EVALUATION: MARCH 2019 PUBLICATION: APRIL 2022

# SUSTAINABLE WATER AND SANITATION IN DRC

## Impact evaluation of the 'Sustainable WASH in Fragile Contexts (SWIFT 1)' project

Effectiveness Review Series 2018/19


Kahindo Kitambala collects water from a water point constructed through the SWIFT programme.

Credit: Jane Beesley/Oxfam GB/SWIFT.

JAYNIE VONK

**Data gathering led by CIS-ISSNT**

**OXFAM GB**

www.oxfam.org.uk/effectiveness

OXFAM

# ACKNOWLEDGEMENTS

The author would like to recognize and thank everyone who participated in and contributed to this Effectiveness Review. Thank you to those at Oxfam in DRC, especially Josue Ibulungu, Pepe Ken Mvulu, Ephrem Kambale, Janvier Ngambwa and Divin Salumu for their engagement in the evaluation process. Also, thank you to Eddy Muhiwa from our partner HYFRO for participating in the evaluation design workshop and supporting data collection as needed.

All data collection was done by the Centre Informatique et Statistique – Institut Supérieur de Statistique et de Nouvelles Technologies (CIS-ISSNT; Centre for Computer Science and Statistics of the Higher Institute of Statistics and New Technology) including Edson Niyonsaba, Célestin Kimanuka, Promesse Kavulirene and a team of 28 interviewers led by supervisors Florence and Clément – 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 who took time out of their day to share information and experiences at part of this evaluation.

Thanks also to Simone Lombardini and Alexia Pretari for their support throughout the evaluation and to the Oxfam GB water team – Joanna Trevor, Jola Miziniak, Louise Medland and Tom Wildman – for input and feedback on the Sustainable Water and Sanitation Index. Also, thank you to the HWISE team, especially Sera Young, Josh Miller and Chad Staddon, for support during this initial collaboration using the HWISE Scale.

Finally, thanks to Sarah Davies (Oxfam GB's Global Adviser – Strategic Learning and Communications) and Oxfam GB's publishing team for publication of this report.

# CONTENTS

# EXECUTIVE SUMMARY

Oxfam GB's Global Performance 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 impact evaluation, known as an 'Effectiveness Review'. During the 2018/19 financial year, one of the projects selected was the 'DRC Sustainable Water, Sanitation and Hygiene (WASH) in Fragile Contexts' project – in short, 'SWIFT 1 in DRC', and for simplicity referred to as SWIFT 1 throughout this report.

The project was carried out through the SWIFT Consortium, which was led by Oxfam with Tearfund and the Overseas Development Institute (ODI) as members working in DRC and Kenya. In DRC, the consortium worked with implementing partners Hydraulique Sans Frontières (HYFRO; Hydraulics without Borders), Centre de Promotion Socio – Sanitaire (CEPROSSAN; Centre for Public Health Promotion) and Programme de Promotion des Soins de Santé Primaires (PPSSP; Primary Health Care Promotion Programme). The project operated in rural and semi-urban areas in three eastern provinces: North Kivu, South Kivu and Maniema. Oxfam led implementation in some areas of North Kivu, while Tearfund led in other areas of North Kivu as well as in South Kivu and Maniema. ODI provided Monitoring, Evaluation, Accountability and Learning (MEAL) support.

The project was implemented in two phases – an outputs phase (April 2014 to March 2016) and an outcomes phase (April 2016 to March 2018). During the outputs phase, in rural areas (approximately 75% of project participants, and the focus of this evaluation) the project expanded on the national 'Villages et Ecoles Assainis' (VEA; Healthy Villages and Schools) approach. In this phase, project implementation followed the VEA approach in coordination with the government. Activities included rehabilitating and constructing water infrastructure, encouraging households to build latrines, helping to form and train local committees, promoting healthy hygiene behaviours and working to achieve a 'Healthy Village Certification'; SWIFT 1 incorporated some additional support, such as providing digging kits for latrines. SWIFT 1 also added the outcomes phase, which was not included in the VEA approach, to carry out additional capacity-building activities with the committees to increase project sustainability. Throughout the duration of the project, radio programming and regional events took place, for example, to promote healthy hygiene behaviours, such as handwashing. Note, the project followed a different approach in semi-urban areas (approximately 25% of project participants, and outside the scope of this evaluation).

# EVALUATION APPROACH

The Effectiveness Review, for which data collection was carried out in March 2019, aimed to evaluate the success of SWIFT 1 in increasing the sustainability of water and sanitation systems and services in DRC, one year after the end of the project. It focuses on measuring benefits attributable to additional activities the project carried out, above and beyond the standard VEA approach (i.e., extra support during the outputs phase plus the outcomes phase).

We used a quasi-experimental evaluation design, purposefully comparing the SWIFT 1 project to the standard VEA approach. We assessed impact among households in rural communities where the project was active compared to households in similar communities where the national VEA approach had been implemented. Therefore, impacts measured with this approach are not reflective of the entire project, but only community-level aspects that were done in addition to the standard VEA approach. This means that impacts of broader activities conducted across the entire area (e.g., radio programming, regional events) are also not within the scope of this evaluation.

The evaluation was carried out in North Kivu in two Health Zones – Kirotshe and Mweso – where Oxfam led the implementation together with PPSSP (outputs phase) and HYFRO (outcomes phase) (i.e., we did not visit areas where Tearfund was the lead implementer). Other project areas where Oxfam led the implementation in North Kivu were excluded because of insecurity and the ongoing Ebola outbreak. Within these two Health Zones, the intervention group included all 14 villages across seven Health Areas where all three services (water, sanitation and hygiene) were completed during SWIFT 1 (according to the payment by results [PbR] contract, at least two services had to be completed in each village; the consortium was already completing the remaining services as part of a follow-on project, SWIFT 2). The comparison group consisted of 20 randomly selected villages (out of 58 possible where the national VEA approach had been implemented) in the same Health Zones (i.e., Kirotshe and Mweso), but not in the same Health Areas as the intervention group or SWIFT 2 (to reduce possible spillover effects).

Households were randomly selected for interviews (among the entire population within each village) using a random walk protocol. To ensure representation of women and men's perspectives, the gender of the interviewee representing each household was randomly allocated by SurveyCTO (Pretari, 2018). A total of 1,073 household interviews were completed – 441 in the intervention group and 632 in the comparison group. During analysis, propensity score matching (PSM) and multivariate regression were used to account for apparent baseline differences (using recalled baseline data) between the groups. Unexpected differences between the intervention and comparison groups required filtering prior to analysis, reducing the sample size to 448 (214 in the intervention group and 234 in the comparison group).

The primary aim of the evaluation was to investigate 'the extent to which inequalities brought about by poor water access, management and governance have been reduced and the significance of the intervention's contribution', as the final outcome indicator for all Sustainable Water Effectiveness Reviews. We measured this concept using a multidimensional Sustainable Water and Sanitation Index comprised of six dimensions – Water Security, Equity, Institutions, Operations, Well-Being and The Environment. We considered access to water and sanitation infrastructure using standard definitions for unimproved and improved, considering piped water both on premises with other types of improved sources and separately (WHO/UNICEF, 2017). We assessed water insecurity using the Household Water Insecurity Experiences (HWISE) Scale (HWISE, 2019).

# RESULTS

In comparing SWIFT 1 to the standard VEA approach in DRC, this evaluation found limited evidence of some positive impacts as well as some negative impacts. In both the intervention and comparison areas, improved water and sanitation access increased substantially, but at similar rates. Improved water access increased by 22% in SWIFT 1 areas and by 17% in VEA areas. Improved sanitation access increased by 9% in SWIFT 1 areas and by 6% in VEA areas. These rates exceed the average rate of improvement seen in rural areas of DRC between 2013 and 2017 (7% for water, 2% for sanitation; most recent data available) (WHO/UNICEF, 2019).

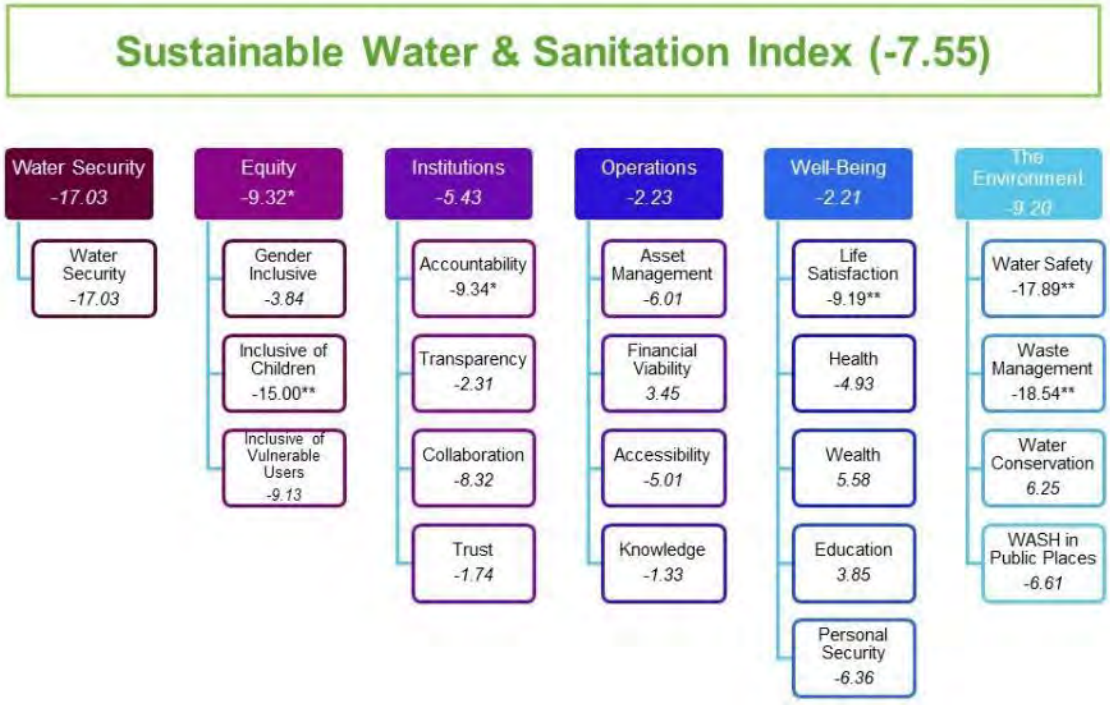
In terms of impact estimates using PSM, the overall Sustainable Water and Sanitation Index and most of the dimensions and indicators showed no significant impact for SWIFT 1 relative to the VEA programme. Positive differences attributable to SWIFT 1 included increased access to piped water (3.26, $p<0.10$) and a higher rate of school-aged children currently attending school (0.10, $p<0.10$). Other dimensions and indicators were found to be less favourable for SWIFT 1 areas in comparison to VEA areas, such as water insecurity (HWISE), Equity, Accountability, Life Satisfaction, Water Safety and Waste Management. Detailed results for these and other indicators are summarized in Table 1. A visual representation of the results for the Sustainable Water and Sanitation Index are shown by dimension and indicator Figure 1.

**Table 1: Summary of Effectiveness Review results**

| Outcome | Evidence of impact? |
|---|---|
|  Improved water access | Access to improved water infrastructure increased similarly in all areas; we see a significant impact specifically for piped water access on premises during the dry season for SWIFT 1 in comparison to VEA. |
|  Improved sanitation access | Access to improved sanitation increased similarly in SWIFT 1 and VEA areas. |
|  Reduced water insecurity | No significant difference for water insecurity between SWIFT 1 and VEA. |
|  Sustainable Water and Sanitation: Overall | No significant impact on the *Sustainable Water and Sanitation* Index overall. |
|  Sustainable Water and Sanitation: Water Security | We see a significant difference in this dimension for four indicators (but not overall), with lower Water Security in SWIFT 1 areas than in VEA areas. |
|  Sustainable Water and Sanitation: Equity | In the Equity dimension the difference between SWIFT 1 and VEA is negative, driven most by more negative perceptions of facilities being *Inclusive of Children* in SWIFT 1 areas. |
|  Sustainable Water and Sanitation: Institutions | The difference between SWIFT 1 and VEA in the Institutions dimension is minimally negative – nothing overall, but *Accountability* is significantly lower for SWIFT 1. |
|  Sustainable Water and Sanitation: Operations | We do not find a significant difference in the Operations dimension. |
|  Sustainable Water and Sanitation: Well-Being | Overall, there is no significant difference in the Well-Being dimension, although there is a significant negative impact for SWIFT 1 relative to VEA for *Life Satisfaction* and a significantly positive impact for school attendance. |
|  Sustainable Water and Sanitation: The Environment | The difference in The Environment dimension is negative, based on lower values for the *Water Safety* and *Waste Management* in SWIFT 1 areas than in VEA areas. |

By assessing differential impacts, we looked at how impacts varied by gender, age, disability of household members and the presence of school-aged children in the household. For example, we found a negative impact among men for Sustainable Water and Sanitation in SWIFT 1 areas versus VEA areas. Similarly, we saw a negative impact among households with members having a disability or chronic illness for water insecurity and access to improved sanitation. Among households with school-aged children we saw a significant positive differential impact on piped water and a negative impact for water insecurity and Sustainable Water and Sanitation Index for those in SWIFT 1 areas compared with VEA areas. We did not find any significant differential impacts by age.

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




This evaluation was limited by inaccessibility at the time of data collection in several areas where SWIFT 1 was implemented because of the ongoing Kivu Ebola outbreak and heightened security risk. Our results may be further biased by the fact that communities participating in SWIFT 1 were less comparable to communities participating in the VEA programme at baseline than expected (e.g., less wealthy, lower rates of improved water and sanitation access). The reasons for this are still unclear and should be examined further (e.g., through follow-up, see Programme Learning Considerations), as we were unable to include further follow up in the scope of this evaluation. Additionally, negative impacts on perception-based indicators, such as Equity, may have been influenced by increased awareness of existing inequities through additional project activities during the SWIFT 1 outcomes phase, although we were unable to confirm this. These and other limitations and possible biases are further explained in the Risk of Bias table (see Appendix 4).

# PROGRAMME LEARNING CONSIDERATIONS

## How change happens, and how fast it happens, is influenced by the unique social characteristics of each community

In seeking to understand the differences in impact between SWIFT 1 and the standard VEA approach, we assumed all the communities included in the evaluation were starting from a similar baseline situation, at least on average between the two groups. However, that proved not to be the case, and we found how important those baseline characteristics may have been for making progress on Sustainable Water and Sanitation. Communities that started out less wealthy, on average, and had lower rates of improved water and sanitation access, may have been able to realize larger improvements (e.g., in infrastructure). However, owing to the baseline situation, it may have been harder for other important aspects of Sustainable Water and Sanitation to stay aligned with that rate of improvement (e.g., Equity, Institutions, Asset Management) in communities with more vulnerable populations overall. With this in mind, it will be important in future programming to learn more about how all aspects of Sustainable Water and Sanitation can increase more together, in communities with different social characteristics.

## Based on the findings in this evaluation, follow up and reflect on possible interpretations together with communities

This evaluation shows that, despite the additional outcomes phase, SWIFT 1 achieved less progress overall in improving Sustainable Water and Sanitation in comparison to areas where the VEA programme was implemented. It would make more sense if adding activities increased impact, but that is not what we find. How might these additional activities have reduced impact? Or was the relatively lower impact due to systematically different community characteristics or other external factors? For example, did the way communities were selected to participate in SWIFT 1 influence project impact? To answer these questions, it would be necessary to follow up with the communities included in this evaluation (intervention and comparison) to discuss possible interpretations, for instance, through community reflection sessions.

## Consider further impact evaluation for the SWIFT programme in DRC

The SWIFT programme has evolved since SWIFT 1 was implemented, now nearing the end of the follow-on project SWIFT 2 and preparing for the next potential phase. The programme has incorporated evaluation, learning and adaptation from the beginning, which means issues identified in SWIFT 1 led to improvements in the design and implementation of SWIFT 2. An impact evaluation has not yet been conducted for SWIFT 2. However, doing so would further help in understanding to what extent those improvements have increased project impacts.

Also, in future projects look for ways to incorporate counterfactual-based and/or comparative approaches into regular programme monitoring, evaluation and learning. Impact evaluation can be incorporated in many ways. One option could be to compare those who participate in the project sooner (e.g., in the first year) to those who will participate later (e.g., in the second year), in cases where the project is rolled out in different areas at different times. Another option could be to systematically compare different versions of the project (e.g., two different types of asset management systems) to better understand which factors lead to greater impacts.

## Whenever possible, include externally validated indicators, such as JMP and HWISE, to measure project outcomes

The project used definitions to indicate whether SWIFT water sources and latrines were being used, which met the payment-by-results requirements. However, we are unable to directly compare such project-specific results to global indicators used in other evaluations, including this one, given the differences in definitions. We suggest gathering data in a way that will allow tracking of both project-specific indicators (e.g., for payment), as needed, and global indicators (e.g., to facilitate comparison across projects and programmes and to national statistics). The Joint Monitoring Programme (JMP) water and sanitation service ladders (WHO/UNICEF, 2017) and the Household Water Insecurity Experiences (HWISE) Scale (HWISE, 2019), both used in this evaluation, are two possible externally validated indicators.

# 1 INTRODUCTION

Every year since 2011, Oxfam Great Britain (GB) has conducted rigorous impact evaluations known as Effectiveness Reviews as part of our Global Performance Framework. For these reviews, we 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 2018/19, we selected projects under five thematic areas – Livelihoods, Women's Empowerment, Resilience, Good Governance, and Sustainable Water and Sanitation. The 'DRC Sustainable Water, Sanitation and Hygiene (WASH) in Fragile Contexts' project – in short, 'SWIFT 1 in DRC', and for simplicity referred to as SWIFT 1 throughout this report – was selected under the thematic area of Sustainable Water and Sanitation.

The SWIFT Consortium, led by Oxfam with Tearfund and the Overseas Development Institute (ODI) as members, carried out SWIFT 1 in DRC and Kenya to provide access to water and sanitation and to promote basic hygiene practices. In DRC, the consortium worked with implementing partners Hydraulique Sans Frontières (HYFRO; Hydraulics without Borders), Centre de Promotion Socio – Sanitaire (CEPROSSAN; Centre for Public Health Promotion), and Programme de Promotion des Soins de Santé Primaires (PPSSP; Primary Health Care Promotion Programme). The project operated in rural and semi-urban areas in three eastern provinces – North Kivu, South Kivu and Maniema. Oxfam led implementation in some areas of North Kivu, while Tearfund led in other areas of North Kivu as well as in South Kivu and Maniema. ODI provided Monitoring, Evaluation, Accountability and Learning (MEAL) support. With a budget of £5,206,147 under a payment by results (PbR) contract, the project reached around 859,000 people with water, 699,000 with sanitation, and 719,000 with hygiene, overachieving project targets for all three services.

The project was implemented in two phases – an outputs phase (April 2014 to March 2016) and an outcomes phase (April 2016 to March 2018). During the outputs phase, in rural areas (approximately 75% of project participants, and the focus of this evaluation) the project expanded on the national 'Villages et Ecoles Assainis' (VEA; Healthy Villages and Schools) approach. In this phase, project implementation followed the VEA approach in coordination with the government. Activities included rehabilitating and constructing water infrastructure, encouraging households to build latrines, helping to form and train local committees, promoting healthy hygiene behaviours and working to achieve a 'Healthy Village Certification'; SWIFT 1 incorporated some additional support, such as providing digging kits for latrines. SWIFT 1 also added the outcomes phase, which was not included in the VEA approach, to carry out additional capacity-building activities with the committees to increase project sustainability. Throughout the duration of the project, radio programming and regional events took place, for example, to promote healthy hygiene behaviours, such as handwashing. Note, the project followed a different approach in semi-urban areas (approximately 25% of project participants, and outside the scope of this evaluation).

The evaluation focuses on rural areas in North Kivu that were accessible at the time of the evaluation where Oxfam and partners PPSSP (outputs phase) and HYFRO (outcomes phase) carried out project implementation. In these areas, households live approximately one hour from the nearest market, on average, and livelihoods are based largely on agriculture and livestock. Safe and affordable access to water and sanitation services is an ongoing challenge. Other infrastructure, such as piped water and the electricity grid, have reached very few households in these areas. The evaluation does not assess impact in urban areas or in areas where Tearfund worked. These decisions are based on substantial operational research already underway in semi-urban areas and a desire to better understand the SWIFT 1 approach in relation to the standard VEA approach. We also restricted the sampling to locations where all three services (water, sanitation, hygiene) had

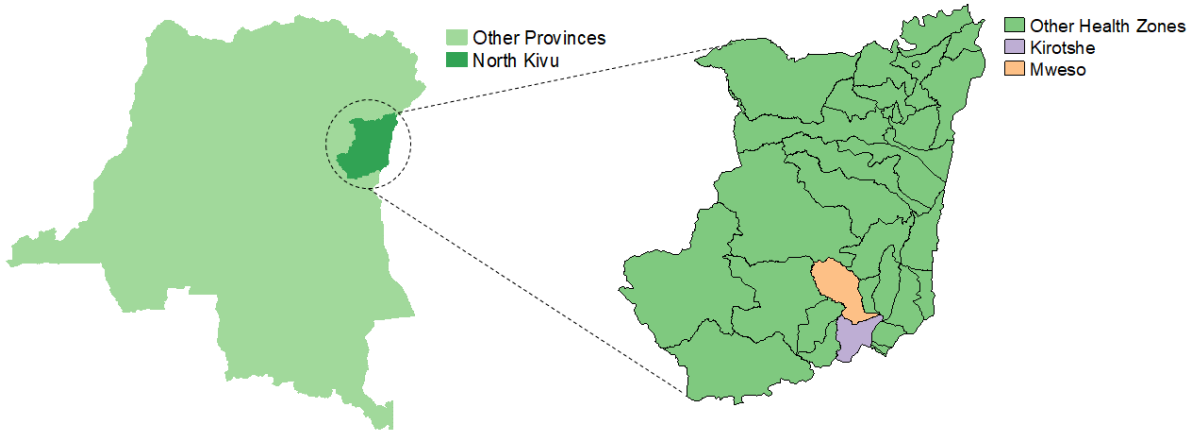
been implemented, to avoid interference with ongoing implementation activities by SWIFT 2 for remaining services.

The data collection phase, in March 2019, took place one year after the end of the project in March 2018. Owing to the ongoing Kivu Ebola outbreak that began in 2018 and increasing insecurity from armed groups and conflict in some project areas, our sampling frame was restricted to areas deemed safe to conduct data collection activities. These areas are in the southern part of the province, namely Kirotshe and Mweso Health Zones (see Figure 1.1.)

The primary aim of the evaluation was to investigate 'the extent to which inequalities brought about by poor water access, management and governance have been reduced and the significance of the intervention's contribution', as the final outcome indicator for all Sustainable Water Effectiveness Reviews. The questions guiding the evaluation were:

1. How did the project impact water and sanitation access and water insecurity at the household level?
2. How did the project impact Sustainable Water and Sanitation at the household level?
3. How do impacts differ:
    a. by gender?
    b. by age?
    c. for households with members who have disabilities or chronic illnesses?
    d. for households with school-aged children?

**Figure 1.1: Location of North Kivu in DRC (left) and the two Health Zones included in this evaluation (OCHA, 2019)**



# 2 PROJECT DESCRIPTION

## 2.1 PROJECT ACTIVITIES

Communities were selected to participate in SWIFT 1 based on discussions between Oxfam, the provincial government, the *Bureau Central de la Zone de Santé* (BCZ; Central Health Zone Office) and other stakeholders. The project was implemented following the VEA approach, but with additional activities. VEA was implemented in all other villages in the same Health Zones where SWIFT 1 was not implemented.

Project activities focused on improving water, sanitation and hygiene (WASH) systems, in terms of upgrading infrastructure and improving hygiene practices (during the outputs phase from April 2014 to March 2016) as well as strengthening management and governance structures (during the outcomes phase from April 2016 to March 2018).

The project activities are listed below. The main differences from VEA were the addition of a two-year period of capacity-building activities with water and sanitation committees, the distribution of digging kits for latrine construction and the addition of radio programming, regional events and two more years of hygiene promotion.

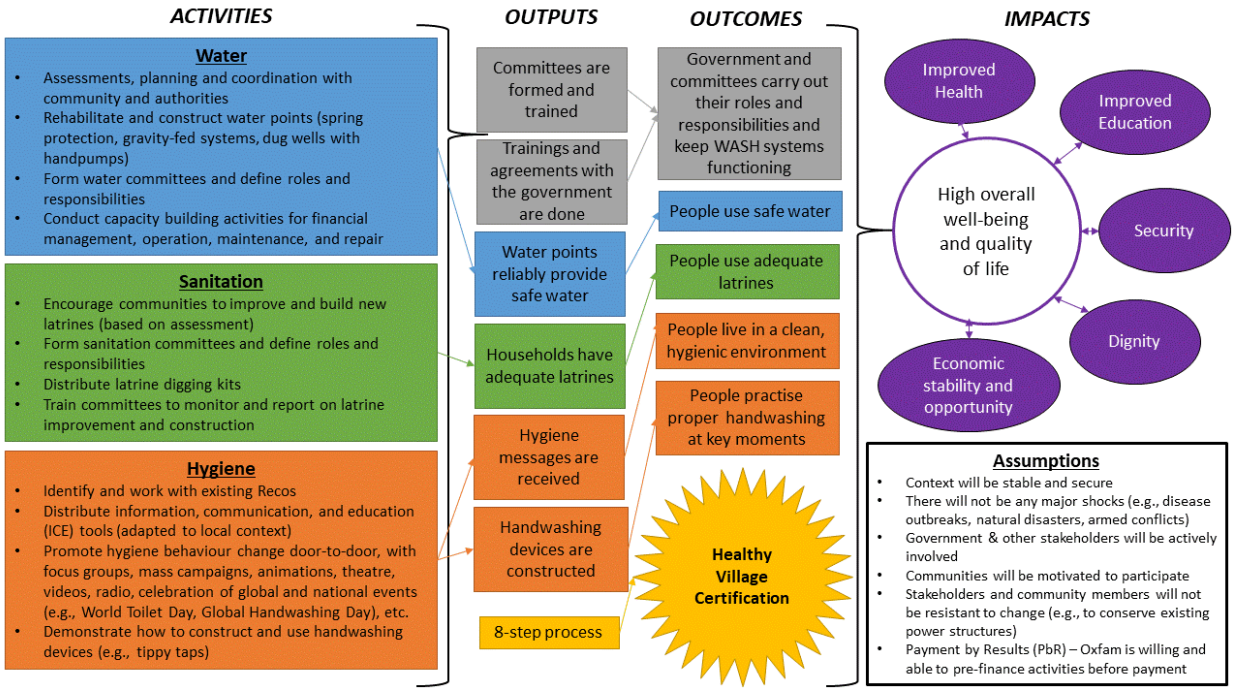
- Water assessments, planning and coordination with communities and authorities
- Rehabilitating or constructing water points (e.g., spring protection, gravity-fed systems, dug wells with handpumps) based on the assessments
- Forming water committees and defining roles and responsibilities
- Conducting capacity building activities with water committees on financial management, operation and maintenance (O&M) and repair
- Encouraging community members to build new latrines or improve existing latrines, based on an assessment of existing sanitation facilities
- Forming sanitation committees and defining roles and responsibilities
- Distributing latrine digging kits, as needed
- Training sanitation committees to monitor and report on latrine improvement and construction
- Identifying and working with *relais communautaires* (ReCos; community health workers)
- Distributing information, communication and education (ICE) tools that had been adapted to the local context
- Promoting hygiene behaviour change through door-to-door messages, focus groups, mass campaigns, animations, theatre, videos, radio and the celebration of national events (e.g., World Toilet Day, Global Handwashing Day, World Water Day), etc.
- Demonstrating how to construct and use handwashing devices (e.g., tippy taps)

All WASH activities were done in alignment with government guidelines for achieving a 'Healthy Village Certification'. Both SWIFT 1 and VEA followed the same eight-step process:

1. Community makes commitment
2. Knowledges, Practices & Attitudes (KPA) assessment of 8 WASH practices and norms
3. Participatory WASH conditions assessment
4. Confirmation that committees are in place
5. Develop a Community Action Plan (CAP) based on the KPA assessment
6. Implementation of the CAP
7. Final KPA assessment to verify 8 WASH practices and norms and the maintenance plan
8. Health Village Certification and flag-raising ceremony

# 2.2 THEORY OF CHANGE

**Figure 2.1 The project's theory of change (recreated with Oxfam and HYFRO during a workshop held 19–20 February 2019 in Goma, DRC)**




The project activities, which fit into the three broad categories of water, sanitation and hygiene, expected outputs regarding the structures and systems being put in place including:

- Committees are formed and trained
- Trainings and agreements with the government are done
- Water points reliably provide safe water
- Households have adequate latrines
- Hygiene messages are received
- Handwashing devices are constructed

In terms of the functioning of those structures and the use of those systems, the following outcomes were expected:

- The government and committees carry out their roles and responsibilities and keep WASH systems functioning
- People use safe water
- People use adequate latrines (according to national standards)
- People live in a clean and hygienic environment
- People practise proper handwashing at key moments
- Communities achieve and maintaining Healthy Village Certification status

These project outcomes aimed for impact – positive change in the lives of individuals and their communities. The expected impacts were: improved health, improved education, security, dignity, and economic stability and opportunity. Together, these impacts were expected to increase overall well-being and quality of life.

The ToC made assumptions about how change would happen. The key assumptions raised during the workshop were as follows:

- The context will be stable and secure
- There will not be any major shocks (e.g., disease outbreaks, natural disasters, armed conflicts)
- The government and other stakeholders will be actively involved
- Communities will be motivated to participate
- Stakeholders and community members will not be resistant to change (e.g., to conserve existing power structures)
- Regarding the PbR contract, Oxfam is willing and able to pre-finance activities before payment is received

Given the fragile context in which the project was implemented, not all assumptions were fully met, which in turn reduced the overall impact. During the workshop, it was discussed that two of these assumptions did not hold true. Namely, the context did not remain stable and secure, and there was a major shock considering the ongoing Kivu Ebola outbreak that began in 2018 shortly after the project ended, together with increased insecurity and conflict across the region. However, the other assumptions were valid for the most part.

# 3 EVALUATION DESIGN

The central problem in evaluating the impact of any programme is understanding what changes are attributable to project activities versus what would have happened otherwise. In this Effectiveness Review, the situation in project areas was examined through quantitative household surveys, but clearly, we could not directly observe what the situation would have been without the project. This 'counterfactual' situation can 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) against those who were not (comparison group). If these two groups were as similar as possible before the project, observing the situation in the comparison group provides a good estimate of what the situation would have been without the project.

In the case of such interventions, an ideal approach from a statistical perspective would be to randomly assign individuals (or households, communities, etc.) to the intervention and comparison groups to minimize the probability of systematic differences between the groups and maximize confidence that any observed impacts were caused by the project. However, this approach is often not ideal for large-scale implementation, and was not possible ex-post. Thus, we adopted a 'quasi-experimental' evaluation design using *propensity score matching (PSM)* to answer the evaluation questions for those in the intervention group in contrast to others who are similar in the comparison group. 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. To ensure sufficient data for the matching process, we interviewed three comparison interviewees for every two intervention interviewees.

For this evaluation, there were no areas similar enough to SWIFT 1 where no WASH activities had taken place. VEA had been implemented in all rural areas within the same Health Zones. The best available option was to purposefully compare SWIFT 1 to VEA, which was also of interest to the programme team, to understand the impact of the additional activities carried out by SWIFT 1 beyond the VEA approach as described in Section 2.1. SWIFT 1 and VEA implementation in Kirotshe and Mweso began at the same time, with the outputs phase running from 2014 to 2016. In both cases, implementation was carried out by the same organization, PPSSP. At the end of this phase, VEA was completed. SWIFT 1 continued from 2016 to 2018 by implementing the outcomes phase of the project with a different organization, HYFRO. Thus, this evaluation compares SWIFT 1 to VEA, showing any differential impacts that would likely be mainly due to the addition of the outcomes phase.

The baseline data needed for matching were not available, so interviewees were asked about their situation in 2013 (the year before project implementation began in 2014), thereby creating recall data. While recall data may not be completely accurate (Nicola & Giné, 2013; Godlonton *et al.*, 2018), 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) when sufficient baseline data are not available.

Other key evaluation design choices for this evaluation were as follows:

- Evaluating at the household level owing to the large number of direct project participants and the programme team's desire to understand impact at this level
- Restricting the sampling to locations where all three services (water, sanitation, hygiene) had been implemented during SWIFT 1, to avoid interference with ongoing implementation activities by the follow-on project, SWIFT 2, which was implementing remaining services in villages where only two were completed during SWIFT 1 (per the PbR contract requirements)
- Focusing on project implementation led by Oxfam, PPSSP and HYFRO in rural areas of North Kivu, and therefore not evaluating impact in urban areas or in areas where Tearfund worked, for the following reasons:
    - 75% of project implementation took place in rural areas
    - Substantial operational research was already underway in urban areas
    - The programme team's interest to understand the added value of the outcomes phase in relation to VEA, which is an approach used only in rural areas
    - Potential issues with including both rural and urban areas in this evaluation, given the substantive differences in project implementation and context, considering resources available (e.g., the added variation would require a larger sample size and probably different indicators)
    - Intent to evaluate impact in areas where Oxfam was lead consortium member

Overall, this evaluation design allows us to see project impacts at the household level. While we can look at differences in water and sanitation access overall, we can only make causal statements about the activities that SWIFT 1 did beyond the VEA approach as described in Section 2.2, which largely consist of the addition of further capacity building with committees during the two-year outcomes phase and additional support provided during the outputs phase (e.g., digging kits for latrine construction). Also, any impacts of broader activities conducted across the entire local area (e.g., radio programming, regional events) are not explicitly evaluated, since we would expect these activities to benefit households in both the intervention and comparison areas, and to do so would require a different evaluation methodology, which was not possible to integrate into this evaluation at the desired level of confidence, given the resources available.

# 4 DATA

## 4.1 INTERVENTION AND COMPARISON

Oxfam implemented SWIFT 1 in DRC in 10 Health Zones in North Kivu. Owing to insecurity and Ebola in the *Grand Nord* (a region located in the northern part of North Kivu), only two Health Zones in the southern part of the province could be accessed by the data collection team. We visited SWIFT 1 project areas (intervention group) and VEA comparison areas (comparison group) in both of these Health Zones, Kirotshe and Mweso.

**Intervention Group:** Within Kirotshe and Mweso, SWIFT worked with 53 communities, 25 of which are in rural areas. The project was implemented as three services (water, sanitation and hygiene), with SWIFT 2 finishing the work started by SWIFT 1 in areas where all three had not yet been completed (e.g., sanitation and hygiene had been implemented during SWIFT 1 and water would be implemented during SWIFT 2). For consistency and clarity, the evaluation focuses on the 14 villages where all three services had been implemented during SWIFT 1. These villages are spread across seven different Health Areas.

**Comparison Group:** Data from the *Division Provincial de la Santé* (DPS; Provincial Health Division) was referenced to generate a list of possible comparison villages. All villages are in one of three categories – (1) SWIFT 1, (2) SWIFT 2, (3) VEA. We have automatically excluded SWIFT 1 and SWIFT 2 villages, along with 14 VEA villages in the same Health Areas as SWIFT 1 and SWIFT 2 (to reduce possible spillover effects). From the remaining 58 possible villages, 20 were randomly selected using stratification (i.e., 10 per Health Zone). These villages are spread across six different Health Areas.

The household and interviewee sampling strategy was the same in the intervention group and the comparison group. In the project areas, we had a stratified random sample of 400 households from 14 villages. In the comparison area, we had a stratified random sample of 600 households from 20 villages. The target sample size was 30 households in each village. For reference, Table 4.1.1 presents the sampling frame by Health Zone and Health Area; village names are not listed to protect the privacy of interviewees.

**Table 4.1.1: Sampling frame**

| Group | Health Zone | Health Area | Number of Villages | Sample Size (Target) |
|---|---|---|---|---|
| Intervention | Mweso | JTN | 1 | 30 |
| | | Kashuga | 1 | 30 |
| | | Mushebere | 3 | 90 |
| | | Rugarama | 2 | 60 |
| | Kirotshe | Kingi Buroha | 3 | 90 |
| | | Kirotshe | 1 | 30 |
| | | Matanda | 3 | 90 |
| Comparison | Mweso | Bukama | 8 | 240 |
| | | Mweso/Rugarama | 2 | 60 |
| | Kirotshe | Bweremana | 4 | 120 |
| | | Kihindo | 2 | 60 |
| | | Mitumbala | 1 | 30 |
| | | Mushaki | 3 | 90 |
| **Totals** | **2 Health Zones** | **13 Health Areas** | **34 Villages** | **1,020 Households** |

## 4.2 HOUSEHOLD SURVEYS

Following a three-day training course, a team of two supervisors and 28 interviewers (contracted and managed by evaluation consultant, CIS-INNST) conducted the household surveys from 28 February to 8 March 2019. The questionnaire was reviewed in French and Swahili 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 back-up. Throughout data gathering, privacy and data protection were prioritized, following principles and protocols based on Oxfam's Responsible Data Policy and GDPR (Vonk, 2019).

In each village, survey interviewees were randomly selected from all current residents using a random walk protocol as follows:

1. From the centre of the village, each interviewer spun a pen to indicate which direction to walk.
2. Each interviewer counted the households according to a pre-determined interval set by the supervisor (based on the target sample size and the number of households in the village).
3. At every $n^{th}$ household, the interviewer asked to interview a household member of a specified gender (randomly indicated by SurveyCTO) who was knowledgeable about the household's water and sanitation situation.
    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 temporarily not available, the interviewer asked if they could come back at a more convenient time.
    c. If a suitable interviewee either did not exist in the household or was away for an extended period, the interviewer proceeded to interview an interviewee of the opposite gender, if possible.
    d. In case of refusal, or if a suitable interviewee could not be found after two attempts, the interviewer recorded the reason for no interview and proceeded to the next household according to the protocol.

We aimed for a sample size of 1,020 (600 comparison, 420 intervention). Interviewee gender was randomly indicated for each household, so we expected 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 interviewers visited 1,097 households, of which eight had no interviewee available (e.g., away from home; 1 intervention, 7 comparison) and four did not consent to be interviewed (2 intervention, 2 comparison), thereby completing 1,085 surveys (623 comparison, 462 intervention). Of those surveys, 12 were dropped from analysis because of irreconcilable errors in the household roster. The final sample of interviewees who consented to and completed the survey is shown in Table 4.2.1.

**Table 4.2.1: Final sample of interviewees by gender**

| Group | Total Sample (N) | Men Interviewees | Women Interviewees |
|---|---|---|---|
| Intervention | 441 | 228 (52%) | 213 (48%) |
| Comparison | 632+ | 324 (51%) | 306 (48%) |
| **Total** | **1,073+** | **552 (51%)** | **519 (48%)** |

+Two interviewees self-identified as Other.

## 4.3 MATCHING PROCESS OVERVIEW

Our expectation that the comparison villages would be similar to the intervention villages at baseline was not met, despite using random selection from the only options available that met pre-set criteria within the two Health Zones. On average, households in the intervention group were much less likely to have had access to improved water for drinking or improved sanitation in 2013. Furthermore, the comparison group had more residents who were wealthier (in the 5th quintile), while the intervention area had more residents who were poorer (in the 1st quintile). These differences were large and found to be associated with village-level water, sanitation, and wealth characteristics.

For good matching, we had to first filter the data. We removed: (1) interviewees who did not live in the same village at baseline (165 dropped), (2) interviewees living in villages with more than 80% (310 dropped) or less than 20% (86 dropped) of households reporting improved water access at baseline, and (3) interviewees living in villages with a median distance to the nearest local market of less than 30 minutes (i.e., rural only) (50 dropped). Further filtering by wealth was not necessary as it was highly correlated with (2) and (3). Additionally, in the matching process 12 intervention and two comparison observations were dropped because no adequate matches existed. The final sample size is underpowered, meaning we are now less likely to find significant impacts that did, in fact, occur (i.e., risk of Type II error), which we would have seen with the planned sample size of 1,020. However, we can still be confident in the significant impacts we do observe. Table 4.3.1 shows the final sample sizes and gender distributions for each group.

**Table 4.3.1: Final sample of interviewees after filtering and PSM by gender**

| Group | Total Sample (N) | Men Interviewees | Women Interviewees |
|---|---|---|---|
| Intervention | 214 | 113 (53%) | 101 (47%) |
| Comparison | 234+ | 123 (53%) | 110 (47%) |
| **Total** | **448+** | **236 (53%)** | **211 (47%)** |

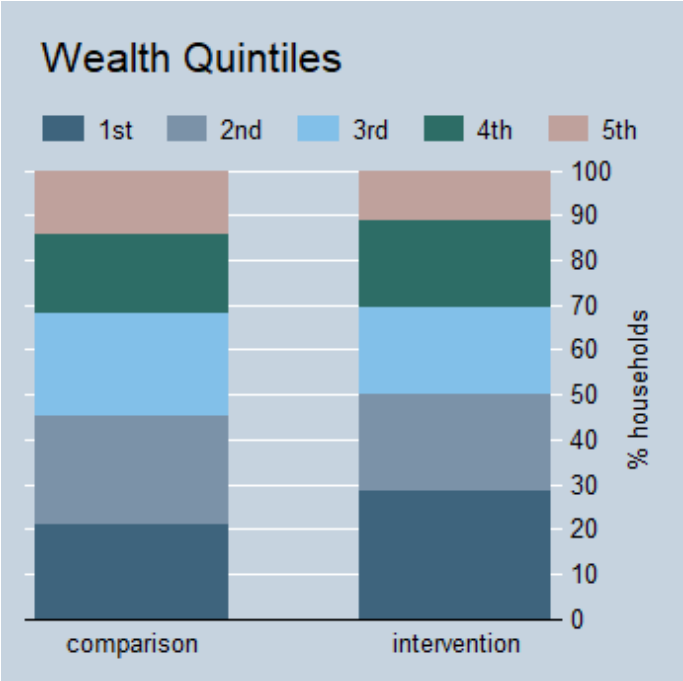
+One interviewee self-identified as Other.

An overview of the most pertinent information from PSM process and other descriptive information is discussed below. Detailed summary statistics for the intervention and comparison group before matching are provided in Appendix 2. Further details on how we did PSM and full specifications for this evaluation are available in Appendix 3. In short, before matching, we found several significant differences between intervention and comparison. By using PSM, we can adjust for these differences when estimating impacts. When we checked the balance variables after matching no significant differences remained.

One key matching variable we used is a wealth index based on household ownership of assets (e.g., furniture, livestock, equipment) and housing conditions in 2013. When generating the index, we first verified internal consistency using Cronbach's alpha, following the guidance of Bland and Altman (1997), and then used 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 ensured comparability of the wealth indexes from 2013 (based on recall data) and 2019 (based on the situation at the time of the survey), by pooling data by time period before undertaking PCA. We used 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 – in the same quintile – along with other matching variables, such as group participation, household head and interviewee gender and age, etc.).

The significant differences before matching tell us a few important things. Households in the intervention group had slightly fewer members and were more likely to own their home in 2013. The interviewees and the household heads in the intervention group were slightly younger, household demographics vary somewhat, and more household members were reported as fit to work. After filtering, those in the intervention group were still less likely to use an improved drinking water source or improved sanitation in 2013. The 2013 wealth quintiles show – again, even after filtering – that we can summarize by saying the intervention area had more poorer residents (in the 1st quintile); see Figure 4.3.1.

**Figure 4.3.1: Wealth quintiles in 2013 (difference for first wealth quintile is statistically significant, p<0.05; differences for other wealth quintiles are not statistically significant)**



# 5 MEASURING SUSTAINABLE WATER AND SANITATION

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'. For Sustainable Water Effectiveness Reviews, our aim is to investigate 'the extent to which inequalities brought about by poor water access, management and governance have been reduced and the significance of the intervention's contribution', as the final outcome indicator.

The SWIFT 1 in DRC project under review was specifically aimed at sustainably improving water and sanitation facilities and services. Building on our approach for Effectiveness Reviews when measuring other complex themes, such as Women's Empowerment (Bishop & Bowman, 2014; Lombardini *et al.*, 2017) and Resilience (Hughes & Bushell, 2013; Fuller & Lain, 2015), we gathered a set of 'dimensions' to create a household-level Sustainable Water and Sanitation Index. The dimensions shown in Table 5.1 are based on (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 frameworks and indexes relevant for measuring Sustainable Water and Sanitation (Banerjee & Morella, 2011; Bartram *et al.*, 2014; Bratton & Gyimah-Boadi, 2016; Giné-Garriga & Pérez-Foguet, 2018; Kayser *et al.*, 2013; Porteous, 2006; 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**

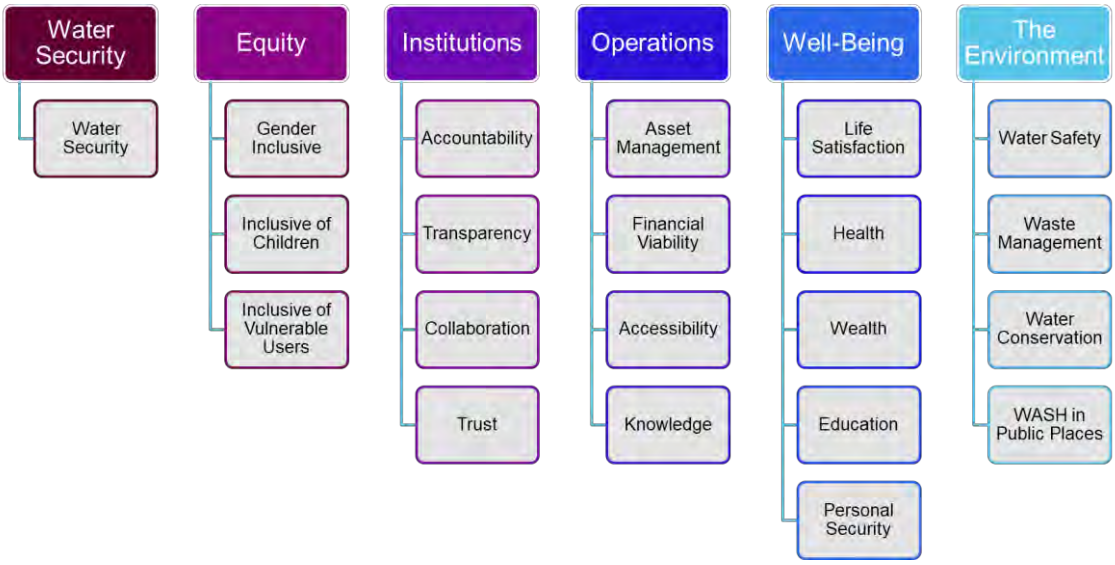
| Dimension | Definition |
|---|---|
| **Water Security*** | People have consistent and sufficient access to preferable and acceptable water; water is not a source of worry, shame, or anger. |
| **Equity** | Water and sanitation (WatSan) facilities, services, and related management systems and decision-making processes are inclusive 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 15 items, as indicated in Appendix 1 (HWISE, 2019; 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 a more holistic measurement of successes and failures (or both).

The index is designed to be adaptable to the given context, while maintaining consistency across contexts to enable meta-analysis. 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 always remain constant, but the set of indicators used to measure each dimension can vary to capture what Sustainable Water and Sanitation means in each context. However, 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 and HYFRO during a workshop held 19–20 February 2019 in Goma, DRC. It is constructed using a total of 21 indicators that are assessed using 143 questions.

**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 (equal to or greater than 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 is provided in Appendix 1. The table also shows whether each indicator is directly related to the project's theory of change.

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, individual 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. It has three indicators – *Gender Inclusive, Inclusive of Children,* and *Inclusive of Vulnerable Users*.

- Gender Inclusive is the average value across the following six variables: *watsancomm_who, womanleaders, equalwat_gender, equalsan_gender, watcollectgender, sanaccess_gender.*
- Inclusive of Children is the average value across the following six variables: *equalwat_children, equalsan_children, schoolwat, schoolsan, schoolhyg, watcollectadults.*
- Inclusive of Vulnerable Users is the average value across the following five variables: *pwdleaders, equalwat_disability, equalsan_disability, feeoptynwat, feeoptynsan.*

Then, **Equity** equals *Gender Inclusive* plus *Inclusive of Children* plus *Inclusive of Vulnerable Users,* divided by three.

Alongside analysis based on the six dimensions of the index, we also consider the following in relation to Sustainable Water and Sanitation in this evaluation:

- **Household water and sanitation access levels (i.e., JMP) (WHO/UNICEF, 2017, p. 9)**

  Access levels for water reflect on whether households report that their main source is unimproved, improved, or piped water on premises. Note that piped water on premises is both considered as an improved water source and separately on its own. 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.

  **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, unprotected 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 Experiences Scale (HWISE, 2019; Young *et al.*, 2020)**

  The Household Water Insecurity Experiences (HWISE) 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 water insecurity. The full list of HWISE Scale questions can be found in Appendix 1 (see the first 12 questions for the Water Security dimension; ignore the threshold column).

These measurement tools all focus on experiences and perspective of individuals and other household-level information. We acknowledge the importance of other system-level, sub-national and national-level indicators key to monitoring the overall sustainability of water and sanitation in these Health Zones and beyond, which fell outside the scope of this evaluation. To evaluate project impact in this case, we were mainly interested in understanding changes in the lives of individuals and their households, recognizing that households are made up of different individuals who have different lived experiences depending on structural inequalities and intersecting identities.

# 6 RESULTS

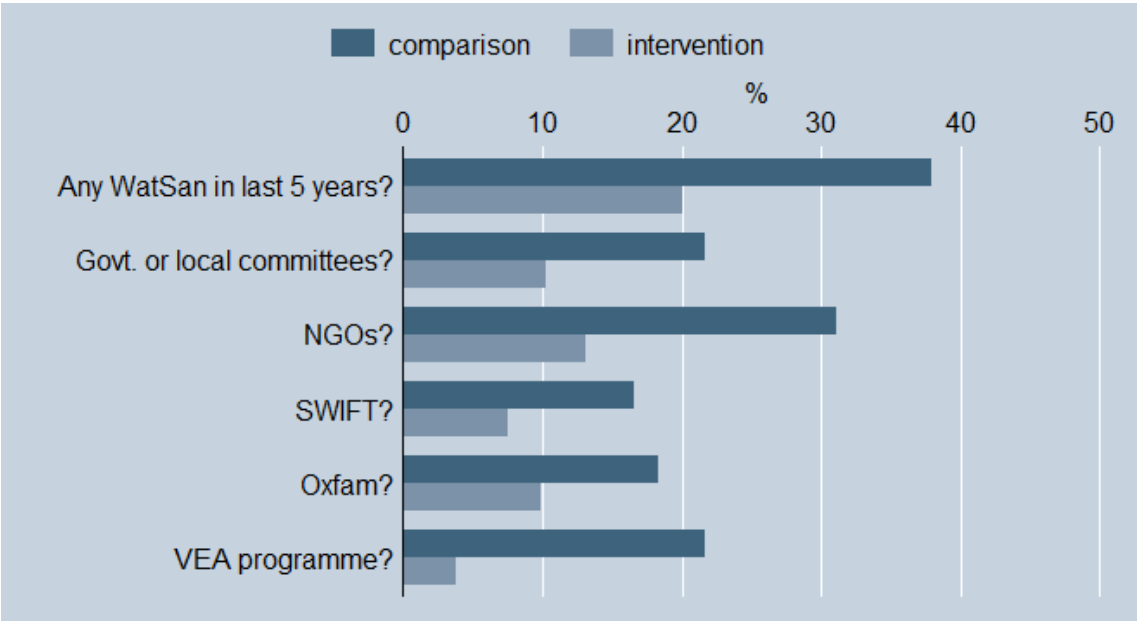
In this section, we present and discuss the results. Throughout this section, statistically significant differences in tables are highlighted in light green if positive, red if negative, and blue otherwise (e.g., descriptive differences rather than impacts). Differences that are not statistically significant are not highlighted. Note that all quantitative information is based on a final dataset from 448 interviewees after filtering and matching, as described in Section 4.3.

## 6.1 PROJECT EXPOSURE

In this section, we look at participation in water and sanitation activities for interviewees in the intervention and comparison groups. Interviewees were asked about their participation in these activities in the last five years (which would include the period of project implementation) and in the last one year (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, for the last five years, project exposure appears to be higher in the comparison group than in the intervention group as can be seen in Figure 6.1.1. Around 38% of comparison group interviewees said they had participated in water and sanitation activities (of any kind) in the last five years versus 20% in the intervention group. Those in the comparison group also said yes more often regarding each type of activity. In consultation with the project team, it was noted that there may have been some confusion between the different types of organizers (i.e., no one in the comparison areas participated in SWIFT, but perhaps said yes because they had participated in similar activities; many of the same Health Zone staff and NGO partners, including PPSSP, implemented the activities for the outputs phase of SWIFT 1 in DRC and the VEA programme). It is also possible that the quality of implementation, or the intensity of activities, varied between SWIFT and VEA areas, although this could not be confirmed.

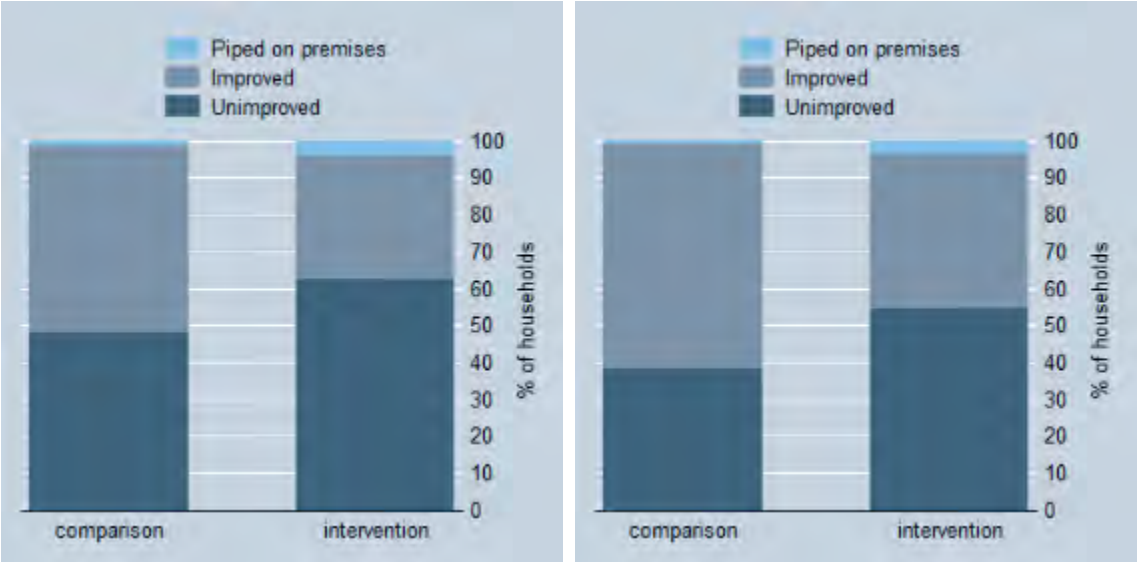
**Figure 6.1.1: Participation in water and sanitation activities in the last five years by group (descriptive only; all differences statistically significant, p<0.01)**



## 6.2 WATER AND SANITATION ACCESS

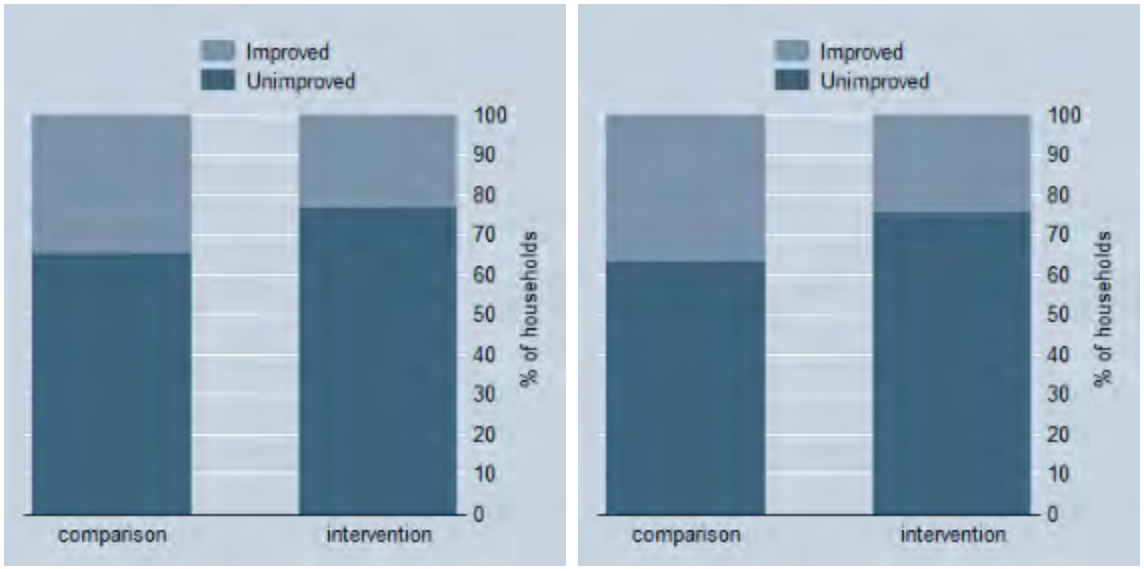
For context, we start with a descriptive overview of the water and sanitation situation based on the categories described in Section 5. Figure 6.2.1 shows the proportion of households by primary drinking water source during the dry season (results during the wet season were mostly the same, so for simplicity we focus on the dry season in this report). We see different levels of improved water access in the intervention and comparison areas before and after the project. In the intervention area, improved access (including piped on premises) went from 37% to 45%, while in the comparison area it went from 52% to 61% (i.e., 22% and 17% increase, respectively). In both cases, large increases in improved water access were achieved.

**Figure 6.2.1: Primary drinking water source (dry season) in 2013 (left) and 2019 (right) (differences in improved and unimproved statistically significant, p<0.01; differences in piped on premises statistically significant, p<0.10 level in 2013 and p<0.05 in 2019)**




Next, Figure 6.2.2 shows the proportion of households by sanitation facility category. Here we also see different levels of access in the intervention and comparison areas before and after the project. In the intervention area, improved access went from 23% to 25%, while in the comparison area it went from around 35% to 37% (i.e., 9% and 6% increase, respectively). Households in both groups were mostly using open/uncovered pit toilet/latrine facilities in 2013 and 2019.

**Figure 6.2.2: Sanitation facility used most in 2013 (left) and 2019 (right) (all differences statistically significant, p<0.01)**




Now, we estimate the impact of the project on water and sanitation access. Table 6.2.1 shows the PSM estimates for improved water access (inclusive of piped water on premises), piped water on premises and improved sanitation access. These estimates account for differences in the baseline characteristics, as explained in Section 6.1. We see a significant 3.26 percentage point increase in access to piped water on premises (during the dry season) due to the project activities under review. We do not observe significant impacts for improved water or sanitation access, which aligns with the similar gains in water and sanitation access in both SWIFT 1 and VEA areas during the project period, as described above.

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

| | Improved water+ (%) | Piped water (%) | Improved sanitation (%) |
|---|---|---|---|
| Intervention group mean | 45.33 | 3.74 | 24.77 |
| Comparison group mean | 50.63 | 0.48 | 29.60 |
| Difference (Impact) | -5.31 | 3.26* | -4.83 |
| Standard error | (8.61) | (1.68) | (8.96) |
| Observations (intervention group) | 214 | 214 | 214 |
| Observations (total) | 448 | 448 | 448 |

+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 1,000 repetitions.

# 6.3 HOUSEHOLD WATER INSECURITY

Next, we look at household water insecurity using the Household Water Insecurity Experiences (HWISE) Scale, which was described in Section 5. Table 6.3.1 shows the PSM estimates for water insecurity. Overall, water insecurity is higher in the intervention group, although the difference is not statistically significant. We also see higher water insecurity across each of the 12 items, but these differences are also not significant (see Table 6.3.2). Thus, the project activities under review did not have a significant impact on water insecurity.

**Table 6.3.1: Impact of the project on water insecurity**

| | Water Insecurity (HWISE) |
|---|---|
| Intervention group mean | 13.91 |
| Comparison group mean | 9.10 |
| Difference (Impact) | 4.66 |
| Standard error | (2.88) |
| Observations (intervention group) | 192 |
| Observations (total) | 407 |

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

**Table 6.3.2: Impact of the project on water insecurity for each item in the scale**

| | HWISE Scale Item |
|---|---|
| HWISE Worry (item 1) | 0.37 |
| HWISE Interrupt (item 2) | 0.39 |
| HWISE Clothes (item 3) | 0.42 |
| HWISE Plans (item 4) | 0.38 |
| HWISE Food (item 5) | 0.24 |
| HWISE Hands (item 6) | 0.22 |
| HWISE Body (item 7) | 0.33 |
| HWISE Drink (item 8) | 0.38 |
| HWISE Angry (item 9) | 0.36 |
| HWISE Sleep (item 10) | 0.35 |
| HWISE No Water (item 11) | 0.19 |
| HWISE Shame (item 12) | 0.26 |

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

# 6.4 SUSTAINABLE WATER AND SANITATION

Now, we move on to 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, which indicates a lower score for the overall index in the intervention group. However, this difference is not significant, meaning the project activities under review did not have a significant impact on Sustainable Water and Sanitation overall. Table 6.4.2 provides the PSM estimates for each dimension. Significant results for each dimension are discussed in more detail, focusing on significant differences observed, with **dimensions in bold** and *indicators in italics*. For a visual overview of the index and impact estimates, see Figure 6.4.1.

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

| | Sustainable Water and Sanitation Index |
|---|---|
| Intervention group mean | 48.27 |
| Comparison group mean | 55.82 |
| Difference (Impact) | -7.55 |
| Standard error | (4.92) |
| Observations (intervention group) | 214 |
| Observations (total) | 448 |

* 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 each dimension in the index**

| | Water Security | Equity | Institutions | Operations | Well-Being | The Environment |
|---|---|---|---|---|---|---|
| Intervention group mean | 34.43 | 49.53 | 48.64 | 59.25 | 50.50 | 47.43 |
| Comparison group mean | 51.51 | 58.85 | 54.07 | 61.48 | 52.71 | 56.63 |
| Difference (Impact) | -17.03 | -9.32* | -5.43 | -2.23 | -2.21 | -9.20 |
| Standard error | (11.98) | (5.35) | (6.89) | (3.41) | (3.10) | (5.78) |
| Observations (Intervention group) | 208 | 214 | 214 | 214 | 214 | 214 |
| Observations (total) | 438 | 448 | 448 | 448 | 448 | 448 |

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

The overall impact for **Water Security** is not significant, although the negative difference aligns with that of the 12-item HWISE Scale, confirming lower water security in the intervention group (recall that the HWISE Scale is the sum of 12 items, while the Water Security dimension of the Sustainable Water and Sanitation Index is an average of 15 items meeting the threshold as described in Appendix 1). Four of the 15 individual questions do show a significant negative impact for the project activities under review as follows: interrupted or limited water supply (*hwise2_i*), not having enough water to wash clothes (*hwise3_i*), having to change schedules or plans due to problems with water (*hwise4_i*) and not having as much water to drink as anyone in the household would like (*hwise8_i*). These differences provide some indication that Water Security increased more in VEA areas than in SWIFT 1 areas.

For **Equity**, we see a significant negative impact (-9.32, $p<0.10$) that is largely driven by a negative result for *Inclusive of Children* (-15.00, $p<0.05$), as well as negative but insignificant results for the other two indicators. The data suggest the difference is due to fewer interviewees in the intervention group agreeing that the sanitation needs of children and people with disabilities and/or chronic illnesses are being met, and fewer saying that schools have a handwashing facility with soap available. This suggests the project activities under review decreased perceptions of equity; however, this finding could also be the result of additional activities during the outcomes phase leading to increased awareness of existing inequities.

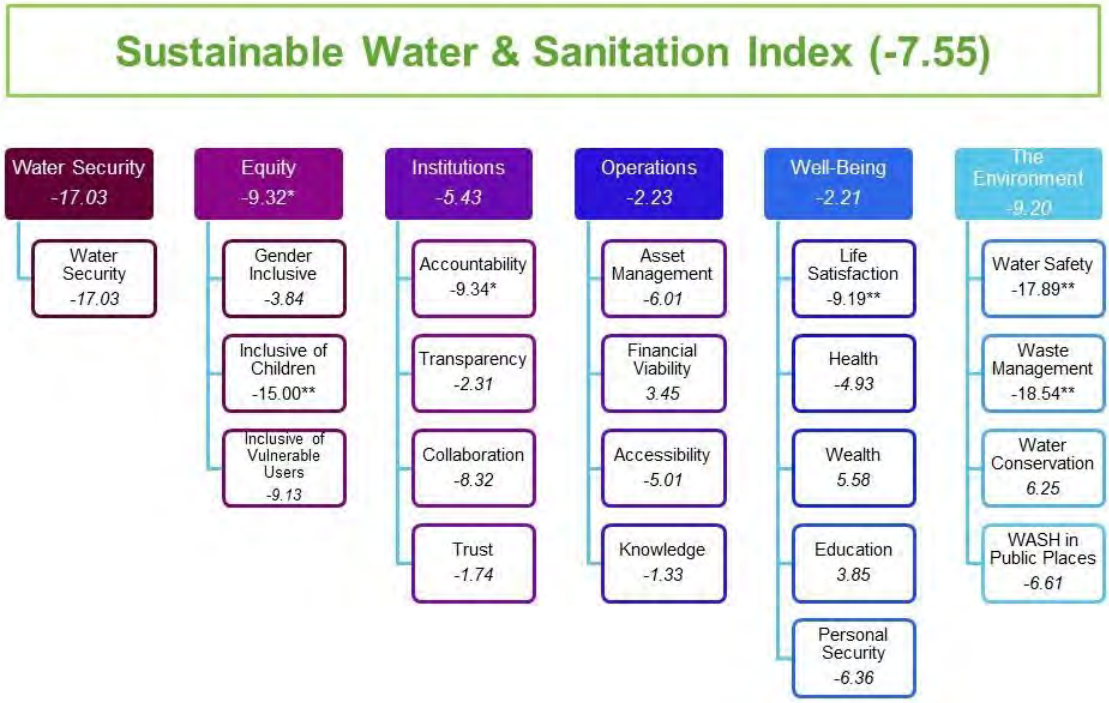
We do not see a significant impact overall for **Institutions**. This null result is driven by insignificant differences between the intervention and comparison for *Transparency, Collaboration* and *Trust*. However, *Accountability* does show a significant negative impact, which appears to be largely because of lower overall satisfaction with water and sanitation services in SWIFT1 areas compared with VEA areas.

We also do not see a significant impact overall for **Operations**. This null result is driven by insignificant results for all four indicators – *Asset Management, Financial Viability, Accessibility* and *Knowledge*. A few individual questions show significant differences indicating that households in SWIFT 1 areas are more likely than those in VEA areas to pay a reasonable amount for water (less than 5% of all household expenses). At the same time, they are less likely to have a functional handwashing device with soap or a water storage lid and they are more likely to have shared rather than private sanitation facilities, and these facilities are more likely to experience frequent breakdowns.

In the **Well-Being** dimension, which shows a significant negative impact overall, one of the indicators – *Life Satisfaction* – is significantly negative (-9.19, $p<0.05$). This result is based on interviewees in the intervention group rating their overall satisfaction with life lower than those in the comparison group. We do not see any significant differences in terms of *Health, Wealth* or *Education*, apart from a higher rate of school-aged children currently attending school in SWIFT 1 areas in comparison with VEA areas (0.10, $p<0.10$).

For **The Environment**, two indicators show negative impacts, including *Water Safety* and *Waste Management*. For *Water Safety,* the data suggests that fewer households in the intervention group are using water protection techniques and fewer are aware of any community-level water safety measures. For *Waste Management,* in the intervention group the presence of solid waste was higher, and fewer interviewees were aware of any community-level sanitation and solid waste management systems/plans or Healthy Village/Open Defecation Free (ODF) certification.

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




## 6.5 IMPACTS FOR DIFFERENT SOCIAL GROUPS

We also looked for differential impacts by subgroup to see who experienced the effects of the project more, or less. We used propensity score weighting, as described at the end of Appendix 3, to understand how impact varies for different social groups. Three characteristics that shape social inequalities and vulnerability in the context of rural DRC are (1) gender, (2) age and (3) whether households have any members with a disability or chronic disease. In addition, because hygiene education is often conducted in schools, we also looked for differential impacts depending on (4) whether households have school-aged children.

For individual-level characteristics – interviewee gender and age – we assessed differential impact for the Sustainable Water and Sanitation Index, which consists of both individual- and household-level indicators. For household-level characteristics – whether a household has members with a disability or chronic illness or has school-aged children – we assessed differential impact for household-level indicators, including water and sanitation access and water insecurity, in addition to the Sustainable Water and Sanitation Index.

First, we looked at the results by gender – did women and men benefit from the project activities differently? Recall from Table 4.3.1 that 47% of interviewees are women in the intervention and comparison groups. In Table 6.5.1, we see a significant negative impact for men in the intervention group for Sustainable Water and Sanitation index. This means that, while we do not see a significant impact overall for the project activities under review, we do see a significantly lower index score for men in SWIFT 1 areas compared with men in VEA areas.

**Table 6.5.1: Impact of the project by interviewee gender**

| | Sustainable Water and Sanitation Index |
|---|---|
| Effect of being a woman in the comparison group | -3.07<br>(2.92) |
| Project impact for men | -9.42***<br>(2.22) |
| Differential project impact for women and men | 4.73<br>(3.17) |

* p < 0.1, ** p < 0.05, *** p < 0.01; Standard errors are in parentheses. PSM estimates are bootstrapped with 1,000 repetitions. +During the dry season; improved water includes piped water on premises.

Next, we assessed the results by age for two age groups – 35 years and older and under 35 years (but at least 18 years, as all interviewees must be adults). In the intervention group, 51% of interviewees are 35 years and older, while in the comparison group that figure is 57%. In Table 6.5.2, we find a negative impact for the Sustainable Water and Sanitation Index for interviewees under 35 years old. This means that, while we do not see a significant impact overall for the project activities under review, we do see a significantly lower index score for younger people in SWIFT 1 areas compared with those in VEA areas.

**Table 6.5.2: Impact of the project by interviewee age**

| | Sustainable Water and Sanitation Index |
|---|---|
| Effect of being 35+ years in the comparison group | -0.71<br>(3.55) |
| Project impact for those <35 years | -7.08***<br>(2.30) |
| Differential project impact for older and younger interviewees | -0.29<br>(3.37) |

* p < 0.1, ** p < 0.05, *** p < 0.01; Standard errors are in parentheses. PSM estimates are bootstrapped with 1,000 repetitions. +During the dry season; improved water includes piped water on premises.

Robustness checks showed a slightly larger effect size (-8.89), but less significance (p<0.01) for 'Project impact for those <35 years', but the overall conclusion remained the same.

Next, we looked at impact for vulnerable groups, namely those with disabilities or chronic illnesses. However, our data are insufficient, as very few households reported having any such member. For this reason, we instead consider households reporting 80-100% of members are 'fit for work' against those with fewer than 80%. 'Fit for work' is indicated by the interviewee for each household member over the age of 5 when asked whether that person is 'fit and able to do domestic or livelihood work now if they wanted to?'

In the intervention group, 50% of interviewees report that less than 80% of their household members are fit to work, while in the comparison group the proportion is 62%. In Table 6.5.3, we see a positive impact for piped water similar to the overall impact. We also see significant impacts for water insecurity and Sustainable Water and Sanitation. The only differential impact that is significant is for access to improved sanitation. These results mean that households having members with a disability had increased access to piped water, lower access to improved sanitation, higher water insecurity and lower Sustainable Water and Sanitation Index scores in SWIFT 1 areas compared with those in VEA areas.

**Table 6.5.3: Impact of the project by household members' disability (based on question about being fit to work)**

| | Improved water+ (%) | Piped water (%) | Improved sanitation (%) | Water Insecurity (HWISE) | Sustainable Water and Sanitation Index |
|---|---|---|---|---|---|
| Effect of having household members with a disability in the comparison group | -3.13<br>(5.12) | -0.54<br>(0.99) | 13.98**<br>(5.46) | -0.29<br>(1.70) | 0.15<br>(2.76) |
| Project impact among households having members with a disability | -7.05<br>(5.32) | 3.39*<br>(2.05) | 0.41<br>(3.76) | 5.63***<br>(1.41) | -9.73***<br>(1.93) |
| Differential project impact for households with or without members with a disability | 4.57<br>(7.48) | 0.08<br>(3.00) | -12.34*<br>(6.47) | -1.12<br>(2.25) | 4.96<br>(3.45) |

* p < 0.1, ** p < 0.05, *** p < 0.01; Standard errors are in parentheses. PSM estimates are bootstrapped with 1,000 repetitions. +During the dry season; improved water includes piped water on premises.

The last characteristic we considered was having school-aged children (7 to 17 years old) in the household (i.e., those with no school-aged children versus those with at least one). Because hygiene education is conducted more often in schools, this analysis seeks to understand what impact this type of learning might have had at the household level. In the intervention group, 60% of households have at least one school-aged child, while in the comparison group 70% of households do. In Table 6.5.4, we find a differential impact for piped water among households with school-aged children compared to those without. We also see significant impacts indicating higher water insecurity and lower Sustainable Water and Sanitation Index scores for households with school-aged children.

**Table 6.5.4: Impact of the project by presence of school-aged children in the household**

| | Improved water+ (%) | Piped water (%) | Improved sanitation (%) | Water Insecurity (HWISE) | Sustainable Water and Sanitation Index |
|---|---|---|---|---|---|
| Effect of having school-aged children in the household in the comparison group | -2.68<br>(6.64) | 0.15<br>(1.26) | 5.29<br>(6.72) | 0.19<br>(1.87) | -1.91<br>(2.84) |
| Project impact among households with school-aged children | -9.78<br>(6.40) | 0.58<br>(1.37) | -0.55<br>(5.38) | 6.03***<br>(1.68) | -8.73***<br>(2.41) |
| Differential impact for households with or without school-aged children | 8.10<br>(7.63) | 4.74*<br>(2.89) | -8.20<br>(7.14) | -1.64<br>(2.19) | 2.29<br>(3.37) |

* p < 0.1, ** p < 0.05, *** p < 0.01; Standard errors are in parentheses. PSM estimates are bootstrapped with 1,000 repetitions. +During the dry season; improved water includes piped water on premises.

# 7 CONCLUSIONS

## 7.1 CONCLUSIONS

This evaluation, in comparing SWIFT 1 to the standard VEA approach in DRC, found limited evidence of some positive impacts as well as some negative impacts. In both the intervention and comparison areas, improved water and sanitation access increased substantially, yet at similar rates. Improved water access increased by 22% in SWIFT 1 areas and by 17% in VEA areas. Improved sanitation access increased by 9% in SWIFT 1 areas and by 6% in VEA areas. These rates exceed the average rate of improvement seen in rural areas of DRC between 2013 and 2017 (7% for water, 2% for sanitation; most recent data available) (WHO/UNICEF, 2019).

In terms of impact estimates using PSM, the overall Sustainable Water and Sanitation Index and most of the dimensions and indicators showed no significant impact for SWIFT 1 relative to the VEA programme. Positive differences attributable to SWIFT 1 included increased access to piped water (3.26, $p<0.10$) and a higher rate of school-aged children currently attending school (0.10, $p<0.10$). Other dimensions and indicators were found to be less favourable for SWIFT 1 areas versus VEA areas such as water insecurity (HWISE), **Equity**, *Accountability*, *Life Satisfaction*, *Water Safety* and *Waste Management*.

By assessing differential impacts, we looked at how impacts varied by gender, age, the disability of household members and the presence of school-aged children in the household. For example, we found a negative impact among men for Sustainable Water and Sanitation in SWIFT 1 areas versus VEA areas. Similarly, we saw a negative impact among households with members having a disability or chronic illness for water insecurity and access to improved sanitation. Among households with school-aged children we see a significant positive differential impact on piped water and a negative impact for water insecurity and Sustainable Water and Sanitation Index for those in SWIFT 1 areas compared with VEA areas. We did not find any significant differential impacts by age.

This evaluation was limited by inaccessibility at the time of data collection in several areas where SWIFT 1 was implemented, due to the ongoing Kivu Ebola outbreak and heightened security risk. Our results may be further biased by the fact that communities participating in SWIFT 1 were less comparable to communities participating in the VEA programme at baseline than expected (e.g., less wealthy, lower rates of improved water and sanitation access). The reasons for this are still unclear and should be examined further (e.g., through follow-up, see Programme Learning Considerations), as we were unable to include further follow-up in the scope of this evaluation. Additionally, negative impacts on perception-based indicators, such as **Equity**, may have been influenced by increased awareness of existing inequities through additional project activities during the SWIFT 1 outcomes phase, although we were unable to confirm this. These and other limitations and possible biases are further explained in the Risk of Bias table (see Appendix 4).

# 7.2 PROGRAMME LEARNING CONSIDERATIONS

## How change happens, and how fast it happens, is influenced by the unique social characteristics of each community

In seeking to understand the differences in impact between SWIFT 1 and the standard VEA approach, we assumed all the communities included in the evaluation were starting from a similar baseline situation, at least on average between the two groups. However, that proved not to be the case, and we found how important those baseline characteristics may have been for making progress on Sustainable Water and Sanitation. Communities that started out less wealthy, on average, and had lower rates of improved water and sanitation access, may have been able to realize larger improvements (e.g., in infrastructure). However, owing to the baseline situation, it may have been harder for other important aspects of Sustainable Water and Sanitation to stay aligned with that rate of improvement (e.g., Equity, Institutions, Asset Management) in communities with more vulnerable populations overall. With this in mind, it will be important in future programming to learn more about how all aspects of Sustainable Water and Sanitation can increase more together, in communities with different social characteristics.

## Based on the findings in this evaluation, follow up and reflect on possible interpretations together with communities

This evaluation shows that, despite the additional outcomes phase, SWIFT 1 achieved less progress overall in improving Sustainable Water and Sanitation in comparison to areas where the VEA programme was implemented. It would make more sense if adding activities increased impact, but that is not what we find. How might these additional activities have reduced impact? Or was the relatively lower impact due to systematically different community characteristics or other external factors? For example, did the way communities were selected to participate in SWIFT 1 influence project impact? To answer these questions, it would be necessary to follow up with the communities included in this evaluation (intervention and comparison) to discuss possible interpretations, for instance, through community reflection sessions.

## Consider further impact evaluation for the SWIFT programme in DRC

The SWIFT programme has evolved since SWIFT 1 was implemented, now nearing the end of the follow-on project SWIFT 2 and preparing for the next potential phase. The programme has incorporated evaluation, learning and adaptation from the beginning, which means issues identified in SWIFT 1 led to improvements in the design and implementation of SWIFT 2. An impact evaluation has not yet been conducted for SWIFT 2. However, doing so would further help in understanding to what extent those improvements have increased project impacts.

Also, in future projects look for ways to incorporate counterfactual-based and/or comparative approaches into regular programme monitoring, evaluation and learning. Impact evaluation can be incorporated in many ways. One option could be to compare those who participate in the project sooner (e.g., in the first year) to those who will participate later (e.g., in the second year), in cases where the project is rolled out in different areas at different times. Another option could be to systematically compare different versions of the project (e.g., two different types of asset management systems) to better understand which factors lead to greater impacts.

## Whenever possible, include externally validated indicators, such as JMP and HWISE, to measure project outcomes

The project used definitions to indicate whether SWIFT water sources and latrines were being used, which met the payment-by-results requirements. However, we are unable to directly compare such project-specific results to global indicators used in other evaluations, including this one, given the differences in definitions. We suggest gathering data in a way that will allow tracking of both project-specific indicators (e.g., for payment), as needed, and global indicators (e.g., to facilitate comparison across projects and programmes and to national statistics). The Joint Monitoring Programme (JMP) water and sanitation service ladders (WHO/UNICEF, 2017) and the Household Water Insecurity Experiences (HWISE) Scale (HWISE, 2019), both used in this evaluation, are two possible externally validated indicators.

# 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. The table also shows if each indicator is directly (highlighted in green) or indirectly (highlighted in yellow) linked to the SWIFT 1 in the DRC project's theory of change (ToC).

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

| Indicator | Variable | Question | Threshold | ToC link? |
|---|---|---|---|---|
| Water Security (Average of 15 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<br>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? | | |
| | *hwise13_i* | In the last month, how frequently have you or anyone in your household drunk water that looked, tasted and/or smelled bad? | | |
| | *hwise14_i* | In the last month, how frequently have you or anyone in your household drunk water that you thought was unsafe? | | |
| | *hwise15_i* | In the last month, how frequently have you or anyone in your household been unable to access the water that you preferred? | | |

**Table A1.2: Equity dimension = (Gender Inclusive + Inclusive of Children + Inclusive of Vulnerable Users)/3**

| Indicator | Variable | Question | Threshold | ToC link? |
|---|---|---|---|---|
| Gender Inclusive (Average of 6 variables x 100) | *watsancom_who* | Calculation of the gender balance of household members that participate in a Water, Sanitation and/or Hygiene Committee and/or a Health Board/Committee | Both genders = 1<br>One gender = 0 | Indirect |
| | *womanleaders* | Which of the two statements do you agree with most:<br>1 – A woman can be a leader in making decisions for water and sanitation in my community.<br>2 – Men are by nature better leaders than women when it comes to making decisions about water and sanitation in my community. | Option 1 = 1<br>Option 2 = 0 | |
| | *equalwat_gender* | The water needs of people of all genders are being met. | True = 1<br>False = 0 | |
| | *equalsan_gender* | The sanitation needs of people of all genders are being met. | | |
| | *watcollectgender* | Who is responsible for collecting water for the household? (Men, Women, Boys, Girls, Other) | Both genders = 1<br>One gender = 0 | |
| | *sanaccess_gender* | Can all the members of your household use the main toilet or sanitation facility (excluding babies)? If not, who cannot use it (Men, Women, Girls, Boys) | | |
| Inclusive of Children (Average of 6 variables x 100) | *equalwat_children* | The water needs of children are being met. | True = 1<br>False = 0 | Indirect |
| | *equalsan_children* | The sanitation needs of children are being met. | | |
| | *schoolwat* | In your community, at schools, do they have safe water? | Yes = 1<br>No = 0 | |
| | *schoolsan* | In your community, at schools, do they have a toilet or sanitation facility available? | | |
| | *schoolhyg* | In your community, at schools, do they have a handwashing facility with soap available? | | |
| | *watcollectadults* | Who is responsible for collecting water for the household? (Men, Women, Boys, Girls, Other) | Adults only/adults and children = 1<br>Children only = 0 | |
| Inclusive of Vulnerable Users (Average of 5 variables x 100) | *pwdleaders* | Which of the two statements do you agree with most:<br>1 – A person with a disability can be a leader in making decisions for water and sanitation in my community.<br>2 – A person with a disability is not able to be a leader when it comes to making decisions about water and sanitation in my community. | Option 1 = 1<br>Option 2 = 0 | Indirect |
| | *equalwat_disability* | The water needs of people with disabilities and/or chronic illnesses are being met. | True = 1<br>False = 0 | |
| | *equalsan_disability* | The sanitation needs of people with disabilities and/or chronic illnesses are being met. | | |
| | *feeoptynwat* | Do you know of any options for those who are not able to pay for water? | Yes = 1<br>No = 0 | |
| | *feeoptynsan* | Do you know of any options for those who are not able to pay for sanitation? | | |

**Table A1.3: Institutions dimension = (Accountability + Transparency + Collaboration + Trust)/4**

| Indicator | Variable | Question | Threshold | ToC link? |
|---|---|---|---|---|
| Accountability (Average of 4 variables x 100) | *holdacc_idea* | What do you think you can do to encourage or pressure the government and service providers to increase the level and quality of water and sanitation services in your community? | At least one idea = 1<br>Don't know/Not sure = 0 | Indirect |
| | *holdacc_conf* | 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<br>Not confident = 0 | |
| | *watsansatisfaction* | Overall, how satisfied are you with your water and sanitation services and systems? | Very/somewhat satisfied = 1<br>Very/somewhat dissatisfied = 0 | |
| | *renewmandate* | Will you choose to renew the mandate of the committee at the end of their contract? | Yes = 1<br>No = 0 | |
| Transparency (Average of 6 variables x 100) | *info_yn* | Do you get any information about the management of your water and/or sanitation systems? | Yes = 1<br>No = 0 | Indirect |
| | *concern_yn* | Do you know who you can go to if you have questions or concerns about your water and/or sanitation systems? | | |
| | *watdecision_who* | Do you know who makes the decisions about your water services/systems? | | |
| | *watdecision_how* | Do you understand how decisions are made for your water services/systems? | | |
| | *sandecision_who* | Do you know who makes the decisions about your sanitation services/systems? | | |
| | *sandecision_how* | Do you understand how decisions are made for your sanitation services/systems? | | |
| Collaboration (Average of 2 variables x 100) | *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<br>No = 0 | Indirect |
| | *collaboration* | If the Water and Sanitation Committee is unable to manage any problems, how confident are you that local authorities will support them? | Very/somewhat confident = 1<br>Not confident = 0 | |
| Trust (Average of 7 variables x 100) | *trustlocgov* | How much would you say you trust: Local government | A lot/Somewhat = 1<br>Not at all/Just a little = 0 | Indirect |
| | *trustlocdev* | How much would you say you trust: Local Development Committee | | |
| | *trustwatsan* | How much would you say you trust: Water and sanitation service providers | | |
| | *trustpropowner* | How much would you say you trust: Property owner (landlord/landlady) | | |
| | *trustngos* | How much would you say you trust: Non-governmental organizations (NGOs) | | |
| | *trustcbos* | How much would you say you trust: Community-based organizations (CBOs) | | |
| | *wattreat* | How often do you treat water for drinking that you get from a public tap, standpipe, kiosk or piped connection? | Never = 1<br>Sometimes/Always = 0 | |

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

| Indicator | Variable | Question | Threshold | ToC link? |
|---|---|---|---|---|
| Asset Management (Average of 11 variables x 100) | *watsancomm* | Does your community have a Water, Sanitation and/or Hygiene Committee? | Yes = 1<br>No/Don't know = 0 | Direct |
| | *confidentoandm* | For water and sanitation systems, how confident are you in this committee to manage routine operation and maintenance to avoid breakdowns? | Very/somewhat confident = 1<br>Not confident = 0 | |
| | *confidentrepair* | For water and sanitation systems, how confident are you in this committee to manage major repairs after a breakdown? | | |
| | *confidentreplace* | 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* | For water and sanitation systems, how confident are you in this committee to manage through a crisis, shock or disaster? | | |
| | *shock_watdisrupt* | Has your community faced any shocks or disasters in the past year (related to water and sanitation)? If yes, did this shock or disaster cause any problems with your main drinking water source? | No, it kept working normally/Yes, it was a problem for less than 1 week = 1<br>Yes, it was a problem for 1 week or more/Yes, it is still a problem = 0 | |
| | *shock_sandisrupt* | Has your community faced any shocks or disasters in the past year (related to water and sanitation)? If yes, did this shock or disaster cause any problems with your main toilet or sanitation facility? | | |
| | *watbdfreq* | How often does it happen that you cannot use your main drinking water source because it is broken down? | Never/Almost never/Annually = 1<br>Monthly/Weekly/Daily = 0 | |
| | *sanbdfreq* | How often does it happen that you cannot use your main toilet or sanitation facility because it is broken down? | | |
| | *watbddur* | When it last happened that you could not use your main drinking water source because it was broken down, about how long was it broken down before it was repaired and functioning properly again? | It has never happened/Less than 1 week = 1<br>1 week or more = 0 | |
| | *sanbddur* | When it last happened that you could not use your main toilet or sanitation facility because it was broken down, about how long was it broken down before it was repaired and functioning properly again? | | |
| Financial Viability (Average of 16 variables x 100) | *feestructwat* | What is the fee structure for use of the main system that you use for water? | Fixed or variable = 1<br>Irregular or free = 0 | Direct |
| | *feestructsan* | What is the fee structure for use of the main system that you use for sanitation? | | |
| | *payinstallwat* | How important do you think it is for you to pay to install your own system for water? | Very/somewhat important = 1<br>Not important at all/Neither important nor unimportant = 0 | |
| | *payinstallsan* | How important do you think it is for you to pay to install your own system for sanitation? | | |

| | | | | |
|---|---|---|---|---|
| | *payfeewat* | How important do you think it is for you to pay for regular costs for your system (service fees, etc.) for water? | | |
| | *payfeesan* | How important do you think it is for you to pay for regular costs for your system (service fees, etc.) for sanitation? | | |
| | *paydifficultwat* | How difficult is process of paying for services for water? | Not difficult at all = 1<br>Very/somewhat difficult = 0 | |
| | *paydifficultsan* | How difficult is the process of paying for services for sanitation? | | |
| | *fairpricewat* | Do you think your main service is fairly priced for water? | Completely/Somewhat fair = 1<br>Completely/Somewhat unfair = 0 | |
| | *fairpricesan* | Do you think your main service is fairly priced for sanitation? | | |
| | *affordwat* | In the past 30 days (1 month), how difficult has it been for your household to come up with the money to pay for this service for water? | Not difficult at all = 1<br>Very/somewhat difficult = 0 | |
| | *affordsan* | In the past 30 days (1 month), how difficult has it been for your household to come up with the money to pay for this service for sanitation? | | |
| | *feeoptwat* | Do you know of any options for those who are not able to pay for water? | Yes = 1<br>No = 0 | |
| | *feeoptsan* | Do you know of any options for those who are not able to pay for sanitation? | | |
| | *expensewat_pct* | 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<br>>5% of total = 0 | |
| | *expensesan_pct* | 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 20 variables x 100) | *watimproved* | What is your main source for drinking water during the dry season? During the wet season? | Piped/Improved in both seasons = 1<br>Unimproved during one or both seasons= 0 | Direct |
| | *sanimproved* | What type of sanitation facilities do members of your household use most often now? | Closed pit/flush toilet = 1<br>Open pit/none = 0 | |
| | *waterdom* | What is the main source you use for other domestic purposes? | Piped/Improved = 1<br>Unimproved = 0 | |
| | *watsourceofficial* | Is your primary drinking water source during the dry season an official or government service? | Yes = 1<br>No = 0 | |
| | *watsource_sp* | Is your primary drinking water source during the dry season private, only for your household, or is it shared? | Private = 1<br>Shared = 0 | |
| | *san_sp* | Is your main toilet or sanitation facility private, only for your household, or is it shared? | | |

| | | | | |
|---|---|---|---|---|
| | *watsourceaccess* | Where is your main drinking water source located?<br>How long does it take to walk there?<br>What is the terrain like on the way?<br>How long do you have to wait in queues at the source before you get water? | On-plot/In the house OR less than 5-minute walk and wait (each) over flat and easy terrain = 1<br>Off-plot walking and/or waiting over 5 minutes (each) or walking over uneven terrain/steep hills = 0 | |
| | *sanloc* | Where is your main toilet or sanitation facility located? | In the house/On-plot = 1<br>Off-plot = 0 | |
| | *wataccess* | When your main drinking water source is functioning normally, how many hours per day are members of your household able to access it? | 24 hours = 1<br><24 hours = 0 | |
| | *sanaccess* | 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* | How satisfied are you with the privacy of this toilet or facility? | Very/somewhat satisfied = 1<br>Very/somewhat dissatisfied = 0 | |
| | *waterbd_alt* | What water source does your household use for drinking water when your main source is broken down or inaccessible? | Piped/improved = 1<br>Unimproved = 0 | |
| | *sanbd_alt* | What type of toilet or sanitation facility do members of your household use when your main one is broken down or inaccessible? | Closed pit/flush toilet = 1<br>Open pit/none = 0 | |
| | *watsourcesafeday* | Do you feel safe accessing this water source during the day? | Yes = 1<br>No = 0 | |
| | *watsourcesafenight* | Do you feel safe accessing this water source at night? | | |
| | *sansafeday* | Do you feel safe accessing this toilet or facility during the day? | | |
| | *sansafnight* | Do you feel safe accessing this toilet or facility at night? | | |
| | *hwdevice* | Do you have a functional handwashing device (tippy tap, etc.)? | | |
| | *hwsoap* | Interviewer: Make observation if possible. Does the household have any soap present for handwashing? | | |
| | *hwsoapbuy* | Is it easy for your household to find soap to buy locally? | | |
| Knowledge (Average of 6 variables x 100) | *watstorageclean* | When was your drinking water container last cleaned? | Within the last week = 1<br>More than 1 week ago/Don't know = 0 | Direct |
| | *watstoragelid* | Interviewer: Make observation if possible. Is the drinking water container covered with a lid? | Yes = 1<br>No = 0 | |
| | *hwknowledge* | When is it important for you to wash your hands? | Names at least 2 key times = 1<br>Names 1 or 0 key times = 0 | |
| | *committeestructure* | What type of water and sanitation committee(s) does your community have? | Structure known = 1 | |

| | | | | |
|---|---|---|---|---|
| | *feestructwatknown* | What is the fee structure for use of the main system that you use for water? | Don't know = 0 | |
| | *feestructsanknown* | What is the fee structure for use of the main system that you use for sanitation? | | |

**Table A1.5: Well-Being dimension = (Life Satisfaction + Health + Wealth + Education + Personal Security)/5**

| Indicator | Variable | Question | Threshold | ToC link? |
|---|---|---|---|---|
| Life Satisfaction (lifesatisfaction x 100) | *lifesatisfaction* | Overall, how satisfied are you with life these days? | Very/Somewhat satisfied = 1<br>Very/Somewhat dissatisfied = 0 | Direct |
| Health (Average of 8 variables x 100) | *health* | Overall, how would you describe your state of health now? | Very good/Good = 1<br>Very poor/Poor = 0 | Direct |
| | *stomachpain* | In the last 30 days (1 month), have you had any stomach pain (excluding menstrual pain, for women)? | No = 1<br>Yes = 0 | |
| | *fever* | In the last 30 days (1 month), have you had a fever? | | |
| | *diarrhoea* | I'm sorry to have to ask this, but in the last 30 days (1 month), have you had any diarrhoea? | | |
| | *unwell* | In the last 30 days (1 month), have you felt so unwell that you had to interrupt your normal daily activities (excluding menstrual pain, for women)? | | |
| | *typhoid* | Have you had typhoid (diagnosed at a health centre or hospital) in the last 1 year? | | |
| | *cholera* | Have you had cholera (diagnosed at a health centre or hospital) in the last 1 year? | | |
| | *expensehealth_pct* | 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<br>>5% of total = 0 | |
| Wealth (Average of 2 variables x 100) | *wealthnorm* | Calculation of wealth relative to other households (construction of the wealth index is described in Section 4.3). | Wealth Index (normalized)>0 = 1<br>Wealth Index (normalized)<0 = 0 | Direct |
| | *wealthincrease* | Calculation of wealth relative to baseline (construction of the wealth index is described in Section 4.3). | Wealth Index increased = 1<br>Wealth Index stayed the same or decreased = 0 | |
| Education (Average of 2 variables x 100) | *schoolnow* | Are all school-aged children in the household attending school? (calculation) | Yes = 1<br>No = 0 | Direct |
| | *absencewatsan* | Have any school-aged children in the household missed school in the past 1 month due to water and sanitation related issues? (calculation) | No = 1<br>Yes = 0 | |

| | | | | |
|---|---|---|---|---|
| Personal Security (Average of 7 variables x 100) | *security1* | In the last month, how frequently have you or anyone in your household felt unsafe or insecure while going to get water? | Never = 1<br>Rarely to Always = 0 | Direct |
| | *security2* | In the last month, how frequently have you or anyone in your household been threatened by someone while going to get water? | | |
| | *security3* | In the last 1 month, how frequently have you or anyone in your household been attacked while going to get water? | | |
| | *watsourcesafeday* | Do you feel safe accessing this water source during the day? | Yes = 1<br>No = 0 | |
| | *watsourcesafenight* | Do you feel safe accessing this water source at night? | | |
| | *sansafeday* | Do you feel safe accessing this toilet or facility during the day? | | |
| | *sansafnight* | Do you feel safe accessing this toilet or facility at night? | | |

**Table A1.6: The Environment dimension = (Water Safety + Waste Management + Water Conservation + WASH in Public Places)/4**

| Indicator | Variable | Question | Threshold | ToC link? |
|---|---|---|---|---|
| Water Safety (Average of 4 variables x 100) | *waterprotect* | What techniques do your household use to protect your main water source from contamination? | Any technique = 1<br>N/A = 0 | Indirect |
| | *environment1* | Water systems in my community are monitored to ensure water is safe to drink. | True = 1<br>False = 0 | |
| | *environment2* | Water safety plans exist and are enforced to protect drinking water sources in my community from contamination. | | |
| | *environment3* | Water points are fenced and animals are contained to keep drinking water safe. | | |
| Waste Management (Average of 5 variables x 100) | *clean* | Interviewer: Make observation if possible. How clean and hygienic is the area around the household. | Very/Somewhat clean = 1<br>Not clean = 0 | Direct |
| | *environment4* | Sanitation systems in my community are monitored to ensure wastewater/sewage is safely managed. | True = 1<br>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 (salongo) in my community. | | |
| | *environment7* | My village is certified as a Healthy Village and/or as Open Defecation Free (ODF). | | |
| Water Conservation (waterconserve x 100) | *waterconserve* | What techniques do your household use to reduce the amount of water you use? | Any technique = 1<br>N/A = 0 | Indirect |
| | *schoolwat* | In your community, at schools, do they have safe water? | Yes = 1 | Indirect |

| | | | | |
|---|---|---|---|---|
| WASH in Public Places (Average of 9 variables x 100) | *schoolsan* | In your community, at schools, do they have a toilet or sanitation facility available? | No = 0 | |
| | *schoolhyg* | In your community, at schools, do they have a handwashing facility with soap available? | | |
| | *hospitalwat* | In your community, at health centres and hospitals, do they have safe water? | | |
| | *hospitalsan* | In your community, at health centres and hospitals, do they have a toilet or sanitation facility available? | | |
| | *hospitalhyg* | In your community, at health centres and hospitals, do they have a handwashing facility with soap available? | | |
| | *marketwat* | In your community, at markets, do they have safe water? | | |
| | *marketsan* | In your community, at markets, do they have a toilet or sanitation facility available? | | |
| | *markethyg* | In your community, at markets, do they have a handwashing facility with soap available? | | |

# APPENDIX 2: SUMMARY STATISTICS BEFORE MATCHING

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

**Table A2.1: Selected balance variables before matching.**

| Variable | Intervention group mean | Comparison group mean | Difference | Standard error |
|---|---|---|---|---|
| Number of HH members | 3.90 | 4.06 | -0.16 | 0.20 |
| Number of HH members in 2013 | 5.01 | 5.51 | -0.50* | 0.21 |
| % of HHs that owned their home in 2013 | 81.42 | 72.88 | 8.53* | 3.90 |
| % of HHs using an improved drinking water source+ in the dry season in 2013 | 35.84 | 52.12 | -16.28*** | 4.57 |
| % of HHs using an improved drinking water source+ in the wet season in 2013 | 38.94 | 53.81 | -14.88** | 4.60 |
| % of HHs with piped water on their premises in the dry season in 2013 | 3.98 | 1.27 | 2.71 | 1.48 |
| % of HHs with piped water on their premises in the wet season in 2013 | 1.77 | 2.54 | -0.77 | 1.36 |
| % of HHs with improved sanitation in 2013 | 22.12 | 35.17 | -13.05** | 4.18 |
| % child HH members (<18) | 44.21 | 49.08 | -4.87* | 2.28 |
| % school-aged HH members (6-18) | 23.30 | 27.58 | -4.28* | 2.05 |
| % youth HH members (<30) | 29.79 | 24.02 | 5.77* | 2.37 |
| % elderly HH members (65+) | 3.03 | 5.51 | -2.48 | 1.55 |
| % women HH members | 50.54 | 50.37 | 0.17 | 1.90 |
| % HH members fit to work | 73.87 | 69.17 | 4.70* | 2.36 |
| % seriously disabled/chronically ill HH members | 0.85 | 1.43 | -0.58 | 0.49 |
| HH head age | 39.58 | 42.66 | -3.08* | 1.39 |
| HH head, % women | 21.78 | 20.94 | 0.84 | 3.83 |
| HH head, % fit for work | 91.15 | 86.02 | 5.13 | 2.96 |
| HH head, % in HH in 2013 | 95.58 | 94.49 | 1.08 | 2.03 |
| HH head, % completed primary school | 69.47 | 70.76 | -1.29 | 4.27 |
| HH head, % literate | 55.75 | 63.56 | -7.81 | 4.56 |
| Interviewee age | 36.41 | 39.69 | -3.28* | 1.39 |
| Interviewee, % women | 48.23 | 47.66 | 0.57 | 4.66 |
| Interviewee, % fit for work | 90.71 | 87.29 | 3.42 | 2.92 |
| Interviewee, % in HH in 2013 | 94.69 | 95.34 | -0.65 | 2.03 |
| Interviewee, % completed primary school | 26.99 | 28.39 | -1.40 | 4.17 |
| Interviewee, % married | 82.30 | 82.63 | -0.33 | 3.55 |
| % with income from salaried employment in government in 2013 | 7.08 | 5.51 | 1.57 | 2.26 |
| % with income from salaried employment in private sector and/or NGOs in 2013 | 1.33 | 1.69 | -0.37 | 1.14 |
| % with income from casual labour in 2013 | 26.99 | 25.85 | 1.14 | 4.11 |
| % with income from own business in 2013 | 7.52 | 7.20 | 0.32 | 2.44 |

| | | | | |
|---|---|---|---|---|
| % with income from farming, agriculture and/or animal husbandry in 2013 | 78.76 | 79.66 | -0.90 | 3.78 |
| % with income from remittances in 2013 | 3.10 | 2.97 | 0.13 | 1.60 |
| % with income from cross-border trading in 2013 | 0.88 | 0.00 | 0.88 | 0.61 |
| % in lowest wealth quintile in 2013 | 30.09 | 20.76 | 9.33* | 4.03 |
| % in second lowest wealth quintile in 2013 | 20.80 | 24.58 | -3.78 | 3.90 |
| % in middle wealth quintile in 2013 | 19.03 | 22.88 | -3.85 | 3.79 |
| % in second highest wealth quintile in 2013 | 18.58 | 17.80 | 0.79 | 3.60 |
| % in highest wealth quintile in 2013 | 11.50 | 13.98 | -2.48 | 3.11 |
| % of interviewees who participated in at least one group in 2013 | 47.79 | 46.61 | 1.18 | 4.66 |
| Number of groups interviewee participated in 2013 | 0.80 | 0.80 | 0.00 | 0.10 |
| **Observations** | **462** | | | |

Construction of the wealth index is described in Section 4.3. Variables dated 2013 are estimates, based on recall data. * p < 0.1, ** p < 0.05, *** p < 0.01

# APPENDIX 3: PROPENSITY SCORE MATCHING

The 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 intervention 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 a 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 due to one or more missing values for four households (1 comparison, 3 intervention). 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 matching with marginal effects, standard errors, and p-values**

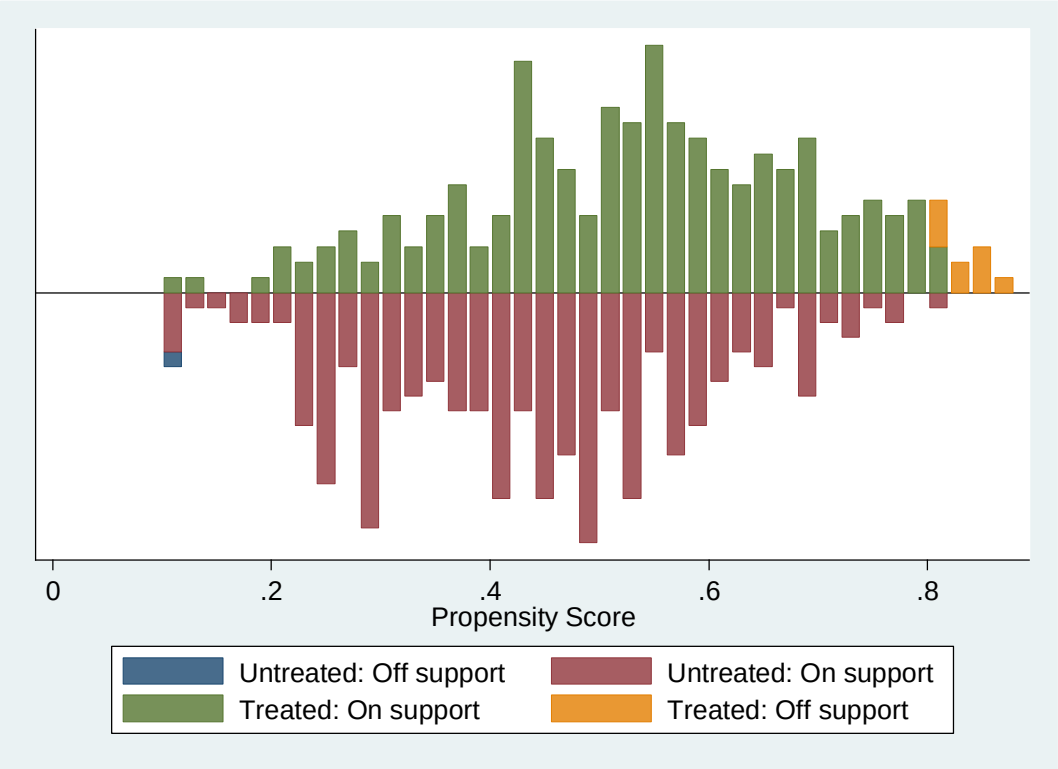
| Variable | Marginal effect | Standard error | p-value |
|---|---|---|---|
| Interviewee is a woman | 0.04 | 0.06 | 0.51 |
| Interviewee age (years) | -0.00 | 0.00 | 0.31 |
| Interviewee completed primary school | -0.02 | 0.07 | 0.72 |
| Interviewee is married | 0.04 | 0.08 | 0.65 |
| HH head is a woman | -0.02 | 0.08 | 0.81 |
| HH head age (years) | -0.00 | 0.00 | 0.66 |
| HH head completed primary school | -0.00 | 0.06 | 0.98 |
| HH head is literate | -0.12* | 0.06 | 0.04 |
| Number of HH members in 2013 | -0.02 | 0.01 | 0.25 |
| HH owned their home in 2013 | 0.17** | 0.06 | 0.01 |
| HH has at least 80% members over 5 years old fit to work | 0.19*** | 0.05 | 0.00 |
| HH earned income from salaried employment in govt. in 2013 | 0.07 | 0.11 | 0.49 |
| HH earned income from salaried employment in private sector and/or NGOs in 2013 | -0.01 | 0.20 | 0.96 |
| HH earned income from casual labour in 2013 | 0.04 | 0.06 | 0.54 |
| HH earned income from their own business in 2013 | -0.01 | 0.10 | 0.90 |
| HH earned income from farming, agriculture, husbandry in 2013 | -0.03 | 0.06 | 0.62 |
| HH earned income from remittances in 2013 | 0.04 | 0.14 | 0.80 |
| HH earned income from cross-border trading in 2013 | | | |
| HH was in the lowest 20% of wealth distribution in 2013 | 0.10 | 0.08 | 0.25 |
| HH was in the second lowest 20% of wealth distribution in 2013 | 0.01 | 0.08 | 0.86 |
| HH was in the second highest 20% of wealth distribution in 2013 | 0.08 | 0.08 | 0.30 |
| HH was in the highest 20% of wealth distribution in 2013 | 0.02 | 0.09 | 0.85 |
| Interviewee participated in a group in 2013 | 0.06 | 0.05 | 0.30 |
| Primary drinking water source (dry season) in 2013 (ordinal 0-2) | -0.15 | 0.08 | 0.06 |
| Primary drinking water source (wet season) in 2013 (ordinal 0-2) | 0.02 | 0.08 | 0.82 |
| HH had an improved sanitation facility (e.g., toilet) in 2013 | -0.12* | 0.06 | 0.03 |
| **Observations** | **458** | | |

Construction of the wealth index is described in Section 4.3. Variables dated 2013 are estimates, based on recall data. The dependent variable is binary, taking 1 for project participant households, and 0 otherwise. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

## 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 A3.1 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 reasonably good area of overlap between the groups. However, in constructing the model for household-level outcomes, 14 observations have been dropped (12 intervention, 2 comparison) 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 large, with the lowest value being 0.46 and most being more than 0.80. 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 | 0.47 | 0.45 | 0.67 |
| Interviewee age (years) | 37.07 | 37.10 | 0.98 |
| Interviewee completed primary school | 0.28 | 0.29 | 0.83 |
| Interviewee is married | 0.81 | 0.82 | 0.88 |
| HH head is a woman | 0.21 | 0.19 | 0.60 |
| HH head age (years) | 40.25 | 40.22 | 0.99 |
| HH head completed primary school | 0.70 | 0.72 | 0.66 |
| HH head is literate | 0.56 | 0.58 | 0.71 |
| Number of HH members in 2013 | 3.97 | 4.03 | 0.79 |
| HH owned their home in 2013 | 0.80 | 0.79 | 0.81 |
| HH has at least 80% members over 5 years old fit to work | 0.50 | 0.51 | 0.78 |
| HH earned income from salaried employment in govt. in 2013 | 0.07 | 0.07 | 0.99 |
| HH earned income from salaried employment in private sector and/or NGOs in 2013 | 0.01 | 0.01 | 0.99 |
| HH earned income from casual labour in 2013 | 0.28 | 0.24 | 0.46 |
| HH earned income from their own business in 2013 | 0.07 | 0.08 | 0.81 |
| HH earned income from farming, agriculture, husbandry in 2013 | 0.79 | 0.80 | 0.96 |
| HH earned income from remittances in 2013 | 0.03 | 0.04 | 0.70 |
| HH earned income from cross-border trading in 2013 | 0.00 | 0.00 | . |
| HH was in the lowest 20% of wealth distribution in 2013 | 0.29 | 0.28 | 0.87 |
| HH was in the second lowest 20% of wealth distribution in 2013 | 0.21 | 0.21 | 0.85 |
| HH was in the second highest 20% of wealth distribution in 2013 | 0.19 | 0.20 | 0.84 |
| HH was in the highest 20% of wealth distribution in 2013 | 0.11 | 0.13 | 0.67 |
| Interviewee participated in a group in 2013 | 0.47 | 0.49 | 0.79 |
| Primary drinking water source (dry season) in 2013 (ordinal 0-2) | 0.37 | 0.38 | 0.92 |
| Primary drinking water source (wet season) in 2013 (ordinal 0-2) | 0.39 | 0.39 | 0.92 |
| HH had an improved sanitation facility (e.g., toilet) in 2013 | 0.23 | 0.23 | 0.86 |
| **Observations** | **448** | | |

Construction of the wealth index is described in Section 4.3. Variables dated 2013 are estimates, based on recall data. The dependent variable is binary, taking 1 for project participant households, and 0 otherwise. * p < 0.1, ** p < 0.05, *** p < 0.01.

# Propensity score weighting

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

**PS-weighted OLS regression to estimate project impact using difference-in-difference for improved water, piped water and improved sanitation**

Where data for both recall and now are available, we used PS-weighted OLS regression to estimate project impact on the difference in each outcome – improved water, piped water, improved sanitation. We calculated this difference by subtracting the 2013 (recall) value from the 2019 (now) value. All three of these outcomes are binary, meaning the difference variables can take the following three values: -1 (the situation got worse), 0 (the situation stayed the same), 1 (the situation got better). The results of the regression tell us the difference in this difference, controlling for the PSM matching variables and with PS-weighting.

**PS-weighted OLS regression with an interaction term to estimate differential impacts for subgroups**

The average overall effect was estimated using PSM. To test for differential impacts, we used PS-weighted OLS regression together with an interaction term (all matching variables used for PSM are 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. **Age:** the interaction term multiplies a variable for interviewee age (equal to 1 if 35 years and older, 0 if under 35 years, but at least 18 years, as all interviewees must be adults) by the intervention variable (equal to 1 if in the intervention group, 0 if in the comparison group). Therefore, we define people 35 years and older as the group of focus and people under 35 years as the reference group.
3. **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 members with disabilities as the reference group.
4. **School-aged children:** the interaction term multiplies a variable indicating whether there are any school-aged children (7 to 17 years old) in the household (equal to 1 if there is at least one school-aged child, equal to 0 if there are none) by the intervention variable (equal to 1 if in the intervention group, 0 if in the comparison group). Therefore, we define households with school-aged children as the group of focus and households without school-aged children as the reference group.

The tables in Section 6.5 provide the results of 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**

| Sustainable Water & Sanitation Index | Example results | Interpretation |
|---|---|---|
| Effect among the group of focus in the comparison group, relative to the reference group | 2.50**<br>(2.10) | The first row shows the coefficient, standard error and significance for *the subgroup variable* (i.e., gender, disability, home ownership status). 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, while the size of the coefficient tells us how large this difference is (e.g., between women and men). The standard error and significance indicate to what extent this difference is statistically meaningful.<br>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***<br>(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 to what extent this difference is statistically meaningful.<br>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<br>(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 to what extent the difference in impact between the group of focus and reference group is statistically meaningful.<br>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 ten 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:<br>Sampling is conducted using probability random sampling methods on a clearly established sample frame.<br><br>Score MEDIUM risk if:<br>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.<br><br>Score HIGH otherwise. | MEDIUM | Interviewee selection was done using stratification by Health Zone and village using a random walk protocol (see Section 4). |
| 2 | Representativeness of project participants | Score LOW risk if:<br>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.<br>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.<br><br>Score MEDIUM risk if:<br>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.<br><br>Score HIGH otherwise. | MEDIUM | Interviewee selection was only done in areas where all three services (water, sanitation and hygiene) were implemented during SWIFT 1 for consistency; however, this was also necessary because the follow-on project (SWIFT 2) was implementing the remaining services that were not yet implemented in these areas. |

| | | | | |
|---|---|---|---|---|
| 3 | Selection survey interviewees | Score LOW risk if:<br>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).<br>The resulting selection of survey interviewees is not affected by project participation (based on observables).<br><br>Score MEDIUM risk if:<br>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).<br>The resulting selection of survey interviewees is affected by project participation (based on observables).<br><br>Score HIGH otherwise. | LOW | All community members were included in the sampling frame. Within each village, households were equally likely to be selected for an interview. |
| *Selecting comparison group* | | | | |
| 4 | Potential for contamination (spillovers) | Score LOW risk if:<br>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.<br>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).<br><br>Score HIGH risk if:<br>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 | The comparison group includes villages within the same Health Zones but not with in the same Health Areas. No comparison villages were bordering intervention villages.<br><br>The project did implement some activities, including radio programming and regional events, that may have also had an impact in the comparison group. The report makes clear that any impact of these activities is not measured in this evaluation (see Section 3). |
| 5 | Self-selection of project participants | Score LOW risk if:<br>The comparison group is exploiting an experiment or natural experiment.<br>Units are randomly selected at community level both in the intervention and comparison group.<br>The selection process for the comparison group is mimicking the same selection process used by the project. | 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 random walk method in the intervention and comparison groups (see Section 4). |

| | | | | |
|---|---|---|---|---|
| | | Score MEDIUM risk if<br>If the self-selection is corrected during the matching procedure (e.g., controlling for group participation at baseline)<br><br>Score HIGH risk if:<br>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:<br>There are no other actors in the area (e.g., INGOs, NGOs, governmental programmes)<br>Other actors are conducting activities which are not linked to the project's theory of change.<br><br>Score MEDIUM risk if:<br>Other actors are conducting similar activities linked to the project's theory of change in both the intervention and the comparison group.<br><br>Score MEDIUM-HIGH risk if:<br>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 group making it clear that the impact is compared with these other activities (e.g., as a natural experiment).<br><br>Score HIGH risk if:<br>Other actors are conducting similar activities in the comparison communities only.<br>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. | MEDIUM-HIGH | The national VEA approach for WASH has been implemented in all comparison areas. In DRC, SWIFT 1 project implementation followed the VEA approach, in coordination with the government, but with additional activities and support as described in Section 2.<br><br>This evaluation purposefully compared SWIFT 1 to the standard VEA approach to measure any added impact attributable to the additional support, activities and, particularly, the outcomes phase of SWIFT 1. In other words, it is a natural experiment. |
| *Analysis* | | | | |
| 7 | Representativeness | Score LOW risk if:<br>During analysis or matching procedure less than 10% of the sample in the intervention group is excluded.<br><br>Score HIGH risk if:<br>During analysis or matching procedure more than 10% of the sample in the intervention group is excluded. | HIGH | During analysis, 58% of the sample in the intervention group was excluded. 57% of the sample was removed by filtering to get satisfactory matching. Of the remaining sample, 5% of the intervention group (i.e., one additional percentage point overall) was excluded by the matching procedure. |

| | | | | |
|---|---|---|---|---|
| 8 | Robustness checks | Score LOW risk if:<br>Magnitude and statistical significance of the results are approximately consistent with different econometric models.<br><br>Score HIGH risk if:<br>Results are not consistent with different econometric models. | HIGH | Without filtering, more results were statistically significant; however, the matching was not good owing to the unexpected lack of comparability between SWIFT 1 and VEA areas at baseline. With filtering (resulting in good matching) the results were not statistically significant but are also underpowered. |
| 9 | Triangulation | Score LOW risk if:<br>Results are triangulated and consistent with other evaluation methods within the same evaluation.<br>Results are triangulated and consistent with other data on the same project but from different evaluations.<br><br>Score MEDIUM risk if:<br>Results are not consistent with other evaluation methods or sources but differences are explained in the report.<br><br>Score HIGH risk if:<br>Results are not consistent or triangulated with other evaluation methods. | MEDIUM | The results on water, sanitation and hygiene do vary from the project's endline evaluation for variables that are available. However, these differences can be explained by different definitions used for water and sanitation access (i.e., SWIFT definitions versus JMP definitions) as well as the 16 months that had elapsed between data collection for the project endline (November 2017) and the Effectiveness Review (March 2019). |
| 10 | Multiple hypothesis testing | Score LOW risk if:<br>Multiple hypothesis tests apply Benjamini or Bonferroni tests.<br>The evaluation drafted a pre-analysis plan prior data analysis, and followed the plan.<br><br>Score MEDIUM risk if:<br>The evaluation drafted a pre-analysis plan prior data analysis and significantly altered the plan, but changes are clearly justified.<br><br>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:<br>Clustering is applied.<br>Clustering was tested but rejected as providing higher standard errors than non-clustering estimates.<br><br>Score HIGH otherwise. | LOW | Village-level clustering was applied. |

| *Other* | | | | |
|---|---|---|---|---|
| 12 | Other | Any other issue reported by the evaluator. | MEDIUM | Many areas of the project were inaccessible during the evaluation owing to Ebola and insecurity. For this reason, only areas in the southern part of the province were sampled, where implementation quality for SWIFT 1 was understood by the programme team to be lower. It is also possible that implementation quality, or the intensity of project activities, varied between the SWIFT 1 outputs phase and VEA, although implementation was done by the same partner. |

# REFERENCES


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

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

Banerjee, S. G. and 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. and 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), 8137–8165.

Bishop, D. and Bowman, K. (2014). Still learning: a critical reflection on three years of measuring women's empowerment in Oxfam. *Gender & Development, 22*(2), 253–269.

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

Bratton, M. and Gyimah-Boadi, E. (2016). *Do Trustworthy Institutions Matter for Development? Corruption, Trust and Government Performance in Africa*. Afrobarometer Dispatch No. 112. https://afrobarometer.org/sites/default/files/publications/Dispatches/ab_r6_dispatchno112_trustworthy_institutions_and_development_in_africa.pdf

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

Filmer, D. and 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. and 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. and Pérez-Foguet, A. (2018). *Measuring sanitation poverty: A multidimensional measure to assess delivery of sanitation and hygiene services at the household level*. OPHI Working Paper No. 116. https://www.ophi.org.uk/wp-content/uploads/OPHIWP116.pdf

Godlonton, S., Hernandez, M. and 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. and Smith, J. (1999). The economics and econometrics of active labor market programs, *Handbook of Labor Economics, 3*(A), 1865–2097.

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

HWISE (2019). *Household Water Insecurity Experiences (HWISE) Scale.* https://sites.northwestern.edu/hwise/

Kayser, G. L., Moriarty, P., Fonseca, C. and 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), 4812–4835.

Leuven, E. and 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. Version 4.0.12.

Lombardini, S., Bowman, K. and 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

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

Mizniak, J., Wildman, T. and 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. and Giné, X. (2013). *How Accurate Are Recall Data? Evidence from Coastal India*. The World Bank. Policy Research Working Paper 6009. https://openknowledge.worldbank.org/handle/10986/19872

OCHA. (2019). *DR Congo – Health Zones*. https://data.humdata.org/dataset/dr-congo-health-0

Porteous, D. (2006). *The enabling environment for mobile banking in Africa.* Report Commissioned by Department for International Development (DFID). https://www.findevgateway.org/paper/2006/05/enabling-environment-mobile-banking-africa

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

Rosenbaum, P. and 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. and Fong, V. (2013). Using indicators to measure water resources sustainability in California. In *World Environmental and Water Resources Congress 2013: Showcasing the Future* (pp. 2708–2715).

Thomas, E., Andrés, L. A., Borja-Vega, C. and 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.

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

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. and 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, *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



The information in this publication is correct at the time of going to press.

ISBN: 978-1-78748-871-7
DOI: 10.21201/2022.8717

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