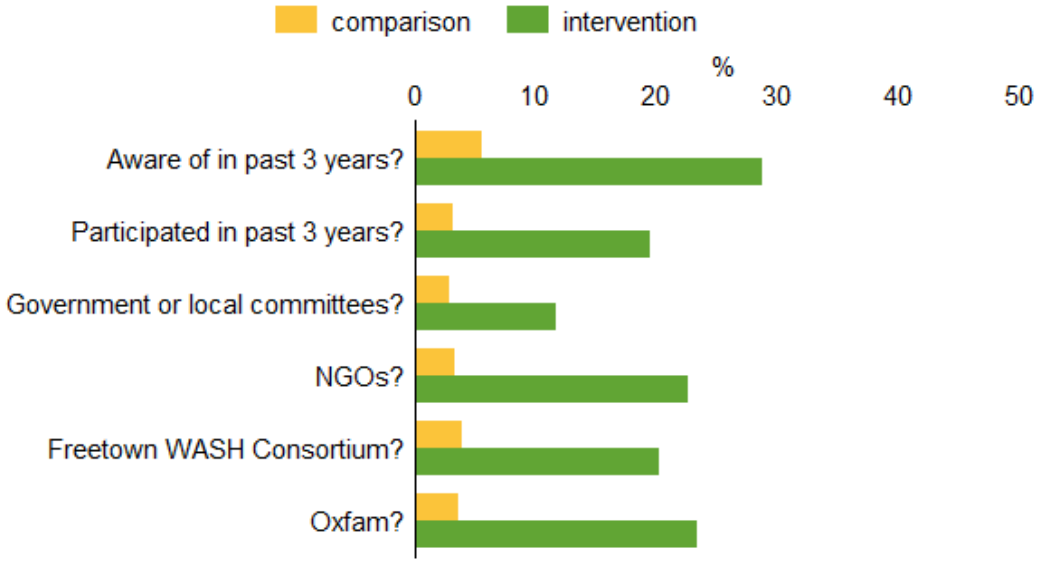
Here we present the combined results from the household surveys and KIIs. As described in Section 4, these results are based on a final dataset from 981 household surveys after doing filtering and propensity score matching (PSM), 347 individual stories (gathered during household surveys) and 15 KIIs with five community leaders and 10 water point attendants. Throughout this section, significant impacts (from quantitative analysis) are highlighted in green if positive and red if negative. Other significant differences are highlighted in blue and insignificant results are not highlighted. Direct quotes from interviews are shown in italics.

6.1 PROJECT EXPOSURE

In this section, we look at exposure to project (or similar) activities, as reported through the household surveys, for both the intervention and comparison groups. We asked about both awareness and participation related to water and sanitation activities, such as projects, meetings, trainings and events in the last three years (which would include the period of project implementation) and in the last six months (after project implementation had stopped). For those who did participate in activities, we asked follow-up questions to better understand project exposure – which institutions organized these activities and did people in the comparison group also participate in activities with these institutions?

Overall, we see much higher levels of awareness and participation in water and sanitation activities in the intervention group, and these are greatest in areas where boreholes were implemented. At the same time, we see a small proportion of people saying yes to these questions in the comparison group, which may relate to project activities that were carried out across the entire district (e.g., BCC, DRR, advocacy). However, overall there is a clear distinction in the level of project exposure between the two groups.

Figure 6.1.1: Exposure to water and sanitation activities in past three years (descriptive only; all differences statistically significant, p<0.001)

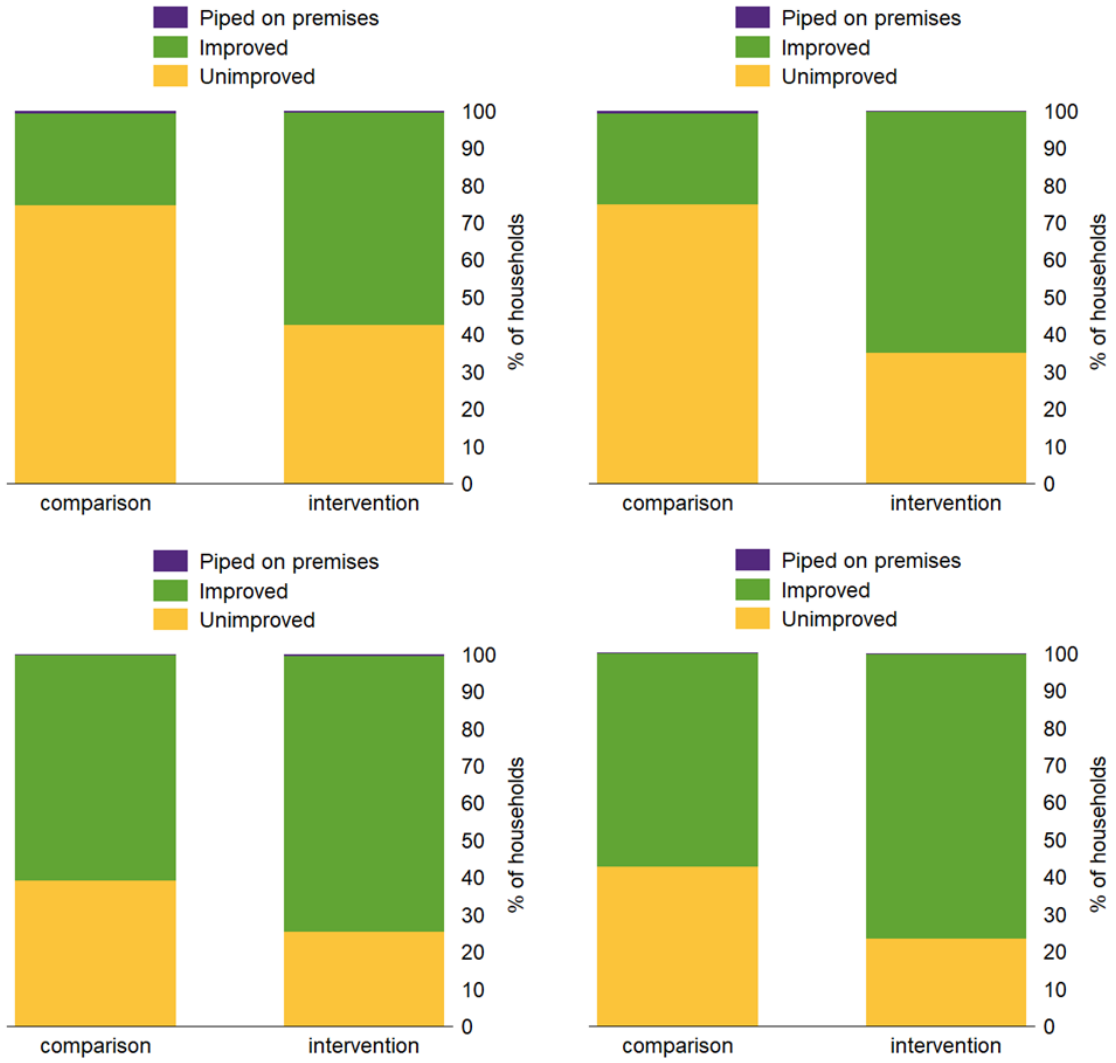


6.2 WATER AND SANITATION ACCESS

To contextualize, we start with a descriptive overview of the water and sanitation situation in the intervention and comparison areas. Access-level definitions are provided in Section 5. Figure 6.2.1 shows the proportion of households using each category of water source for drinking during the dry and wet seasons in 2016 and 2019. In all cases, more were using improved sources during the wet season than in the dry season, likely due to rainwater availability.

In the intervention area, improved access (including piped on premises) went from 58% to 65% during the dry season, from 2016 to 2019, and 75% to 77% during the wet season. This translates to a 12% increase during the dry season and a 2% increase during the wet season between 2016 and 2019. In the comparison area, figures decreased slightly, from 26% to 25% in the dry season and 61% to 57% during the wet season.

Figure 6.2.1: Main drinking water source during the dry season (top row) and wet season (bottom row) in 2016 (left) and 2019 (right) (descriptive only)

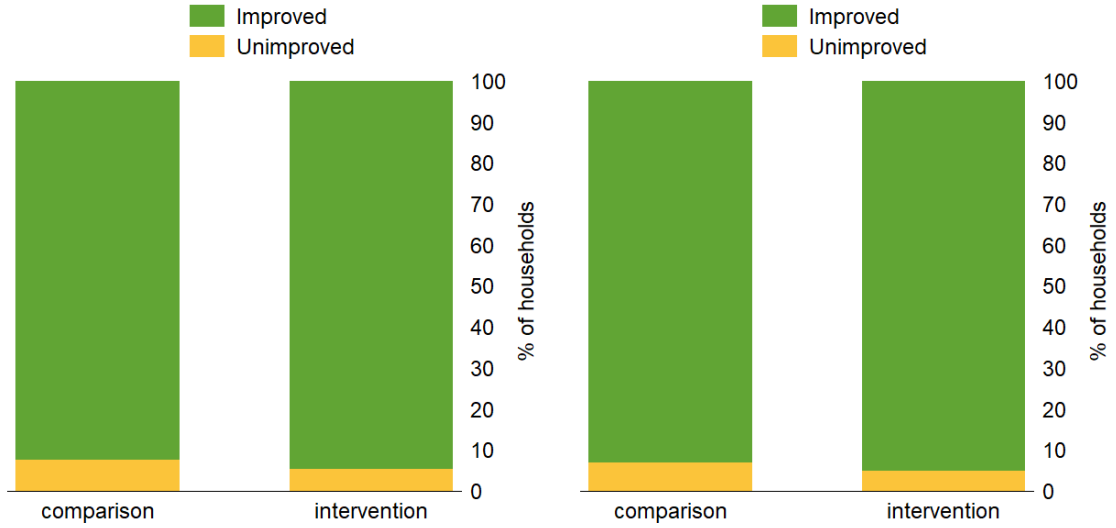


Beyond the categories for drinking water, households listed all sources they use for all purposes. In the intervention area, the most common unimproved source was bottle or sachet water (78%) and the most common improved sources were protected dug wells (70%), rainwater collection (56%), tube wells or boreholes (52%) and public taps, standpipes or kiosks (49%). In the comparison area the most common unimproved sources were bottle or sachet water (82%) and unprotected dug

wells (46%) and the most common improved sources were protected dug wells (56%) and rainwater collection (53%).

Next, Figure 6.2.2 shows the proportion of households by sanitation facility category. Like those for water access, these graphs show different levels of improved sanitation access in the intervention and comparison areas in 2016 and 2019. In the intervention area, improved access remained constant at 95%, while in the comparison area it increased slightly from 92% to 93%. In both areas, most households were using closed/covered pit toilet/latrine facilities (58% on average) or flush toilet/latrine facilities (36% on average).

Figure 6.2.2: Main sanitation facility in 2016 (left) and 2019 (right) (descriptive only)



Now, we estimate the impact of the project on water and sanitation access, shown in Table 6.2.1. These PSM estimates account for differences in baseline characteristics, and particularly differences in access to improved water, which were very different between the two groups in 2016, as explained in Section 4.4. We see the project has led to a significant 20% increase in the proportion of households having access to improved water during the dry season. However, we do not estimate any significant impact on the proportion of households having access to improved water during the wet season, piped water or improved sanitation.

Table 6.2.1: Impact of the project on water and sanitation access

	Improved water⁺ dry season (%)	Improved water⁺ wet season (%)	Piped water dry season (%)	Piped water wet season (%)	Improved sanitation (%)
Intervention group mean	65.12	76.59	0.24	0.24	95.12
Comparison group mean	54.37	74.06	2.19	0.68	95.35
Difference (Impact)	10.76***	2.53	-1.95	-0.44	-0.23
Standard error	(3.48)	(3.40)	(1.25)	(0.75)	(1.22)
Observations (intervention)	410	410	410	410	410
Observations (total)	981	981	981	981	981

⁺Improved water includes piped water on premises. * p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions.

6.3 HOUSEHOLD WATER INSECURITY

Next, we look at household water insecurity using the Household Water Insecurity Experiences (HWISE) Scale as described in Section 5. Table 6.3.1 shows the PSM estimates for the HWISE Scale. We see that the project has reduced water insecurity by 54% in the intervention group. Furthermore, using the recommended HWISE score threshold of 12, we also find the proportion of water insecure households has decreased by 62%. Both impacts are statistically significant.

Table 6.3.1: Impact of the project on water insecurity

	HWISE Scale (0–36)	Water Insecure Households (%)
Intervention group mean	6.82	24.88
Comparison group mean	14.83	66.13
Difference (Impact)	-8.01***	-41.26***
Standard error	(1.20)	(5.23)
Observations (intervention)	410	410
Observations (total)	981	981

* p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions.

The EvalC3 configurational analysis shows us which attributes are good predictors of water insecurity. Table 6.3.2 provides the list of attributes, noting whether each is Necessary and/or Sufficient and its level of Accuracy in predicting water secure communities (average HWISE score in the community below the recommended threshold of 12). Eight single attributes are Sufficient, highlighted in purple, but none are Necessary.

We always see low water insecurity in communities with any of the following: the borehole intervention, the public toilet intervention, women members, women leaders, a bank account, a WASH Management Team (WMT) that was already established before the project, a Water Safety plan in place, plans clearly consider equity. In terms of accuracy, there are four single attributes that stand out (accuracy of 70% or more) – collecting fees at 86%, the borehole intervention at 82%, women members at 77% and a Water Safety plan in place at 73%.

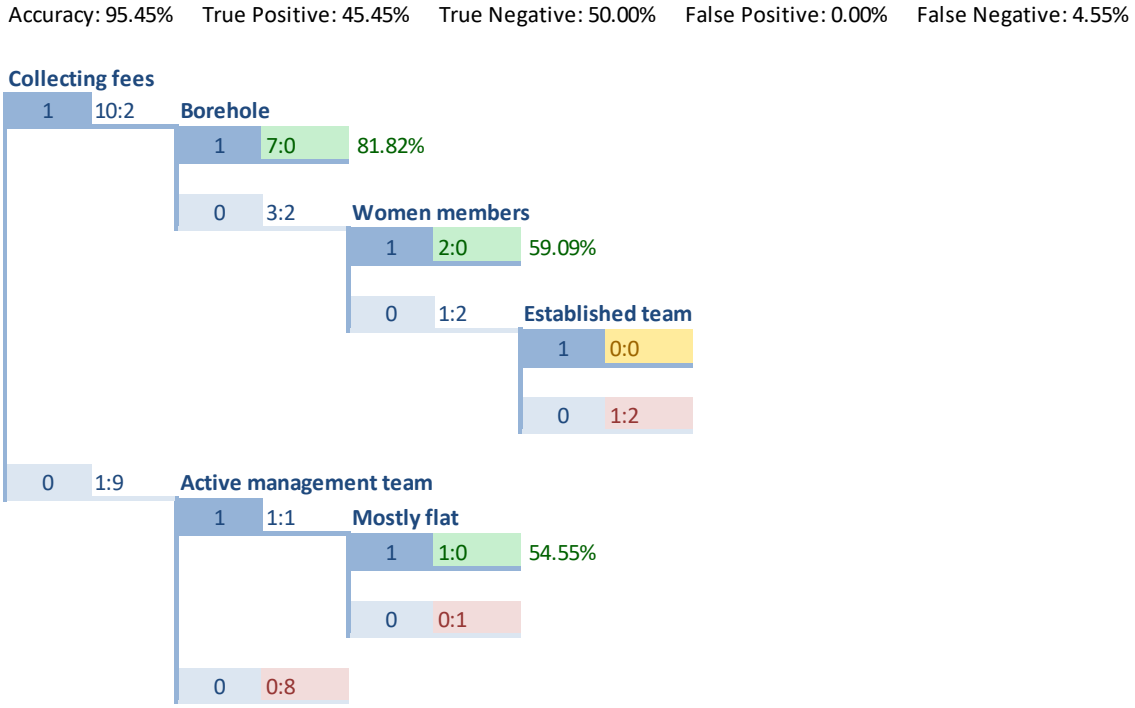
Table 6.3.2: Single attributes predicting a community is water secure (on average)

#	Attribute	Sufficient	Necessary	Accuracy
1	Borehole	Yes	No	82%
2	Kiosk	No	No	59%
3	Public Toilet	Yes	No	59%
4	Paved road	No	No	45%
5	Mostly flat	No	No	55%
6	Active WMT	No	No	68%
7	Women members in WMT	Yes	No	77%
8	Women leaders in WMT	Yes	No	68%
9	Collecting fees	No	No	86%
10	Bank account	Yes	No	59%
11	Established WMT	Yes	No	59%
12	Water Safety plan	Yes	No	73%
13	Equity plans	Yes	No	64%

We also assessed various configurations of attributes to maximize Accuracy. Figure 6.3.1 shows the optimal decision tree with six attributes, resulting in 95% Accuracy. From this decision tree, the most important attributes for lower water insecurity within a community are fee collection followed by the borehole intervention. In the absence of fee collection, the most important attributes are

having an active WMT and mostly flat terrain. In the absence of the borehole intervention, the next most important attribute is having women members. In the absence of women members, the most important factor is that the WMT was already established before the project.

Figure 6.3.1: EvalC3 decision tree with the six attributes that maximize Accuracy for the outcome of water insecurity (i.e., average community-level HWISE score below 12)



6.4 SUSTAINABLE WATER AND SANITATION

In this section, we look at the impact of the project on the Sustainable Water and Sanitation Index, as described in Section 5. Table 6.4.1 shows the PSM estimates for the overall index. The results indicate a significant positive impact, with an overall increase of 24% due to the intervention. Further details on each dimension are provided in Table 6.4.2. Each dimension is then discussed further, focusing on significant differences observed. For a visual overview of the index and impact measurements for each dimension and indicator, see Figure 6.4.1.

Table 6.4.1: Impact of the project on Sustainable Water and Sanitation

	Sustainable Water and Sanitation Index
Intervention group mean	70.08
Comparison group mean	56.40
Difference (Impact)	13.69***
Standard error	(1.64)
Observations (intervention)	410
Observations (total)	981

* p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions.

Table 6.4.2: Impact of the project on Sustainable Water and Sanitation dimensions

	Water Security	Equity	Institutions	Operations	Well-Being	The Environment
Intervention group mean	59.25	77.34	55.89	68.30	83.62	76.10
Comparison group mean	27.02	62.59	46.18	59.64	80.93	62.03
Difference (Impact)	32.23***	14.76***	9.71***	8.66***	2.69	14.07***
Standard error	(4.32)	(3.31)	(2.89)	(1.21)	(1.66)	(3.63)
Observations (intervention)	410	410	410	410	410	410
Observations (total)	981	981	981	981	981	981

* p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions.

The overall impact for Water Security is significantly positive (32.23, p<0.01). This result aligns with the findings on water insecurity in Section 6.3 (both are based on HWISE). It appears that the largest improvements relate to people not having to interrupt their plans or change what foods they eat because of issues with water.

For Equity, the overall impact is significantly positive (14.76, p<0.01) as well. Both indicators in this dimension show positive impacts, including Gender Inclusive (15.73, p<0.01) and Inclusive of Vulnerable Users (13.78, p<0.01). The data suggest that the largest differences come from interviewees' opinions on needs of all genders, people with disabilities and/or chronic illnesses, and children being met, as well as the role of feedback and input from both women and men in local water and sanitation management.

Similarly, for Institutions, we see a significant positive impact overall (9.71, p<0.01) and for all three indicators – Accountability (11.28, p<0.01), Transparency (10.33, p<0.05) and Trust (8.09, p<0.05). These results are driven by differences in overall satisfaction with water and sanitation systems and services, access to information about the management of water and sanitation services, and increased trust in institutions, including local development committees, community leaders, NGOs, and the Government of Sierra Leone.

For Operations, again we see that overall the impact is significantly positive (8.66, p<0.01). Each of the indicators is significantly positive as well, including Asset Management (15.39, p<0.01), Financial Viability (5.70, p<0.01) and Accessibility (4.89, p<0.01). The data suggest this impact is due to the presence of water and sanitation committees, water and sanitation pricing that is perceived to be fair, the use of official water sources, and an increased proportion of interviewees feeling safe accessing their water source at night.

In the Well-Being dimension, we do not find a significant impact overall. Among the three indicators, only Quality of Life shows a significant positive impact (5.66, p<0.1), mainly due to increased overall life satisfaction, while there is no significant impact for indicators of Health or Education.

For The Environment, the overall impact is positive and significant (14.07, p<0.01) along with that of both indicators, Water Safety (17.58, p<0.01) and Waste Management (10.56, p<0.01). It appears that this impact is due to increased water safety plans and quality monitoring as well improved drainage (to reduce the risk of flooding), sanitation systems and solid waste management planning.

Figure 6.4.1: Sustainable Water and Sanitation Index dimensions (top row) and indicators showing project impact (PSM estimates; * p <0.1, ** p <0.05, *** p <0.01)

The EvalC3 configurational analysis shows us which attributes are good predictors of Sustainable Water and Sanitation. Table 6.4.3 provides the list of attributes, noting whether each is Necessary and/or Sufficient and its level of accuracy in predicting Sustainable Water and Sanitation within communities (average community-level Sustainable Water and Sanitation Index score above the mean score of 62.47). Four single attributes are Sufficient and one is Necessary, highlighted in purple.

We always see above average Sustainable Water and Sanitation in communities with any of the following: the public toilet intervention, women leaders, a bank account, a WASH Management Team (WMT) that was already established before the project. We only see above average Sustainable Water and Sanitation in communities where the WMT is collecting fees. In terms of accuracy, there are five single attributes that stand out (accuracy of 70% or more) – collecting fees at 86%, the borehole intervention at 82%, women members at 77%, women leaders at 77% and a Water Safety plan in place at 73%.

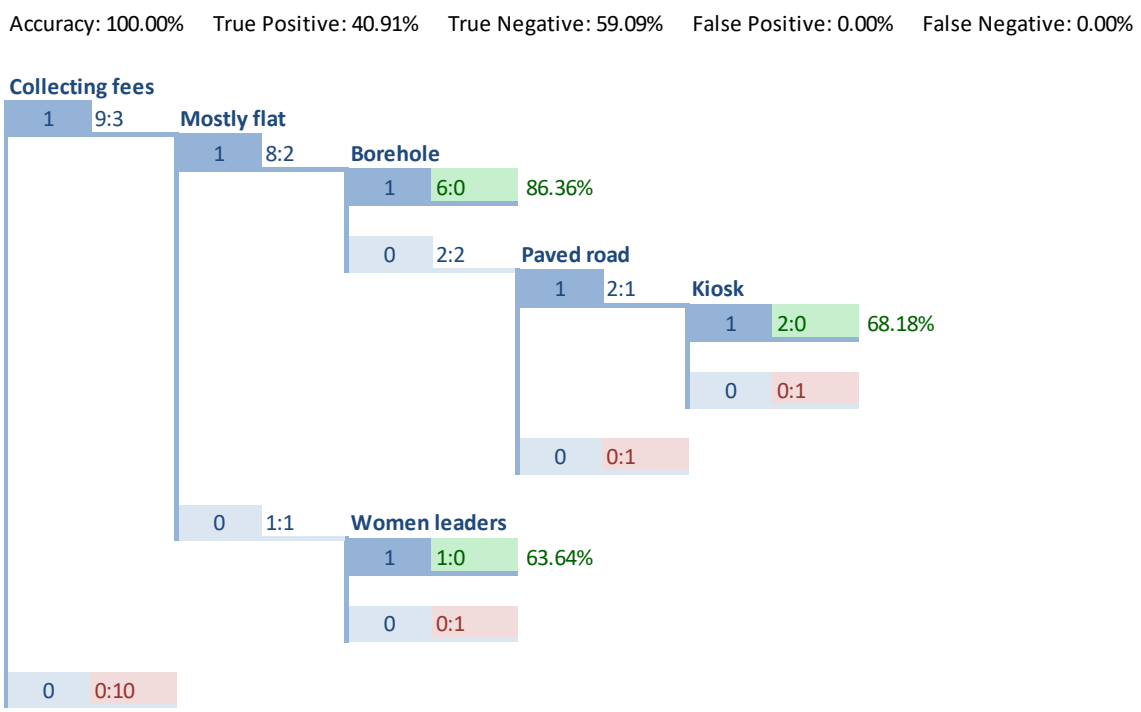
Table 6.4.3: Single attributes predicting Sustainable Water and Sanitation within a community (on average)

#	Attribute	Sufficient	Necessary	Accuracy
1	Borehole	No	No	82%
2	Kiosk	No	No	68%

3	Public Toilet	Yes	No	68%
4	Paved road	No	No	55%
5	Mostly flat	No	No	55%
6	Active WMT	No	No	59%
7	Women members in WMT	No	No	77%
8	Women leaders in WMT	Yes	No	77%
9	Collecting fees	No	Yes	86%
10	Bank account	Yes	No	68%
11	Established WMT	Yes	No	68%
12	Water Safety plan	No	No	73%
13	Equity plans	No	No	64%

We also assessed various combinations of attributes to maximize Accuracy. Figure 6.4.2 shows the optimal decision tree with six attributes, resulting in 100% Accuracy. From this decision tree, the most important attributes for above average Sustainable Water and Sanitation within a community are fee collection followed by mostly flat terrain and the borehole intervention. In the absence of mostly flat terrain, the next most important attribute is having women leaders. In the absence of the borehole intervention, the next most important attributes are being along a paved road and the kiosk intervention.

Figure 6.4.2: EvalC3 decision tree with the six attributes that maximize Accuracy for the outcome of Sustainable Water and Sanitation (above average community-level score)



Through personal stories (inspired by SenseMaker®), interviewees could choose to share further detail through stories about personal experiences with water and sanitation. They were asked, 'Can you tell me about something important that happened in your life, where water and sanitation played a role?' Overall 35% (n=347) of interviewees shared a story. Slightly more women shared stories (38%) than men (32%), but otherwise storytellers were not systematically different from the overall population interviewed in terms of group, wealth quintile, or age. Because more women shared stories and we interviewed approximately 50% more people in the comparison group, we recognize that the analysis of the stories is biased in this way.

Based on their own qualitative coding of the stories, we quantify some information while also sharing direct quotes from a random selection of stories by topic. As shown in Figure 6.4.3, the stories were more about water than sanitation and were more negative than positive, on average. The focus on water versus sanitation did not vary between stories in the intervention and comparison groups. We do see significantly more positive stories in the intervention group (23%) than in the comparison group (10%), based on PSM estimates ($p<0.01$).

Figure 6.4.3: How interviewees categorized their stories (descriptive only)

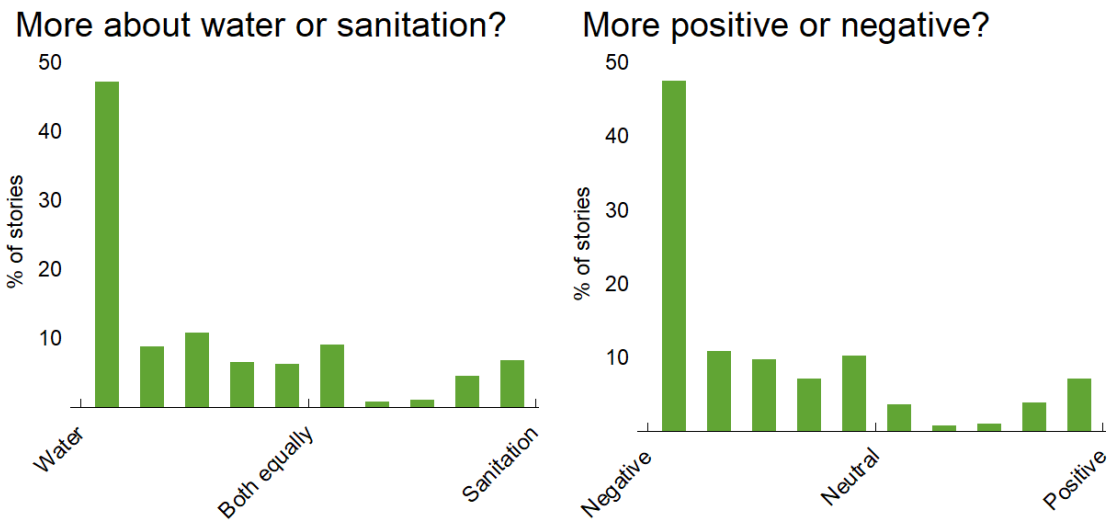
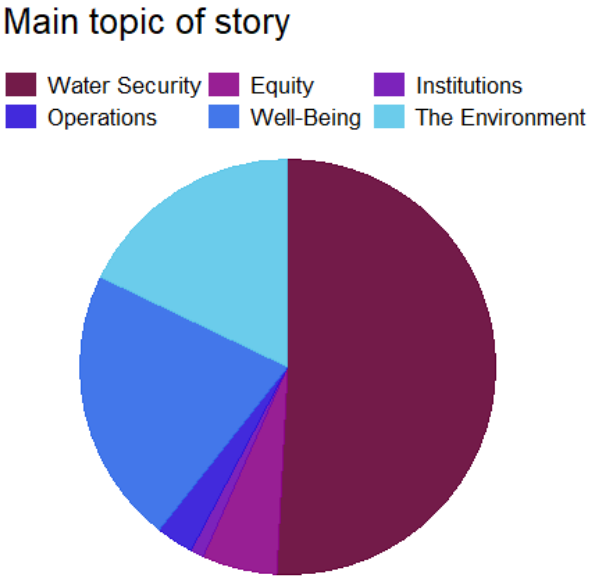


Figure 6.4.4: How interviewees categorized their stories (descriptive only)



Interviewees selected the main topic of their story from the six dimensions of the Sustainable Water and Sanitation Index (interviewers provided translated definitions for each). They could additionally give their story a title and provide their own topic (both open-ended). The distribution of stories across the dimensions is shown in Figure 6.4.4. This proportion of stories under each topic did not vary significantly by gender for the most part, except more men than women said their stories were about The Environment.

Below we share a selection of 27 stories with titles, each of which have been chosen randomly within each dimension, stratified by household wealth quintile. Two initially selected stories were replaced for being unrelated to water or sanitation (one about Equity and one about Well-Being; reselection was done randomly). In each case, for context, we provide the storyteller's age, wealth quintile of their household and whether they were from an intervention or comparison community. Direct quotes of the stories are shown in italics; stories have been minimally edited for grammar and clarity. ***Please be aware that some of the stories may be distressing to read, particularly those that talk about instances of gender-based violence (GBV).***

For Water Security (51% of stories), by far the most common topic, the random selection of stories highlights many key aspects of the definition shared in Table 5.1 – 'People have consistent and sufficient access to preferable acceptable water; water is not a source of worry, shame, or anger.' Most evident, is that two of these five randomly selected stories are about GBV in relation to water insecurity. Women in particular face sexual harassment and, in the worst cases, rape, while bearing the burden of water collection for their households. It is promising to hear in the Dangers of water story that GBV has reduced in at least one intervention community, but further work still needs to be done.

Box 1: Stories about Water Security (randomly selected)

Water greedy: At one time I went to fetch water in a neighbour's compound, unfortunately for me I was not given access. Immediately they saw, and they went and closed the gate. I felt so bad on that day.

27-year-old man in the lowest wealth quintile, from a comparison community

Dangers of water: Before most young girls were getting pregnant just in the name of fetching water. Most men were taking advantage of these young girls. It has now reduced due to the water kiosk we are having here.

36-year-old man in the second wealth quintile, from an intervention community

Water confusion: Presently I am in court with someone concerning a water issue, we are fighting seriously.

45-year-old man in the third wealth quintile, from an intervention community

Sexual harassment and water: The boys around sometimes take advantage of girls going in search of water. My girl child has made that complaint to me on several occasions.

46-year-old man in the fourth wealth quintile, from a comparison community

No title: Back in 2016 we were ashamed in my house during my wedding because water was not enough in the house to be used by guests. This embarrassed me so much that I decided to buy a tank and dig a well with submersible pump to supply water in our home also.

33-year-old woman in the highest wealth quintile, from an intervention community

For Equity (6% of stories), the stories from those in lower wealth quintiles talk about struggling to access water because of lack of money to pay, as well as arguing over water. Those in higher wealth quintiles shared stories about building their own toilet to have more privacy, when before they had been sharing a toilet with their neighbours. Being unable to access water without paying can be an issue of Equity, especially for poorer households, and is why it is important to have a payment structure that still allows affordable access for the most vulnerable users. As for building private toilets, in the index this would normally fall under the Operations dimension, but those telling these stories clearly felt there is an element of Equity to it as well. By having their own sanitation facility, it is more likely that all household members can access it as needed.

Box 2: Stories about Equity (randomly selected)

The struggle for water: Most times our neighbours are greedy and won't give us water even if we have money to pay. They will allow people to fetch water at their own well.

26-year-old man in the lowest wealth quintile, from a comparison community

Argued for water: There was a time when I got into an argument at the queue with another community member.

21-year-old man in the second wealth quintile, from a comparison community

Greediness in water: One day I was without money, so I went to fetch water with the hope that the owner would pity me. But unfortunately, I was not allowed to fetch the water because I couldn't afford to pay. So, for that I went without water.

33-year-old woman in the third wealth quintile, from a comparison community

Owning a toilet: At the initial stage, my family and I were using our neighbour's toilet whilst I was trying to construct mine. However, I sped up my own construction because there was no privacy in my neighbour's toilet.

39-year-old woman in the fourth wealth quintile, from a comparison community

Insufficient toilet: We were sharing one toilet with five households, but now we have constructed our own toilet. There was no privacy before when using the toilet.

29-year-old man in the highest wealth quintile, from an intervention community

For Institutions (1% of stories), each story is quite different. No one in the first and second wealth quintiles said their story was about institutions, so we share only three stories. In the first story, again we see the link between GBV and water with the link to the police as the institution. The second story is about neighbourhood theft of water, which could also be related to a lack of security. The third story is about a former water committee that no longer exists in a comparison community and the storyteller is not sure why it was dissolved. While different, these stories have all been categorized under the same topic by those who shared them and highlight a broader issue with trust and insecurity, as well as the dissolution of a water committee.

Box 3: Stories about Institutions (randomly selected)

Lack of security: Last year, a girl at the age of 17 years was raped by unknown men whilst trying to fetch water in the middle of the night. The matter was reported to the police, but the perpetrators were nowhere to be found.

43-year-old woman in the third wealth quintile, from an intervention community

Neighbourhood theft: There was a time when my father had an encounter with our neighbour because she steals water from our bucket.

40-year-old man in the fourth wealth quintile, from a comparison community

Dissolved committee: There used to be a water committee, but it has been dissolved, reason being I don't know.

23-year-old man in the highest wealth quintile, from a comparison community

For Operations (3% of stories), the stories revolve around issues of safe access to water and sanitation. No one in the third wealth quintile said their story was about operations, so we share only four stories about long queues at water sources and lack of privacy at sanitation facilities. As we see in the stories, these can have unacceptable consequences – a child being late for their school exam, a pregnant woman standing under the sun for an hour. In the last story, a father shares that his daughter became pregnant due to issues related to water access, which forces her to gather water for her household late in the evening when it is unsafe to do so.

Box 4: Stories about Operations (randomly selected)

My child was late for school: I went to the tank to get water. There were so many people and it was in the morning hours. I had to prepare my child to go to school and she was on exam. Because of the many people at the tank, my child was late to go to school and was late for the exam.

28-year-old woman in the lowest wealth quintile, from an intervention community

Poor sanitation: There is little privacy when using our toilet. I remember one day I was at the toilet easing myself, and all of a sudden someone came and knocked on the door. I guess he was so pressed. So, I had to get out of the toilet without finishing just to give him a chance.

25-year-old woman in the second wealth quintile, from a comparison community

Effect of water: Last year March when I was pregnant, I went to fetch water, but I met a lot of people there, so I tried to beg them to give me a chance to get water, but they totally refused. I was under the sun for about an hour it was so disgusting as a pregnant woman.

30-year-old woman in the fourth wealth quintile, from an intervention community

Effect of water: My first daughter got pregnant while she was going out to look for water. Because of the water crisis, they go out by 7 in the evening and they may normally return by 9 to 10 at night. It is not really easy for us.

40-year-old man in the highest wealth quintile, from a comparison community

For Well-Being (22% of stories), the stories are about health issues, including snakebites and increased water and food insecurity during the Ebola outbreak. Well-Being is also linked to stress, which can be related to arguments over water, like those described in the fourth and fifth stories. In the extreme, in the third story a woman tells of her eldest daughter being tragically killed in an accident as she crossed the road while going to get water. These stories highlight just some of the ways that water and sanitation effect Well-Being, which go far beyond more commonly associated health problems like diarrhoeal disease.

Box 5: Stories about Well-Being (randomly selected)

Dangers of fetching water at night: One time a lady who was my neighbour was bitten by snake fetching water at night. She was very fortunate to survive it because people were quick to respond to her.

23-year-old woman in the lowest wealth quintile, from a comparison community

Ebola crises: During the Ebola outbreak, I had a very severe headache and it was hurting me so bad and making me feel unwell. I got sick and could not go out of my house because I was scared that if someone around me found out that was not well they might call the Ebola medical centre line 117 to come and take me and that might be the end of me because I was sure those were not Ebola symptoms. So, I had no choice but to stay indoors for 5 days without any food or water. It was a terrible moment for me and only God got me out of that situation.

20-year-old woman in the second wealth quintile, from an intervention community

Effect of water: Two years ago, I lost my daughter during March. She went to fetch water across the street on her way she got a car accident she was my first born.

35-year-old woman in the third wealth quintile, from an intervention community

No title: I fought with the chairlady because she has this habit of fetching water to her satisfaction before allowing others to fetch and also taking the advantage or claiming ownership because the water facility is close to her house.

31-year-old woman in the fourth wealth quintile, from an intervention community

Cleaning toilet problem: So many stories about water and toilets. Even yesterday I had to withdraw a case from the police station to settle it at home. Two of my tenants have been quarrelling over who scrubs the toilet.

47-year-old man in the highest wealth quintile, from a comparison community

For The Environment (18% of stories), the first three stories are about safety. Two different men share stories of violence perpetrated against their sisters. In the third story, another man tells about his daughter getting injured while getting water due to the rough terrain between their home and the source. The remaining stories are about issues between landlord and tenant related to water facilities and food insecurity because of water problems. The storytellers have interpreted this dimension differently than the original definition in Table 5.1, which is about water safety, management and conservation. Instead their stories focus on the general environment in terms of the situation in their communities.

Box 6: Stories about The Environment (randomly selected)

My sister was raped on her way from the well: One day at around 5 am my sister went to get water. On her way, she was attacked by some unknown men and she was raped.

38-year-old man in the lowest wealth quintile, from an intervention community

My sister was pregnant: One day I found out that my sister was pregnant, and when I asked her she said she was impregnated by one of the boys at night when collecting water at the tap. Because of that my sister had to drop out of school.

36-year-old man in the second wealth quintile, from a comparison community

The struggle for water: One day my daughter was going to fetch water when she knocked her feet along the way due to the bad road. She had some cuts on her feet which took some time to heal. Up until now she has that scar.

40-year-old man in the third wealth quintile, from a comparison community

Effect on water: I was given a notice just because I told my landlord the truth about the water facility in the compound.

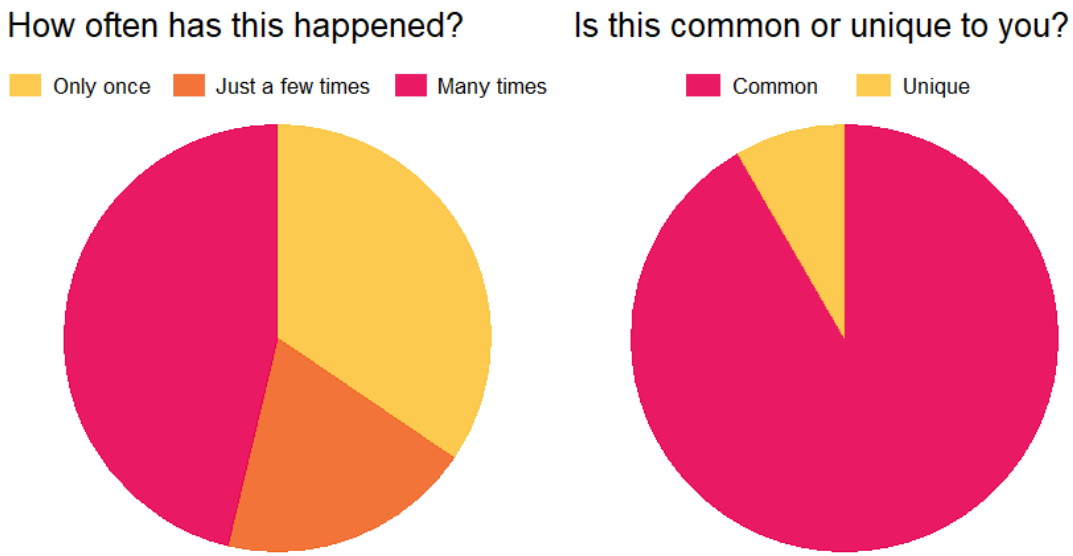
40-year-old man in the fourth wealth quintile, from a comparison community

Scarcity of water: It has been two days now that I cannot cook because of water. My family and I have been eating street foods which is very bad.

38-year-old woman in the highest wealth quintile, from a comparison community

The intensity of the stories people chose to share following a simple prompt to, 'tell me about something important that happened in your life, where water and sanitation played a role,' show how water and sanitation issues touch all aspects of their lives. Throughout the six dimensions, instances of GBV, including sexual harassment and rape, and teenage pregnancy stand out, and are found in more than 10% of stories in each case. The stories quoted above are only a random selection, a small representation of far more untold stories. When asked how often what they described in their stories had happened and whether it was something common among others, 46% said many times and 92% said it was common (see Figure 6.4.5).

Figure 6.4.5: How interviewees categorized their stories (descriptive only)



6.5 EMPOWERMENT IN WASH

In this section we review findings from an adapted version of the Empowerment in WASH Index (EWI) as described in Section 5. Overall, we see higher Empowerment in WASH for men than women, with most of this difference coming from indicators in the relational level as well as a small difference in the environmental level. In the relational level, which is mainly about intra-household dynamics, we note the largest differences are that women spend more time on water collection and have less input into decisions about involvement in community WASH activities. In the environmental level, we find that more women participate in groups but feel less comfortable speaking up about WASH implementation and accountability. Among those who are disempowered, we see fewer men and women not meeting the threshold for ownership of assets and input into household decision-making on WASH expenditures and on WASH roles and responsibilities. See Table 6.5.1 for summary descriptive statistics by gender. In Table 6.5.2, we do not find any significant impacts of the project on Empowerment in WASH.

Table 6.5.1: Empowerment in WASH (EWI) summary (descriptive only)

	Women (n=472)	Men (n=509)	Difference
% Empowered	51%	58%	-6.70** (3.18)
Level of achievement by disempowered+	55%	58%	-2.54** (1.03)
EWI score	0.78	0.82	N/A

+ Among those not empowered, the proportion of indicators for which they met the empowerment threshold, on average. * p <0.1, ** p <0.05, *** p <0.01; Standard errors are in parentheses.

Table 6.5.2: Impact of the project on Empowerment in WASH (individual scores)

	EWI score	Individual level	Household level	Community level	% Empowered
Intervention group mean	0.72	0.87	0.74	0.61	57.80
Comparison group mean	0.70	0.84	0.72	0.61	54.44
Difference (Impact)	0.01	0.03	0.02	-0.00	3.36
Standard error	(0.02)	(0.03)	(0.02)	(0.04)	(5.30)
Observations (intervention)	410	410	410	410	410
Observations (total)	981	981	981	981	981

* p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions.

The EvalC3 configurational analysis shows us which attributes are good predictors of Empowerment in WASH. Table 6.5.3 provides the list of attributes, noting whether each is Necessary and/or Sufficient and its level of accuracy in predicting Empowerment in WASH within communities (average community-level EWI score above the mean value of 0.80). Only one single attribute is Sufficient, highlighted in purple, while none are Necessary.

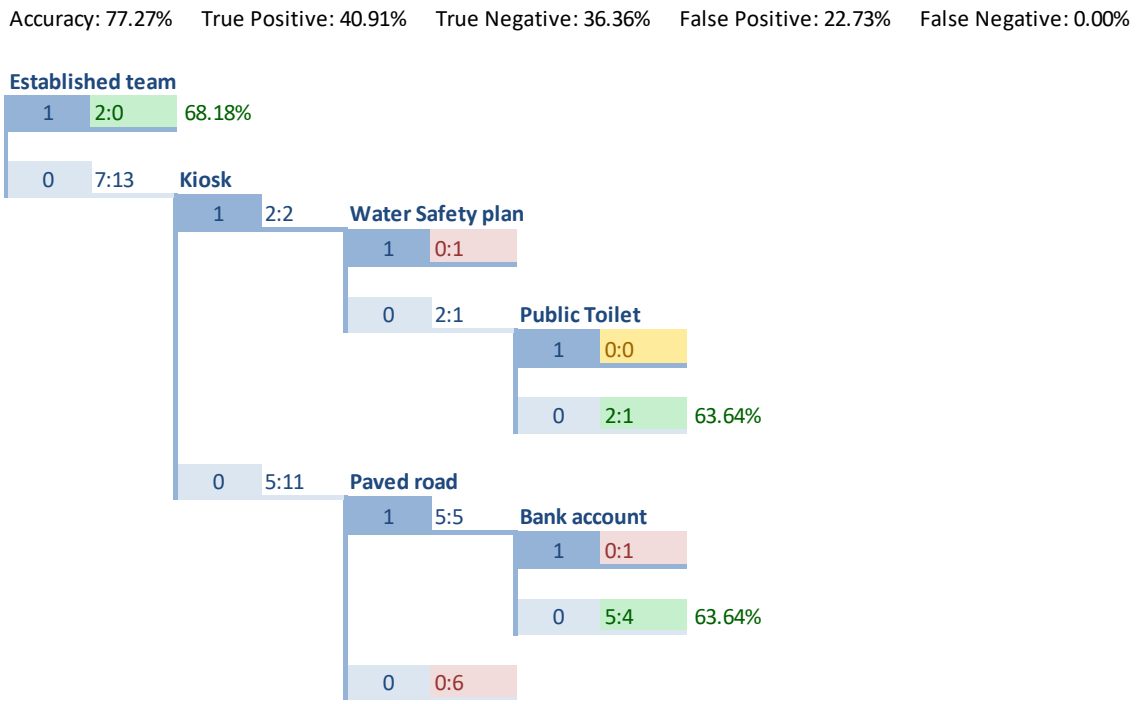
We always see above average Empowerment in WASH (EWI) in communities with a WASH Management Team (WMT) that was already established before the project. In terms of accuracy, there are no single attributes that stand out (accuracy of 70% or more).

Table 6.5.3: Single attributes predicting Empowerment in WASH within a community (on average)

#	Attribute	Sufficient	Necessary	Accuracy
1	Borehole	No	No	55%
2	Kiosk	No	No	59%
3	Public Toilet	No	No	59%
4	Paved road	No	No	64%
5	Mostly flat	No	No	55%
6	Active WMT	No	No	59%
7	Women members in WMT	No	No	59%
8	Women leaders in WMT	No	No	68%
9	Collecting fees	No	No	59%
10	Bank account	No	No	50%
11	Established WMT	Yes	No	68%
12	Water Safety plan	No	No	55%
13	Equity plans	No	No	55%

We also assessed various combinations of attributes to maximize Accuracy. Figure 6.5.1 shows the optimal decision tree with six attributes, resulting in 77% Accuracy. From this decision tree, the most important attribute for above average Empowerment in WASH within a community is having an established WMT. In the absence of that, the most important attributes are the kiosk intervention and having a Water Safety plan. In the absence of the kiosk intervention, the next most important attributes are being along a paved road and having a bank account. In the absence of a Water Safety plan, the next most important attribute is the public toilet intervention.

Figure 6.5.1: EvalC3 decision tree with the six attributes that maximize Accuracy for the outcome of Empowerment in WASH (EWI) (above average community-level score)



6.6 IMPACTS FOR DIFFERENT INTERVENTIONS AND SOCIAL GROUPS

We also look for differential impacts by subgroup to see who experienced the effects of the project more and who less. We use propensity score weighting, as described at the end of Appendix 3, to understand how impact varies by intervention type and for different social groups, based on two characteristics that shape social inequalities and vulnerability in the context of urban Freetown – gender and having a disability. Therefore, we consider relevant differential impacts for three subgroups – (1) by intervention type, (2) by interviewee gender and (3) whether households have any members with disabilities. Differential impacts, where available, are highlighted in blue.

First, overall significant positive impacts are found for both the kiosk and borehole intervention types in terms of improved water, water insecurity and sustainable water and sanitation. The main differential impact we see is for water insecurity, that the borehole intervention decreased water insecurity significantly more than the kiosk intervention. This is likely because the boreholes more easily provide a continuous, reliable flow of water, while the kiosks rely on water trucking. We see a similar differential impact for the Water Security dimension of the Sustainable Water and Sanitation Index (table not shown, available on request), but do not find other significantly different impacts between the borehole and kiosk interventions.

Next, we look for impacts of the public toilet rehabilitation intervention, which was implemented alongside the borehole intervention in two communities. We do find significant impacts of the public toilet intervention within the Sustainable Water and Sanitation Index including in the Equity dimension (6.59, $p<0.05$) and the Well-Being dimension (3.93, $p<0.05$) (table not shown, available on request). These results indicate that the public toilet intervention contributed to positive impacts, but we do not see significant impacts for improved sanitation or the Sustainable Water and Sanitation Index overall and the impacts of the public toilet intervention are not significantly different from the borehole intervention alone.

Table 6.6.1: Impact of the project by intervention type

	Improved water+ (%)	Improved sanitation (%)	HWISE Scale (0 to 36)	Sustainable Water and Sanitation Index
Impact of the borehole intervention relative to the comparison group	10.34*** (2.66)	N/A	-9.33*** (0.71)	14.04*** (1.19)
Impact of the kiosk intervention relative to the comparison group	12.87*** (2.91)	N/A	-5.92*** (0.84)	13.22*** (1.38)
Impact of the public toilet intervention relative to the comparison group	N/A	-1.34 (1.72)	N/A	2.39 (1.82)

+During the dry season; improved water includes piped water on premises. * p <0.1, ** p <0.05, *** p <0.01; Standard errors are in parentheses. Impacts of water interventions on sanitation outcomes were not estimated, and vice versa.

Next, we look at the results by gender – do women and men experience impacts differently? Recall from Table 6.1 that 48% of all interviewees are women. In Table 6.6.2, we do not see any significant differential impacts by gender. However, among the six dimensions of the Sustainable Water and Sanitation Index, we find a significant differential impact for Well-Being (-3.84, $p<0.1$) – there is a significant impact for men (4.51, $p<0.01$) but not for women or overall (table not shown, available on request).

Table 6.6.2: Impact of the project by interviewee gender

	Sustainable Water and Sanitation Index	% Empowered
Effect of being a woman in the comparison group (compared to being a man)	-1.53 (1.62)	-0.50 (5.30)
Project impact for men	14.94*** (1.37)	7.30 (4.50)
Differential impact for women and men	-1.22 (1.97)	-7.40 (6.60)

* p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions. Standard errors are in parentheses.

Next we consider households with members who have disabilities, assessed using the Washington Group Short Set (Washington Group on Disability Statistics, 2020). Very few households reported having members with severe disabilities (based on the recommended criteria of 'Yes, a lot of difficulty' or 'Can't do it at all'). To enable subgroup analysis, we included those with less severe disabilities ('Yes, some difficulty'), resulting in 13% of intervention households having at least one member with any disability, and 15% in the comparison group.

Two significant differential impacts are found for improved sanitation and Empowerment in WASH, as shown in Table 6.6.3. For those reporting no disabilities, the project had a small but significant negative impact on improved sanitation access (-1.95, p<0.1). However, households that did report having at least one member with a disability, had a positive differential impact (5.89, p<0.1; overall estimated impact for these households is 3.94). They were disadvantaged before the project according to the comparison (-5.62, p<0.05), but after the intervention they report similar levels of improved sanitation access regardless of whether any members have a disability. For Empowerment in WASH, the proportion of individuals considered empowered (based on EWI scores) decreased for those in households having any member with a disability, while that figure increased for those in households not having any member with a disability.

Table 6.6.3: Impact of the project by household members' disability status

	Improved water+ (%)	Improved sanitation (%)	HWISE Scale (0 to 36)	Sustainable Water and Sanitation Index	% Empowered
Effect of having household members with a disability in the comparison group	-0.11 (5.40)	-5.62** (2.18)	2.63 (1.64)	-1.78 (2.32)	8.90 (6.70)
Project impact among households having members with a disability	12.24*** (2.20)	-1.95* (1.13)	-7.91*** (0.65)	14.61*** (1.03)	5.90* (3.50)
Differential project impact for households with or without members with a disability	0.38 (6.85)	5.89* (3.35)	-1.68 (1.91)	-2.20 (3.15)	-17.20* (9.00)

+During the dry season; improved water includes piped water on premises. * p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1000 repetitions. Standard errors are in parentheses.

In summary, we find a few differences in project impacts by implementation type, interviewee gender, and whether any household members have a disability. The borehole intervention led to greater impact than the kiosk intervention, especially in reducing water insecurity. The addition of the public toilet intervention to the borehole intervention significantly increased project impact in terms of Equity and Well-Being. Differential impacts by gender are minimal overall, but men benefited more with increased Well-Being than did women. Households having any member with a disability benefited more in terms of improved sanitation. However, they did not benefit in terms of Empowerment in WASH, while households not having any member with a disability did.

6.7 ACCOUNTABILITY

At the end of the household survey we had a module on accountability. The purpose was twofold – to systematically assess Oxfam's accountability mechanisms across all 2019/20 Effectiveness Reviews and impact evaluations, and to help inform which results would be of most interest for sharing back with communities. Here we share descriptive results (see Figure 6.7.1 and Table 6.7.1) and summarize what people would be most interested to hear back (see Table 6.7.2 and 6.7.3).

Only 10% of interviewees said they had been part of a similar survey in the last three years; of the 94 interviewees who said yes, only 15% had heard back about the results of that survey. We see that around half (46%) of people familiar with the project said they had experienced a problem at some point, and in a majority of those cases (79%) did share feedback on the problem or raise the issue. Of those who did raise the issue, almost all received a response (91%), although just over half (64%) were satisfied with the response they received. Of those who did not, the most common reason was not knowing who to approach. Only 35% said they knew how to provide feedback or make a complaint. The top three ways people said they prefer to share feedback were face-to-face, hotline and in a group.

Figure 6.7.1: How interviewees responded to accountability questions (descriptive only)

Table 6.7.1: Accountability module descriptive statistics

Question	Total N	Overall	Women	Men
Have you been part of a similar survey in the last 3 years?	983	10%	8%	11%
If yes, have you heard back about that results of the survey, evaluation, or research?	94	15%	18%	13%
Think about all the time that you have been part of projects implemented by the Freetown WASH Consortium and Oxfam. Now please think about whether you experienced any problems with this specific project. Think of one specific example of a problem you experienced. If you have experienced several problems, choose the one that was most significant to you. Do you have one example in mind?	124	46%	52%	41%
At the time when the problem happened, did you want to raise the issue or share feedback relating to the issue you are thinking about?	57	88%	83%	93%
Did you raise the issue or share feedback relating to the issue you are thinking about?	57	79%	75%	83%
If not, why? I didn't feel safe to do so	12	8%	14%	0%
If not, why? I didn't know who to approach	12	50%	29%	80%
If not, why? I didn't know how to give feedback	12	0%	0%	0%
If not, why? I didn't trust that my feedback would be handled	12	17%	14%	20%
If not, why? I didn't have time	12	17%	29%	0%
If not, why? Other	12	8%	14%	0%
If yes, did you receive a response?	45	91%	86%	96%
If you received a response, were you satisfied with the response you received?	45	64%	67%	63%
In general, do you know how to provide feedback or make a complaint to the Freetown WASH Consortium and Oxfam?	124	35%	30%	39%
Would you feel safe providing feedback to the Freetown WASH Consortium and Oxfam?	114	93%	90%	95%
How much would you say you believe in the Freetown WASH Consortium and Oxfam's reliability and ability to listen and act in your best interest? 1 = Not at all, 2 = Just a little, 3 = Somewhat, 4 = A lot	113	3.8	3.8	3.8
Overall, how satisfied are you with how your feedback and ideas have been taken into account by the Freetown WASH Consortium and Oxfam? 1 = Very dissatisfied, 2 = Somewhat dissatisfied, 3 = Somewhat satisfied, 4 = Very satisfied	124	3.5	3.5	3.6
In general, do you feel that you were able to contribute to making decisions on how Freetown WASH Consortium and Oxfam projects were implemented?	124	68%	69%	67%

Overall, the majority of people (93%) said they feel safe providing feedback, and levels of trust and satisfaction are also relatively high (3.8 and 3.5, respectively, both on a scale of 1–4). Many people (68%) also felt they were able to contribute to decision-making in the project's implementation.

Within the accountability module, we do see some differences by gender. For example, women's reasons for not raising an issue are more varied in comparison to men and include 'I didn't feel safe to do so' and 'I didn't have time', while no men gave these reasons. Also, women more often said they experienced a problem (52%, compared to 41% of men), but fewer women raised the issue (75%, compared to 83% of men), perhaps at least partially because a lower proportion of women said they know how to provide feedback or make a complaint (30%, compared to 39% of men).

Interviewees also shared what they would be most interested to hear back from this evaluation (n=969). We used emergent coding to categorize their responses by topic, with a second round of

coding to further categorize the most common topic – which was about future improvements – into subtopics. In each case, it was possible for a given response to be coded with multiple topics (i.e., topics are not mutually exclusive). Topic and subtopic frequencies and examples are shown in Table 6.7.2 and Table 6.7.3.

Table 6.7.2: Topics interviewees would be most interested to hear back from the survey

Response topic	Frequency	Response subtopic	Frequency
Future improvements	76%	Water	75%
		Sanitation	11%
		Jobs	1%
		Health	<1%
		Roads	<1%
		General (not specified)	24%
Feedback	19%		
Good news	13%		
Influence	2%		
Food	<1%		
Water pricing	<1%		
Monitoring	<1%		
School attendance	<1%		
Nothing	1%		

Table 6.7.3: Examples of what interviewees would be most interested to hear back from the survey, by topic and subtopic

Feedback 'The final report of the survey' 'Report back to us' 'Feedback from Oxfam' 'Yes I am interested' 'A briefing on sanitation'	**Good News** 'Good news' 'Anything good' 'Positive feedback and quick response'	**Influence** 'To submit the final report to the ministry of planning' 'To share this report with government and city council so they can be in the know about the challenges we have…'
Water Pricing '…the price for the water is too much they have increased it.' '…engage the caretakers of the well to reduce the price…'	**Monitoring** 'Monitoring the water system' '…one of those monitoring the borehole [said] they will close because it is broken…'	**Nothing** 'Nothing' 'I don't have anything to suggest'
Future Improvements		
Water 'Development in this community including access to water' 'I hope that there will be more access to water because water is the main issue' 'Help us with water facility'	***Sanitation*** 'More emphasis should be placed on sanitation…' 'I will be interested to hear that there will be a public latrine in my community' 'Help us with a public toilet'	***Jobs*** 'I am expecting better things will come to us after this survey. Better things like water, food and jobs for the youth' 'We need in areas like food, water and jobs'
Health 'I am expecting this message will reach the president to help with water and health care' 'I hope development in water and health in this community will be established by Oxfam'	***Roads*** 'More development in our community apart from water [such as] roads, sanitation' 'More improvement in terms of supply of water and also help with our roads'	***General*** 'I want to know if there will be any development after this survey' 'More improvement needed' 'Help us with more projects'

In summary, people would most like to hear back about plans for future water and sanitation improvements in their communities. They also would like to hear more about plans for using the results for influencing. Many raised issues of water accessibility, especially during the dry season in March and April each year, but also the opening hours each day and long queues, which also came through in the personal stories. People requested upgrades to their water systems to make it safe for drinking and to address issues of water pricing affordability, with differential pricing reported in some cases. Finally, a few people wanted to know more about any positive changes found by the survey, such as school attendance of children, well-being and any other good news. Recall from Figure 6.7.1 that among the people who had taken part in a similar survey in the past, only 15% had heard back about the results, representing an overall lack of accountability in survey research (Kashurha, 2019). Within the scope of this evaluation, we were able to share summary information back with community leaders and water point attendants addressing these requests but were unable to reach all community members who had participated in the survey due to various resource limitations. We acknowledge this is a critical limitation of this work.

7 CONCLUSIONS

7.1 CONCLUSIONS

This evaluation sought to address three main questions, as follows:

1. What is the impact of FWC Phase III for individuals and their households in terms of Sustainable Water and Sanitation?

2. What community-level factors are necessary and sufficient for sustainability of FWC Phase III?

3. How do community-level factors for Sustainable Water and Sanitation relate to the household-level outcomes observed?

To answer these questions, we interviewed 1,079 people through household surveys and 15 key informants to gather community-level information from intervention and comparison communities. These data were then analysed using mixed methods including quasi-experimental methods using propensity score matching (PSM), qualitative coding of personal stories (inspired by SenseMaker®), and configurational analysis, using EvalC3. Below we summarize and reflect on the findings for each evaluation question, considering the first question by itself and then the second and third questions together, followed by a summary on accountability.

WHAT IS THE IMPACT OF FWC PHASE III FOR INDIVIDUALS AND THEIR HOUSEHOLDS IN TERMS OF SUSTAINABLE WATER AND SANITATION?

The project implemented water and sanitation infrastructure, while also establishing WASH Management Teams (WMTs) and carrying out behaviour change communication (BCC), disaster risk reduction (DRR) and advocacy activities. These activities varied by community, and in this evaluation we considered intervention communities that had (1) new boreholes installed, (2) new water kiosks installed, or (3) new boreholes installed and public toilets rehabilitated.

We noted a significant 20% increase in access to improved water during the dry season in the intervention group, on average (65% versus 54%, $p<0.01$). We did not find an impact during the wet season, when improved access rates are relatively higher in both the intervention and comparison groups (around 75%). We also did not find an impact on improved sanitation, which is relatively high in both groups (around 95%). Overall, the project significantly reduced water insecurity (by 54%), reducing the proportion of water insecure households by 62% (from 66% water insecure to 25%, $p<0.01$). The results indicate that the borehole intervention led to greater impact than the kiosk intervention especially in reducing water insecurity. Households with members having any disability benefited more in terms of improved sanitation.

The intervention had a significant positive impact on the Sustainable Water and Sanitation Index of 13.69 percentage points ($p<0.01$). This positive trend continues with significant impacts across five of the six dimensions – Water Security (31.99, $p<0.01$), Equity (14.87, $p<0.01$), Institutions (9.59, $p<0.01$), Operations (8.64, $p<0.01$), and The Environment (14.26, $p<0.01$). The addition of the public toilet intervention to the borehole intervention significantly increased project impact in terms of Equity and Well-Being. Differential impacts by gender are minimal, but men benefited more with increased Well-Being. Among the indicators, 12 of the 14 show significantly positive impact. The largest gains (more than 10 percentage points) were seen for Water Security, Gender Inclusive,

Inclusive of Vulnerable Users, Accountability, Asset Management, Water Safety, and Waste Management. The only two indicators without any significant impact were Health and Education.

In terms of Empowerment in WASH, at the time of the survey and among all survey participants, there were differences in empowerment by gender, with a lower proportion of women being considered empowered (based on EWI scores) than the proportion of men (51% versus 58%, $p<0.05$). Empowerment in WASH was not directly related to the project's theory of change and we did not find any significant impacts of the project overall. Although, there was a significant positive impact for people in households that did not have any member with a disability, the proportion of individuals considered empowered decreased for people in households reporting any member with a disability (i.e., the project had a negative impact for them).

Personal stories (inspired by SenseMaker®) allowed additional information to emerge in relation to Sustainable Water and Sanitation. A subset of people interviewed (35%) chose to share stories about something important that had happened in their lives, where water and sanitation played a role. Most of these stories more about water than sanitation and more negative than positive. Storytellers categorized their stories within the six dimensions of the Sustainable Water and Sanitation Index, with a majority about Water Security, Well-Being and The Environment. The main themes that emerged beyond these dimensions were gender-based violence (GBV) – sexual harassment and rape – and teenage pregnancy, illustrating the link between these issues and problems with water and sanitation.

WHAT COMMUNITY-LEVEL FACTORS CONTRIBUTE TO THE SUSTAINABILITY OF FWC PHASE III? HOW DO COMMUNITY-LEVEL FACTORS CONTRIBUTE TO INDIVIDUAL AND HOUSEHOLD OUTCOMES OF INTEREST?

Focusing on Sustainable Water and Sanitation, EvalC3 analysis of the KIIs highlights the importance of the borehole intervention. We also see that is it necessary to collect fees for water and important to have women leaders in the WMT. Other community-level factors also contribute to more successful outcomes. All EvalC3 results are summarized in Table 7.1.1.

In detail, we always see above average Sustainable Water and Sanitation in communities with any of the following: the public toilet intervention, women leaders, a bank account, a WASH Management Team (WMT) that was already established before the project (i.e., these are Sufficient). We only see above average Sustainable Water and Sanitation in communities where the WMT is collecting fees (i.e., this is Necessary).

Furthermore, if we look at the household-level outcomes of water insecurity and empowerment in WASH, we find that it is also important to have women members in the WMT and to have a well-established WMT (in this case before the project started, meaning success is more likely if it is possible to work with an existing team rather than establishing a new one).

Table 7.1.1: Community-level attributes leading to better household-level outcomes

#	Attribute	Sustainable Water and Sanitation	Water Insecurity	Empowerment in WASH
1	Borehole	A, DT3	S, A, DT2	
2	Kiosk	DT5		DT2
3	Public Toilet	S	S	DT4
4	Paved road	DT4		DT3
5	Mostly flat	DT2	DT3	
6	Active WMT		DT2	
7	Women members in WMT	A	S, A, DT3	
8	Women leaders in WMT	S, A, DT3	S	
9	Collecting fees	N, A, DT1	A, DT1	
10	Bank account	S	S	DT4
11	Established WMT	S	S, DT4	S, DT1
12	Water Safety plan	A	S, A	DT3
13	Equity plans		S	

S = Sufficient, N = Necessary, A = Accuracy is 70% or more, DT# = Decision Tree order of appearance

WHAT REFLECTIONS ON ACCOUNTABILITY EMERGE FROM THIS EVALUATION?

Finally, we consider Oxfam's accountability mechanisms and discuss which results would be of most interest for sharing back with communities. Based on information from an accountability module included within the household surveys, among those familiar with the project who had experienced a problem at some point, only 35% said they knew how to provide feedback or make a complaint (30% of women and 39% of men). The top three ways people said they prefer to share feedback were face-to-face, hotline and in a group. People would most like to hear back about plans for future water and sanitation improvements in their communities and plans for using the results for influencing. Many raised issues of water accessibility, especially during the dry season in March and April each year, but also the opening hours each day and long queues. People requested upgrades to their water systems to make it safe for drinking and to address issues of water pricing affordability, with differential pricing reported in some cases. Finally, a few people wanted to know more about any positive changes found by the survey, such as school attendance of children, well-being and any other good news. Note that among the people who had taken part in a similar survey in the past, only 15% had heard back about the results, representing an overall lack of accountability in survey research (Kashurha, 2019). Within the scope of this evaluation, we were able to share summary information back with community leaders and water point attendants addressing these requests but were unable to reach all community members who had participated in the survey due to various resource limitations. We acknowledge this is a critical limitation of this work.

7.2 PROGRAMME LEARNING CONSIDERATIONS

These learning considerations were developed through discussion of the evaluation findings with the Programme Manager and Oxfam GB's Gender Integrated Programming Lead.

ENSURE GENDER MAINSTREAMING IN ALL FUTURE WASH PROGRAMMING, INCLUDING ADDRESSING WOMEN'S LEADERSHIP AND GENDER-BASED VIOLENCE (GBV)

Findings from the analysis of stories about personal experiences with water and sanitation reflected the emergence of key themes around gender-based violence and teenage pregnancy. Future programming should ensure robust gender mainstreaming using Oxfam's body of knowledge and building on Oxfam in Sierra Leone's learning from gender justice programmes. In addition to following guidance and best practice on gender mainstreaming, the country team can support this process by partnering with women's rights organizations to establish referral mechanisms while also following Oxfam's safe programming approach. Programmes could also work more closely with WMT women leaders; we do see that stories about GBV were shared less often in communities where the WMT has women leaders (6% of stories, compared to 14% in communities without WMT women leaders; the proportion of stories about teenage pregnancy did not vary substantially for these two groups), although there could be other reasons for this difference that we have not considered. In any case, WMT women leaders and members could benefit from additional training on how they can contribute to gender justice in relation to their role within the WMT. Broader awareness-raising activities could also help, as well as consulting women when designing and implementing infrastructure to ensure their needs are being met, particularly for safety (e.g., better lighting around water points and toilets).

CONSIDER ASSESSING AND IMPROVING WATER INFRASTRUCTURE TO MAINTAIN ADEQUATE SERVICE LEVELS DURING THE DRY SEASON

In their comments, people in both intervention and comparison areas raised ongoing issues with water access during the dry season, especially from March to May. For those in the intervention group, some reported that the boreholes provide less water during this time, and the queues also increase as others travel from further away to access that water when their own sources dry up. An infrastructure assessment is needed, which would have to be tailored to the design and context of each system. Then, based on the findings from that assessment, necessary improvements would need to be made to make it through dry season in all communities. Any assessments and subsequent improvements would likely need to be done within the existing governance structures.

WASH MANAGEMENT TEAMS NEED FURTHER TRAINING AND SUPPORT, PARTICULARLY ON IMPLEMENTATION OF WATER PAYMENT SYSTEMS

Through configurational analysis of community-level factors, this evaluation confirmed that collecting fees is necessary for increased sustainability of water systems, and it is also critical for achieving lower water insecurity. However, when asked what they would like to hear back from the survey, several people said that those collecting fees had raised the price and they wanted to find ways to reduce the price, or to keep the pricing fair (e.g., clearer or more trans-parent rules for any payment waivers or fee reductions given to specific households). Future programming should work to ensure payments happen regularly, to streamline the payment process, and to increase

transparency and accountability related to the payment systems. For example, what else could be learned from existing invoice-based payment systems and/or mobile money payment systems in other countries? How can Guma Valley Water Company increase transparency for WMTs to know what happens to the funds they send? And how can WMTs increase transparency for water users to understand what happens with the fees they pay? Finally, some WMTs were not very active any more at the time of this evaluation. It is important to better understand the reasons for disengagement, in order to identify what steps could be taken to support them in continuing to carry out their mandate.

APPENDIX 1: DETAILED INDICATORS, QUESTIONS, AND THRESHOLDS

The following set of tables provides the detailed indicators, questions, and thresholds for each dimension of the Sustainable Water and Sanitation Index followed by the Empowerment in WASH Index (EWI). The table also shows if each indicator is directly (highlighted in green) or indirectly (highlighted in yellow) linked to the FWC Phase III project's theory of change (ToC).

Table A1.1: Water Security dimension = Water Security indicator

Indicator	Variable	Question	Threshold	ToC link?
Water Security (Average of 12 variables x 100)	hwise1_i	In the last month, how frequently did you or anyone in your household worry you would not have enough water for all of your household needs?	Never = 1 Rarely to Always = 0	Indirect
	hwise2_i	In the last month, how frequently has your main water source been interrupted or limited (water pressure, less water than expected, dried up)?		
	hwise3_i	In the last month, how frequently has there not been enough water to wash clothes?		
	hwise4_i	In the last month, how frequently have you or anyone in your household had to change schedules or plans due to problems with your water situation? Activities that may have been interrupted include caring for others, doing household chores, agricultural work, income-generating activities, etc.		
	hwise5_i	In the last month, how frequently have you or anyone in your household had to change what was being eaten because there were problems with water (for washing foods, cooking)?		
	hwise6_i	In the last month, how frequently have you or anyone in your household had to go without washing hands after dirty activities (defecating or changing diapers, cleaning animal dung) because of problems with water?		
	hwise7_i	In the last month, how frequently have you or anyone in your household had to go without washing their body because of problems with water (not enough water, dirty, unsafe)?		
	hwise8_i	In the last month, how frequently has there not been as much water to drink as you would like for you or anyone in your household?		
	hwise9_i	In the last month, how frequently did you or anyone in your household feel angry about your water situation?		
	hwise10_i	In the last month, how frequently have you or anyone in your household gone to sleep thirsty because there wasn't any water to drink?		
	hwise11_i	In the last month, how frequently has there been no useable or drinkable water whatsoever in your household?		
	hwise12_i	In the last month, how frequently have problems with water caused you or anyone in your household to feel ashamed, excluded and/or stigmatized?		

Table A1.2: Equity dimension = (Gender Inclusive indicator + Inclusive of Vulnerable Users indicator)/2

Indicator	Variable	Question	Threshold	ToC link?
Gender Inclusive (Average of 5 variables x 100)	equity1	The water and sanitation needs of people of all genders are being met equally.	True = 1 False = 0	Indirect
	equity4	Women and men share equal responsibility in ensuring that the water and sanitation needs of their households are met.		
	equity5	Women and men are equally aware of how to provide feedback and submit complaints about local water and sanitation management.		
	equity6	Feedback and input from women and men about local water and sanitation management is equally valued.		
	equity7	Women and men are equally aware of their water and sanitation rights.		
Inclusive of Vulnerable Users (Average of 4 variables x 100)	equity2	The water and sanitation needs of people with disabilities and/or chronic illnesses are being met.	True = 1 False = 0	
	equity3	The water and sanitation needs of children are being met.		
	feeoptynwat_i	Do you know of any options for those who are not able to pay for water?	Yes = 1 No or Water is Free = 0	
	feeoptynsan_i	Do you know of any options for those who are not able to pay for sanitation?		

Table A1.3: Institutions dimension = (Accountability indicator + Transparency indicator + Trust indicator)/3

Indicator	Variable	Question	Threshold	ToC link?
Accountability (Average of 3 variables x 100)	watsansatisfaction_i	Overall, how satisfied are you with your water and sanitation services and systems?	Very/somewhat satisfied = 1 Very/somewhat dissatisfied = 0	Indirect
	holdacc_yn	Have you done anything in the past year to try to encourage or pressure the government and service providers to provide better water and sanitation services?	Yes = 1 No = 0	
	holdacc_conf_i	How confident do you feel in your ability to encourage or pressure government and service providers to provide better water and sanitation services?	Very/somewhat confident = 1 Not confident = 0	
Transparency (Average of 2 variables x 100)	info_yn	Do you get any information about the management of your water and sanitation systems?	Yes = 1 No = 0	
	info_enough	Do you feel that you get enough information about the management of your water and sanitation services?		
Trust (Average of 7 variables x 100)	How much would you say you trust the following in relation to water and sanitation laws, systems and services?			
	trust1_i	Local Government (FCC/WARDC)	A lot/Somewhat = 1 Not at all/Just a little = 0	
	trust2_i	Local Development Committee		
	trust3_i	Guma Valley Water Company		
	trust4_i	Property owner (landlord/landlady)		
	trust5_i	Non-governmental organizations (NGOs)		
	trust6_i	Community leaders		
	trust7_i	National government		

Table A1.4: Operations dimension = (Asset Management indicator + Financial Viability indicator + Accessibility indicator)/3

Indicator	Variable	Question	Threshold	ToC link?
Asset Management (Average of 7 variables x 100)	watsancomm_i	What type of water and sanitation team(s), committee(s) or group(s) does your community have?	Committee exists = 1 No committee = 0	Direct
	confidentoandm_i	For water and sanitation systems, how confident are you in this committee to manage routine operation and maintenance to avoid break downs?	Very/somewhat confident = 1 Not confident = 0	
	confidentrepair_i	For water and sanitation systems, how confident are you in this committee to manage major repairs after a break down?		
	confidentreplace_i	For water and sanitation systems, how confident are you in this committee to manage a full replacement if the current system cannot be repaired?		
	confidentshock_i	For water and sanitation systems, how confident are you in this committee to manage through a crisis, shock or disaster?		
	watbd_i	How often can you not use your main water source because it is broken down?	Never = 1 Otherwise = 0	
	sanbd_i	How often can you not use your main toilet or sanitation facility because it is broken down?		
Financial Viability (Average of 12 variables x 100)	payinstallwat_i	How important do you think it is for you to pay to install your own system for water?	Very/somewhat important = 1 Not important at all/Neither important nor unimportant = 0	
	payinstallsan_i	How important do you think it is for you to pay to install your own system for sanitation?		
	payfeewat_i	How important do you think it is for you to pay for regular costs for your system (service fees, etc.) for water?		
	payfeesan_i	How important do you think it is for you to pay for regular costs for your system (service fees, etc.) for sanitation?		
	feestructwat_i	What is the fee structure for use of the main system that you use for water?	Fixed or variable = 1 Irregular, free or don't know = 0	
	feestructsan_i	What is the fee structure for use of the main system that you use for sanitation?		
	paydifficultwat_i	How difficult is process of paying for services for water?	Not difficult at all = 1 Very/somewhat difficult = 0	
	paydifficultsan_i	How difficult is the process of paying for services for sanitation?		
	fairpricewat_i	Do you think your main service is fairly priced for water?	Completely/Somewhat fair = 1 Completely/Somewhat unfair = 0	
	fairpricesan_i	Do you think your main service is fairly priced for sanitation?		
	expensewat_pct_i	Calculation of how much the household spends on water (fees, bottled water, chlorine, etc.), as a percentage, based on an accounting of monthly expenses.	<=5% of total = 1 >5% of total = 0	
	expensesan_pct_i	Calculation of how much the household spends on sanitation and personal hygiene (toilet, soap, etc.), as a percentage, based on an accounting of monthly expenses.		

Accessibility (Average of 14 variables x 100)	watimproved_i	What is your main source for drinking water during the dry season? During the wet season? What is the main source of water for all your other household needs (bathing, handwashing, cooking, cleaning, etc.)?	Piped/Improved in all cases = 1 Unimproved in any case = 0	
	sanimproved_i	What type of sanitation facilities do members of your household use most often now?	Closed pit/flush toilet = 1 Open pit/none = 0	
	watsource_sp_i	Is your main water source private, only for your household, or is it shared?	Private = 1 Shared = 0	
	san_sp_i	Is your main toilet or sanitation facility private, only for your household, or is it shared?		
	watsourceofficial_i	Is your main water source an official or government service?	Yes = 1 No/Don't know = 0	
	watsourceloc_i	Where is your main water source located? How long does it take to walk there? What is the terrain like on the way? How long do you have to wait in queues at the source before you get water?	On-plot/In the house OR no more than 5-minute walk over flat and easy terrain and 5-minute wait in queue = 1 Off-plot, walking and/or waiting over 5 minutes (each) or walking over uneven terrain/steep hills = 0	
	sanloc_i	Where is your main toilet or sanitation facility located? If 'off-plot', is it a public toilet facility? How long does it take to walk there (one way)?	In the house/On-plot or Public toilet no more than 5-minute walk away = 1 Off-plot or Public toilet more than 5-minute walk away = 0	
	wataccess_i	When your main water source is functioning normally, how many hours per day are members of your household able to access it?	24 hours = 1 <24 hours = 0	
	sanaccess_i	When your main toilet or sanitation facility is functioning normally, how many hours per day are members of your household able to access it?		
	sanprivacy_i	How satisfied are you with the privacy of this toilet or facility?	Very/somewhat satisfied = 1 Very/somewhat dissatisfied = 0	
	watsourcesafeday_i	Do you feel safe accessing this water source during the day?	Yes = 1 No = 0	
	watsourcesafenight_i	Do you feel safe accessing this water source at night?		
	sansafeday_i	Do you feel safe accessing this toilet or facility during the day?		
	sansafnight_i	Do you feel safe accessing this toilet or facility at night?		

Table A1.5: Well-Being dimension = (Quality of Life indicator + Health indicator + Education indicator)/3

Indicator	Variable	Question	Threshold	ToC link?
Quality of Life (Average of 2 variables x 100)	lifesatisfaction_i	Overall, how satisfied are you with life these days?	Very/Somewhat satisfied = 1 Very/Somewhat dissatisfied = 0	Direct
	stress_i	Based on Perceived Stress Scale (PSS-4 score).	PSS-4 is 8 or more (low stress) = 1 PSS-4 is less than 8 (high stress) = 0	
Health (Average of 8 variables x 100)	health_i	Overall, how would you describe your state of health now?	Very good/Good = 1 Very poor/Poor = 0	
	stomachpain_i	In the last week, have you had any stomach pain (excluding menstrual pain, for women)?	No = 1 Yes = 0	
	fever_i	In the last week, have you had a fever?		
	diarrhoea_i	I'm sorry to have to ask this, but in the last week, have you had any diarrhoea?		
	unwell_i	In the last week, have you felt so unwell that you had to interrupt your normal daily activities (excluding menstrual pain, for women)?		
	typhoid_i	Have you had typhoid (diagnosed at a health centre or hospital) in the last year?		
	cholera_i	Have you had cholera (diagnosed at a health centre or hospital) in the last year?		
	expensehealth_pct_i	Calculation of how much the household spends on health and medicine, as a percentage, based on an accounting of monthly expenses.	<=5% of total = 1 >5% of total = 0	
Education (Average of 3 variables x 100)	schoolnow_i	Calculation from household roster of whether all school-aged children in the household are attending school (calculation).	Yes = 1 No = 0	
	absence_any_i	Calculation from household roster of whether any school-aged children in the household missed school in the past 1 month for any reason.		
	absence_watsan_i	Calculation from household roster of whether any school-aged children in the household missed school in the past 1 month due to water and sanitation related issues.	No = 1 Yes = 0	

Table A1.6: The Environment dimension = (Water Safety indicator + Waste Management indicator)/2

Indicator	Variable	Question	Threshold	ToC link?
Water Safety (Average of 3 variables x 100)	environment1	Water systems in my community are monitored to ensure water is safe to drink.	True = 1 False = 0	Direct
	environment2	Water safety plans exist and are enforced to protect drinking water sources in my community from contamination.		
	environment3	Drainage systems in my community are kept clean to reduce the risk of flooding.		
Waste Management (Average of 3 variables x 100)	environment4	Sanitation systems in my community are monitored to ensure wastewater/sewage is safely managed.	True = 1 False = 0	
	environment5	Solid waste management plans exist and are enforced to keep my community clean from refuse.		
	environment6	People participate in a monthly or weekly cleaning day in my community (community cleaning campaign, etc.)		

Table A1.7: Adapted Empowerment in WASH Index (EWI) score = average of 12 indicators

Level	Indicator	Variable	Question	Threshold	Modified?	ToC link?
Personal (Individual)	Intrinsic attitudes about WASH roles and responsibilities	ewi_agency	If you want to, to what extent do you feel you can make your own personal decisions about each of the following? • Water collection • Managing water use in the household • Maintenance of household toilet or sanitation facilities • Household expenditures for water (fees, bottled water, chlorine, etc.) • Household expenditures for sanitation (building facilities, emptying pits, maintaining facilities, etc.) • Participating in community water planning? (e.g., planning of water facilities or resolution of problems)	To some or a large extent for at least 4 of the 6 topics listed = 1 Otherwise = 0	No	Indirect
Relational (Household)	Input into decisions about WASH roles and responsibilities	ewi_hhdm_roles	In your household, who normally takes the decision (A), and how much input do you have in making decisions (B), about each of the following? • Water collection • Managing water use in the household • Maintenance of household toilet or sanitation facilities • Household expenditures for water (fees, bottled water, chlorine, etc.) • Household expenditures for sanitation (building facilities, emptying pits, maintaining facilities, etc.) • Participating in community water planning? (e.g., planning of water facilities or resolution of problems)	Makes the decision (self) (A) or has input in most or all decisions (B) for at least 2 of the 6 topics listed = 1 Otherwise = 0	No	
	Input into decisions about WASH expenditures	ewi_hhdm_expenses	In your household, who normally takes the decision (A), and how much input do you have in making decisions (B), about each of the following? • Household expenditures for water (fees, bottled water, chlorine, etc.) • Household expenditures for sanitation (building facilities, emptying pits, maintaining facilities, etc.)	Makes the decision (self) (A) or has input in most or all decisions (B) for at least 1 of the 2 topics listed = 1 Otherwise = 0	No	
	Input into decisions about involvement in	ewi_hhdm_activities	In your household, who normally takes the decision (A), and how much input do you have in making decisions (B), about participating in community water planning? (e.g., planning of water facilities or resolution of problems)	Makes the decision (self) (A) or has input in most or all decisions (B) related to this topic = 1	No	

	community WASH activities			Otherwise = 0		
	Control over household assets	ewi_hhdm_assets	For each asset that your household has now, do you own it? (based on predefined list of 15 possible assets)	Owns solely or jointly at least 50% of household assets = 1 Otherwise = 0	Yes; list of assets adapted to context	
	Work balance	ewi_workbalance	Calculation based on hourly accounting of primary and secondary activities over the last 24 hours, categorized as work (e.g., care work, agriculture, income-generating activities), leisure (e.g., eating, social activities), other (e.g., sleeping, personal care) AND Overall, how satisfied are you with the time you have to do the things you enjoy?	Worked less than 10.5 hours in the previous day AND is satisfied = 1 Otherwise = 0	Yes; 1-hour time blocks used (rather than 30-minute)	
	Time for water collection	ewi_watercollection	Calculation based on hourly accounting of primary and secondary activities over the last 24 hours	Spent a maximum of 1 hour on water collection in the previous day = 1 Otherwise = 0	Yes; 1-hour time blocks used (rather than 30-minute)	
	Access to and sharing of WASH practices information	ewi_infopractices	How often did you hear about the following in the last 12 months? Did you share this information with members of your household? • Good water practices (e.g., safe water storage) • Good sanitation practices (e.g., safe toilet use and maintenance) • Good hygiene practices (e.g., handwashing with soap)	Heard information and shared it for at least 1 of the 3 topics listed = 1 Otherwise = 0	No	
	Access to and sharing of information on WASH rights and responsibilities	ewi_inforights	How often did you hear about the following in the last 12 months? Did you share this information with members of your household? • Citizens' rights related to water, sanitation and hygiene services • The duties of water service providers • Rules that water and sanitation services must follow • Possible consequences for water and sanitation services that did not follow these rules	Heard information and shared it for at least 1 of the 4 topics listed = 1 Otherwise = 0	No	

Environmental (Community)	Group membership	ewi_groups	Are you an active member of any of the following groups? • Water, Sanitation, and/or Hygiene team/committee • Health board/committee (HMC, volunteer group) • School board/committee (SMC, PTA) • Religious group (youth group, men's group, women's group) • Women's group, not religious (mothers' club, etc.) • Men's group, not religious (fathers' group, etc.) • Youth group, not religious • Savings/credit/ finance group • Other community group or committee (not previously indicated)	Is an active member in at least 1 group = 1 Otherwise = 0	Yes; list of groups adapted to context	
	Leadership in WASH implementation	ewi_implementation	Do you feel comfortable speaking up in public (e.g., in a meeting) to help decide on each of the following topics? • Installing, operating and maintaining water systems in your community • Installing, operating and maintaining sanitation systems in your community	Feels comfortable speaking up on at least 1 of the 2 topics = 1 Otherwise = 0	No	
	Leadership in WASH accountability	ewi_accountability	Do you feel comfortable taking each of the following actions? • Speaking up in public (e.g., in a meeting) to give feedback or raise issues with water systems (like breakdowns, inadequate service) in your community? • Giving feedback or raising issues to community leaders or local authorities about water systems (like breakdowns, inadequate service) in your community?	Feels comfortable doing at least 1 of the 2 actions = 1 Otherwise = 0	No	

APPENDIX 2: SUMMARY STATISTICS BEFORE MATCHING

For reference, Table A2.1 below shows various summary statistics for the intervention and comparison. These data represent the sample before propensity score matching (PSM) as described in Section 4.3. The difference column indicates several significant differences between the intervention and comparison group before matching. The purpose of PSM is to balance these differences during analysis (see Appendix 3).

Table A2.1: Selected balance variables with significant differences before matching

Variable	Intervention group mean	Comparison group mean	Difference	Standard error
% of HHs using an improved drinking water source+ in the dry season in 2016	57.66	25.52	32.14***	2.99
% of HHs using an improved drinking water source+ in the wet season in 2016	74.70	61.01	13.68***	3.02
% of HHs with improved sanitation in 2016	94.65	92.31	2.34	1.62
Number of HH members in 2016	4.48	4.10	0.38**	0.15
% of HHs that owned their home in 2016	40.88	36.71	4.16	3.15
% child HH members (<18)	26.26	25.91	0.35	1.43
% school-aged HH members (6–18)	15.33	14.40	0.93	1.16
% youth HH members (<30)	34.74	33.08	1.67	1.79
% elderly HH members (65+)	1.73	1.42	0.31	0.40
% women HH members	49.72	48.11	1.61	1.71
% HH members fit to work	78.11	77.43	0.67	1.65
% of HHs having any member with a disability	12.65	15.21	-2.56	2.25
% of HHs, two spouses without children	7.79	5.24	2.54	1.57
% of HHs, two spouses and children	46.96	53.85	-6.89*	3.23
% of HHs, one person without children	9.25	8.04	1.20	1.81
% of HHs, one person with children	8.27	9.62	-1.34	1.86
% of HHs, other – family (e.g., siblings)	18.00	12.94	5.07*	2.31
% of HHs, other – non-family (e.g., roommates)	2.92	4.37	-1.45	1.23
% of HHs, other (e.g., mixed family/non-family)	6.81	5.94	0.87	1.57
% of HHs with a woman decision-maker	63.26	66.26	-3.00	3.09
% of HHs with a man decision-maker	80.78	80.59	0.18	2.56
Average age of HH decision-makers (years)	40.16	38.80	1.36	0.72
% of HHs, anyone completed secondary school	74.94	74.65	0.29	2.81
% of HHs, anyone completed tertiary school	31.39	31.12	0.27	3.00
Interviewee age (years)	36.55	36.42	0.13	0.73
Interviewee, % women	50.98	45.96	5.01	3.23
Interviewee, % fit for work	89.05	89.34	-0.28	2.01
Interviewee, % in HH in 2016	95.38	94.41	0.97	1.44
Interviewee, % completed secondary school	57.66	61.01	-3.35	3.17
Interviewee, % completed tertiary school	19.71	22.90	-3.19	2.66
Interviewee, % married	54.26	56.47	-2.21	3.22
% with income from salaried employment in government in 2016	31.63	32.52	-0.89	3.02
% with income from salaried employment in private sector and/or NGOs in 2016	17.76	16.96	0.80	2.45

% with income from casual labour in 2016	27.49	32.34	-4.85	2.97
% with income from own business in 2016	75.18	73.43	1.76	2.83
% with income from farming, agriculture, and/or animal husbandry in 2016	3.41	2.62	0.78	1.09
% with income from remittances in 2016	35.04	32.17	2.87	3.05
% with income from cross-border trading in 2016	13.14	10.14	3.00	2.05
% in lowest wealth quintile in 2016	15.57	19.41	-3.83	2.47
% in second lowest wealth quintile in 2016	19.71	20.63	-0.92	2.60
% in middle wealth quintile in 2016	24.33	18.01	6.32*	2.61
% in second highest wealth quintile in 2016	20.44	21.15	-0.72	2.63
% in highest wealth quintile in 2016	19.95	20.80	-0.85	2.61
% of HH participated in any group in 2016	72.99	74.30	-1.31	2.85
Number of groups HH participated in 2016	1.53	1.33	0.20*	0.08
Observations	**983**			

* p <0.1, ** p <0.05, *** p <0.01. Variables dated 2016 are estimates, based on recall data. The construction of the wealth index is described in Section 4.4.

APPENDIX 3: PROPENSITY SCORE MATCHING

The quantitative results presented in Section 6 of this report have been estimated using propensity score matching (PSM). PSM is a statistical technique that allows the effect of an intervention to be estimated by accounting for other factors that predict receiving the intervention, or 'treatment'. The idea behind PSM is to match households in the intervention group to similar households in the comparison group, based on baseline characteristics. After each participant is matched with a non-participant, the average treatment effect on the treated (those who benefited from the intervention) is equal to the difference in average outcomes of the intervention and the comparison groups after project completion. This appendix describes and tests the specific matching procedure employed in this Effectiveness Review. The approach follows the guidance provided by Caliendo and Kopeinig (2008).

ESTIMATING PROPENSITY SCORES

Finding an exact match for treated households based on various baseline characteristics is very hard to implement in practice. Rosenbaum and Rubin (1983) demonstrate that a 'propensity score' can summarize all this information in one single variable. The propensity score is defined as the conditional probability of receiving the intervention given background variables. Specifically, propensity scores are calculated using a statistical probability model (e.g., probit or logit) to estimate the probability of participating in the project based on a set of characteristics.

Table A3.1 shows the variables used to estimate the propensity score in this report, alongside marginal effects at the mean, standard errors, and p-values. Note that the propensity score could not be calculated because of one or more missing values for one household in the comparison group. Following Caliendo and Kopeinig (2008), only variables that influence the participation decision, but which are not affected by participation in the project, have been included in the matching model. In the table, the dependent variable corresponds to whether the household received the intervention (i.e., it is equal to one if the household participated in the project, and zero otherwise). The coefficients in the table correspond to the marginal effects, which are the change in the probability of receiving the intervention if the independent variable is increased by one. Significant effects are highlighted in blue.

Table A3.1: Variables used for PSM with marginal effects, standard errors, and p-values

Variable	Marginal effect	Standard error	p-value
Interviewee is a woman	0.04	0.04	0.35
Interviewee age (years)	0.00	0.00	0.77
Interviewee completed secondary school	-0.02	0.05	0.74
Interviewee completed tertiary school	-0.07	0.07	0.30
Interviewee is married	-0.03	0.05	0.48
HH has at least one decision-maker who is a woman	-0.02	0.04	0.63
HH has at least one decision-maker who is a man	0.08	0.06	0.17
Average age of HH decision-makers (years)	0.00	0.00	0.34
Any HH member completed secondary school	-0.00	0.06	0.98
Any HH member completed tertiary school	0.09	0.06	0.15
HH composition is two spouses without children	0.15*	0.07	0.04
HH composition is one person without children	0.13	0.07	0.08
HH composition is one person with children	-0.03	0.07	0.65
HH composition is other – family (e.g., siblings)	0.10	0.05	0.05
HH composition is other – non-family (e.g., roommates)	-0.02	0.09	0.84
HH composition is other (e.g., mixed family/non-family)	0.11	0.07	0.13
Number of HH members in 2016	0.03**	0.01	0.00

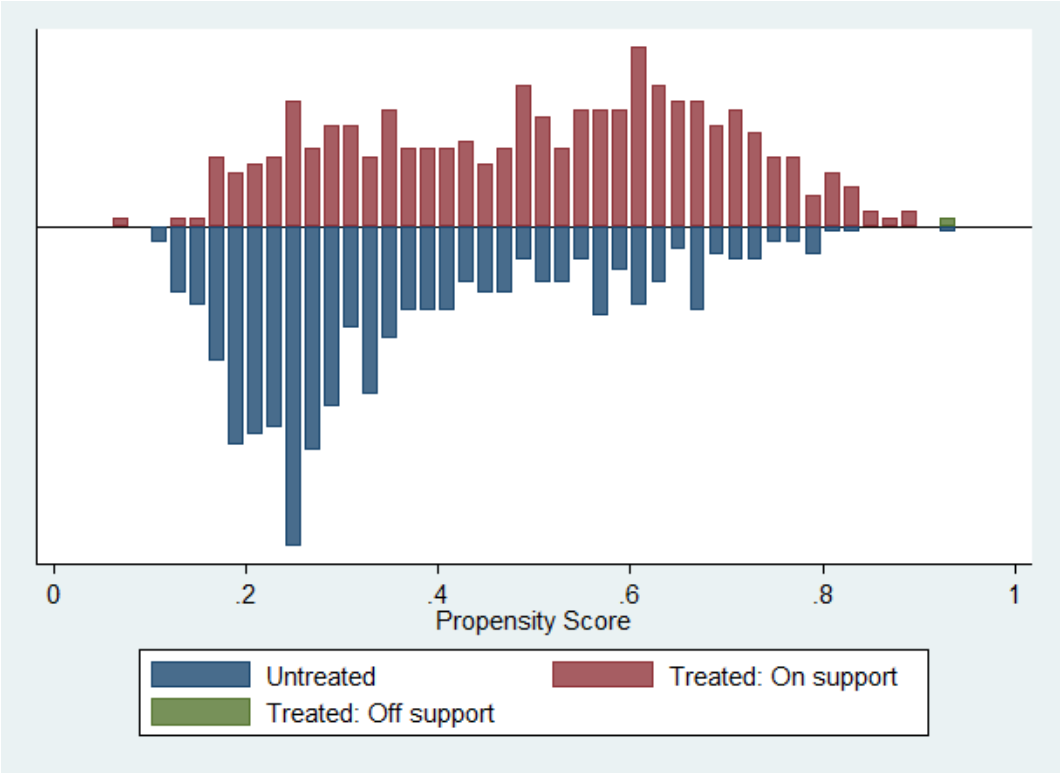
HH owned their home in 2016	0.07	0.04	0.11
HH has at least 80% members over 5 years old fit to work	-0.04	0.04	0.27
Any disability is reported among HH members	-0.09	0.05	0.07
HH earned income from salaried employment in government in 2016	-0.03	0.05	0.56
HH earned income from salaried employment in private sector and/or NGOs in 2016	0.02	0.05	0.61
HH earned income from casual labour in 2016	-0.07	0.04	0.08
HH earned income from farming, agriculture, husbandry in 2016	-0.02	0.10	0.82
HH earned income from remittances in 2016	0.02	0.04	0.54
HH earned income from cross-border trading in 2016	0.06	0.06	0.29
HH was in the lowest 20% of wealth distribution in 2016	-0.03	0.06	0.55
HH was in the second lowest 20% of wealth distribution in 2016	-0.07	0.05	0.20
HH was in the second highest 20% of wealth distribution in 2016	-0.15**	0.05	0.00
HH was in the highest 20% of wealth distribution in 2016	-0.21***	0.05	0.00
Any HH member participated in a community group in 2016	-0.01	0.04	0.81
HH used improved primary drinking water source (dry season) in 2016	0.34***	0.03	0.00
HH used improved sanitation facility (e.g., toilet) in 2016	0.05	0.07	0.49
Observations	**982**		

* p <0.1, ** p <0.05, *** p <0.01. Variables dated 2016 are estimates, based on recall data. The construction of the wealth index is described in Section 4.4. The dependent variable is binary, taking 1 for project participant households, and 0 otherwise.

DEFINING THE REGION OF COMMON SUPPORT

After estimating the propensity scores, it is necessary to verify that potential matches exist for the observations in the intervention group with those from the comparison group – checking that there is common support. The area of common support is the region where the propensity score distributions of the intervention and comparison groups overlap. The common support assumption ensures that each 'treatment [intervention] observation has a comparison observation "nearby" in the propensity score distribution' (Heckman, LaLonde & Smith, 1999). Figure A2.1 shows the propensity score density plots for both groups. It can be observed that, although the distributions of propensity scores are clearly different between the intervention and comparison groups in each case, there is a good area of overlap between the groups. In constructing the model for household-level outcomes, only one observation has been dropped (1 intervention) for lacking a suitable match.

Figure A3.1: Common support histogram of propensity scores for intervention ('Treated') and comparison ('Untreated') households



MATCHING INTERVENTION HOUSEHOLDS TO COMPARISON HOUSEHOLDS

Following Rosenbaum and Rubin (1983), households are matched based on propensity scores using a kernel matching algorithm. Kernel matching assigns more weight to the closest comparison group observations that are found within a selected 'bandwidth'. Thus 'good' matches are given greater weight than 'poor' matches. The psmatch2 module in Stata (Leuven & Sianesi, 2003) was used with a bandwidth of 0.06 and the analysis was restricted to the area of common support. When using PSM, standard errors of the estimates were bootstrapped using 1,000 repetitions to account for the additional variation caused by the estimation of the propensity scores.

CHECKING BALANCE

For PSM to be valid, the intervention group and the matched comparison group need to be balanced. In other words, the intervention and comparison groups need to be similar in terms of their observed characteristics. The most straightforward method of doing this is to test whether there are any statistically significant differences in baseline covariates between both groups in the matched sample. The balance of each of the matching variables after kernel matching is shown in Table A3.2. No significant differences remain. For all the variables, the p-values for the difference in means tests are fairly large, with the lowest value being 0.17 and most being more than 0.50. It can therefore be concluded in each case that a satisfactory match has been found for the intervention group in the sample, according to this set of matching variables.

Table A3.2: Variable balance check after propensity score matching

Variable	Intervention group mean	Comparison group mean	p-value
Interviewee is a woman	0.51	0.51	0.99
Interviewee age (years)	36.57	36.10	0.60
Interviewee completed secondary school	0.58	0.58	0.95
Interviewee completed tertiary school	0.20	0.16	0.17
Interviewee is married	0.54	0.55	0.92
HH has at least one decision-maker who is a woman	0.63	0.62	0.71
HH has at least one decision-maker who is a man	0.81	0.79	0.66
Average age of HH decision-makers (years)	40.17	40.07	0.91
Any HH member completed secondary school	0.75	0.76	0.77
Any HH member completed tertiary school	0.31	0.27	0.21
HH composition is two spouses without children	0.08	0.08	0.92
HH composition is one person without children	0.09	0.08	0.42
HH composition is one person with children	0.08	0.10	0.38
HH composition is other – family (e.g., siblings)	0.18	0.17	0.70
HH composition is other – non-family (e.g., roommates)	0.03	0.04	0.43
HH composition is other (e.g., mixed family/non-family)	0.07	0.08	0.50
Number of HH members in 2016	4.00	3.86	0.42
HH owned their home in 2016	0.41	0.41	0.98
HH has at least 80% members over 5 years old fit to work	0.56	0.51	0.20
Any disability is reported among HH members	0.13	0.14	0.57
HH earned income from salaried employment in government in 2016	0.32	0.29	0.49
HH earned income from salaried employment in private sector and/or NGOs in 2016	0.18	0.16	0.42
HH earned income from casual labour in 2016	0.28	0.32	0.17
HH earned income from farming, agriculture, husbandry in 2016	0.03	0.03	0.80
HH earned income from remittances in 2016	0.35	0.35	0.88
HH earned income from cross-border trading in 2016	0.13	0.11	0.39
HH was in the lowest 20% of wealth distribution in 2016	0.16	0.18	0.37
HH was in the second lowest 20% of wealth distribution in 2016	0.20	0.22	0.46
HH was in the second highest 20% of wealth distribution in 2016	0.20	0.20	0.74
HH was in the highest 20% of wealth distribution in 2016	0.20	0.18	0.44
Any HH member participated in a community group in 2016	0.73	0.71	0.50
HH used improved primary drinking water source (dry season) in 2016	0.58	0.59	0.63
HH used improved sanitation facility (e.g., toilet) in 2016	0.95	0.94	0.62
Observations	**981**		

* p <0.1, ** p <0.05, *** p <0.01. Variables dated 2016 are estimates, based on recall data. The construction of the wealth index is described in Section 4.4. The dependent variable is binary, taking 1 for project participant households, and 0 otherwise.

PROPENSITY SCORE WEIGHTING

We use propensity score (PS) weighting for subgroup analysis. The PS-weighted models used in this report are described below.

PS-weighted OLS regression to estimate impact for different implementation types

We used PS-weighted OLS regression to estimate project impact for different implementation types, including the borehole, kiosk and public toilet interventions. The model estimates the relative effect of each implementation type using three separate binary variables (one for each implementation type, equal to 1 when the intervention is present and 0 when the intervention is absent) with all matching variables used for PSM also included as control variables. Table 6.6.1 in Section 6.6 provides the results of these PS-weighted regressions.

PS-weighted OLS regression with an interaction term to estimate differential impacts for different social groups

Average overall effects were estimated using PSM. To test for differential impacts, we used PS-weighted OLS regression together with an interaction term, with all matching variables used for PSM also included as control variables. Based on how each subgroup variable is defined, those equal to 1 can be considered as the 'group of focus' and those equal to 0 as the 'reference group'. In this report we considered three subgroups with interaction terms as follows:

1. **Gender:** the interaction term multiplies a variable for interviewee gender (equal to 1 if a woman, 0 if a man) by the intervention variable (equal to 1 if in the intervention group, 0 if in the comparison group). Therefore, we define women as the group of focus and men as the reference group.

2. **Disability of household members:** the interaction term multiplies a variable indicating household members' disabilities (equal to 1 if 80% of members over five years old are 'fit to work', 0 otherwise) by the intervention variable (equal to 1 if in the intervention group, 0 if in the comparison group). Therefore, we define households with fewer members with disabilities as the group of focus and households with more with disabilities as the reference group.

Tables 6.6.2 and 6.6.3 in Section 6.6 provide the results of these PS-weighted regressions, each with the relevant interaction term. Table A.3.3 provides tips for interpreting these results.

Table A3.3: Example results and interpretation of PS-weighted regression with an interaction term for the Sustainable Water and Sanitation Index

Sustainable Water and Sanitation Index	Example results	Interpretation
Effect among the group of focus in the comparison group, relative to the reference group	2.50** (2.10)	The first row shows the coefficient, standard error and significance for the subgroup variable (i.e., implementation type, gender, disability). It indicates the difference between the group of focus and reference group, irrespective of the intervention. The sign of the coefficient tells us whether the average is higher (if positive) or lower (if negative) for the group of focus, relative to the reference group. The size of the coefficient tells us how large this difference is (e.g., between women and men). The standard error and significance indicate how statistically meaningful this difference is. The example results would mean that the average value for the group of focus is 2.50 higher than it is for the reference group for the Sustainable Water and Sanitation Index, and this difference is statistically significant ($p<0.05$).
Project impact for the reference group	5.09*** (4.02)	The second row shows the coefficient, standard error and significance for the intervention (i.e., project impact) for the 'reference' group (i.e., men, households with fewer members with disabilities and tenants). The sign of the coefficient tells us whether the project impact is positive or negative, while the size of the coefficient tells us how large this impact is for the 'reference' group. The standard error and significance indicate how statistically meaningful this difference is. The example results would mean that the project had a statistically significant ($p<0.01$) positive impact of 5.09 for the reference group for the Sustainable Water and Sanitation Index.
Differential project impact for the group of focus and the reference group	1.08 (1.02)	The third row shows the coefficient, standard error and significance for the interaction term. The sign and size of the coefficient can be interpreted together with the two other coefficients for the subgroup and intervention variables to calculate the project impact for the group of focus. The standard error and significance indicate how statistically meaningful the difference in impact between the group of focus and reference group is. The example results would mean that the project also had a positive impact for the group of focus for the specified outcome, and that impact is slightly larger (e.g., 1.08 + 5.09 = 6.17). However, this difference in impact between the group of focus and the reference group is not statistically significant (i.e., there is no differential impact by gender for the Sustainable Water and Sanitation Index).

* p <0.1, ** p <0.05, *** p <0.01; PSM estimates are bootstrapped with 1,000 repetitions. Standard errors are in parentheses.

APPENDIX 4: RISK OF BIAS

Not all quasi-experimental impact evaluations are the same. Choices made during sampling, selection of the comparison group, and at the analysis stage are crucial in assessing overall confidence in the results. Table A4.1 uses our standard framework to assess the risk of bias against 10 predetermined parameters for this Effectiveness Review. This framework is specifically for ex-post quasi-experimental impact evaluations. Lower overall risk provides higher confidence in the results.

Table A4.1: Risk of bias table

No.	Title	Description	Assessment	Description
Sampling				
1	Random sampling	**Score LOW risk if:** Sampling is conducted using probability random sampling methods on a clearly established sample frame. **Score MEDIUM risk if:** Sampling is conducted using probability random sampling methods at geographical level (e.g., village level), and uses random sampling to select interviewees within the geographical area. **Score HIGH otherwise.**	MEDIUM	Interviewee selection was done using stratification by community using a random walk protocol (see Section 4.2 Household Surveys for details).
2	Representativeness of project participants	**Score LOW risk if:** Project participants have been involved for the entire duration of the project and have been involved in the project with the same level of exposure. Project participants have been exposed to a variety of different activities, some may have dropped out from some activities, but sampling is conducted on the entire list of project participants. **Score MEDIUM risk if:** Project participants have been exposed to a variety of different activities. Sampling is conducted only among those project participants that have been enrolled for the entire duration of the project or that have been enrolled in all the activities. These are not less than 80% of the entire list of project participants OR it is clear the results apply only to a particular group of project participants. **Score HIGH otherwise.**	MEDIUM	Different interventions were implemented in different communities and by different consortium members. The following were selected: (a) all (5) communities with new boreholes implemented by Oxfam, (b) a random selection (4 out of 12) of communities with kiosks implemented by Oxfam, and (c) two (2) communities with new boreholes AND public toilets rehabilitated by Concern Worldwide. The report has made clear that the results apply to this set of project activities, and subgroup analysis has been done to the extent possible (e.g., to understand any differential impacts for boreholes vs. kiosks and any additional impacts related to public toilet rehabilitation).
3	Selection survey interviewees	**Score LOW risk if:**	LOW	All community members within pre-determined sections (i.e., near enough to access the systems

		Identification of survey interviewees is not determined by project participation (the same protocol to identifying the interviewee(s) within the household is applied in intervention and comparison groups). The resulting selection of survey interviewees is not affected by project participation (based on observables). **Score MEDIUM risk if:** Identification of survey interviewees is not determined by project participation (the same protocol to identifying the interviewee(s) within the household is applied in intervention and comparison group). The resulting selection of survey interviewees is affected by project participation (based on observables). **Score HIGH otherwise.**		implemented by the project for the intervention group and not near enough to access any similar systems built in the past 3 years for the comparison group) were included in the sampling frame. In each of these areas, households were equally likely to be selected for an interview.
Selecting comparison group				
4	Potential for contamination (spillovers)	**Score LOW risk if:** The units for comparison group are selected in geographical areas where it is not reasonable to expect for the project to have had spillover effects. The project also implemented some activities (which are not considered the most relevant under analysis) which are expected to have had an impact also in the comparison group (e.g., the project implemented campaigns using radio and other digital media, but these are only a minor component of the activities implemented). The report makes clear which impact is assessed (added-value of other components, taking into account exposure to those minor components). **Score HIGH risk if:** Units for the comparison group are selected within the same geographical area as the intervention group, and it is reasonable to expect that project activities had spillover effects (e.g., comparison observations within the same village, for awareness-raising projects).	LOW	Comparison areas were carefully checked by the consultant to ensure they were sufficiently far away from any of the systems implemented by the project. The project did implement some activities for hygiene promotion, DRR, and WASH advocacy at the district level and with the Freetown City Council (FCC) that may have also had an impact in the comparison group.
5	Self-selection of project participants	**Score LOW risk if:** The comparison group is exploiting an experiment or natural experiment. Units are randomly selected at community level both in the intervention and comparison groups. The selection process for the comparison group is mimicking the same selection process used by the project. **Score MEDIUM risk if:** The self-selection is corrected during the matching procedure (e.g., controlling for group participation at baseline).	LOW	Project activities were largely conducted at the community level and there were no participant lists at the household level. In each community, households were randomly selected using the same method in the intervention and comparison groups (random walk).

		Score HIGH risk if: Project participants were selected or self-selected based on idiosyncratic or unobservable characteristics, and the selection of comparison interviewees is done randomly from neighbouring geographical sites.		
6	Other interventions in the comparison group	**Score LOW risk if:** There are no other actors in the area (e.g., INGOs, NGOs, governmental programmes). Other actors are conducting activities that are not linked to the project's theory of change. **Score MEDIUM risk if:** Other actors are conducting similar activities linked to the project's theory of change in both the intervention and the comparison groups. **Score MEDIUM-HIGH risk if:** Other actors are conducting similar activities linked to the project's theory of change in the comparison group only, but the evaluation purposefully chooses to compare these activities to the intervention making it clear that the impact is compared with these other activities (e.g., as a natural experiment). **Score HIGH risk if:** Other actors are conducting similar activities in comparison communities only. Other actors are conducting activities in the comparison communities, which are not the same, but are partially related to the project's theory of change.	LOW	Comparison areas were carefully checked by the consultant to ensure that no similar systems had been implemented by other organizations in the past 3 years (during the project).
Analysis				
7	Representativeness	**Score LOW risk if:** During analysis or matching procedure less than 10% of the sample in the intervention group is excluded. **Score HIGH risk if:** During analysis or matching procedure more than 10% of the sample in the intervention group is excluded.	LOW	Less than 1% of the sample is not on common support (1 intervention, 0 comparison).
8	Robustness checks	**Score LOW risk if:** Magnitude and statistical significance of the results are approximately consistent with different econometric models. **Score HIGH risk if:** Results are not consistent with different econometric models.	LOW	Different econometric model specifications – PSM and PS-weighted OLS regression – produced approximately consistent results.

9	Triangulation	**Score LOW risk if:** Results are triangulated and consistent with other evaluation methods within the same evaluation. Results are triangulated and consistent with other data on the same project but from different evaluations. **Score HIGH risk if:** Results are not consistent or triangulated with other evaluation methods.	LOW	Results are consistent between quantitative and qualitative methods used in this evaluation and with the project's final evaluation
10	Multiple hypothesis testing	**Score LOW risk if:** Multiple hypothesis tests apply Benjamini or Bonferroni tests. The evaluation drafted a pre-analysis plan before analysis and followed the plan. **Score MEDIUM risk if:** The evaluation drafted a pre-analysis plan prior data analysis and significantly altered the plan, but changes that are clearly justified. **Score HIGH otherwise.**	LOW	This evaluation drafted a pre-analysis plan prior to data analysis and followed the plan.
11	Clustering	**Score LOW risk if:** Clustering is applied. Clustering was tested but rejected as providing higher standard errors than non-clustering estimates. **Score HIGH otherwise.**	LOW	Community-level clustering is applied.
Other				
12	Other	Any other issue reported by the evaluator.		

APPENDIX 5: EVALC3 CONFIGURATIONAL ANALYSIS

Table A5.1 shows the 'truth table' that was input into EvalC3 for configurational analysis. Each case represents one community. Attributes are highlighted in grey and outcomes are highlighted in green.

Table A5.1: Truth table for input into EvalC3

Case	Intervention	Bore hole	Kiosk	Public toilet	Paved road	Mostly flat	Active WMT	Women members in WMT	Women leaders in WMT	Collecting fees	Bank account	Established WMT	Water Safety plan	Equity plans	Sustainable Water and Sanitation Index	HWISE	EWI
1	1	1	0	0	0	1	0	1	1	1	0	0	0	0	1	1	0
2	1	1	0	0	0	1	1	1	1	1	0	1	0	0	1	1	1
3	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0
4	1	1	0	0	1	1	1	1	1	1	0	1	1	0	1	1	1
5	1	1	0	0	0	0	1	1	0	1	0	0	1	1	0	1	0
6	1	0	1	0	0	0	1	1	1	1	0	0	0	1	1	1	1
7	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0
8	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	0
9	1	0	1	0	1	1	0	0	0	1	0	0	0	0	1	1	1
10	1	1	0	1	1	1	0	0	0	1	0	0	1	0	1	1	1
11	1	1	0	1	1	1	0	0	0	1	1	0	1	0	1	1	0
12	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1
13	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0
14	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1
15	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1

REFERENCES

Abadie, A., & Imbens, G.W. (2008). On the Failure of the Bootstrap for Matching Estimators. Econometrica, 76(6), 1537–57.

Alkire, S., & Foster, J. (2011). Counting and Multidimensional Poverty Measurement. Journal of Public Economics, 95(7–8), 476–87.

Banerjee, S. G., & Morella, E. (2011). Africa's water and sanitation infrastructure: access, affordability, and alternatives. The World Bank.

Bartram, J., Brocklehurst, C., Fisher, M. B., Luyendijk, R., Hossain, R., Wardlaw, T., & Gordon, B. (2014). Global monitoring of water supply and sanitation: history, methods and future challenges. International journal of environmental research and public health, 11(8), 8137–65.

Bishop, D., & Bowman, K. (2014). Still learning: a critical reflection on three years of measuring women's empowerment in Oxfam. Gender & Development, 22(2), 253–69.

Bland, M., & Altman, D.G. (1997). Statistics Notes: Cronbach's Alpha. British Medical Journal, 314(7080), 572.

Bratton, M., & Gyimah-Boadi, E. (2016). Do Trustworthy Institutions Matter for Development? Corruption, Trust and Government Performance in Africa.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. Journal of Economic Surveys, 21:31–72.

Cognitive Edge. (2020). SenseMaker®. https://cognitive-edge.com/sensemaker/

Davies, R. (2020). EvalC3: Tools for developing, exploring and evaluating predictive models of expected outcomes (Excel app). https://evalc3.net/

Dickin, S., & Bisung, E. (2019). Empowerment in WASH Index. SEI Discussion Brief. https://www.sei.org/wp-content/uploads/2019/10/dickin-ewi-wash-db.pdf

Filmer, D., & Pritchett, L.H. (2001). Estimating Wealth Effects without Expenditure Data – Or Tears: An Application to Educational Enrollments in States of India. Demography, 38(1), 115–32.

Fuller, R., & Lain, J. (2015). Measuring resilience: Lessons learned from measuring resilience in Oxfam's large-N Effectiveness Reviews. https://policy-practice.oxfam.org.uk/publications/measuring-resilience-lessons-learned-from-measuring-resilience-in-oxfams-large-583601

Giné-Garriga, R., & Pérez-Foguet, A. (2018). Measuring sanitation poverty: A multidimensional measure to assess delivery of sanitation and hygiene services at the household level. OPHI Working Papers, (116), 1-18.

Godlonton, S., Hernandez, M., & Murphy M. (2018). Anchoring Bias in Recall Data: Evidence from Central America. American Journal of Agricultural Economics, 100(2), 479–501.

Heckman, J., LaLonde, R., & Smith, J. (1999). The economics and econometrics of active labor market programs, Handbook of Labor Economics, 3(A), 1865–2097.

Hughes K., & Bushell, H. (2013). A multidimensional approach to measuring resilience. https://policy-practice.oxfam.org.uk/publications/a-multidimensional-approach-to-measuring-resilience-302641

HWISE. (2020). Household Water Insecurity Experiences (HWISE) Scale. https://sites.northwestern.edu/hwise/

Kayser, G. L., Moriarty, P., Fonseca, C., & Bartram, J. (2013). Domestic water service delivery indicators and frameworks for monitoring, evaluation, policy and planning: a review. International journal of environmental research and public health, 10(10), 4812–35.

Kashurha, C.C. (2019). "When will we get a report on your findings?": reflections on researcher accountability from DRC. https://oxfamblogs.org/fp2p/when-will-we-get-a-report-on-your-findings-reflections-on-researcher-accountability-from-drc/

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. http://ideas.repec.org/c/boc/bocode/s432001.html

Lombardini, S., Bowman, K., & Garwood, R. (2017) A 'How To' Guide To Measuring Women's Empowerment: Sharing experience from Oxfam's impact evaluations. https://policy-practice.oxfam.org.uk/publications/a-how-to-guide-to-measuring-womens-empowerment-sharing-experience-from-oxfams-i-620271

Mager, F., & Lombardini, S. (2019). Livelihoods in the Za'atari Camp: Impact evaluation of Oxfam's Cash for Work activities in the Za'atari camp (Jordan). https://policy-practice.oxfam.org/resources/livelihoods-in-the-zaatari-camp-impact-evaluation-of-oxfams-cash-for-work-activ-620883/

Mager, F., Smith, B., & Guijt, I. (2018). How decent is decent work? Using SenseMaker to understands workers' experiences. https://policy-practice.oxfam.org/resources/how-decent-is-decent-work-using-sensemaker-to-understand-workers-experiences-620476/

Medland, L. (2018). Outcomes-Based Programming for Water and Sanitation: Oxfam's vision and monitoring framework. https://policy-practice.oxfam.org.uk/publications/outcomes-based-programming-for-water-and-sanitation-oxfams-vision-and-monitorin-620573

Miziniak, J., Vonk, J., & Young, S. (2020). Podcast: The journey and challenges in measuring sustainable water in Oxfam's impact evaluations with the HWISE Network. https://views-voices.oxfam.org.uk/2020/03/the-journey-and-challenges-in-measuring-sustainable-water-in-oxfams-impact-evaluations-with-the-hwise-network/

Mizniak, J., Wildman, T., & Medland, L. (2017). A Global Water and Sanitation System That Works for All: Oxfam GB's Sustainable Water and Sanitation Strategy. https://policy-practice.oxfam.org.uk/publications/a-global-water-and-sanitation-system-that-works-for-all-oxfam-gbs-sustainable-w-620333

Nicola, F., & Giné, X. (2013). How Accurate Are Recall Data? Evidence from Coastal India. The World Bank. Policy Research Working Paper 6009.

OCHA. (2020). Sierra Leone – Subnational Administrative Boundaries. https://data.humdata.org/dataset/sierra-leone-all-ad-min-level-boundaries

Porteous, D. (2006). The enabling environment for mobile banking in Africa.

Pretari, A. (2018). Sampling strategies for gendered impact evaluations. https://views-voices.oxfam.org.uk/2018/08/whose-voice-sampling-strategies/

Pretari, A., & Gaboune, A. (2019). Podcast: How to measure resilience capacities – Experience from Oxfam's impact evaluations. https://views-voices.oxfam.org.uk/2019/12/podcast-how-to-measure-resilience-experience-from-oxfams-impact-evaluations/

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70:41–55.

Shilling, F., Khan, A., Juricich, R., & Fong, V. (2013). Using indicators to measure water resources sustainability in California. In World Environmental and Water Resources Congress 2013: Showcasing the Future (pp. 2708–15).

Thomas, E., Andrés, L. A., Borja-Vega, C., & Sturzenegger, G. (Eds.). (2018). Innovations in WASH Impact Measures: Water and Sanitation Measurement Technologies and Practices to Inform the Sustainable Development Goals. The World Bank.

Torre, M., & Lombardini, S. (2019). Podcast: The challenges of measuring women's empowerment. https://views-voices.oxfam.org.uk/2019/10/podcast-the-challenges-of-measuring-womens-empowerment/

UN. (2018). Sustainable Development Goal (SDG) 6: Ensure availability and sustainable management of water and sanitation for all. https://sustainabledevelopment.un.org/sdg6

Vonk, J. (2019). Going Digital: Privacy and data security under GDPR for quantitative impact evaluation. https://policy-practice.oxfam.org.uk/publications/going-digital-privacy-and-data-security-under-gdpr-for-quantitative-impact-eval-620884

Vonk, J., & Lombardini, S. (2020). Podcast: GDPR and the right to privacy in practice for impact evaluations. https://views-voices.oxfam.org.uk/2020/02/podcast-gdpr-and-the-right-to-privacy-in-practice-for-impact-evaluations/

Washington Group on Disability Statistics. (2020). WG Short Set on Functioning (WG-SS). https://www.washingtongroup-disability.com/question-sets/wg-short-set-on-functioning-wg-ss/

WHO. (2020). Ebola virus disease. https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease

WHO/UNICEF. (2017). JMP Methodology: 2017 Update & SDG Baselines. https://washdata.org/sites/default/files/documents/reports/2018-04/JMP-2017-update-methodology.pdf

WHO/UNICEF. (2019). JMP Data. https://washdata.org/data

Wilbur, J., & Danquah, L. (2015). Undoing inequity: water, sanitation and hygiene programmes that deliver for all in Uganda and Zambia–an early indication of trends. In 38th WEDC International Conference, Loughborough University, UK.

Young, S.L., Collins, S.M., Boateng, G.O., et al., (2019). Development and validation protocol for an instrument to measure household water insecurity across cultures and ecologies: the Household Water InSecurity Experiences (HWISE) Scale. BMJ open, 9(1), bmjopen-2018.

Oxfam Effectiveness Reviews

For more information, or to comment on this report, email policyandpractice@oxfam.org.uk

ISBN: 978-1-78748-841-0
DOI: 10.21201/2021.8401

Oxfam GB, Oxfam House, John Smith Drive, Cowley, Oxford, OX4 2JY, UK.

OXFAM

Oxfam is an international confederation of 21 organizations, working with its partners and allies, reaching out to millions of people around the world. Together, we tackle inequalities to end poverty and injustice, now and in the long term – for an equal future. Please write to any of the agencies for further information or visit www.oxfam.org.

Oxfam America (www.oxfamamerica.org)
Oxfam Aotearoa (www.oxfam.org.nz)
Oxfam Australia (www.oxfam.org.au)
Oxfam-in-Belgium (www.oxfamsol.be)
Oxfam Brasil (www.oxfam.org.br)
Oxfam Canada (www.oxfam.ca)
Oxfam Colombia (lac.oxfam.org/countries/colombia)
Oxfam France (www.oxfamfrance.org)
Oxfam Germany (www.oxfam.de)
Oxfam GB (www.oxfam.org.uk)
Oxfam Hong Kong (www.oxfam.org.hk)
Oxfam IBIS (Denmark) (http://oxfamibis.dk/)
Oxfam India (www.oxfamindia.org)
Oxfam Intermón (Spain) (www.oxfamintermon.org)
Oxfam Ireland (www.oxfamireland.org)
Oxfam Italy (www.oxfamitalia.org)
Oxfam Mexico (www.oxfammexico.org)
Oxfam Novib (Netherlands) (www.oxfamnovib.nl)
Oxfam Québec (www.oxfam.qc.ca)
Oxfam South Africa (http://www.oxfam.org.za/)
KEDV (www.kedv.org.tr)

OXFAM