

From Shelfware to Action: Unpacking the Utilization of Project Evaluation Reports in Zimbabwe's Development Sector

Albert Munyanyi

Development Studies, African Leadership Institute, Masvingo, Zimbabwe, Zimbabwe

DOI: <https://doi.org/10.51244/IJRSI.2025.12050096>

Received: 07 May 2025; Accepted: 11 May 2025; Published: 12 June 2025

ABSTRACT

This paper critiques the use of project evaluation reports in Zimbabwe's development sector. The goal is to explore the relationships between findings and decision-making, perceptions regarding evaluation purpose and credibility, and strategies to promote a transition from shelf ware to insights that are used in decision-making. A mixed-methods sequential explanatory research design was employed. Initially, qualitative data was gathered through case studies of development organizations (NGOs, government-linked organizations, donor-funded projects), involving semi-structured interviews with stakeholders (Commissioner representatives, evaluators, project managers, M&E officers) and thematic analysis of project evaluation reports and related organisational documents. This was followed by a quantitative survey administered to a broader sample of practitioners to ascertain the prevalence of identified themes and patterns. The Weiss's models of knowledge utilization and Patton's Utilization-Focused Evaluation principles guide the theoretical framework. The findings, integrating both qualitative depth and quantitative breadth, indicate that while evaluation is a widely-practiced routine, its utilization is highly heterogeneous. Conceptual and symbolic uses are reportedly more frequent than direct instrumental use. Key factors influencing utilization include perceptions of credibility and relevance, timeliness, communication clarity, organizational culture, stakeholder involvement, and the political/funding environment. Quantitative data further highlights the significant impact of specific barriers such as resource limitations and lack of formal follow-up mechanisms. This study describes specific actionable strategies for commissioners, evaluators and implementing organizations to improve evaluation utility, supported by both qualitative narratives and quantitative prioritizations.

Keywords: Evaluation utilization, knowledge utilization, evidence-informed decision-making, shelf ware, project evaluation, utilization-focused evaluation, mixed-methods research.

INTRODUCTION

Background: The Zimbabwean Development and Evaluation Landscape

Zimbabwe has a complex and dynamic socio-economic environment, its national development goals are framed in the National Development Strategy 1 (NDS1), significant resources are committed to pursuing these investments from government, international donors and non-governmental organizations (NGOs) (Government of Zimbabwe, 2020). In this regard, ensuring the effectiveness, efficiency, and accountability of development interventions is of utmost importance. Locally and internationally, project evaluation has developed into an important mechanism for performance assessment, learning generation, resource allocation and experience representation to stakeholders (Abbott & Makumbe, 2010; GoZ OPC, 2015). The implementation of Integrated Results-Based Management (IRBM) at the government level provides an additional indication of the professed commitment to measuring performance and using evidence (Mpfu, 2018; Mamimine & Zebron, 2021a). Professional organisations such as the Zimbabwe Evaluation Association (ZEA) also seek to promote quality evaluation practice (ZEA, 2017).

The “Shelf ware” Dilemma

Amidst the proliferation of evaluation activities and with it, the normative push towards evidence-informed decision-making, one issue continues to challenge the development sector globally and anecdotally within Zimbabwe: the production of shelfware. This refers to evaluation reports that are produced despite significant investment in their production, but then remain largely unused, gathering dust on shelves (physical or digital) – not informing practice, policy, or organizational learning (Patton, 2008; Bamberger et al., 2012). Reports might be commissioned to meet donor needs or organizational mandates, but their results do not penetrate decision-making processes or shape future programming. This is a large waste of resource and a real opportunity cost for improving development effectiveness. Knowing what causes this gap and how to make evaluation both available and useful is another important part of this discussion.

Aim, Objective and Research Questions

Based on anecdotal accounts, the study seeks to go beyond the missives reporting on dead-ended development in Zimbabwean development literature by reflecting empirically on what current practice, perceptions and pathways are around project evaluation reporting in Zimbabwe. Rooted in the current context (Masvingo, Zimbabwe, April 2025), this research is attempting to learn:

1. How well are project evaluation reports being used in Zimbabwe’s development sector decision making (strategic, programmatic, operational)?
2. How do key stakeholders (commissioners, evaluators, project managers, M&E officers, policymakers, where relevant) perceive the purpose, effectiveness, credibility, and overall value of project evaluation reports?
3. What enables or blocks the use of evaluation findings in this context?
4. How can the diversity of actors produce concrete strategies for making project evaluation reports part of their actionable insights and back to informing practice and policy, instead of falling in dormant shelf ware to collect dust?

Importance and Contribution

While much evaluation utilization research takes place at the global level (not least through international initiatives such as the Global Evaluation Agenda 2020), little has been grounded in well-tested evidence from a targeted developing country context. This study, through its mixed-methods approach, offers both in-depth qualitative understanding and broader quantitative validation of utilization patterns, enablers, and barriers. It has practical implications for development practitioners, policymakers, donors, evaluators, and organizations working in Zimbabwe and wanting to increase both the value and the impact of their evaluation investments. By examining the enablers and barriers reported by staff, organizations could design more utilization-focused evaluations also create contexts more conducive to learning and evidence use, and lead to better development outcomes.

Structure of the Paper

Following this introduction, Section 2 reviews relevant literature and outlines the conceptual framework guiding the study that, in particular, draws on theories of knowledge utilization and factors influencing evaluation use. Section 3 outlines the mixed-methods methodology applied in this research, detailing both the qualitative and quantitative phases of the study, including research design, sampling, data collection, and analysis techniques. Thematic analysis of interview and documentary data, complemented by descriptive statistics from the survey data, is reported in Section 4 and highlights the main findings. Section 5 discusses these findings against the literature and conceptual framework and in response to the research questions. Finally, Section 6 summarises the conclusions, and limitations and makes practical recommendations to enhance evaluation utilization in the development sector in Zimbabwe.

LITERATURE REVIEW AND METHODOLOGICAL FRAMEWORK

The Nature of Utilization of Evaluation

Evaluation utilization means the use of evaluation processes and findings by users intended to facilitate decisions, improve a program, create knowledge, and stimulate learning (Patton, 2008; Alkin & Taut, 2003). Its first task is to expand beyond simply wrapping up a given evaluation report, its second to encompass the actual way evaluation matters for people, organizations and policy. The idea is complex, acknowledging that use is not necessarily direct or immediate.

Understanding How and Why Evaluations Are (or Aren't) Used

Weiss's Models of Knowledge Utilization

The pioneering work of Carol Weiss (1979) has served as a foundational typology for understanding the mechanisms through which research and evaluation findings may inform organizational action, moving beyond simplistic direct impact assumptions:

- **Instrumental Use:** This is the most straightforward type of use, in which evaluation findings directly inform specific, identifiable decisions or actions (e.g., adjusting a program component, reallocating resources by effectiveness data). Although often regarded as the gold standard, some research indicates it is a relatively rare occurrence (Weiss, 1979; Cousins & Leithwood, 1986).
- **Conceptual Use:** gradual impact of evaluation findings on how those interested in the evaluation understand, think about, see, or conceptualize problems or solutions. This process is typically indirect and incremental, informing future considerations and not decisions (Weiss, 1979; Shulha & Cousins, 1997). This type of use is more ubiquitous than instrumental use.
- **Symbolic Use (Political/Legitimative Use):** In this mode, evaluation findings are simply manufactured, as it were, to justify or legitimize decisions made in advance, to bolster pre-existing positions, to postpone action, or simply to show that an evaluation has been done (thus satisfying bureaucratic or donor demands), without any real intention of using the findings to achieve improvement or change (Weiss, 1979; Patton, 2008). This is very similar to "shelf ware", where the process of evaluation has meaning but that the content does not.

This framework is useful in the Zimbabwean context as it enables us to understand not only if reports are used, but how and why.

Michael Quinn Patton's Utilization Focused Evaluation (UFE)

Patton (2008) takes a more practical view that evaluations should be purposefully designed to meet their use. UFE is not a methodology per se, but rather a process principle grounded in:

- **Laying out Primary Intended Users:** A plan for who the intended users will be
- **Emphasizing Intended Uses:** Working directly with users to determine the essential questions needing answers and intended uses of the information.
- **Facilitating the Process:** The evaluator will act as a facilitator who guides users on the relevant questions to be defined, how to interpret findings, and how to consider implications.

UFE principles underscore essential aspects in tackling shelf ware: relevance, user involvement, and defining communication based on user requirements.

Supporting Lens: Institutional Theory

According to institutional theory, organizations mirror the structures and practices (e.g., evaluation) of peer organizations and adapt to the pressures of the external environment to be legitimate and accepted

(DiMaggio & Powell, 1983) through mimetic isomorphism, coercive isomorphism (e.g., donor mandates), and normative isomorphism. This lens can help us to understand why evaluations might be regularly commissioned in Zimbabwe's development sector (as a result of donor requirements or as a perceived best practice), but to find their use in symbolic rather than instrumental or conceptual terms, if internal capacity or commitment to instrumental or conceptual use is not prioritized.

Factors that Effect Evaluation Utilization

Various factors impact the effective use of evaluation findings, which have been identified across multiple areas of study and synthesized into frameworks based on the work of scholars like Cousins & Leithwood (1986), Shulha & Cousins (1997), Alkin & Taut (2003), Patton (2008), Bamberger et al. (2012), and Donaldson (2007). These relate to multiple major areas.

At the heart of utilization are the features of the evaluation itself (process and product). The literature, including much of the work reflected by these scholars, stresses two key factors: relevance (e.g., whether findings speak to the relevant questions users have) and credibility (i.e., whether the evaluation has been conducted using rigorous methodology, by qualified evaluators, and without bias in research question formulations). Timeliness is repeatedly pointed out as critical. In addition, timely communication and reporting (e.g. clear, concise, accessible language, appropriate formats and actionable recommendations) are crucial which is echoed if not further addressed in Patton's (2008) focus on utilization-focused evaluation. It also underscores the importance of stakeholder and end-user engagement throughout, starting with evaluation design and extending through interpretation, which concurs with participatory approaches put forth by authors such as Cousins & Leithwood (1986) and Shulha & Cousins (1997); embracing participatory approaches fosters local ownership and skill development.

Organizational factors play a major role in mediating the acquisition of evaluation findings. Alkin & Taut (2003) and others have investigated the role of organizational context. A clear need for information in the organization prepares the ground. The organizational culture itself is also key; cultures that prioritize learning, inquiry, evidence, and tolerance for counterintuitive findings enable greater uptake of findings. Leadership buy-in, which involves leaders advocating for evaluation — and driving its use — offers much-needed momentum. Researchers like Bamberger et al (2012) have emphasized the practical necessity of capacity, i.e. staff skills and resources to interpret and act on findings for organizations. At last, formalize use by internal processes, such as management responses or learning workshops, to analyse and discuss findings, and to follow up on them.

Individual user characteristics also come into play, identifying the attributes of the person intended to use the evaluation. Also, the person's authority and power in the organization affect their ability to act on findings. Their own value and utility of evaluation belief draws its receptivity. Finally, the investment of an individual in the program under evaluation may help improve feelings of buy-in for the use of findings, something that Patton (2008) associates with building on primary intended users.

Finally, larger contextual factors create the backdrop for utilisation. The political environment, including support or opposition for the program or evaluation as well as general stability, can significantly impact outcomes, a fact that has been much discussed in development evaluation contexts (e.g., Bamberger et al., 2012). The funding environment, governed by donor priorities and the availability of resources to implement recommendations, affects their implementation. decisions cycles (such as, budget or strategic planning) is also an important consideration; ensuring findings will be up for consideration when another relevant decision is made.

Use of evaluations in developing countries and Zimbabwe

Though these general factors apply to a wide variety of contexts, research identifies particular nuances within developing countries. Challenges include lower institutional capacity for M&E, data scarcity and poor

quality, increased reliance on donor agendas (which may focus on accountability more than on learning), political instability hindering uptake of evidence in policy, and lack of resources to implement recommendations (Bamberger et al. 2012; Holvoet & Renard 2007; Goldman et al. 2018). African Studies highlight the need for context sensitivity, participatory methodologies and capacity building for increased use (Abbott & Makumbe, 2010; Uwizeyimana, 2020).

There seems to be limited specific research conducted on evaluation utilisation in Zimbabwe. Nevertheless, research on M&E capacity (Abbott & Makumbe, 2010; Zhou, 2013) and research on RBM implementation (Mpfu, 2018; Mamimine & Zebron, 2021b) reveal many of the potential barriers identified above including: capacity constraints; challenges relating to the merging of data; absence of buy-in; and political factors. The present study seeks to address this gap by directly examining utilization patterns and perceptions in Zimbabwe's development sector.

METHODOLOGY

Research Philosophy and Approach

This study employed a mixed-methods sequential explanatory research design. According to Creswell & Plano Clark (2017), this approach involves initial collection and analysis of qualitative data, followed by the collection and analysis of quantitative data. The collection and analysis of quantitative data was designed to test and generalize the initial qualitative findings. The interpretivist philosophy guided the initial qualitative phase, aiming to understand the subjective experiences and meanings stakeholders attribute to evaluation utilization. The subsequent quantitative phase leaned towards a post-positivist stance, seeking to measure the prevalence and strength of observed patterns.

Qualitative Phase

Research Strategy

The study employed a multiple case study strategy (Yin, 2018) for the qualitative phase. The design permits the exploration of evaluation use and utilization across diverse organizational contexts in the development sector of Zimbabwe, allowing for comparison and strengthening of findings. Cases were chosen to reflect a variety of types of organizations (e.g., international NGO, local NGO, government-linked projects, donor agency perspective) that commission, conduct, or use project evaluations.

Sampling Strategy

At two levels purposive sampling (Patton, 2015) was used for the qualitative phase:

Case Selection: Organizations were chosen based on their involvement in development projects in Zimbabwe, history of commissioning or use of evaluations, and willingness to participate, with the goal of diversity with respect to sector, size, and sources of funding. A total of five case study organizations were selected.

Participant selection: The study purposively selected key informants within each case organisation (and among independent evaluators) based on their roles and experience in relation to project evaluation. A total of 25 key informants were interviewed (3-6 per case organization/evaluator group). This included:

- Evaluation Commissioners: Those who commission and manage evaluations (e.g., senior managers, donor representatives).
- Program/Project Managers: Persons responsible for implementing the projects being evaluated as well as persons who may use findings.
- M&E Officers/Specialists: Staff with responsibility for the internal M&E systems, and often work to support or liaise on evaluations.

- Consultants or Internal staff who conducted relevant evaluations.
- Policymakers or strategic decision-makers. The purpose of this effort was to gain diverse perspectives on the evaluation use process.

During data collection, saturation was assessed.

Data Collection Methods

There were two main ways qualitative data collection was done:

Semi-structured Interviews: Key informants were interviewed using semi-structured instruments. An interview guide was constructed aligned to the research questions and conceptual framework, with questions exploring topics including: experiences with specific evaluations carried out, perceptions of purpose and quality of evaluations, processes for receiving and discussing reports, examples of findings being used (or not used) and why, perceptions of barriers/enablers to utilization, and suggestions for improvement. The format was semi-structured allowing freedom to follow up on emergent themes (Bryman, 2016). Interviews were undertaken in person and remotely. Interview audios were recorded with consent and were transcribed verbatim.

Document analysis: Relevant documents were collected and analysed to corroborate interview data and provide context. These included: Project evaluation reports (terms of reference, inception reports, final reports); Management response to evaluation reports (if incorrect there will be a separate document that will follow); Project Proposals and design documents (Logframes); Organizational strategic plans or M&E policies; Agendas and minutes of meetings where evaluation findings were discussed (where available).

Data Analysis

Qualitative data were analysed using thematic analysis (Braun & Clarke, 2006, 2021) of the transcribed interview data and relevant sections of documents. The process involved: Familiarising; Initial Coding; Searching for Themes; Reviewing Themes; Defining and Naming Themes; and Reporting. We used the NVivo qualitative data analysis software to organize and aid and the process of coding and theme development.

Quantitative Phase

Survey Instrument Development

A structured questionnaire was developed based on the emergent themes from the qualitative phase. The questionnaire incorporated sections on the demographic information of respondents and their organizations, the frequency of different types of evaluation use (instrumental, conceptual, symbolic) based on Weiss's (1979) typology, the perceived importance of factors influencing evaluation credibility (e.g., timeliness, evaluator expertise, methodological rigor) rated on a 5-point Likert scale as well as the perceived impact of various barriers and enablers to evaluation utilization, which were also rated on a 5-point Likert scale (1=Not at all significant, 5=Very significant).

Sampling and Data Collection

A larger sample of development practitioners (N=134) from a diverse range of NGOs (local and international), government-linked projects, and donor agencies operating in Zimbabwe was targeted for the quantitative survey. A combination of simple random sampling of development organizations from the National Association of Non-Governmental Organizations (NANGO) directory with convenience and snowball sampling was used, leveraging professional networks and contacts from the qualitative phase. The survey was administered online using Google forms over a period of six weeks.

Quantitative Data Analysis

The collected quantitative data was cleaned and statistically analysed using Pydroid 3. Descriptive statistics (frequencies, percentages, means, standard deviations) were computed in order to summarize the responses. This process allowed for the quantification of the prevalence of different utilization patterns and the relative importance of various factors influencing utilization, as perceived by a broader group of stakeholders.

Ethical Implications

A number of important ethical related issues were covered as part of the research process. Participants in both phases were informed about the study's objective, methods, potential advantages, and risks before providing written consent; participants were also made aware that they could opt out at any stage of the research. In all accounts and reporting, individual participant identity (and specific organizations, beyond their general type) were kept confidential; data were further anonymized during transcription and analysis to ensure confidentiality and anonymity. Moreover, all audio recordings and transcripts were securely stored on password-protected devices and in cloud storage, accessible only to the researcher. This survey was entirely optional and voluntary, and no coercion was used. The different case organizations were contacted for verbal approval before their involvement. Notably, as this was a personal professional research project, institutional review board approval was not applicable.

Rigor and Trustworthiness / Validity and Reliability

For the qualitative phase, in order to increase the robustness and credibility of the findings, strategies identified by Lincoln and Guba (1985) were applied. We sought credibility through prolonged involvement within the cases that allowed enough time to be spent in the data collection process, and through triangulation, using both interviews and documents. In addition, member checking was conducted through explaining and discussing with some participants the early findings and interpretations for feedback. Transferability was enhanced by delivering rich, thick descriptions of the cases and their specific contexts thereby allowing readers to make informed judgments about the potential transferability of the findings to other settings. Trustworthiness was based on keeping a clear audit trail by documenting the research process, including fact sheets, methodology outline and data analysis steps. Finally, conformability was enhanced by ensuring the findings were well grounded in the data, a process aided both by the researcher's reflexive journaling.

For the quantitative phase, content validity of the survey instrument was enhanced through expert review by two senior M&E practitioners and a pilot test with 10 professionals not included in the final sample. Internal consistency of Likert scale items was assessed using Cronbach's alpha. The use of a larger sample size in the quantitative phase aimed to increase the generalizability of these specific findings.

Integration of Qualitative and Quantitative Data

The integration of qualitative and quantitative data occurred at the interpretation and reporting stage. The quantitative results were used to confirm, generalize, and provide statistical significance to the themes that emerged from the qualitative data. Conversely, the qualitative data provided context, depth, and illustrative examples to help explain the quantitative findings.

FINDINGS

Findings This section outlines the empirical results drawn from both the qualitative thematic analysis and the quantitative survey data. The results are structured according to the overall aims of the study, highlighting patterns of use, stakeholder perceptions, and what affects the journey of evaluation reports from the time of production to potential use or abandonment.

Patterns of Evaluation Utilisation

Qualitative insights indicated that while evaluation commissioning is routine, actual use varies. Instrumental use was described as infrequent, with conceptual and symbolic uses being more common. As one M&E officer qualitatively shared, “The report didn’t make us change right away, but it made us think critically about our approach...in the future proposal,” highlighting conceptual use. Symbolic use was often linked to fulfilling donor mandates, a “tick-box exercise.”

The quantitative survey data (N=134) provides a broader perspective on these patterns, as shown in Figure 1.

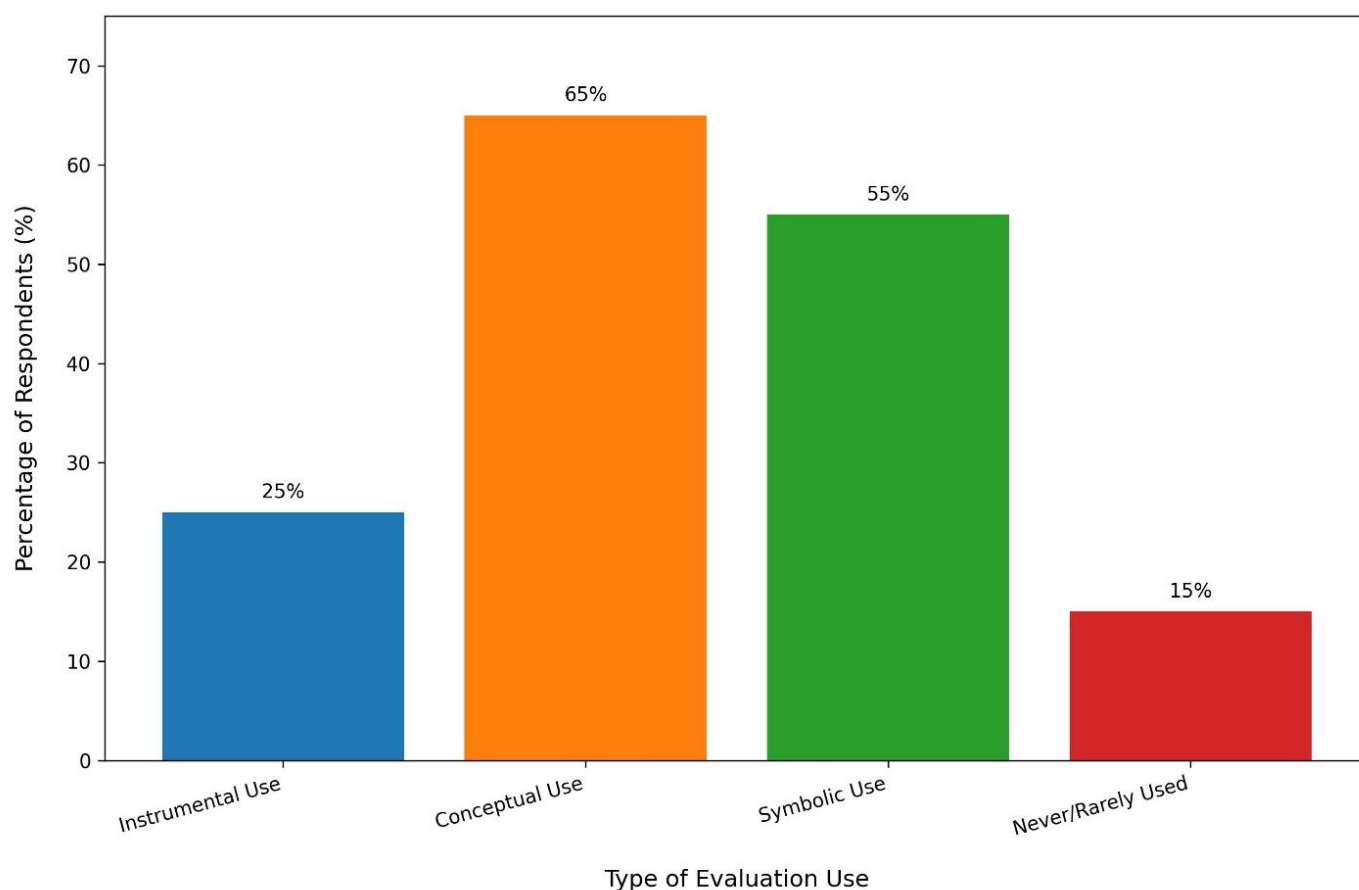


Figure 1: Reported Frequency of Evaluation Use Types by Practitioners (N=134)

The survey confirms the qualitative finding that conceptual use (65%) is most frequently reported, followed by symbolic use (55%). Instrumental use, while present, is reported as "Often/Very Often" by a smaller proportion (25%) of practitioners. This suggests that while evaluations do influence thinking and fulfil obligations, their direct translation into concrete actions is less consistent across the sector. The 15% reporting that reports are "Never/Rarely Used" for any substantive purpose underscores the persistence of the 'shelf ware' problem.

Stakeholder Perceptions of Value, Credibility, and Purpose

Qualitatively, stakeholders expressed varied views. Commissioners often prioritized accountability, while implementers hoped for learning, though some saw evaluations as mere obligations. The credibility of an evaluation was paramount, influenced by evaluator expertise, methodological rigor, and participatory approaches. Lack of timeliness and overly academic reporting styles were common complaints.

The quantitative survey further explored factors influencing perceived credibility. Table 1 presents mean scores for the perceived importance of selected credibility factors.

Table 1: Perceived Importance of Factors Influencing Evaluation Credibility (N=134)

(Scale: 1=Not at all important, 5=Very important)

Credibility Factor	Mean Importance Score	Standard Deviation
Relevance to information needs	4.65	0.55
Clarity and actionability of recommendations	4.58	0.60
Evaluator's technical expertise	4.52	0.68
Methodological rigor and transparency	4.45	0.70
Timeliness of the report	4.30	0.85
Independence/objectivity of evaluator	4.25	0.75
Stakeholder participation in process	4.10	0.80

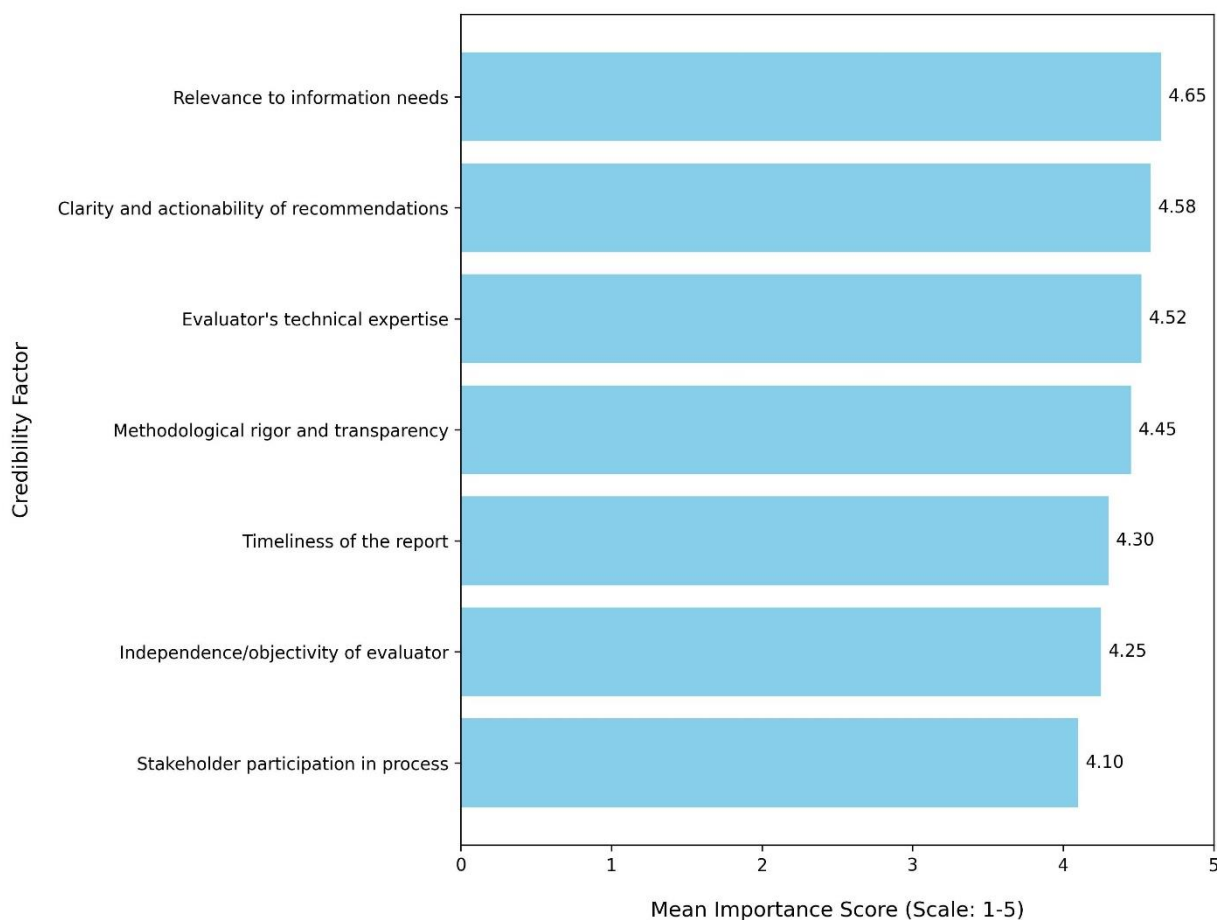


Figure 2: Perceived Importance of Factors Influencing Evaluation Credibility

The survey data reinforces qualitative themes, with "Relevance to information needs" (Mean=4.65) and "Clarity and actionability of recommendations" (Mean=4.58) rated as most important for credibility. While "Timeliness" (Mean=4.30) is still considered very important, its slightly lower mean and higher standard deviation might suggest more varied experiences or expectations regarding this factor compared to the top-rated items, aligning with qualitative complaints about delays.

Factors for Encouraging Use of The Evaluation: What Works?

Qualitative findings strongly emphasized user involvement throughout the evaluation lifecycle, active senior leadership engagement, an organizational culture open to learning, and clear, actionable recommendations as key enablers of utilization. Participatory approaches were frequently lauded.

The quantitative survey asked respondents to rate the impact of various enablers. Table 3 summarizes the top-rated enablers.

Table 2: Top-Rated Enablers of Evaluation Utilization (N=134)

(Scale: 1=No impact, 5=Very high positive impact)

Enabler	Mean Impact Score	Percentage Rating as High/Very High Impact (4 or 5)
Clear, actionable recommendations	4.60	90%
Strong leadership commitment & demand for use	4.55	88%
Stakeholder involvement in evaluation process	4.40	85%
Timely delivery of findings	4.35	82%
Organizational culture valuing learning	4.25	78%

The quantitative results in Table 2 strongly support the qualitative insights. "Clear, actionable recommendations" (Mean=4.60) and "Strong leadership commitment" (Mean=4.55) are ranked highest, with 90% and 88% of respondents respectively viewing them as having a high or very high positive impact on utilization. This underscores the critical role of both the evaluation product and the organizational environment in fostering use.

Barriers to evaluation utilization

Qualitatively, numerous inter-related barriers were identified, including organizational cultures resistant to change, lack of formal follow-up mechanisms, resource limitations, poor quality or irrelevant reports, delayed delivery, poor communication, lack of user involvement, and impractical recommendations. Contextual factors like donor pressures and the socio-economic environment also played a role.

The survey data provides a ranked perspective on the most significant barriers, as shown in Table 3.

Table 3: Top-Rated Barriers to Evaluation Utilization (N=134)

(Scale: 1=No impact, 5=Very high negative impact)

Barrier	Mean Impact Score	Percentage Rating as High/Very High Impact (4 or 5)
Lack of financial resources to implement recommendations	4.50	87%
Absence of formal follow-up/action planning mechanisms	4.42	84%
Reports delivered too late for decision-making	4.38	80%
Competing priorities within the organization	4.30	79%
Recommendations are too vague or impractical	4.15	75%
Lack of organizational learning culture	4.05	70%

The quantitative findings highlight "Lack of financial resources to implement recommendations" (Mean=4.50) and "Absence of formal follow-up mechanisms" (Mean=4.42) as the most impactful barriers, with 87% and 84% of respondents respectively citing them as having a high or very high negative impact. This strongly resonates with qualitative accounts of practical constraints hindering the translation of findings into action, even when reports are otherwise well-received.

DISCUSSION

This section draws out the implications of the findings reported in section 4 above, relating them to the theoretical framework set out in Section 2 (Weiss 1979; Patton 2008; DiMaggio & Powell 1983) and the literature.

The Dominance of Conceptual and Symbolic Use

The mixed-methods findings clearly indicate that while evaluation is a common practice, its direct application to decision-making (instrumental use) is less prevalent than its influence on thinking (conceptual use) or its function as a legitimizing ritual (symbolic use). The qualitative data provided rich examples of these dynamics, while the quantitative survey (Figure 1) confirmed this pattern across a broader sample, with 65% reporting frequent conceptual use and 55% frequent symbolic use, compared to 25% for instrumental use. This pattern aligns with Weiss's (1979) early observations and subsequent research (Cousins & Leithwood, 1986) suggesting instrumental use is often the least common form. In Zimbabwe's complex development context, instrumental use can be perceived as more direct and potentially threatening, often requiring significant operational adjustments (Davis et al., 2008).

The prevalence of symbolic use warrants particular attention. The 'shelfware' phenomenon—reports being commissioned and filed with little substantive engagement—directly reflects Weiss's (1979) category of symbolic or political/legitimative use. The quantitative data showing 55% of respondents frequently engaging in symbolic use provides a stark measure of this phenomenon. Evaluations may serve primarily to demonstrate accountability to external actors (donors, headquarters), legitimize existing practices, or meet bureaucratic demands without a genuine intent to use the findings for learning or change. Institutional Theory (DiMaggio & Powell, 1983) offers a valuable lens here. Many organizations in Zimbabwe's development sector operate within an institutional field where evaluation is increasingly mandated due to coercive pressures (e.g., donor requirements, government RBM directives) and normative pressures (e.g., professional standards, perceived 'best practice'). To gain legitimacy and resources, organizations adopt evaluation structures. However, if internal drivers for evidence use—such as a strong learning culture (rated as a key enabler by 78% in Table 2, but its lack cited as a barrier by 70% in Table 3), adequate M&E capacity, and leadership demand—are weak, organizations may 'decouple' the formal structure (evaluation) from actual practice (utilization). Consequently, despite efforts to promote M&E and RBM, evaluations often primarily fulfil external expectations rather than driving internal change, leading to symbolic use and the persistence of shelf ware. External isomorphic pressures appear to be more potent determinants of evaluation practice than the intrinsic value of evidence-based decision-making within these organizations.

UFE Principles as Enablers

The findings on facilitators of utilization strongly align with key principles of Patton's (2008) Utilization-Focused Evaluation (UFE). Early and sustained stakeholder involvement, clear specification of intended users and uses, strong commitment by leadership, and effective, tailored communication were all highlighted by participants qualitatively and reinforced by the quantitative data (Table 2 showing "Clear, actionable recommendations," "Strong leadership commitment," and "Stakeholder involvement" as top enablers) as crucial for overcoming shelf ware. This underscores a central tenet of UFE: the process of conducting an evaluation can be as, if not more, important than the technical quality of the final report for ensuring utilization. Engaging intended users in shaping the evaluation fosters ownership and increases the perceived relevance and credibility of findings (Table 1 highlighting relevance and clarity as key credibility factors), thereby enhancing the likelihood of both conceptual and instrumental use.

Participatory methods, frequently mentioned alongside user involvement, reinforce this. The emphasis on user involvement and readiness for evaluation also brings to mind the principles of Evaluability Assessment (EA) (Wholey, 1979), which focuses on ensuring a program is fit for evaluation. While UFE ensures relevance and ownership (Patton, 2008), EA principles could address some underlying barriers found in this

study, such as data gaps or unclear program logic, thereby creating a more solid foundation for UFE to be effective, particularly in contexts like Zimbabwe with known capacity and data challenges (Bamberger et al., 2012). The quantitative finding that "Recommendations are too vague or impractical" is still a significant barrier for 75% of respondents (Table 3) suggests that more attention to evaluability and recommendation formulation is needed.

Evaluation Practice and Theory Implications

The findings carry significant implications. For evaluation practice in Zimbabwe, there's a clear call to shift from merely producing technically sound reports to adopting a more process-oriented, contextualized, and utilization-focused approach. This involves emphasizing the facilitative role of evaluators, engaging users more thoroughly, communicating strategically, and ensuring recommendations are realistic given organizational contexts and capacities. Commissioners, particularly donors and government departments, should critically examine their own requirements and funding modalities that might inadvertently incentivize symbolic use, and instead explore ways to promote genuine learning and adaptive management. The high impact of "Lack of financial resources to implement recommendations" (Table 2) also points to a need for commissioners and organizations to consider resource implications during the evaluation planning and recommendation phases. This also necessitates strengthening the national M&E ecosystem and addressing systemic RBM implementation challenges.

Theoretically, this study validates Weiss's models for classifying use and illustrates the practical utility of Patton's UFE principles in enhancing it. Furthermore, it strongly underscores the explanatory power of Institutional Theory in understanding the persistence of symbolic evaluation practices, especially through the concept of decoupling. In the Zimbabwean context, when organizations face strong external pressures for evaluation but lack internal capacity or genuine commitment, they may adopt evaluation practices primarily for conformity. Here, the drive for legitimacy (isomorphic forces) can overshadow the rational goals of learning and improvement, leading to superficial evaluation activities. Future theoretical work could fruitfully explore the interplay between institutional pressures and organizational learning capabilities in shaping evaluation use trajectories in similar development settings.

CONCLUSION AND RECOMMENDATIONS

Summary of Key Findings

Drawing on a mixed-methods study incorporating qualitative case studies, stakeholder interviews, and a quantitative survey, this research analysed the use of project evaluation reports in Zimbabwe's development sector. The results present a mixed picture: though evaluation is a widespread activity, often motivated by external demands, utilising it in a substantive way is less consistent. Quantitative data confirmed that conceptual and symbolic uses are more frequently reported than direct instrumental use. The 'shelf ware' phenomenon remains a significant challenge. Key enablers identified and quantified include strong stakeholder engagement, dedicated leadership, an organizational culture open to learning, clear and accessible communication, and relevant, actionable recommendations. Conversely, significant barriers include a lack of financial resources for implementation, the absence of formal follow-up mechanisms, limited organisational M&E and RBM capacity, resource constraints, lack of timeliness in reporting, issues with report quality and communication, resistance to change, conflicting priorities, and Zimbabwe's overall difficult political and socio-economic environment.

Contribution to Knowledge (Originality/Value)

The study adds context-specific empirical evidence to the field of evaluation utilization, especially in the under-researched Zimbabwean development sector. By employing a mixed-methods design, it offers both the depth of qualitative insights into the 'how' and 'why' of utilization (or non-utilization) and the breadth of quantitative data on the prevalence and perceived importance of these factors. Building from concerns of

‘shelf ware’ that often dominate anecdotal discussions of evaluation reports, this work advances robust findings on both the mechanisms and explanations for evaluation report use, based on the lived experiences and broader perceptions of various stakeholders. The study provides a theoretically robust inquiry, setting the analysis of evaluation practice and uptake against the backdrop of established utilization theories (Weiss, Patton) and informed by Institutional Theory. It provides rich insights into the political dynamics of evaluation practice and use through both lenses of persistent challenges and enabling opportunities towards greater use of evidence in a specific developing context.

Research Limitations

The findings of the study should be interpreted in the context of its limitations. While the mixed-methods approach enhances overall validity, the qualitative phase, drawn on multiple case studies, provides depth and contextual understanding but is not statistically generalizable beyond those cases. The quantitative survey, while broader, relied on convenience and snowball sampling, which may introduce selection bias and limit generalizability to the entire development sector in Zimbabwe. Self-reported data in the survey is also subject to potential biases. As an interview-based study relying on participant recall and perceptions, there is the potential for subjective bias despite our mitigation of this through triangulation with documentary evidence and a focus on seeking diverse perspectives. Qualitative analysis includes researcher interpretation; conformability was sought through an audit trail and reflexive practice.

Recommendations to Improve Use of Evaluations

Drawing from our findings and discussion, we propose some recommendations to move evaluation from being an exercise in compliance — in Zimbabwe and the development sector more broadly — towards being used as a tool for action and learning.

For Evaluation Commissioners (Govt. Donors, NGO Depts, NGO Leaders)

- Focus on ensuring evaluation is commissioned only on well-articulated internal learning questions and decision-making needs, rather than primarily meeting requirements of external stakeholders.
- Integrate planning for the evaluation and considerations around its potential use from the design stage of the project, including using Evaluability Assessment (EA) principles where applicable to ensure readiness. Consider budgeting for the implementation of key recommendations from the outset.
- Promote or require the meaningful participation of intended users (program managers, staff) throughout the evaluation lifecycle – from defining questions to interpreting findings and developing recommendations. Terms of References (ToRs) should explicitly allocate time and resources for this process.
- Develop formal, systematic, and time-bound management response systems for receiving, discussing, and responding to evaluation reports. This was identified as a critical gap quantitatively (Table 4). Record decisions taken regarding recommendations, along with clear accountability and timelines for follow-up actions. Monitor progress of agreed actions.
- Invest in department-specific capacity to support M&E and to build staff capacity to commission, manage, interpret and use evaluations. Tackle systemic issues that prevent effective RBM implementation to foster a more results-oriented context. Foster internal ownership by balancing efforts to engage external consultants with opportunities for capacity building and retraining of internal M&E staff.
- Leadership needs to visibly commit to placing high value on evaluation by not only responding to findings, but also, encouraging open and honest discussion (even of negative results) and hold their managers to account for using evidence to inform decision-making and improve performance.

For Evaluators (Internal and External):

- Early detection of major intended users and what they want to know. They should facilitate and walk the stakeholders through the evaluation process, and check that it remains relevant at every stage.

- Start debunking lengthy reports. Work on reports so that they are clear, crisp, and free of jargon and compatible with the interests and capacities of different audiences. Use executive summaries, briefs, visual and presentations. Make sure recommendations are feasible, high-impact, actionable, and linked back to findings. The quantitative data (Table 1 & 2) strongly supports the importance of clarity and actionability.
- A deep understanding of the specific organizational, political and socio-economic context of Zimbabwe. Frame the recommendations realistically, given capacity and resource constraints.
- The evaluators role does not end with the submission of the report. When appropriate and where possible, volunteer to co-facilitate dissemination workshops, lead action planning sessions, and be available to assist users in unpacking their implications and applying findings.

For Implementing Organizations / Project Managers:

- When evaluations are commissioned, actively participate in defining questions, providing information, validating findings, and discussing recommendations. Directly communicate information requirements so that it ultimately aligns.
- Ensure project teams set aside time to look at evaluation findings and consider implications for practice together. Embed evaluation learning into usual project review meetings, work planning, and adaptive management processes.
- After recommendations are accepted, champion internally for the required resources (time, budgets, expertise) needed for implementation. Implement assigned duties pertaining to evaluation action plans.

To Policymakers (e.g., OPC, Relevant Ministries, Parliament)

- Examine and amend the national M&E policies and the IRBM system so that they are realistic, well resourced and make evidence use for learning from and improving practice a reality, rather than a box-ticking exercise.
- Set example by showing demand for and use of credible evaluation evidence in policy formulation, strategic planning and resource allocation decisions. Enhance tracking of evaluation findings against key public sector processes e.g. the Public Sector Investment Programme (PSIP) and national/local budget cycles.
- Create the space for long-term investments to help develop a cadre of public sector M&E professionals. Incorporate M&E into the curriculum of national and tertiary institutions.

Directions for Future Research

This mixed-methods study provides a foundation for further inquiry. Building on these findings, future research could focus on:

- Intervention studies that design and test the effectiveness of specific strategies (e.g., enhanced management response systems, UFE coaching) on improving evaluation utilization rates within Zimbabwean organizations.
- Deeper exploration of the political economy of evaluation in Zimbabwe, examining how power dynamics and stakeholder interests influence the entire evaluation cycle, from commissioning to selective use of findings.
- Comparative mixed-methods studies across different sectors (e.g., health vs. agriculture vs. governance) within Zimbabwe, or with similar countries in the SADC region, to identify sector-specific nuances in utilization challenges and facilitators.
- Longitudinal studies that track how the use (or non-use) of a specific set of evaluations influences organizational learning, adaptation, and ultimately, development outcomes over an extended period.

REFERENCES

1. Abbott, G., & Makumbe, R. (2010). Monitoring and evaluation capacity development in Zimbabwe: Issues and opportunities.
2. African Evaluation Association (AfrEA). (2007). African Evaluation Guidelines. AfrEA.
3. African Evaluation Association (AfrEA). (2021). The African Evaluation Principles. AfrEA.
4. Alkin, M. C., & Taut, S. M. (2003). Unpacking evaluation use: Aspects of the process. *Studies in Educational Evaluation*, 29(3), 183–208.
5. Bamberger, M., Rugh, J., & Mabry, L. (2012). *RealWorld evaluation: Working under budget, time, data, and political constraints* (2nd ed.). Sage Publications.
6. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
7. Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage Publications.
8. Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.
9. Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
10. Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage Publications.
11. DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.
12. Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. Lawrence Erlbaum Associates Publishers.
13. Goldman, I., Byamugisha, A., Gounou, A., Smith, L. R., & Ntakumba, S. (2018). National monitoring and evaluation systems in Africa: A review of the literature. *African Evaluation Journal*, 6(1), a286.
14. Government of Zimbabwe. (2005). *Integrated Results Based Management (IRBM) Manual*. Office of the President and Cabinet.
15. Government of Zimbabwe. (2015). *National Monitoring and Evaluation Policy for Zimbabwe*. Office of the President and Cabinet.
16. Government of Zimbabwe. (2020). *National Development Strategy 1 (NDS1) 2021-2025*. Ministry of Finance and Economic Development.
17. Holvoet, N., & Renard, R. (2007). Monitoring and evaluation under the PRSP: Solid rock or quicksand? *Evaluation and Program Planning*, 30(1), 66–81.
18. Kusek, J. Z., & Rist, R. C. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. The World Bank.
19. Leviton, L. C., Kettel Khan, L., Rog, D., Dawkins, N., & Cotton, D. (2010). Evaluability assessment to improve public health policies, programs, and practices. *Annual Review of Public Health*, 31, 213–233.
20. Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.
21. Muguti, A., Chiunye, T., Mutongi, C., & Kandufa, P. (2022). Unpacking integrated results-based management (IRBM) in Zimbabwean local authorities: A case of Harare City Council. *International Journal of Multidisciplinary Research and Publications*, 4(6), 58–68.
22. Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage Publications.
23. Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Sage Publications.
24. Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice*, 18(3), 195–208.
25. Uwizeyimana, D. E. (2020). Monitoring and evaluation capacity building in the African public sector: Lessons from Rwanda. *Administratio Publica*, 28(1), 154–175.
26. Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426–431.
27. Wholey, J. S. (1979). *Evaluation: Promise and performance*. The Urban Institute.
28. Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.