

# Real-Time Evaluation of the Water Supply Improvement Project, Guyana

FINAL REPORT



OFFICE OF  
INDEPENDENT  
EVALUATION

# Table of Contents

<b>1</b>	<b>Introductory note to evaluation</b>	<b>1</b>
<b>2</b>	<b>Background information</b>	<b>2</b>
	2.1 Water sector context in Guyana	2
	2.2 Guyana Water Incorporated	3
	2.3 The Water Supply Improvement Project	3
	2.4 The Caribbean Development Bank’s water strategy	4
<b>3</b>	<b>Evaluation methodology and tools</b>	<b>5</b>
	3.1 Purpose and scope of the evaluation	5
	3.2 Evaluation questions	6
	3.3 Methodology and tools	7
	3.4 Limitations	8
<b>4</b>	<b>Key findings</b>	<b>9</b>
	4.1 Answers to the evaluation questions	9
	4.1.1 EQ1. To what extent is WSIP implementation on track relative to its planned timeline, outputs, and intermediate results?	9
	4.1.2 EQ2. What factors are enabling or constraining WSIP delivery?	10
	4.1.3 EQ3. What early signs of intended or unintended outcomes—positive or negative—are emerging from WSIP interventions?	10
	4.1.4 EQ4. How well are institutional arrangements and stakeholder coordination mechanisms functioning to support effective implementation?	10
	4.1.5 EQ5. What insights from WSIP’s design, management, and delivery can inform similar projects in other contexts?	10
	4.1.6 EQ6. What enabling conditions, incentives, and capacities are needed to replicate or adapt real-time evaluation practices?	11
	4.1.7 EQ7. How can real-time evidence be better integrated into decision-making cycles?	12
	4.2 Overall assessment	12
	4.3 Risk/Challenge analysis	14

<b>5 Conclusions and lessons learned</b>	<b>17</b>
<b>5.1 Conclusions</b>	<b>17</b>
5.1.1 Conclusion 1: Infrastructure delivery is progressing, but impact is contingent on non-technical conditions	17
5.1.2 Conclusion 2: Late integration of communities limits behavioural learning and adaptive management	17
5.1.3 Conclusion 3: Trust, not infrastructure, is the binding constraint to behavioural change	18
5.1.4 Conclusion 4: The pace of infrastructure delivery is misaligned with the time required for trust formation	18
5.1.5 Conclusion 5: Gendered impacts are central to outcomes but remain weakly operationalised	18
5.1.6 Conclusion 6: Weak MEAL integration constrains learning, adaptation, and causal understanding	21
5.1.7 Conclusion 7: The principal risk to WSIP is not technical failure, but delayed institutional response to learning	21
5.1.8 Conclusion 8: There are early indications that real-time evaluation is already influencing implementation dynamics and relationships	21
<b>5.2 Lessons learned</b>	<b>21</b>
5.2.1 Lesson 1: Infrastructure is a necessary condition for impact, but behaviour is the decisive one	21
5.2.2 Lesson 2: Early and continuous community integration is a prerequisite for effective learning and adaptation	21
5.2.3 Lesson 3: Trust formation follows a different logic and timeline than construction	21
5.2.4 Lesson 4: Strong MEAL systems are central to institutional performance, not ancillary reporting tools	21
5.2.5 Lesson 5: Real-time evaluation can actively shape outcomes when embedded as a learning practice	21
5.2.6 Lesson 6: Coordinated supervision and evaluation missions enhance learning while preserving accountability	21
<b>6 Recommendations</b>	<b>26</b>

# Acronyms

<b>ADB</b>	Asian Development Bank
<b>AfDB</b>	African Development Bank
<b>CDB</b>	Caribbean Development Bank
<b>EIB</b>	European Investment Bank
<b>EQ</b>	Evaluation Question
<b>GOI</b>	Government of Italy
<b>GRM</b>	Grievance Redress Mechanism
<b>GWI</b>	Guyana Water Incorporated
<b>IADB</b>	Inter-American Development Bank
<b>IWRM</b>	Integrated Water Resources Management
<b>LL</b>	Learning Loop
<b>M&amp;E</b>	Monitoring and Evaluation
<b>M&amp;E/MIS</b>	Monitoring and Evaluation/Management Information System
<b>MEAL</b>	Monitoring, Evaluation, Accountability, and Learning
<b>NRW</b>	Non-Revenue Water
<b>O&amp;M</b>	Operations and Maintenance
<b>OIE</b>	Office of Independent Evaluation
<b>RMF</b>	Results Monitoring Framework
<b>RTE</b>	Real-Time Evaluation
<b>SEA/SH</b>	Sexual Exploitation, Abuse, and Harassment
<b>ToC</b>	Theory of Change
<b>WRC</b>	Water Research Commission
<b>WSIP</b>	Water Supply Improvement Project
<b>WTP</b>	Water Treatment Plants

# Acknowledgement

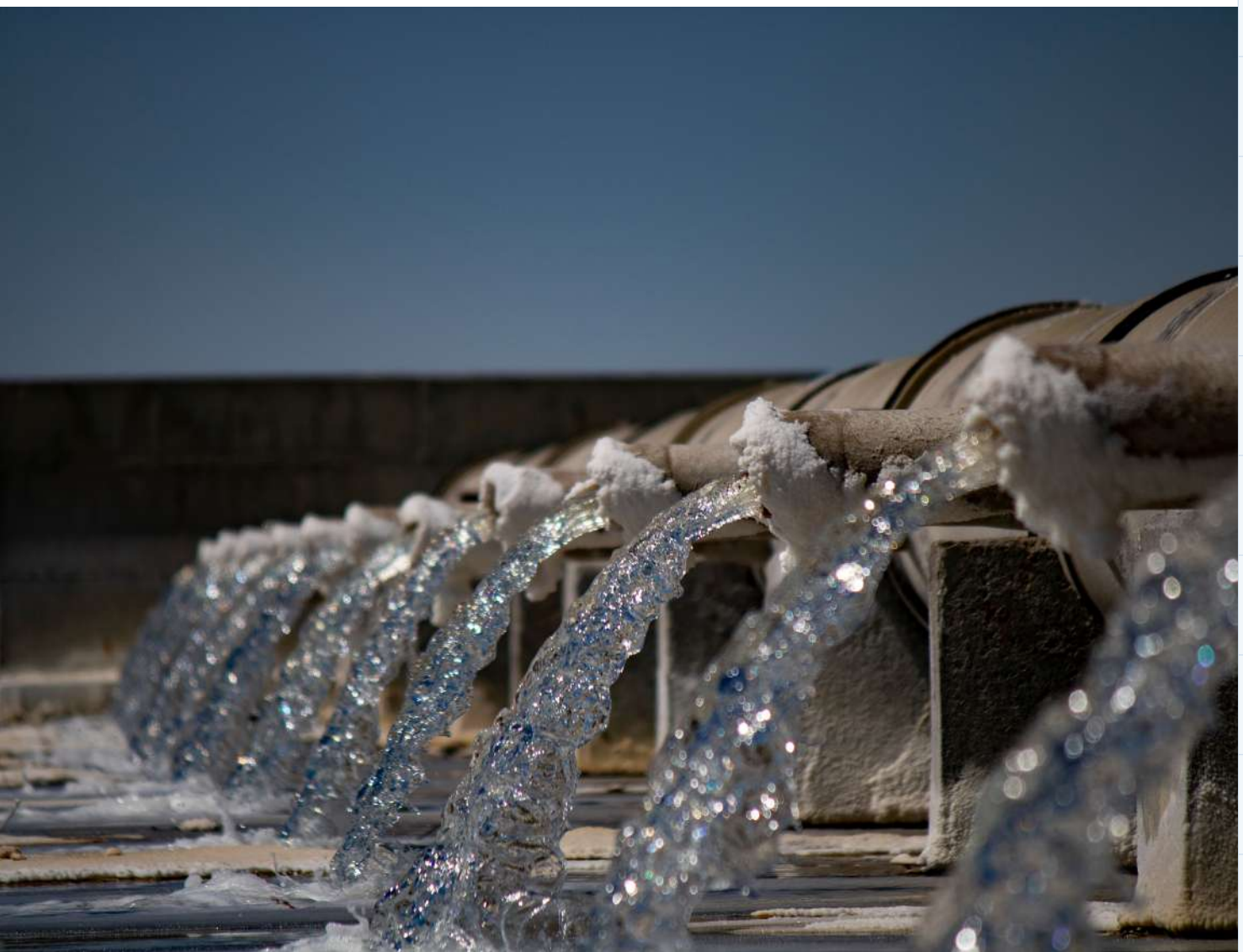
The Office of Independent Evaluation (OIE) thanks the Ministry of Finance, Guyana, for accommodating and supporting this evaluation and its field visit. Staff from the Caribbean Development Bank (CDB), including focal point Zayid Alleyne from the Projects Department, provided support throughout the process. We are also grateful for the country involvement by Arshad Yacoub and Mel Williams, and all those who provided valuable inputs during the field visits and workshops. I acknowledge Serena Rossignoli, Evaluation Manager, and Denise Padmore, from the OIE. I also acknowledge the two consultancy

firms, Technopolis Group, and Integrated Sanitation Solutions for Urban Development (I-San) for conducting the evaluation and their practical field engagement: Juan Carlos Salazar, Team Leader; Rafa Vila, Deputy Team Leader and affiliate consultants Adrien Mazeau, Prit Salian, and Flavio Carrera.

**Roberto La Rovere**

Head of Evaluation  
Office of Independent Evaluation

*The opinions expressed in this report are those of the external evaluation team and do not necessarily reflect those of CDB. Responsibility for the opinions expressed in this report rests solely with the authors. Publication of this document and the opinions expressed within does not imply endorsement by CDB.*



# 1 Introductory note to evaluation

The Real-Time Evaluation (RTE) of the Water Supply Improvement Project (WSIP) was designed as a learning-oriented, decision-support exercise embedded in the project's implementation phase.

This exercise was commissioned independently by the OIE several years after project initiation and was not originally envisaged in the WSIP design. Its **primary purpose** is to generate timely, credible, and actionable evidence to inform adaptive management, support mid-course corrections, and strengthen the likelihood that WSIP's infrastructure investments translate into sustained behavioural and social outcomes. Beyond supporting WSIP implementation in Guyana, the evaluation seeks to generate transferable lessons on applying RTE in infrastructure and utility reform projects, strengthening the CDB capacity to integrate timely evidence into management and decision-making across its portfolio.

Rather than assessing final results or impacts, the evaluation focuses on how the project is unfolding in practice, whether its underlying assumptions—particularly those related to trust, service adoption, and behavioural change—are holding, and what conditions are emerging to enable or constrain impact. In doing so, the RTE supports both project-level decision-making in Guyana and broader institutional learning for CDB on the use of RTE approaches in complex infrastructure operations.

The report is structured to progressively link context, evidence, and evaluative judgment. It begins with *background information* on the water sector in Guyana, the institutional role of Guyana Water Incorporated (GWI), the design and financing of WSIP, and CDB's strategic approach to the water sector. This is followed by a detailed presentation of the *evaluation methodology*, including the purpose and scope of the RTE, the evaluation questions (EQs), and the data-collection and analysis tools. The core of the report presents *key findings* and answers to the EQs, drawing on triangulated evidence from fieldwork, institutional interviews, community interactions, and document review. The *final sections* synthesise

conclusions, identify transferable lessons, analyse risks and challenges, and present targeted recommendations to strengthen sustainability, ensure impact consistency, and support institutional learning across current and future project cycles.

At the time of the RTE, the WSIP is in the early to mid-stage of implementation, with construction of several treatment plants underway. Still, many complementary components—such as network upgrades, metering systems, and sludge management arrangements—remain in progress or under preparation. This timing makes the project particularly suitable for a RTE approach, as key design and implementation choices remain open to adjustment.<sup>1</sup>

The evaluation is subject to important limitations inherent to its timing and context. WSIP is in the early to mid-stages of implementation, with most interventions still focused on infrastructure delivery; therefore, long-term outcomes and impacts cannot yet be assessed. Fragmented and incomplete data systems continue to constrain analytical depth. Baseline information is limited, particularly regarding behavioural indicators such as trust, usage patterns, and willingness to pay. In parallel, the monitoring and evaluation (M&E) function within GWI is still consolidating its capabilities and remains primarily focused on compliance and reporting requirements rather than on outcome-oriented analysis. Separately, tracing behavioural change at the community level presents additional challenges. Long-standing distrust in piped water services affects perception dynamics, and systematic community engagement mechanisms were introduced relatively late in the project cycle, thereby reducing the availability of structured, longitudinal evidence on perceptions and uptake.

At the same time, the RTE presents significant strengths. Its real-time, theory-based approach enables the identification of emerging risks, behavioural dynamics, and institutional constraints while corrective action remains possible. The use of mixed methods and systematic triangulation enhances the credibility of findings despite data limitations. Importantly, the RTE has also functioned as a **catalyst** for dialogue and reflection among stakeholders,

increasing visibility of behavioural and trust-related issues that are central to impact but often under-examined in infrastructure projects. Together, these strengths position the evaluation as both an analytical and a practical contribution to improving WSIP's development effectiveness. This experience also suggests that the RTE model could be adapted and applied to other complex infrastructure or programmatic operations within the CDB portfolio.

## 2 Background information

The RTE of the WSIP is situated at the intersection of three interrelated dynamics. Guyana's evolving water sector context, the institutional role and reform trajectory of GWI, and the strategic priorities of CDB in relation to water, sanitation, and resilient infrastructure.

Understanding this broader backdrop is essential to interpreting both the rationale for the evaluation and the nature of the findings it has generated.

### 2.1 Water sector context in Guyana

Guyana faces persistent, structurally rooted challenges in providing safe, reliable, and affordable water services. Despite the country's relative abundance of water resources, access to treated potable water remains uneven, particularly across hinterland, island, and peri-urban communities. Many areas rely on groundwater sources characterised by high iron content and, in some locations, elevated ammonia levels, resulting in poor organoleptic quality, corrosion of household assets, and health-related concerns. As a consequence, widespread reliance on bottled water, rainwater harvesting, or untreated sources has become a coping strategy for households, with disproportionate impacts on women and low-income families.

Infrastructural deficits, ageing networks, high levels of non-revenue water (NRW), and limited public trust in

pipled water services compound these challenges. Historical experiences with incomplete projects, inconsistent service quality, and unsuccessful metering initiatives have shaped community perceptions and behaviours, creating a context in which infrastructure investments alone are insufficient to trigger sustained behavioural change. Addressing these constraints requires not only technical upgrades but also institutional strengthening, improved communication, and mechanisms to rebuild trust between service providers and users.

### 2.2 Guyana Water Incorporated

GWI is the national public utility responsible for providing water services nationwide. Operating as a state-owned entity, GWI balances commercial, operational, and social mandates within a policy environment that treats access to safe water as a public good. The utility is organised through a hybrid institutional model, combining central functional departments—such as operations, finance, procurement, customer service, and M&E—with decentralised regional management units responsible for day-to-day service delivery.

Over the past decade, GWI has undertaken a gradual modernisation and expansion, supported by government investment and financing from development partners. This process has focused on expanding treatment capacity, improving water quality, reducing system losses, and strengthening internal systems. However, capacity constraints persist, particularly in relation to integrated data management, M&E, and systematic engagement with communities. The M&E function, while operational, remains limited in scale and is primarily oriented toward reporting rather than adaptive learning. Similarly, mechanisms for tracking user perceptions, trust, and behavioural responses to service improvements remain nascent.

Within this institutional context, WSIP represents both a significant investment opportunity and a test of GWI's ability to translate infrastructure delivery into tangible improvements in service use and social outcomes. The project has also highlighted the importance of strengthening cross-functional coordination within GWI, particularly between technical units, social and community-facing roles, and emerging monitoring and learning functions.

### 2.3 The Water Supply Improvement Project

The WSIP is a multi-site infrastructure investment aimed at improving access to safe, treated water in selected regions of Guyana. The project focuses primarily on the construction of five new Water Treatment Plants (WTPs) at Maria's Delight, Adventure, Bath, Leguan, and Wake-naam, along with associated works intended to improve water quality, reliability, and system performance. Innovative technical solutions—such as on-site chlorine generation, biofiltration systems, and air-stripping technologies—are being deployed to address site-specific water-quality challenges, including elevated iron and ammonia concentrations.

The project is financed through a sovereign loan provided to the Government of Guyana, with co-financing from the European Investment Bank (EIB) and the Government of Italy (GOI) under a blended financing framework. The CDB contributes through its Ordinary Capital Resources and Special Development Fund. At the same time, the EIB and the GOI provide concessional funding via the Italian Climate Credit and the Climate Action Line of Credit II mechanisms. The total investment amounts to approximately USD 76.2 million. It has the GWI acting as the executing agency. The loan is structured under standard

CDB terms for public infrastructure investments, combining long-term maturity with concessional elements to ensure fiscal sustainability while enabling the delivery of essential services.

Disbursements are linked to eligible expenditures and compliance with agreed procurement, financial management, and safeguard requirements, with implementation oversight exercised through regular supervision missions and reporting. The financing framework reflects the project's public-good orientation; while operational efficiency and revenue performance are relevant considerations, the loan does not hinge on short-term cost recovery but rather on achieving development outcomes related to access, service quality, and resilience.

The WSIP is designed and implemented in accordance with CDB's environmental and social safeguard policies, recognising both the moderate environmental risks associated with water treatment infrastructure and the significant social benefits expected from improved access to safe water. Environmental considerations include groundwater abstraction, sludge and wastewater management, construction-related impacts such as noise and dust, and compliance with national permitting requirements overseen by the Environmental Protection Agency.

Social dimensions are central to the project, given its direct interface with communities. These include stakeholder engagement; management of expectations regarding water quality and billing; gender-differentiated impacts on household water burdens; and the need to protect confidentiality and ensure inclusion in community consultations. While safeguard instruments are in place, the evaluation highlights that their effectiveness depends on consistent operationalisation, clear communication, and the strengthening of social and community liaison functions to ensure that environmental and social risks are actively managed throughout implementation.

Beyond physical infrastructure, the WSIP is designed to contribute to broader development objectives, including improved public health, reduced household expenditure on alternative water sources, and enhanced quality of life. Implicit in the project's Theory of Change (ToC) is the assumption that improved water quality and reliability will lead to increased adoption of piped water services, reduced reliance on bottled water, and, over time, more sustainable service delivery outcomes. However, these outcomes depend on a range of enabling conditions, in-

cluding effective stakeholder engagement, trust in service quality, and institutional capacity to monitor and respond to user behaviour.

## 2.4 The Caribbean Development Bank's water strategy

The CDB plays a central role in supporting water and sanitation investments across its borrowing member countries. CDB's strategic approach to the water sector emphasises universal access, service quality, climate resilience, and institutional sustainability. Recognising the vulnerability of Caribbean states to climate change, environmental degradation, and fiscal constraints, the Bank increasingly promotes integrated solutions that combine infrastructure

financing with capacity building, governance strengthening, and knowledge generation.

Within this framework, CDB has also sought to strengthen its evaluative and learning practices, moving beyond ex post assessments toward approaches that support real-time learning and adaptive management. The use of RTE reflects this shift, positioning evaluation as a tool to improve implementation performance and inform decision-making while projects are still active. WSIP has been identified as a suitable pilot for this approach within the water sector, given its complexity, behavioural dimensions, and potential to yield transferable lessons for other infrastructure operations.



# 3 Evaluation methodology and tools

## 3.1 Purpose and scope of the evaluation

The RTE was designed as a learning-oriented, decision-support exercise embedded within the project's implementation phase. Its primary purpose is to provide timely, credible, and actionable evidence to inform adaptive management, support mid-course corrections, and strengthen the likelihood that the project achieves its intended outcomes under real-world institutional, social, and operational conditions. Unlike summative or compliance-driven evaluations, the RTE is explicitly forward-looking, focusing on how the project is unfolding in practice and how emerging risks, opportunities, and behavioural dynamics can be addressed while implementation is still underway.

The evaluation seeks to address a core strategic question facing the project and its partners: under what conditions can infrastructure-led investments in water supply translate into sustained improvements in service use, trust, and quality of life? In this sense, the RTE goes beyond assessing physical progress to examine the causal mechanisms through which WSIP interventions are expected to generate behavioural change, institutional learning, and social outcomes. Particular attention is given to the assumptions underpinning the project's ToC, including the expectation that improved water quality and reliability will lead to increased adoption of piped water, reduced reliance on bottled or alternative sources, and gradual improvements in public health and household welfare.

The evaluation's scope encompasses both the operational and systemic dimensions of WSIP implementation. Geographically, the RTE spans multiple regions and comprises a range of project sites, reflecting diverse technical, environmental, and sociocultural contexts. These sites were examined through field visits, community engagement, and direct observation of construction progress, complemented by interviews with institutional stakeholders at national, regional, and community levels. The evaluation does not aim to assess final outcomes or long-term impacts, which remain beyond the temporal scope of an RTE, but rather to identify early signals of change,

bottlenecks, and unintended effects that may influence the project's trajectory.

Analytically, the RTE addresses a set of interrelated dimensions. These include: the extent to which implementation is progressing in line with planned outputs and timelines; the institutional and coordination factors that enable or constrain delivery; early evidence of behavioural and perceptual change among users; the functioning of stakeholder engagement and communication mechanisms; and the adequacy of Monitoring, Evaluation, Accountability and Learning (MEAL) systems to support adaptive management. Cross-cutting considerations such as gender, inclusion, trust, and political economy dynamics are integrated throughout the analysis, recognising their central role in shaping the adoption and sustainability of water services.

While the broader RTE framework and its conceptual components are fully presented in the Annexes, the analysis at this stage of implementation deliberately concentrated on a limited set of causal pathways and operational issues deemed most critical for immediate learning and decision-making. Consistent with the real-time nature of the exercise, the evaluation prioritised those mechanisms where evidence could meaningfully inform ongoing implementation, rather than attempting to assess the full range of theoretical elements simultaneously.

The integration of a MEAL approach into the RTE process is critical to ensuring that evidence generation translates into timely learning and adaptive decision-making. In the context of the WSIP, the MEAL system provides the structural backbone through which diverse data streams—technical, financial, social, and perceptual—can be systematically collected, triangulated, and interpreted against the project's ToC. By embedding MEAL functions within the RTE, the evaluation moves beyond episodic assessment toward a continuous feedback mechanism that supports course correction during implementation. This integration enables the early identification of emerging risks, behavioural signals, and implementation bottlenecks, strengthens accountability to communities and institutional stakeholders, and ensures that learning is not

only produced but also actively used to inform operational and strategic decisions throughout the project lifecycle.

Finally, the evaluation also has a broader institutional purpose. Beyond supporting WSIP implementation in Guyana, it aims to generate transferable lessons on the use of RTE approaches in infrastructure and utility reform projects. By testing learning loops (LLs), process tracing methods, and rapid feedback mechanisms, the RTE supports CDB's evaluative practice and its capacity to integrate real-time evidence into project management and decision-making cycles across its portfolio.

### 3.2 Evaluation questions

This RTE is guided by a set of EQs designed to support adaptive management, generate timely learning, and inform decision-making during implementation. Rather than treating these questions as isolated lines of inquiry, the evaluation treats them as interlinked analytical lenses applied through a theory-based, mixed-methods approach, anchored in the project's ToC. The following are the EQs:

1. To what extent is the WSIP implementation on track relative to its planned timeline, outputs, and intermediate results?
2. What factors are enabling or constraining WSIP delivery?
3. What early signs of intended or unintended outcomes—positive or negative—are emerging from WSIP interventions?
4. How well are institutional arrangements and stakeholder coordination mechanisms functioning to support effective implementation?
5. What insights from WSIP's design, management, and delivery can inform similar water supply and infrastructure projects in other borrowing member countries?
6. What enabling conditions, incentives, and capacities are required to replicate or adapt RTE practices across CDB-supported projects?
7. How can real-time evidence be more effectively integrated into CDB's project management and evaluation decision-making cycles?

The *first question* examines the extent to which WSIP implementation is on track relative to its planned timeline, outputs, and intermediate results. This question is addressed through a systematic review of construction progress, procurement status, and the sequencing of

complementary infrastructure, complemented by site visits and discussions with technical teams and contractors. The focus is not limited to physical completion rates; it also considers whether critical enabling components—such as boreholes, metering systems, and sludge management arrangements—are progressing in a manner consistent with the project's intended outcomes.

The *second question* explores the factors that enable or constrain project delivery. This is addressed through institutional interviews, field-level observations, and political economy analyses, enabling the evaluation to identify coordination bottlenecks, procurement delays, regulatory interfaces, and capacity constraints. Particular attention is paid to how roles and responsibilities are distributed across GWI, contractors, regulators, and community-level actors, and how these relationships shape implementation performance.

The *third question* focuses on early signs of intended or unintended outcomes. Given the project's stage, this question is addressed through proxy indicators and qualitative evidence rather than outcome measurement. Community engagement, user perceptions of water quality, trust in piped services, and early behavioural signals—such as willingness to use and pay for treated water—are examined to assess whether the project's causal pathways are beginning to activate as anticipated.

The *fourth question* assesses the effectiveness of institutional arrangements and stakeholder coordination mechanisms. This is addressed by analysing governance structures, communication practices, and decision-making processes across project and institutional levels. Evidence from coordination meetings, community consultations, and internal reporting systems is used to identify strengths, gaps, and risks affecting collective action and responsiveness.

The *fifth question* synthesises lessons from WSIP's design and implementation that may inform similar interventions in other contexts. This comparative and reflective dimension is addressed by distilling cross-cutting insights on infrastructure delivery, behavioural change, and real-time learning, with a focus on transferability rather than replication.

Finally, the *sixth and seventh questions* address the enabling conditions for scaling RTE practices and integrating real-time evidence into institutional decision-making.

These questions are addressed through analysis of the MEAL system, LIs, and feedback mechanisms, with an assessment of how evidence is generated, shared, and used. Together, the EQs ensure that the RTE not only diagnoses current performance but actively supports learning, adaptation, and institutional strengthening throughout the WSIP implementation cycle.

### 3.3 Methodology and tools

The RTE adopts a theory-based, utilisation-focused methodology designed to generate timely, credible, and actionable evidence during project implementation. The methodological approach is intentionally adaptive, recognising that WSIP is a complex infrastructure intervention operating within dynamic institutional, social, and environmental contexts.

The project's ToC serves as the primary analytical framework. Rather than treating the ToC as a static planning artefact, the RTE uses it as a living hypothesis that is continuously tested against observed implementation realities. The evaluation examines whether key causal pathways—linking infrastructure delivery to improved water quality, user trust, behavioural change, and longer-term social outcomes—are operating as expected, and under what conditions these pathways may weaken or break down. This theory-based orientation allows the evaluation to move beyond descriptive monitoring toward explanatory analysis.

A full implementation of RTE entails employing a mixed-methods design that integrates qualitative and quantitative tools. Qualitative methods are central to understanding behavioural, institutional, and perceptual dynamics that cannot be captured through administrative data alone. These include semi-structured interviews with national and regional stakeholders, focus group discussions and community consultations, and direct observation during field visits to project sites. These methods enable the evaluation to capture lived experiences, uncover informal practices, and identify political economy factors influencing service adoption, billing behaviour, and trust in water quality.

Quantitative data sources are used to complement and triangulate qualitative findings. These include administrative and operational data from GWI, such as construction progress reports, water quality measurements, service coverage records, billing and collection data, and NRW

indicators. While the evaluation acknowledges limitations in data completeness and baseline availability, these sources provide essential signals on system performance and enable trend analysis where feasible. The RTE also draws on historical survey data, where available, and explores opportunities for light-touch perception tracking to fill critical information gaps.

A defining methodological feature of the RTE is the continuous use of process tracing to test selected causal chains within the ToC. Rather than attempting to attribute long-term outcomes, the evaluation focuses on verifying whether specific mechanisms—such as improved water treatment leading to increased trust or enhanced communication leading to behavioural change—are plausibly at work. For each selected causal chain, the evaluation identifies expected intermediate steps, observable manifestations, and alternative explanations, allowing for structured causal inference even in the absence of experimental designs.

The integration of MEAL tools is central to the RTE methodology. The evaluation does not operate as a parallel system but works with existing GWI and project-level monitoring structures to assess data flows, indicator relevance, and feedback mechanisms. Tools such as indicator mapping, data flow analysis, and learning loop design are used to identify where evidence is generated, how it is used, and where gaps or bottlenecks exist. This approach strengthens institutional learning and supports the co-development of a more fit-for-purpose MEAL system aligned with the project's adaptive needs.

Ethical considerations are embedded throughout the methodological design. The evaluation applies principles of informed consent, confidentiality, and do-no-harm, particularly in community engagement and perception-based data collection. Special attention is given to gender and inclusion, ensuring that women's perspectives and the experiences of vulnerable groups are actively sought and safely incorporated into the analysis.

Finally, the RTE methodology is iterative by design. Data collection, analysis, and feedback occur in cycles rather than linear phases, enabling emerging findings to be shared with project stakeholders in real time. This iterative approach ensures that the evaluation functions not only as an assessment mechanism but as an integral component of adaptive project management and institutional learning within the WSIP.

### 3.4 Limitations

This Evaluation is subject to several limitations that reflect both the evaluation approach and the project's contextual realities. First, the evaluation is being conducted at an early stage of implementation, when most interventions are still focused on infrastructure construction and final outputs and outcomes have not yet materialised. As a result, the evaluation does not seek to assess long-term results or impacts, but rather to identify early signals, emerging risks, and implementation dynamics. Many outcome-level changes—particularly those related to sustained behaviour change, public health, and environmental effects—remain necessarily prospective.

Second, the evaluation is constrained by the lack of an integrated, fully operational data ecosystem. Project-relevant information is distributed across multiple administrative, technical, and financial systems, largely managed in standalone spreadsheets, with limited interoperability and real-time functionality. This fragmentation restricts the ability to systematically triangulate data, link financial flows to physical progress, or track service performance and user behaviour coherently. The lack of

up-to-date baseline data further limits the precision of quantitative analysis.

Third, given that this evaluation found that the role of M&E within the project—and within GWI more broadly—remains underdeveloped and under-resourced, it constrained the intrinsic capacity and reach of this type of evaluation at this time. The M&E function operates with limited staffing, tools, and institutional authority, and is primarily oriented toward compliance reporting rather than learning or adaptive management. This constrains the extent to which evaluation findings can be immediately embedded within routine decision-making processes. Similarly, the evaluation faces inherent challenges in tracing behavioural change among the communities served. Historical distrust in piped water quality, limited prior exposure to perception-based data collection, and the absence of established feedback mechanisms make it difficult to systematically capture shifts in attitudes, usage patterns, and willingness to pay. Building a culture of behavioural data collection and use requires sustained engagement that extends beyond the temporal scope of this RTE.



# 4 Key findings

This section presents the conclusions of the RTE based on a systematic process of triangulation across multiple sources of evidence.

Findings and judgments draw on the combined analysis of field observations, community interactions, semi-structured interviews with institutional and operational stakeholders, project documentation, and emerging monitoring and MEAL data. Rather than relying on any single data source, the evaluation cross-checks perceptions, reported practices, and observed implementation dynamics against the project's ToC to assess the plausibility of causal chains and emerging results. This triangulated approach strengthens the credibility of the conclusions, enables the explicit identification of inconsistencies and gaps, and provides a more nuanced understanding of how infrastructure delivery, institutional arrangements, and behavioural dynamics interact in practice at this stage of implementation.

The findings and learning generated through this evaluation operate at three distinct but interrelated levels. First, a set of conclusions directly responds to the WSIP-specific EQs (Q1–Q4), focusing on implementation performance, emerging outcomes, and operational implications for the project in Guyana. Second, a broader layer of lessons (notably Q4–Q5) speaks to the design and management of similar infrastructure and utility reform projects, highlighting transferable insights on behavioural uptake, coordination, and adaptive management. Third, subsequent question 6 addresses institutional implications for the wider CDB and its Board, drawing out considerations for evaluative practice, supervision models, and the strategic integration of real-time evidence across the portfolio.

While the evaluation underscores that infrastructure delivery alone does not guarantee development impact, it is equally important to recognise the substantial technical and operational progress achieved to date. Procurement processes have advanced in line with agreed procedures, detailed designs have been completed and refined to

address local conditions, and construction activities are progressing according to established milestones. These achievements reflect significant coordination effort and technical diligence by GWI and its partners. This progress underscores the importance of current implementation performance and anchors the emphasis on behavioural and institutional conditions alongside engineering accomplishments, as complements to ensure that technical gains are converted into sustained service use and long-term impact.

## 4.1 Answers to the evaluation questions

### 4.1.1 EQ1. To what extent is WSIP implementation on track relative to its planned timeline, outputs, and intermediate results?

**The WSIP is broadly on track with respect to its core physical outputs, particularly the construction of WTPs, but progress toward intermediate outcomes is uneven and uncertain.** While infrastructure delivery is advancing broadly as planned, the conditions required for these assets to translate into improved service use and social outcomes are not yet entirely in place.

Field evidence confirms steady construction progress and the deployment of technically appropriate solutions to address water quality challenges. However, complementary components—such as network extensions, metering systems, sludge management, and user-facing service readiness—are still under development. Crucially, limited early engagement with communities means that intermediate results related to trust, adoption, and behavioural change cannot yet be reliably observed or validated. As a result, physical progress is not yet matched by progress along the project's behavioural and social pathways.

#### 4.1.2 EQ2. What factors are enabling or constraining WSIP delivery?

**Delivery is enabled by strong political commitment, adequate financing, and technical capacity, but constrained by a MEAL system -both at the project and institutional levels- that is underdeveloped and under-resourced, which contributes to institutional fragmentation and the late integration of community engagement and learning functions.**

High-level institutional support and contractor performance have facilitated infrastructure delivery. At the same time, the delayed operationalisation of social, gender, and community liaison roles has constrained the project's ability to anticipate, understand, and manage user perceptions and behavioural responses. This has limited community feedback during critical early phases of implementation, reducing the project's capacity to adjust its design and communication strategies in real time.

#### 4.1.3 EQ3. What early signs of intended or unintended outcomes—positive or negative—are emerging from WSIP interventions?

**Early signals indicate strong latent demand for improved water services and cautious optimism among communities, but behavioural change remains largely untested due to limited community integration to date.**

Community interactions provide strong emerging signals of widespread dissatisfaction with current (before WTPs start operating) water quality and a high reliance on bottled water, suggesting strong potential for the uptake of improved services. Many residents expressed willingness to use and pay for treated water if quality improvements are credible. However, because communities have not been consistently engaged throughout implementation, behavioural dynamics—such as shifts in trust, usage patterns, or willingness to pay—are only partially visible and cannot yet be systematically traced. Historical experiences with failed infrastructure and metering initiatives continue to shape scepticism, underscoring the need for more deliberate behavioural learning.

#### 4.1.4 EQ4. How well are institutional arrangements and stakeholder coordination mechanisms functioning to support effective implementation?

**Institutional arrangements are adequate for managing infrastructure delivery but insufficiently configured to support behaviour-sensitive implementation and learning.**

Formal governance and coordination mechanisms between GWI, contractors, and CDB are in place and functional. However, weak horizontal integration within GWI—particularly between technical units and community-facing functions—limits the project's responsiveness to social and behavioural risks. The late inclusion of communities has further reduced opportunities for iterative learning and adaptation, weakening the feedback loops required for effective coordination around service adoption and trust-building.

#### 4.1.5 EQ5. What insights from WSIP's design, management, and delivery can inform similar projects in other contexts?

WSIP highlights the critical importance of integrating community engagement and behavioural learning early in infrastructure projects, rather than treating them as downstream or complementary activities. It also demonstrates that large-scale infrastructure investments in the water sector must be accompanied by early and sustained investment in MEAL systems as a core component of institutional strengthening, rather than as a secondary or compliance-oriented function.

The project demonstrates that technically sound infrastructure does not automatically generate behavioural change or service adoption. Delayed engagement with communities limits the ability to test assumptions, tailor communication, and build trust incrementally. For similar projects, early investment in understanding user behaviour, perceptions, and incentives is essential to protect the impact of significant capital investments.

The project also shows that without an integrated MEAL system capable of linking physical progress, financial execution, and behavioural outcomes, it is difficult to trace whether infrastructure investments are generating the conditions required for impact. In WSIP, limitations in data integration, perception tracking, and LIs have constrained the project's ability to test assumptions related to trust, service adoption, and willingness to pay. This experience underscores the importance of configuring a solid Results Monitoring Framework (RMF) from the outset—one that

goes beyond output indicators and systematically captures intermediate outcomes, behavioural signals, and institutional responses. For similar projects, investing in MEAL capabilities, including digital data integration, clear indicator ownership, and embedded learning processes, is essential not only for accountability but for enabling adaptive management and protecting the long-term value of infrastructure investments.

#### 4.1.6 EQ6. What enabling conditions, incentives, and capacities are needed to replicate or adapt real-time evaluation practices?

**For GWI, the evidence indicates that, although several foundational conditions for RTE are present in the WSIP and GWI, they remain insufficiently institutionalised to support systematic behavioural learning and adaptive management at scale.** The experience of this RTE suggests that replicating or adapting RTE practices requires a deliberate shift from evaluation as an episodic, externally driven activity to an embedded institutional function. At the same time, these findings point to a wider opportunity for CDB: a more comprehensive review of M&E practices at the design and planning stages of other operations could help identify where real-time, behaviour-sensitive approaches may add value. Such an analysis would support more deliberate replication or adaptation of RTE principles across the portfolio, particularly in operations where outcomes depend on institutional coordination and user uptake rather than on infrastructure delivery alone.

At present, GWI's M&E function provides a basic operational platform for tracking outputs and selected performance indicators. However, it is constrained by fragmented data architectures, limited analytical tooling, and the absence of routine mechanisms to capture and interpret behavioural and perceptual data from served communities. These limitations constrain the organisation's ability to test causal assumptions in real time, particularly those related to trust, service adoption, and willingness to pay—dimensions central to the WSIP's ToC.

Beyond technical constraints, the evaluation highlights the importance of institutional incentives and role clarity. Learning responsibilities are not yet clearly distributed across operational, technical, and community-facing teams, and evaluative evidence is not systematically embedded in management routines or decision-making forums. As a result, real-time insights risk being perceived

as evaluative add-ons rather than as inputs that inform operational choices. For RTE to drive adaptation, staff must have both the mandate and the incentives to engage with evidence, reflect on emerging findings, and adjust practices accordingly.

Capacity development is therefore required at multiple levels. At the systems level, investment is needed in integrated data platforms that link physical progress, financial execution, service performance, and user perceptions within a coherent analytical framework. At the human level, staff require skills in data interpretation, behavioural analysis, and learning facilitation, not only within the M&E unit but across operational teams. At the organisational level, leadership commitment is needed to normalise LRs, protect time for reflection, and legitimise evidence-based course correction.

The RTE experience indicates that without these enabling conditions—integrated systems, aligned incentives, and distributed learning capacity—RTE is likely to remain an external diagnostic exercise with limited institutional uptake. Conversely, when these elements are in place, RTE practices can be adapted and scaled as part of a broader shift toward evidence-informed, behaviour-sensitive infrastructure delivery across GWI and similar institutions.

**In turn, for the CDB, the experience of applying RTE to the WSIP demonstrates clear potential for replication and adaptation within the CDB's operational and evaluative practices, provided that RTE is positioned as a management-support instrument rather than a parallel evaluative product.** The WSIP case shows that RTE is most effective when it is embedded early in the project cycle, aligned with supervision missions, and explicitly linked to decision points where adaptive action is possible. Replication within CDB would therefore require integrating RTE principles into existing project management, supervision, and portfolio review processes, rather than introducing them as stand-alone exercises.

From an institutional perspective, adapting RTE practices implies a shift toward lighter, more iterative evaluative tools that prioritise learning over exhaustiveness. This includes the systematic use of LRs, rapid triangulation of qualitative and quantitative evidence, and selective process tracing focused on high-risk or behaviour-sensitive assumptions. For CDB, this approach could be particularly valuable in infrastructure and service-delivery

projects where outcomes depend on user adoption, trust, and institutional coordination rather than on asset completion alone.

The WSIP experience also highlights the importance of aligning RTE with CDB's Results-Based Management and RMFs. Replication would be strengthened by ensuring that RTE outputs feed directly into supervision reports, mid-term reviews, and management decision forums, thereby reinforcing the use of real-time evidence across operational units.

Finally, scaling RTE practices within CDB would benefit from a portfolio-level learning strategy in which selected projects serve as pilots to refine tools, templates, and guidance. By institutionalising RTE as a flexible, context-sensitive practice—rather than a fixed methodology—CDB can enhance its capacity to manage complexity, respond to emerging risks, and protect the development impact of its investments across diverse country and sector contexts.

#### 4.1.7 EQ7. How can real-time evidence be better integrated into decision-making cycles?

**Real-time evidence can be more effectively integrated into decision-making cycles by grounding project design and implementation in a deeply articulated ToC, explicitly incorporating causal mechanisms and contribution analysis from the outset, and embedding robust MEAL systems as core operational components of complex projects.**

The WSIP experience demonstrates that when causal assumptions—particularly those related to behavioural change, trust, and service adoption—are not sufficiently unpacked early on, the capacity to generate and use real-time evidence during implementation is significantly weakened.

The RTE has generated valuable insights into emerging behavioural risks and opportunities, particularly regarding community perceptions of water quality, trust in the service provider, and willingness to adopt piped water services. However, integrating these insights into decision-making depends on the presence of clearly defined causal mechanisms linking project activities to intermediate outcomes and on management structures capable of acting on this evidence. When the ToC remains high-level

or implicitly infrastructure-centric, real-time evidence is often interpreted retrospectively, thereby limiting its value for adaptive management.

Introducing causal mechanisms into project design clarifies how and why change is expected to occur, thereby enabling the identification of critical assumptions, behavioural triggers, and institutional responses that require active management. This, in turn, allows real-time data to be mapped directly to specific points in the causal chain, enabling project teams to test whether mechanisms function as intended and to adjust strategies accordingly. Contribution analysis further strengthens this approach by providing a structured method for assessing the project's plausible influence on observed changes in complex environments where attribution is neither feasible nor appropriate. Rather than asking whether WSIP alone caused behavioural change, the evaluation can examine whether the project is making a meaningful contribution alongside other contextual and institutional factors.

The effective use of causal mechanisms and contribution analysis depends on the availability of integrated MEAL systems. Without MEAL platforms capable of capturing physical progress, financial execution, and behavioural and perceptual data in a coherent framework, evidence remains fragmented and difficult to translate into decisions. Embedding LRs that explicitly track community perceptions and behavioural responses—and aligning them with routine management forums, supervision missions, and coordination meetings—creates predictable entry points for evidence to inform operational choices.

The WSIP RTE illustrates that when real-time evidence is anchored in causal reasoning and supported by fit-for-purpose MEAL systems, project teams are better positioned to shape the conditions for behavioural change actively rather than assess them after the fact. For CDB and executing agencies, integrating causal mechanisms and contribution analysis into decision-making processes enhances the strategic use of evidence, strengthens adaptive management, and ultimately increases the likelihood that complex infrastructure investments deliver their intended development outcomes.

## 4.2 Overall assessment

A consolidated and interpretative response to the EQs guiding the RTE, in addition to addressing each ques-

tion separately, weaves them into a single narrative that reflects how the project is unfolding in practice, how its underlying assumptions are being tested, and what this implies for the likelihood that the investments will deliver their intended outcomes. **Given the project's early stage of implementation, the conclusions focus less on final results and more on the conditions that are emerging—or failing to emerge—for meaningful behavioural and social change.**

**At the time of the evaluation, WSIP implementation is advancing primarily through its infrastructure components.** Field evidence confirms that the construction of WTPs is progressing broadly in line with planned schedules and that the technical solutions deployed are well aligned with the specific water quality challenges faced across regions. From an engineering perspective, the project demonstrates a solid capacity to deliver phys-

ical assets. However, this progress remains only partially connected to the system-level conditions required to generate impact. Several complementary elements, including network upgrades, metering systems, and sludge management arrangements, remain under development, limiting the extent to which the infrastructure can yet translate into consistent, reliable service improvements for users.

**More critically, the evaluation finds that communities—the ultimate beneficiaries and central agents of change—are not yet sufficiently integrated into the project's implementation logic.** In several sites, residents reported limited awareness of the project's objectives and timelines. They expected benefits, and in some cases, the evaluation mission itself represented the first substantive engagement with the project. This gap has important implications for how behavioural change is understood and managed. The WSIP's ToC assumes



that improved water quality and reliability will lead to increased adoption of piped water, reduced reliance on bottled or alternative sources, and greater willingness to pay for services. However, the current level of community integration does not yet permit these behavioural mechanisms to be systematically observed, tested, or influenced.

**Despite this limitation, early signals emerging from community interactions are instructive. Across sites, dissatisfaction with current water quality—particularly related to high iron content and associated health and household burdens—is widespread. This creates strong latent demand for improvement and suggests that the project addresses a genuine, deeply felt need.** Many residents expressed openness, and in some cases explicit willingness, to use and pay for treated piped water if quality improvements are tangible and sustained. At the same time, historical experiences with failed infrastructure, unreliable service, and unsuccessful metering initiatives continue to shape scepticism and caution.

**Trust in the service provider, rather than infrastructure availability alone, emerges as the decisive factor governing behavioural change.** Institutionally, WSIP benefits from strong political support and alignment with national development priorities, as well as adequate financing and the capacity of technical contractors. These enabling factors have contributed to steady physical progress. However, constraints are evident in the project's social and organisational dimensions. The late establishment of community liaison, social, and gender functions, combined with limited coordination between technical and social teams, has reduced the project's ability to engage communities proactively and to learn from them in real time. As a result, the project is currently better equipped to build infrastructure than to shape the behavioural and perceptual dynamics on which its long-term impact depends.

**The evaluation also highlights limitations in the project's monitoring, evaluation, and learning arrangements.** While GWI has a functioning monitoring system, it remains fragmented and weakly oriented toward capturing user perceptions, trust, and behavioural responses. This constrains the project's capacity to trace whether the conditions for impact are materialising and to adapt implementation strategies accordingly. In this context, the RTE itself becomes a primary source of insight into behavioural dynamics, underscoring both its value

and the risks associated with delayed institutional uptake of its recommendations.

Several unexpected results emerge from this analysis. Positively, communities displayed greater optimism and readiness to embrace improved services than initially anticipated, suggesting a window of opportunity for trust-building and behavioural change. In some cases, even actors who might be expected to resist change, such as bottled water vendors, did not perceive the project as a threat. Conversely, the depth of community disengagement at this stage—and the extent to which behavioural assumptions remain untested—was more pronounced than expected, revealing a critical vulnerability in the project's impact pathway.

**Overall, the evaluation concludes that WSIP is delivering infrastructure effectively but has not yet created the social, behavioural, and learning conditions required to ensure that these investments translate into sustained outcomes.** Bridging this gap will require the timely and deliberate integration of communities into implementation, stronger behavioural insight, and the operationalisation of learning mechanisms that guide adaptive action before critical opportunities are lost.

### 4.3 Risk/Challenge analysis

**The RTE of the WSIP identifies a set of interrelated risks and challenges that extend beyond technical delivery and speak directly to the project's capacity to translate infrastructure investments into sustained development outcomes.** These risks are not exceptional in the context of complex public infrastructure projects; instead, they reflect structural tensions between engineering-led implementation, institutional capacity, and the behavioural dynamics that ultimately determine impact. Understanding these risks as part of a connected system—rather than as isolated issues—is essential for effective mitigation and adaptive management.

**At the operational level, the most significant risk relates to procurement delays and sequencing challenges affecting complementary infrastructure components.** While treatment plants are advancing, delays in borehole installation, network upgrades, metering systems, and sludge management arrangements create uncertainty about the system's overall functionality. This misalignment poses both operational and reputational risks. Communities may perceive partial service

improvements as project failure, even when construction is ongoing, thereby undermining trust and acceptance. Environmentally, unresolved sludge management solutions pose a sustainability risk if not addressed in a timely, coordinated manner.

**Moreover, from the impacts’ perspective, the first and central risk concerns the disconnect between infrastructure delivery and behavioural change.** WSIP has made tangible progress in constructing WTPs and deploying technically appropriate solutions to address iron and ammonia contamination. However, the evaluation finds that community integration into the project has occurred later than required to actively shape user perceptions, trust, and adoption behaviours. As a result, the behavioural assumptions embedded in the project’s ToC—namely that improved water quality will automatically lead to increased use of piped water and willingness to pay—remain only partially tested. Without timely and deliberate engagement, there is a risk that infrastructure assets become underutilised or fail to displace entrenched practices such as reliance on bottled water, rainwater harvesting, or untreated sources.

**Closely related is the risk of persistent trust deficits between communities and the service provider.** Historical experiences with unreliable services, rust-contaminated water, and failed metering initiatives have shaped deeply rooted scepticism in several project areas. While communities express strong latent demand for better water quality, trust in the durability and consistency of improvements remains fragile. The evaluation indicates that trust is not a passive outcome of infrastructure provision, but a condition that must be actively cultivated through transparent communication, credible demonstrations of quality, and responsive service management. If trust-building measures are not prioritised alongside construction, behavioural change may lag behind technical progress, reducing the project’s overall effectiveness.

**A third challenge lies in the late operationalisation of social, gender, and community liaison functions.** Although safeguard instruments and engagement plans exist, their practical implementation has been uneven. In some locations, community members reported limited awareness of the project’s objectives, timelines, and implications, and the RTE mission itself served as the first substantive point of engagement. This delay reduces opportunities for iterative learning and weakens feedback loops between users and implementers. It also constrains

the project’s ability to identify and manage differentiated impacts, particularly those affecting women, who bear a disproportionate burden related to water access, household management, and health outcomes.

**Institutional risks within GWI further compound these challenges. While GWI demonstrates technical competence and a strong commitment to service delivery, the evaluation highlights functional fragmentation and limited integration of learning into operational decision-making.** M&E capacities exist but are constrained by siloed data systems, reliance on manual tools, and limited focus on behavioural and perceptual data. This restricts the organisation’s ability to trace causal mechanisms in real time and to adapt implementation strategies based on emerging evidence. Without targeted investment in MEAL systems and clear assignment of learning responsibilities, there is a risk that evaluative insights remain peripheral to core management processes.

**From an evaluative perspective, a further challenge concerns the difficulty of tracing behavioural change in the absence of baseline data and established feedback mechanisms.** The lack of systematic perception tracking and up-to-date demographic data limits the precision with which changes in attitudes, usage patterns, or willingness to pay can be measured. This creates a risk that adaptive decisions are made on the basis of partial or anecdotal evidence. While the RTE mitigates this through triangulation and qualitative insights, sustained learning will require institutionalisation of behavioural data collection as part of routine operations.

**Finally, at the institutional level, there is also a risk of delayed uptake of RTE recommendations.** GWI may face challenges in integrating and operationalising the RTE recommendations within the required timeframe. Given the project’s implementation pace, the fragmentation of existing data systems, and the still-evolving role of monitoring, evaluation, and learning functions, there is a risk that critical recommendations—particularly those related to behavioural change, stakeholder engagement, and adaptive management—may not be embedded early enough to effectively trace and influence the conditions necessary for positive investment outcomes. If LLs, coordination mechanisms, and decision protocols are not strengthened in a timely manner, the project may continue to deliver infrastructure as planned while missing opportunities to shape user adoption, trust, and sustainability,

thereby reducing the overall return and the transformative potential of the investments. Also, given competing operational priorities and capacity constraints, GWI may struggle to integrate evaluative insights quickly enough to influence critical implementation stages. If LRs are not operationalised through concrete pilots and embedded into management forums, the RTE risks remaining an external diagnostic exercise rather than a catalyst for internal adaptation. This timing risk is particularly acute in relation to behavioural change, where missed early opportunities can be difficult to recover later.

This risk can be managed by prioritising the rapid implementation of targeted MEAL pilots that translate the RTE recommendations into practical, testable actions within the WSIP implementation cycle. By piloting a limited number of LRs—focused, for example, on user perceptions, trust in water quality, and early behavioural responses—GWI can generate timely evidence while incrementally building internal capacity. Assigning clear ownership for each MEAL pilot, aligning them with existing man-

agement and coordination forums, and embedding their outputs into routine decision-making processes will help ensure that learning informs action without overburdening the institution. These pilots can serve as low-risk, high-value mechanisms to trace whether the conditions for positive investment outcomes are materialising and to adjust implementation strategies accordingly.

**Taken together, these risks and challenges point to a central conclusion: the success of WSIP is not primarily threatened by technical failure, but by the possibility that infrastructure delivery outpaces the social, behavioural, and institutional conditions required for impact.** Managing these risks requires a deliberate rebalancing of attention—from assets to adoption, from outputs to mechanisms, and from retrospective assessment to real-time learning. If addressed proactively, these challenges also represent opportunities to strengthen institutional capacity, deepen community trust, and enhance the long-term value of the investments made under the WSIP.



# 5 Conclusions and lessons learned

## 5.1 Conclusions

This section presents the main conclusions of the RTE, synthesising evidence from across fieldwork, institutional interviews, document review, and iterative analytical processes. The conclusions reflect a triangulated assessment of how the WSIP is progressing in practice, the extent to which its underlying assumptions are holding, and the conditions that will determine whether infrastructure investments translate into sustained behavioural and social outcomes. Rather than reiterating findings, this section distils evaluative judgments that highlight what is working, what remains fragile, and where timely institutional action is required to protect development effectiveness and maximise the return on public investment.

Consistent with the structure described above, both the section on answers to the EQs and the conclusions and lessons section generate knowledge at three interconnected levels: findings specific to WSIP, lessons applicable to similar CDB operations, and broader institutional implications for CDB management and the Board of Directors.

### 5.1.1 Conclusion 1: Infrastructure delivery is progressing, but impact is contingent on non-technical conditions

**The evaluation concludes that WSIP is performing effectively as an infrastructure delivery operation, but the realisation of its intended development impact depends on conditions that extend well beyond technical execution.** While treatment plants are being constructed as planned, infrastructure alone is insufficient to ensure improved service uptake, behavioural change, or social outcomes. The project's impact pathway relies on a series of enabling conditions—trust, service reliability, institutional responsiveness, and user adoption—that are not automatically generated by asset completion.

Field visits confirm steady progress in the construction of water treatment facilities and in the deployment of technically appropriate solutions to address iron and ammonia contamination. Contractors demonstrate adequate

capacity, and the project benefits from strong political backing and secure financing. However, complementary components such as network upgrades, metering systems, sludge management arrangements, and service-readiness measures are either lagging or not yet fully aligned. More importantly, early interactions with communities reveal that awareness, trust, and understanding of the project remain uneven, limiting the activation of behavioural mechanisms assumed in the ToC.

This conclusion is critical because it reframes WSIP from a predominantly engineering challenge to a systemic transformation effort. If non-technical conditions are not actively managed, the project risks delivering functional infrastructure without achieving proportional social or behavioural returns. Recognising this dependency allows project stakeholders to reorient attention and resources toward the factors that ultimately determine whether infrastructure investments translate into sustained improvements in quality of life.

### 5.1.2 Conclusion 2: Late integration of communities limits behavioural learning and adaptive management

**A central conclusion of the evaluation is that the late and uneven integration of communities into WSIP implementation has constrained the project's ability to understand, test, and influence behavioural change—one of the core drivers of impact.** Communities are not yet positioned as active agents in the project's learning and adaptation processes, limiting the project's capacity to validate its behavioural assumptions in real time.

In several sites, community members reported limited prior engagement with the project and incomplete information about its objectives, timelines, and implications. In some cases, the RTE mission itself represented the first substantive interaction between residents and project stakeholders. While communities consistently expressed dissatisfaction with existing water quality and optimism about improvement, their perceptions of trust, willingness to adopt piped water, and attitudes toward billing and

metering remain only partially understood and weakly documented. The absence of systematic perception tracking further reinforces this gap.

Behavioural change—particularly shifts in trust, usage, and willingness to pay—is central to the WSIP’s long-term success. Without early and continuous community integration, these dynamics are learned too late to meaningfully influence implementation. Addressing this gap is essential not only for WSIP but for future infrastructure projects, as it highlights the cost of treating communities as downstream beneficiaries rather than as core partners in achieving impact.

### 5.1.3 Conclusion 3: Trust, not infrastructure, is the binding constraint to behavioural change

**The evaluation finds that trust in water quality and service reliability is the primary constraint governing behavioural change, outweighing technical availability or physical access. Infrastructure improvements create the possibility of change, but trust determines whether that possibility is realised in practice.**

Across all sites, residents described long-standing experiences with rust-coloured water, unreliable service, and failed past interventions, particularly related to metering. These experiences have shaped deep-seated scepticism toward piped water services. At the same time, communities demonstrated a strong latent demand for safe water and a willingness to shift away from bottled water if improvements are credible and sustained. This coexistence of demand and distrust illustrates that behavioural change hinges on confidence, transparency, and consistent service performance rather than on infrastructure presence alone.

This conclusion has significant implications for project strategy. It suggests that communication, quality demonstration, and responsiveness to user concerns are not peripheral activities but core delivery functions that must be addressed before, during, and after the infrastructure work. Failure to actively manage trust risks, underutilisation of infrastructure, continued reliance on alternative water sources, and erosion of public confidence in the service provider. Conversely, targeted trust-building interventions can unlock the full developmental value of the investments.

### 5.1.4 Conclusion 4: The pace of infrastructure delivery is misaligned with the time required for trust formation

**There is a structural misalignment between the pace of infrastructure delivery under the WSIP and the slower, nonlinear process by which trust in water services is built among communities.** While construction follows defined timelines and milestones, trust develops incrementally through repeated positive experiences, visibility of service reliability, and consistent institutional behaviour. This temporal mismatch represents a material risk to early service uptake and behavioural change.

Community interactions indicate that residents are generally unwilling to change long-established water-use practices in response to single or short-term improvements. Even where optimism about the project exists, trust is conditional and provisional, shaped by a history of past service interruptions and quality failures. Stakeholders noted that communities often adopt a “wait and see” posture, testing water quality over extended periods before adjusting consumption patterns or accepting billing arrangements. The project’s current sequencing prioritises asset completion over sustained engagement and demonstration, limiting opportunities to consolidate trust during the early operational phases.

This conclusion is critical because it highlights that early operational periods are not neutral but formative. If service readiness precedes user readiness, infrastructure may be underutilised, reinforcing scepticism rather than overcoming it. Recognising trust formation as a time-dependent process allows project managers to adjust expectations, phase engagement strategies, and protect early gains. Failure to do so risks locking in low adoption patterns that are costly to reverse later.

### 5.1.5 Conclusion 5: Gendered impacts are central to outcomes but remain weakly operationalised

**Gendered dimensions of water access and use are central to the WSIP’s intended outcomes. Yet they have not yet been sufficiently operationalised in project implementation, monitoring, or learning processes.** While gender considerations are acknowledged at a conceptual level, they have not been systematically translated into targeted strategies or indicators that could strengthen adoption, trust, and well-being out-

comes. This conclusion should be read as a forward-looking risk and opportunity identified through early signals in the evidence base, rather than as a fully evidenced outcome at this stage of implementation

Across sites, women consistently emerged as the primary managers of household water—responsible for cooking, cleaning, health-related decisions, and coping with poor water quality. They articulated more acute concerns about iron contamination, household impacts, and health risks. However, engagement activities, monitoring tools, and decision-making processes do not yet sufficiently differentiate between men’s and women’s experiences, nor do they track gender-specific behavioural responses. This limits the project’s ability to leverage women’s central role as agents of change within households and communities. Given the current limitations in gender-disaggregated

monitoring and the still-developing, qualitative evidence base, the findings point to an early indication that gendered roles and perceptions may shape behavioural responses, reinforcing the need to strengthen how gender is captured and used in engagement and learning.

The under-operationalisation of gender is not a marginal issue but a missed opportunity for impact. By failing to systematically integrate gender-responsive insights into engagement and learning strategies, the project risks overlooking one of the most powerful entry points for behavioural change. Strengthening gender-responsive MEAL and engagement practices could accelerate trust formation, improve service uptake, and enhance health outcomes—thereby increasing the overall effectiveness of the investment.

### Box 1 Best practices of gender integration on complex projects

Gender integration is essential for making complex projects (multi-actor, multi-sector, high-risk, long timelines) more effective, equitable, and politically sustainable. When gender is treated as a core design variable rather than a secondary consideration, projects tend to better anticipate real-world use patterns, reduce implementation risks (including safeguarding), and deliver benefits that reach those most affected by service gaps. However, results depend heavily on context, incentives, and sustained institutional capacity; without these, “gender integration” can become a compliance exercise with limited operational value.

#### 1. Early and Continuous Gender Analysis (as a design input, not a baseline annex)

Embedding gender analysis at concept and feasibility stages helps identify who uses services, who bears unpaid care burdens, who faces safety barriers, and who controls resources and decisions. On complex projects, the biggest payoff comes when findings directly shape technical choices (routing, service standards, O&M models, staffing plans) and when teams revisit assumptions during implementation as conditions shift. Without continuous iteration, projects can “lock in” inequitable designs, especially when engineering, procurement, and safeguards run on separate tracks. Example (Peru – Rural Roads): World Bank operational guidance on mainstreaming gender in road transport describes Peru’s Second Rural Roads project as using participatory workshops to capture women’s transport needs, improving not only roads but also non-motorised tracks often used by women. Reported survey results included 70% of surveyed women travelling more and further, 67% reporting they travelled more safely, and 43% reporting that the project enabled additional income linked to road maintenance and micro-enterprises. The same guidance highlights the enabling role of staff training, gender-focused initiatives, and the tracking of gender-related indicators.

#### 2. Clear Accountability Mechanisms (quotas, roles, budgets, and decision rights)

Complex projects often dilute accountability, and gender is “owned by everyone,” which can mean “owned by no one.” Strong practice includes: (i) named gender focal points with authority, (ii) gender-responsive procurement/contract clauses, (iii) minimum participation thresholds in community governance, and (iv) dedicated budgets for

participation, training, and monitoring. The risk is that quotas become tokenistic if participation is not meaningful (e.g., women present but excluded from decision-making), or if meeting and time demands are incompatible with care workloads. Example (Nepal – Rural Water Supply & Sanitation): An implementation status report for Nepal’s rural water and sanitation programme notes that Water Supply and Sanitation User Committees have at least three women members and reports that the minimum requirement was met in all schemes, with the target exceeded in hundreds of schemes. It also points to the development of a sectoral M&E/MIS system to ensure sustained accountability beyond a single project cycle.

### **3. Safety, Safeguarding, and “Do No Harm” (especially in civil works and labour influx settings)**

In complex projects with contractors, labour influx, and community interfaces, gender integration must include robust safeguarding—particularly around sexual exploitation, abuse, and harassment (SEA/SH) and workplace harassment. Best practice typically combines risk screening, codes of conduct, contractor obligations, safe and confidential reporting, survivor-centred referral pathways, and GRMs that are accessible to women and marginalised groups. A standard failure mode is the creation of “paper systems” without trusted reporting channels, protection from retaliation, and effective enforcement capacity. Good practice guidance: Both the World Bank and Asian Development Bank (ADB) have issued detailed notes on preventing and managing SEA/SH risks in investment projects, including advice on integrating reporting and case handling into grievance mechanisms for projects with civil works.

### **4. Gender-Responsive Service Design (infrastructure + operations + culture)**

Gender integration is often most visible and most effective when it reshapes the service experience, including safety, accessibility, schedules, staffing, complaint handling, and institutional norms. For complex urban systems (metro/BRT/water utilities), this means combining “hard” measures (design, lighting, CCTV, facilities) with “soft” measures (protocols, training, response systems, workforce inclusion). The challenge is sustaining quality and responsiveness at scale, beyond pilot corridors or flagship assets. Example (Ecuador – Quito Metro): The World Bank’s “She Drives Change” results summary reports that the Quito Metro project implemented measures addressing sexual harassment (including protocols and staff training). In early implementation, it reports that 96% of women felt safe on the metro, with 7 harassment cases recorded over ~20 million trips, and 40% of women’s participation in the metro workforce (exceeding a 20% target)—including a major increase in women in management.

### **5. Gender-Responsive Monitoring, Learning, and Adaptation (sex-disaggregated + decision-relevant)**

On complex projects, gender outcomes often depend on implementation details that only become visible through monitoring: who attends, who benefits, who complains, who drops out, and why. Best practice is to define a small set of decision-relevant indicators (sex-disaggregated and intersectional where feasible), link them to management actions (what changes when an indicator moves), and routinely feed findings into operational decisions. The risk is collecting “nice-to-have” indicators that are not used or relying on digital tools that exclude groups with lower access or literacy.

Based on ADB/WB/OECD toolkits<sup>2</sup>

In practical terms, this is a recommendation-oriented conclusion: it highlights a time-sensitive window to operationalise gender earlier and more systematically (e.g., through gender-responsive indicators, engagement segmentation, and routine gender analysis) to mitigate implementation risks and maximise the likely outcome as evidence accumulates.

### **5.1.6 Conclusion 6: Weak MEAL integration constrains learning, adaptation, and causal understanding**

**Limitations in the integration and maturity of MEAL systems—both at the project level and institutionally within GWI—constrain the ability to trace causal mechanisms, learn from behavioural signals, and adapt implementation in a timely manner.** While basic monitoring functions exist, they are not yet configured to support behaviour-sensitive, real-time learning.

The evaluation finds that existing MEAL arrangements, while functional for compliance, safeguards monitoring, and results reporting, are not yet fully configured to support real-time learning, adaptive management, and detailed causal analysis. Current systems were primarily designed to meet accountability and reporting requirements, and they have performed that role appropriately. However, the application of a behaviour-sensitive and theory-based RTE places additional demands on data integration, feedback cycles, and analytical depth—par-

ticularly with respect to trust, usage patterns, and coordination dynamics.

GWI's M&E function operates with limited staffing, fragmented data systems, and a strong orientation toward output and compliance reporting. There is no integrated platform linking physical progress, financial execution, service performance, and user perceptions. Baseline data for behavioural indicators are weak or absent, and perception tracking is sporadic. As a result, the RTE itself has become a primary source of behavioural insight, rather than building on an institutionalised learning system.

Without robust MEAL systems, real-time evidence risks remaining external and episodic. Strengthening MEAL integration is therefore not a technical upgrade but an institutional necessity. Doing so enables causal analysis, supports adaptive management, and protects the long-term value of infrastructure investments by ensuring that learning informs decisions while change is still possible.

### **5.1.7 Conclusion 7: The principal risk to WSIP is not technical failure, but delayed institutional response to learning**

**The most significant risk facing the WSIP is not related to engineering quality or financial viability, but to the possibility that institutional responses to emerging evidence occur too slowly to influence outcomes.**



While the project has generated timely and credible insights through the RTE, the window for translating learning into action—particularly in relation to behavioural change, trust-building, and service adoption—is finite. If institutional decision-making does not absorb and operationalise this learning during implementation, the project risks consolidating an infrastructure-first trajectory that is difficult to correct later.

The RTE has identified clear signals regarding trust deficits, community disengagement, behavioural uncertainty around metering and payment, and weaknesses in MEAL integration. These findings emerged early enough to inform corrective action. However, the evaluation also observes capacity constraints, fragmented responsibilities, and limited institutional routines for absorbing real-time evidence into operational decisions. In the absence of clear ownership, timelines, and incentives to act on recommendations, learning risks remain advisory rather than directive. Experience in similar infrastructure projects suggests that once construction advances beyond certain thresholds, opportunities to reshape user behaviour and institutional practices diminish significantly.

This conclusion has direct implications for development effectiveness. Infrastructure investments of this scale generate their highest returns when institutional systems adapt alongside physical delivery. Delayed response to learning increases the risk that WSIP achieves output completion without securing proportional behavioural and social outcomes, such as sustained service use, reduced reliance on bottled water, and improved trust in public provision. Recognising institutional responsiveness as a core determinant of success reframes accountability: not only for building assets, but also for ensuring that evidence informs decisions while adaptation remains possible.

#### **5.1.8 Conclusion 8: There are early indications that real-time evaluation is already influencing implementation dynamics and relationships**

**An important and somewhat unexpected conclusion of the evaluation is that the RTE itself is already influencing project dynamics, relationships, and perceptions.** Through its presence, questioning, and facilitation of reflection, the RTE has begun to function not only as an assessment tool but also as an intervention that shapes awareness, dialogue, and responsiveness among stakeholders.

Field visits and interviews provide early indications that the RTE has increased the project's visibility in certain communities and has prompted more open discussions among implementers, service providers, and users. In some instances, community members reported feeling “heard” for the first time. At the same time, institutional actors acknowledged that the evaluation surfaced issues—particularly around trust and behaviour—that had not been sufficiently foregrounded previously. These effects occurred despite the RTE's limited formal authority, suggesting that its relational and reflective dimensions are generating influence beyond data collection.

This conclusion underscores the strategic value of RTE in complex projects. When conducted as a learning-oriented, dialogic process, RTE can catalyse behavioural and institutional change even before formal recommendations are implemented. Recognising this effect strengthens the case for embedding RTE earlier and more systematically, not only to generate evidence but to actively support adaptive management, legitimacy, and trust across the project lifecycle.

## **5.2 Lessons learned**

This section distils the key lessons learned from the RTE of the WSIP, drawing on a systematic analysis of implementation experience, behavioural insights, and institutional dynamics. The lessons go beyond project-specific findings to highlight transferable insights relevant to the design, delivery, supervision, and evaluation of complex infrastructure and service-delivery interventions. They are intended to inform future operations, support cross-project learning, and contribute to broader institutional dialogue on strengthening development effectiveness through improved integration of technical delivery, community engagement, and adaptive learning.

### **5.2.1 Lesson 1: Infrastructure is a necessary condition for impact, but behaviour is the decisive one**

The WSIP demonstrates that infrastructure investments create the potential for development impact, but behavioural change determines whether that potential is realised. Improvements in water treatment and service availability do not, on their own, guarantee adoption, reduced reliance on alternative sources, or willingness to pay. These outcomes depend on user perceptions, trust in service quality, and the lived experience of reliability over

time. This lesson is transferable across infrastructure sectors: projects that focus predominantly on asset delivery, without explicitly managing behavioural pathways, risk underperforming relative to their development objectives. Designing projects with behavioural outcomes treated as first-order results—supported by indicators, engagement strategies, and adaptive mechanisms—substantially increases the likelihood of sustained impact.

### **5.2.2 Lesson 2: Early and continuous community integration is a prerequisite for effective learning and adaptation**

The evaluation highlights that community engagement is not simply a safeguard requirement or a communication exercise, but a core learning function. Delayed or episodic engagement limits the ability to test assumptions, understand behavioural drivers, and adapt implementation strategies when it still matters. In WSIP, the late integration of communities constrained insight into trust dynamics and user readiness. For future projects, integrating communities early and maintaining structured feedback loops throughout implementation, enable real-time learning, strengthen legitimacy, and reduce the risk of misalignment between service provision and user expectations. Community engagement should therefore be designed as an iterative process that informs decision-making rather than as a downstream activity. See Box 1 below for good practices in community engagement.

### **5.2.3 Lesson 3: Trust formation follows a different logic and timeline than construction**

A key lesson from WSIP is the structural mismatch between construction timelines and the slower, cumulative process of building trust. While infrastructure delivery is milestone-driven and predictable, trust develops through repeated positive experiences, transparency, and institutional responsiveness. Communities often adopt a “wait-and-see” approach, testing improvements over time before changing entrenched behaviours. This temporal gap is frequently underestimated in project design. Recognising trust as a time-dependent mechanism allows implementers to adjust sequencing, manage expectations, and protect early operational phases from underutilisation or reputational damage. Projects that align engagement and demonstration strategies with the realities of trust formation are better positioned to secure early and sustained uptake.

### **5.2.4 Lesson 4: Strong MEAL systems are central to institutional performance, not ancillary reporting tools**

The WSIP experience reinforces that MEAL systems are most valuable when conceived as institutional infrastructure rather than compliance mechanisms. Fragmented data systems and output-focused monitoring limit the ability to understand causal dynamics, particularly in projects where outcomes depend on behavioural and institutional change. Investing early in integrated MEAL systems—linking physical progress, financial execution, service performance, and user perceptions—enhances adaptive management and strengthens accountability. This lesson is highly transferable: robust MEAL systems enable organisations to learn systematically, respond to emerging risks, and protect the long-term value of complex investments.

### **5.2.5 Lesson 5: Real-time evaluation can actively shape outcomes when embedded as a learning practice**

An important insight from WSIP is that RTE can serve as more than an assessment tool. When embedded as a learning-oriented, dialogic process, it can influence relationships, awareness, and decision-making during implementation. The RTE created spaces for reflection, surfaced behavioural risks early, and increased visibility of community perspectives. This demonstrates that evaluation, when integrated into management cycles, can act as a catalyst for adaptation rather than a retrospective judgment. For development institutions, embedding RTE practices early and aligning them with operational decision points strengthen responsiveness, enhance legitimacy, and improve development effectiveness across portfolios.

### **5.2.6 Lesson 6: Coordinated supervision and evaluation missions enhance learning while preserving accountability**

The WSIP experience illustrates the added value of coordinating supervision and evaluation missions in complex projects, provided that institutional roles, independence, and objectives are clearly maintained. While supervision and evaluation serve different purposes—one focused on fiduciary oversight and implementation progress, the other on learning, causal understanding, and effectiveness—their coordination can create powerful oppor-

tunities for mutual learning. Joint or closely sequenced missions allow technical, operational, and evaluative perspectives to inform one another, reducing duplication, improving contextual understanding, and enhancing the relevance of findings.

At the same time, the evaluation underscores the importance of safeguarding evaluative independence. Coordination should not imply convergence of mandates or dilution of critical distance. Instead, it should enable

“cross-pollination” of insights, where supervision benefits from deeper behavioural and institutional analysis, and evaluation gains from up-to-date operational knowledge and constraints. When managed well, this approach strengthens the overall evidence base, accelerates learning, and improves decision-making without compromising accountability or objectivity. For development institutions, designing structured interfaces between supervision and evaluation functions represents a pragmatic way to enhance effectiveness while respecting their distinct roles.



## Box 2 Good practices of community engagement on water projects

Effective community engagement is essential for creating responsive, inclusive, and resilient water systems. By integrating local knowledge, fostering transparency, and building partnerships, water utilities and governments can address inequities, improve service delivery, and enhance system sustainability. However, the success of these approaches depends on context, implementation, and ongoing adaptation, as challenges such as scalability, equity, and institutional capacity often persist.

### 1. Early and Continuous Community/Customer Involvement

Engaging communities from the planning stage and throughout project implementation helps ensure that water systems are functional, culturally appropriate, and sustainable. While this approach can foster ownership and improve outcomes, its effectiveness depends on meaningful participation, trust-building, and adaptive management. Without these elements, engagement risks becoming superficial or tokenistic, failing to address deeper structural issues.

In Kampala, Uganda, the National Water and Sewerage Corporation (NWSC) involves communities in planning, tariff design, and infrastructure maintenance, reducing connection fees and adapting technical requirements to local realities. This approach has increased customer satisfaction to 85% and improved service reliability in informal settlements. However, challenges remain in scaling participation equitably across all neighbourhoods, as some areas still face barriers such as limited awareness, uneven access to decision-making processes, and persistent infrastructure gaps. While the programme demonstrates the value of early engagement, it also highlights the need for continuous effort to ensure inclusivity and address systemic inequities (African Development Bank, 2021).

### 2. Transparency and Data Accessibility

Transparency and open data platforms can empower communities to monitor progress, hold utilities accountable, and participate in decision-making. While these tools build trust and enable evidence-based management, their impact depends on accessibility, digital literacy, and institutional responsiveness. Without addressing these factors, transparency initiatives may exclude marginalised groups or fail to translate data into actionable change.

Johannesburg Water's "Water Ambassadors" programme uses a mobile app for real-time leak reporting, making leak and repair data accessible to residents. This initiative has reduced NRW by 20% and cut repair times from 7 days to 48 hours in pilot areas. However, digital divides and uneven internet access limit participation in some communities, particularly in low-income and peri-urban areas. Additionally, while the app improves transparency, its effectiveness relies on the utility's capacity to respond promptly to reports and maintain open communication with residents. The programme illustrates how transparency can enhance accountability but also underscores the need for complementary measures to ensure equitable access and responsiveness (Water Research Commission, 2022).

### 3. Public-Community Partnerships and Integrated Urban Water Management (IWRM)

Public-community partnerships (PCPs) and IWRM frameworks support more efficient, equitable, and sustainable water management by integrating community knowledge with utility expertise. While these approaches can improve service delivery and resilience, their success depends on strong governance, equitable resource allocation, and long-term commitment. Without these conditions, partnerships risk becoming uneven or unsustainable, failing to deliver on their potential.

In Santiago, Chile, Aguas Andinas collaborates with local municipalities and NGOs through the "Agua para

Todos” initiative. Communities co-design tariffs, monitor water quality, and maintain infrastructure, thereby reducing inequities in pilot zones such as Lo Hermida and La Pintana. These efforts have improved service reliability and reduced water loss in areas previously plagued by illegal connections and ageing networks. However, broader equity gaps persist due to privatisation challenges, uneven investment across neighbourhoods, and limited scalability of participatory models. The initiative demonstrates how partnerships can enhance local water management but also highlights the need for broader systemic changes to address structural inequities (Inter-American Development Bank, 2023).

Based on cases from AfDB, WRC and IADB<sup>3</sup>

## 6 Recommendations

The following recommendations are derived from the evaluation’s conclusions and transferable lessons.

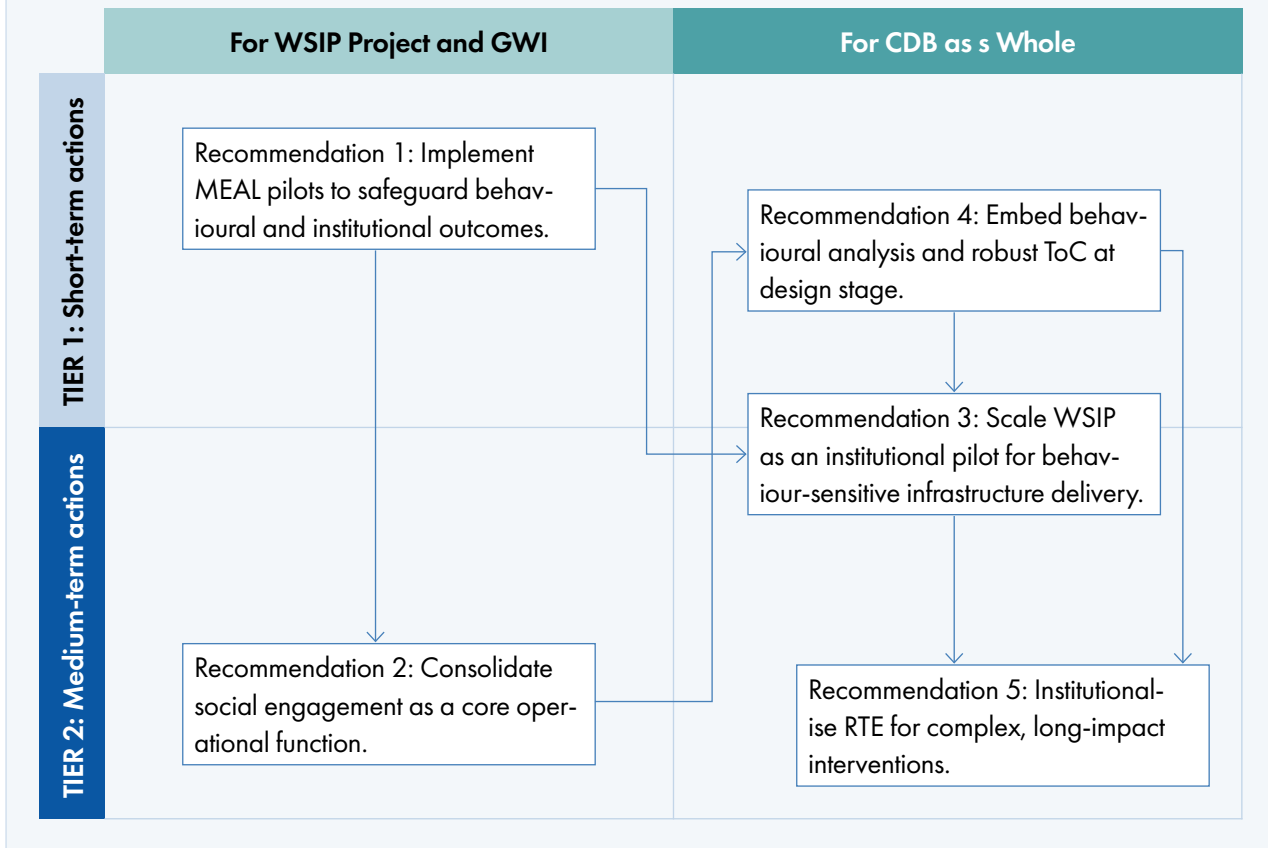
They are designed to strengthen the sustainability of results and ensure consistency between implementation outcomes and the intervention’s intended long-term impacts. They are framed for different institutional levels—project, executing agency, operational teams, and evaluation functions—and emphasise the timely integration of behavioural insight, structured learning, and adaptive management. Together, they aim to ensure that current and future investments consolidate durable value and that learning informs decision-making while meaningful adjustment remains possible.

These recommendations first call for immediate operational measures within WSIP to protect impact during commissioning—specifically, activating behavioural monitoring, clarifying accountability for LLs, and strengthening coordination between operations, community liaison, and MEAL functions. After that, they attend actions for GWI and CDB as a whole to strengthen their portfolios. They propose consolidating these efforts through modest system adjustments, including integrating behavioural indicators into dashboards, formalising communication protocols, and reinforcing staff capacities. Finally, they outline longer-term institutional steps, such as investing in integrated data orchestration, embedding RTE principles into supervision practices, and scaling behaviour-

sensitive MEAL approaches across the CDB portfolio. See *figure 1*.

A pragmatic prioritisation would sequence the recommendations into the presented implementation horizons. First, immediate actions that can be undertaken within existing structures and budgets should focus on activating behavioural monitoring within WSIP, clarifying internal accountability for LLs, integrating perception and service-use indicators into current reporting routines, and strengthening coordination between operations, community liaison, and MEAL staff. These measures primarily require management direction and disciplined follow-through rather than new systems. Second, a consolidation phase would involve modest additional resources and procedural adjustments, including formalising communication protocols linked to commissioning milestones, upgrading dashboard functions to integrate behavioural and operational data, and providing targeted capacity development for MEAL and social teams. Third, longer-term institutional reforms—such as investing in integrated data orchestration platforms, embedding RTE principles into supervision guidance, or expanding behaviour-sensitive monitoring across the portfolio—would require dedicated budget lines, leadership endorsement, and potential adjustments to internal policies.

Figure 1 Map of recommendations



Source: Real-Time Evaluation of the Water Supply Improvement Project, Guyana

### Recommendation 1: Implement MEAL pilots to safeguard behavioural and institutional outcomes.

Ensure that infrastructure investments translate into sustained service adoption and trust by rapidly prioritising targeted MEAL pilots that turn the RTE recommendations into practical, testable actions within the WSIP implementation cycle.

#### Actions:

- Operationalise the MEAL pilots defined in the MEAL Architecture Report as mandatory implementation components, not as parallel analytical exercises.
- Pilot structured tools to track user perceptions, trust in water quality, service reliability, and behavioural responses to metering and billing, ensuring regular data collection at key service milestones (pre-commissioning, early operation, stabilisation).
- Integrate pilot findings into routine project management, supervision meetings, and contractor coordination forums, with clear protocols for translating insights into corrective actions.
- Assign clear ownership within the WSIP team for each MEAL pilot, including responsibilities for data collection, analysis, reporting, and follow-up actions.
- Use pilot results to validate or adjust behavioural assumptions in the ToC and to refine communication, engagement, and sequencing strategies during implementation.

**Responsible Unit:** WSIP Project Management Unit, with support from GWI's MEAL function and coordination with the CDB Task Team.

**Timeline:** Immediate initiation; pilots operational within 3–6 months and sustained throughout the remaining implementation period.

**Recommendation 2: Consolidate social engagement as a core operational function.**

Strengthen the effectiveness and sustainability of WSIP outcomes by consolidating social engagement as a core operational function, with reinforced community liaison capacity, empowered social specialists, and close integration with the MEAL unit to actively manage behavioural change, trust, and service adoption.

**Actions:**

- Formalise social engagement as an integral part of WSIP operations, with clear mandates, decision rights, and accountability alongside technical delivery.
- Strengthen the community liaison function by ensuring consistent presence in project areas, standardising engagement protocols, adequate training to all consultants and aligning liaison activities with key construction, commissioning, and early operation milestones.
- Urgently strengthen and formalise the role of social specialists as a critical operational function, recognising that their current limited integration represents one of the weakest links in WSIP implementation and requires immediate corrective action to manage trust-building and behavioural change effectively.
- Empower social specialists to play a proactive role in shaping communication strategies, identifying behavioural risks, and advising operational teams on trust-building and adoption dynamics, rather than limiting their role to safeguards compliance.
- Integrate community liaison and social teams closely with the MEAL unit to ensure that community feedback, perception data, and behavioural signals are systematically captured, analysed, and fed into management decisions.
- Establish regular internal coordination routines between operations, social engagement, and MEAL teams to review emerging evidence, agree on corrective actions, and track follow-up, ensuring learning translates into operational adjustments in real time.

**Responsible Unit:** WSIP Operations Team, in close coordination with GWI Community Liaison, Social Development, and MEAL units.

**Timeline:** Immediate initiation; strengthened coordination structures and protocols established within 3 months, with ongoing implementation throughout the remaining project lifecycle.

**Recommendation 3: Scale WSIP as an institutional pilot for behaviour-sensitive infrastructure delivery.**

Use WSIP as a learning and demonstration platform to strengthen GWI’s broader portfolio of infrastructure projects where behavioural change is central to impact.

**Actions:**

- Formally designate WSIP as an institutional learning pilot within GWI, with explicit objectives related to integrating infrastructure delivery and behavioural change management.
- Develop and test early-stage community communication protocols that explain service improvements, timelines, quality expectations, and user responsibilities before commissioning.
- Design and pilot incentive and feedback mechanisms (e.g. service guarantees, quality disclosures, responsive complaint resolution) that reinforce trust and adoption rather than relying solely on enforcement or billing measures.
- Standardise the use of integrated MEAL systems across WSIP and progressively across other projects, linking technical performance data with customer service and behavioural indicators.
- Document lessons learned, tools, and decision templates from WSIP and translate them into internal guidance, training modules, and standard operating procedures for future projects.

**Responsible Unit:** GWI Senior Management, in coordination with Operations, Customer Service, and M&E units.

**Timeline:** Short to medium term (6–12 months), with progressive scaling across the project portfolio.

**Recommendation 4:** Embed behavioural analysis and robust ToC at design stage.

Strengthen the coherence between expected results and long-term impacts by integrating social and behavioural dimensions into project design beyond safeguards compliance.

**Actions:**

- Require project design teams to conduct structured behavioural analysis during preparation, identifying key user groups, trust dynamics, incentives, and risks to adoption.
- Develop behaviour-explicit ToC that articulate causal mechanisms linking infrastructure, institutional actions, and behavioural outcomes, including intermediate states and assumptions.
- Ensure that RMF include indicators that capture behavioural and perceptual change, not only physical outputs and service coverage.
- Integrate behavioural risks and mitigation measures into project risk matrices, supervision plans, and implementation arrangements from the outset.
- Strengthen collaboration between technical, social, and results specialists during design to ensure coherence between engineering solutions, engagement strategies, and monitoring systems.

**Responsible Unit:** GWI Senior Management, in coordination with Operations, Customer Service, and M&E units.

**Timeline:** Short to medium term (6–12 months), with progressive scaling across the project portfolio.

**Recommendation 5:** Institutionalise RTE for complex, long-impact interventions.

Strengthen institutional learning and adaptive management by adopting RTE as a standard practice for complex interventions with long impact horizons.

**Actions:**

- Develop internal guidance defining when and how RTE should be applied, focusing on projects characterised by high complexity, behavioural dependence, or long impact horizons relative to implementation timelines.
- Establish criteria for selecting RTE projects and for aligning RTE timing with key decision points in the project lifecycle (e.g. early implementation, mid-course correction windows).
- Promote the use of theory-based approaches, including causal mechanisms and contribution analysis, as core elements of RTE practice.
- Create structured interfaces between RTE outputs and supervision, management, and learning processes, ensuring findings inform decisions without compromising evaluative independence.
- Pilot RTE applications across selected sectors (e.g. infrastructure, policy-based operations) and consolidate lessons into an institutional learning agenda for scaling and adaptation.

**Responsible Unit:** CDB OIE, in coordination with operational and supervision units.

**Timeline:** Medium term (12–18 months), starting with pilot applications and progressive institutionalisation.



# Endnotes

- 1 According to the ToRs for this evaluation “the RTE of the WSIP in Guyana has been included in the Work Plan as a strategic pilot to test CDB’s capacity for agile, learning-oriented evaluation approaches. WSIP is a flagship, multi-million-dollar investment co-financed by CDB and the Government of Guyana, covering the construction of five water treatment plants, deep wells, and transmission mains across several regions. The project is already under implementation, with all major contracts awarded. This makes it highly evaluable and an ideal case for real-time learning and operational feedback.”
- 2 Caballero, L., & Alcahuasi, N. (2007, October). Gender in Peru: Can women be integrated into transport projects? (En Breve, No. 112). World Bank, World Bank. (2012, July 29). Implementation Status & Results Report: Nepal Second Rural Water Supply & Sanitation Project (P071285). (Public Disclosure Copy), World Bank. (2025, July 10). She Drives Change: Empowering Women in Transport. (Results feature—includes Quito Metro), World Bank. (n.d.). Urban Transport – She Drives Change. (Quito Metro results page), Asian Development Bank. (2023). Good Practice Note on Addressing SEA/SH in ADB-Financed Projects with Civil Works, Asian Development Bank. (2023). Integrating SEA/SH Reporting and Case Handling into Project Grievance Redress Mechanisms: Good Practice Note for ADB-Financed Projects with Civil Works and World Bank. (2025, October 29). Procurement: SEA/SH Prevention and Response Measures – Contractor/Sub-Contractor Disqualification for Breach of SEA/SH Contractual Obligations: Questions and Answers for Borrowers and Contractors.
- 3 African Development Bank (2021). NWSC Uganda: Community Engagement in Water Projects, Water Research Commission (2022). Transparency and Data Accessibility in Water Management and Inter-American Development Bank (2023). Public-Community Partnerships in Water Management.



