

Methodology Article

Mixed Methods Design for the Evaluation of Development Projects Across Baseline, Midline, and Endline

Peshal Kumar Puri* 

Development Studies, Kathmandu University, Lalitpur, Nepal

Abstract

The evaluation of national and international development projects faces increasing pressure to provide robust evidence of impact while simultaneously offering rich, contextualized explanations of *how* and *why* change occurs. Traditional mono-method approaches, whether purely quantitative or qualitative, often fall short of meeting this dual mandate, particularly in complex social interventions. This article aims to address this gap by providing a scholarly yet practical guide for evaluators on the systematic application of mixed methods research (MMR) designs across the entire evaluation lifecycle of the project, i.e., baseline, midline, and endline. We define the conceptual foundations of MMR in development contexts, detail a typology of designs (convergent, sequential, embedded, longitudinal), and offer a stage-specific framework for their application. Emphasis is placed on the critical process of integration from data collection to analysis, using techniques such as joint displays and narrative causal explanations. Furthermore, we provide in-depth guidance on tool development, sampling strategies, and the integrated reporting of findings. This study discusses the challenges of conducting longitudinal Mixed Methods Evaluations and the ethical issues surrounding these evaluations. We will also look at some of the implications that this type of evaluation has for donors, evaluators, and future research in methodology. This study is designed to provide practical applications and examples to help improve the quality, relevance, and usefulness of those who evaluate development work.

Keywords

Mixed-methods Evaluation, Longitudinal Evaluation Design, Development Project Evaluation, Evaluation Lifecycle (or Project Lifecycle), Data Integration

1. Introduction and Rationale for Mixed Methods in Development Evaluation

A multitude of interrelated variables may influence each other at different levels or in different time frames. Therefore, it is critical to employ evaluation techniques that account for both how significant changes happen to a community and what ultimately causes the change that occurs as a result of a Project [1]. Due to their multifaceted nature regarding economic development, many development Projects will have

multiple areas of intervention, e.g., poverty alleviation, educational achievement, housing and health improvement, etc., which are intertwined in such a way that they depend upon each other for success at every level. Evaluating the success of such projects, determining whether they work, for whom, and under what conditions, requires a methodological toolkit capable of addressing both the breadth of impact and the depth

*Corresponding author: puri.peshal2@gmail.com (Peshal Kumar Puri)

Received: 30 December 2025; Accepted: 9 January 2026; Published: 23 January 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

of context [2, 9]. It is within this context that the rationale for mixed methods of research (MMR) in development evaluation becomes not merely an option, but an essential methodological imperative.

For decades, evaluation practice has been dominated by a tension between quantitative and qualitative paradigms. Quantitative approaches, such as randomized controlled trials (RCTs) or quasi-experimental designs, excel at establishing statistical causality and measuring the average treatment effect, providing the necessary evidence for accountability and scaling [3, 54]. However, quantitative evaluations are sometimes criticized for providing only a general overview of what happened, rather than providing detailed explanations of the underlying reasons *why* or *how* an intervention led to an outcome, while not including the unintended consequences from that intervention [4, 3, 41].

Qualitative methods (e.g., ethnographic research, case studies, group discussion, in-depth interviewing) are very useful for getting at the context of the evaluation, developing an understanding of how the people involved in the evaluation perceive the outcome, and generating an abundance of rich explanatory narratives based on what the evaluations tell us. On the downside, some donor groups question whether qualitative evaluations generate valid and reliable data and therefore hesitate to use qualitative methods for *Decision-Making and Generalizability* [5, 23]. The challenges of trying to assess the effectiveness of a program, especially in the context of complex social tasks, through only quantitative or only qualitative eval-

uations is most pronounced in the developing world. For example, a quantitative survey may show a statistically significant increase in school enrollment (the *what*), but it does not provide any insight into the cultural barriers that prevented poor families from enrolling in school, or the community factors that contributed to the success or failure of program implementation from one village to another (the *why* and *how*). Mixed-methods evaluations (MME) intentionally combine or integrate both quantitative and qualitative data and analysis within a single study or series of related studies to provide a complete and more thorough picture of the evaluation issue [6, 10, 20].

This article serves as a resource for evaluators on how to systematically choose and implement suitable mixed methods designs for the three critical stages in a programme's evaluation lifecycle: baseline, midline, and endline. The authors emphasize the importance of using MME in practical terms, as opposed to using MME solely in theoretical debates, and instead concentrate on assisting evaluators with actual design and analysis decisions. We believe that a well-designed and implemented MME strategy, when implemented over time, creates a stronger and more actionable evidence base for development practitioners and policymakers [12, 35]. Additionally, using a MME (Mix Method Evaluation) strategy enables evaluators to confirm their findings through triangulation, develop instruments based on the local context, explain the statistical findings through narrative data, and ultimately to enhance the validity and utility of the evaluation [7, 14, 17].

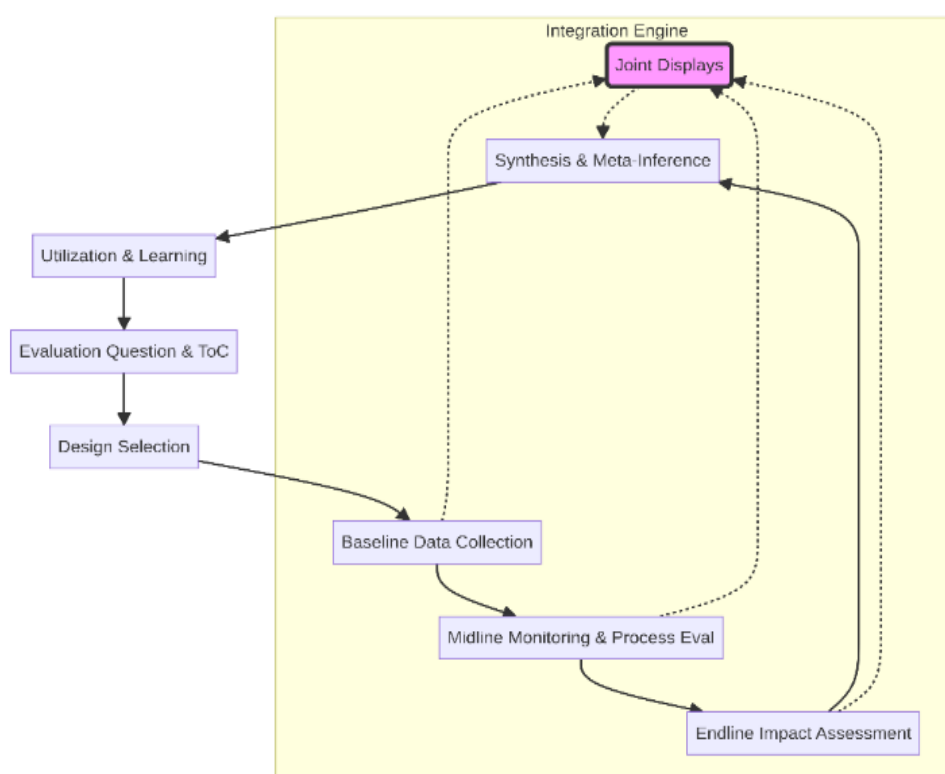


Figure 1. Mixed Methods Evaluation Lifecycle Framework depicting the cyclical nature of MME and the central role of the integration engine.

1.1. Conceptual Foundations of Mixed Methods for Development Projects

In order to integrate Mixed Methods in Development Evaluation, the evaluator needs to have a good understanding of the context in which they are working. So, developing a good conceptual base, which reflects the development environment, is essential to using Mixed Methods Research effectively. Mixed Methods Research/Evaluation is based on the belief that a better understanding of the Research Questions can be achieved through the "marriage" of both quantitative and qualitative research than either of these types of research done in isolation [8, 32]. Development projects, when evaluated, are concerned with measurable results (e.g., reduction of poverty, education achievement, a decrease of disease/illness) and the experiences of and interpretations of changes that are influenced by context (i.e. Empowerment, Social Capital, etc.).

The rationale for Mixed Methods Evaluation in the Development Sector is based on the characteristics of Development Programs - Development Programs are complex programs that often exist in dynamic systems, where the nature of the programs' causal pathways is nonlinear and subject to external shocks [9, 27]. MME (Mix Method Evaluation) allows evaluators to use quantitative data to map the system's outputs and outcomes, while using qualitative data to explore the underlying mechanisms and contextual factors influencing those results. Furthermore, development projects are highly sensitive to context, including local culture, political economy, and institutional capacity. Qualitative methods are essential for capturing this context, which can then be used to interpret quantitative findings, such as why a statistically successful intervention in one region failed to replicate in another [13, 15].

Core mixed methods principles guide the integration process. Integration is the defining feature of MMR, referring to the mixing of the two datasets at various stages of the evaluation, from design to interpretation [10, 14]. This is distinct from simply collecting both types of data. In Complementarity, two methods are used to study different but overlapping aspects of an event, e.g., using a quantitative survey to measure income, while qualitative interviews are used to assess a person's understanding of poverty [16, 43]. Triangulation is both methods being applied to the same research query, where the findings from both methods cross-check each other to improve the quality of the results [17, 46]. And finally, the development principle allows methodology one to inform Design/Implement methodology two, using existing qualitative findings to refine a methodological instrument [15, 38].

The information requirements for an evaluation vary considerably throughout the project's lifecycle, thereby demanding a flexible and adaptive mixed methods approach. First, a detailed description of the situation before the intervention is needed, including a complete look at the problem and the essential background information. This stage necessitates data to establish a counterfactual and to guide the refinement of the

project's Theory of Change (ToC) [19, 58]. Consequently, the midline phase emphasizes both process evaluation and performance monitoring. Evaluators necessitate certain data to determine the project's compliance with its initial design, the validity of its Theory of Change (ToC) assumptions, and the development of initial results. This data pertains to fidelity of implementation, the extent of reach, and the initial indicators of change [11, 64]. Ultimately, at the endline phase, the primary informational requirement centers on impact assessment and sustainability. The evaluation's objectives are to ascertain the final outcomes and impacts, attribute observed changes to the intervention when feasible and evaluate the probability of enduring benefits following the project's conclusion [29, 61].

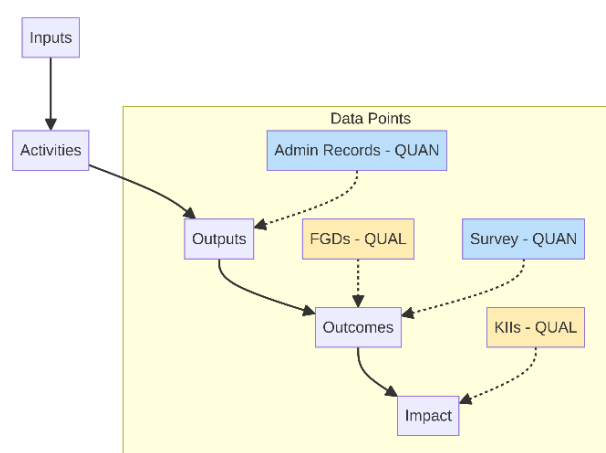


Figure 2. Theory of Change with Embedded Mixed Methods Data Points, mapping quantitative and qualitative data sources to causal links.

1.2. Typology and Selection of Mixed Methods Designs in Evaluation

The strength of mixed methods lies in its flexibility, which is operationalized through a variety of established designs. Evaluators must move beyond a simple choice between quantitative and qualitative methods to a deliberate selection of a design that aligns with the evaluation's purpose, resources, and timeline [11, 47]. Mixed methods designs are usually classed based on two main factors: the importance given to the quantitative (QUAN) or qualitative (QUAL) parts, and the order in which these parts are used, whether at the same time or one after the other [32].

The Convergent (Parallel) Design (QUAN + QUAL) involves collecting both quantitative and qualitative data concurrently and then merging the two datasets during the interpretation phase. Both methods are given equal priority. This design is highly efficient in terms of time, as data collection occurs simultaneously [6, 50]. In a development evaluation setting, a convergent design might involve conducting a household survey (QUAN) and in-depth interviews (QUAL)

on the same topic (e.g., food security) during the same field visit. The limitation is that the two datasets must be collected independently, which can make the subsequent integration and comparison challenging if the findings diverge [10, 14].

The Explanatory Sequential Design (QUAN → QUAL) is characterized by two distinct phases. Quantitative data is collected and analyzed first, followed by the qualitative data collection and analysis. The purpose of the qualitative phase is to help explain or elaborate on the initial quantitative results [42, 49]. For example, if a quantitative endline survey shows that a microfinance program had a statistically insignificant impact on women's business profits, the subsequent qualitative phase would involve in-depth interviews with a subset of women to understand the contextual reasons for this finding, such as market saturation or cultural constraints on women's mobility. This design is particularly strong for impact evaluations as it provides a narrative explanation for measured outcomes [3, 60].

The Exploratory Sequential Design (QUAL → QUAN) reverses the usual order. It starts with a qualitative phase, which is used to explore a phenomenon, create hypotheses, or develop a tool that fits the specific context. The results from this

qualitative phase then guide the design and sampling for the following quantitative phase [15, 38]. This approach is often used at the beginning of development projects. For example, initial focus group discussions (QUAL) might identify local factors related to well-being or specific obstacles to participation. These findings are then used to create a culturally relevant and valid quantitative survey (QUAN). This design is particularly useful for evaluations in situations with limited data or diverse cultures, where standardized tools might not be valid [12, 23].

The Embedded Design (QUAN-qual) = (QUAL-quant)) involves one dataset being nested within a larger, primary study. The secondary, embedded method addresses a different question or provides a supplementary perspective to the dominant method [6, 10]. For example, in a large-scale RCT (dominant QUAN), a small-scale process evaluation (embedded qual) might be conducted to monitor implementation fidelity and understand the intervention's delivery mechanisms. This design is common in large donor-funded evaluations where the primary requirement is a quantitative impact estimate, but contextual data is needed to enhance interpretation [51, 62].

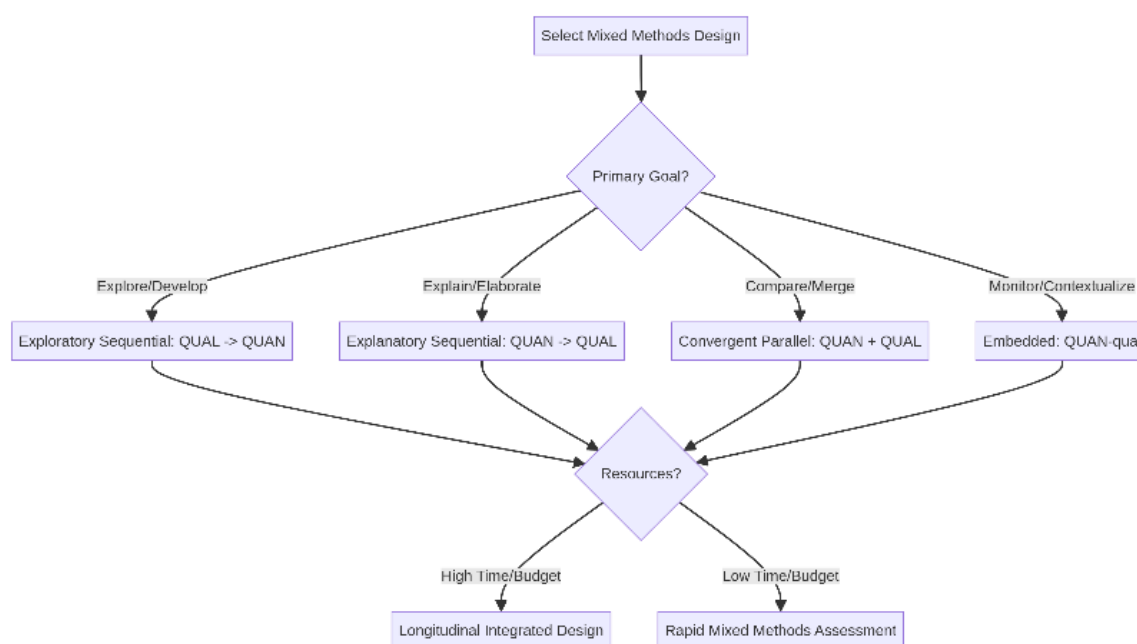


Figure 3. Decision Tree for Selecting Mixed Methods Designs in Development Evaluation based on primary goals and resource constraints.

2. Strategic Sampling Frameworks for Longitudinal Mixed Methods Transitions

The success of a longitudinal mixed methods evaluation depends on the strategic coordination of sampling decisions across the baseline, midline, and endline evaluation stages.

Evaluators must move beyond static sampling plans to a dynamic framework that allows for panel tracking while maintaining the flexibility to capture emerging contextual shifts [35, 48].

During the Baseline phase, a dual-track sampling strategy is implemented. The quantitative aspect necessitates probability sampling methods, such as stratified random sampling, to guarantee that the baseline data accurately reflects the target population and furnishes a strong counterfactual [26, 34].

Concurrently, the qualitative element utilizes purposive sampling techniques, including maximum variation or snowball sampling, to pinpoint information-rich cases capable of offering profound understanding of the local context and aiding in

the validation of the quantitative instruments [8, 31]. At this point, integration often uses a nested sampling method. In this approach, the qualitative participants are a carefully chosen group within the larger quantitative survey sample [10, 14].

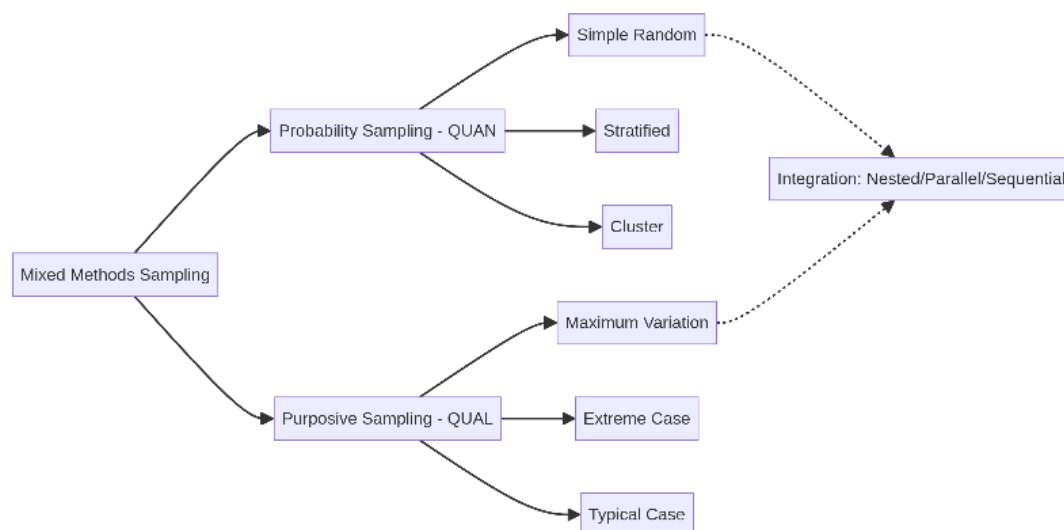


Figure 4. Longitudinal Sampling Transition Framework showing the evolution of QUAN and QUAL samples across BL-ML-EL stages.

Panel maintenance presents the principal difficulty at the midline. The quantitative aspect necessitates monitoring the initial baseline participants to assess temporal shifts, thereby demanding stringent attrition management strategies [64, 40]. Nevertheless, the mixed methods framework facilitates adaptive sampling at the midline. Should the quantitative findings indicate unanticipated patterns or anomalous project sites, the qualitative element can adjust to theoretical sampling, incorporating additional cases or sites to investigate these newly identified occurrences [27, 55]. Consequently, the midline evaluation functions not merely as a progress assessment but as a formative learning process, capable of guiding project adjustments [11, 64].

At the Endline, the sampling strategy culminates in a meta-inference sample. While the quantitative component completes the final panel survey to measure impact, the qualitative component often shifts to extreme case sampling or critical case sampling [26, 34]. By focusing on the most successful and least successful cases identified in the quantitative endline data, evaluators can generate powerful explanations of the causal mechanisms that drove the observed impacts [57, 65]. This sequential integration of sampling where endline quantitative results drive the final qualitative case selection is a hallmark of high-quality explanatory sequential designs [42, 49].

3. Applying Mixed Methods at Baseline, Midline, and Endline Stages

The application of mixed methods is not static; it must be

dynamically tailored to the specific informational needs and constraints of the baseline, midline, and endline stages of a development project. This stage-specific application ensures that the mixed methods approach contributes maximally to the project's learning and accountability objectives across its lifecycle [12, 35].

In the Baseline phase, the principal aim is to ascertain the preliminary conditions that will serve as the benchmark for subsequent evaluation, while simultaneously deepening the project's comprehension of the problem and its surrounding environment. The Exploratory Sequential Design (QUAL → QUAN) frequently proves to be the most suitable approach at this juncture. The preliminary qualitative phase utilizes techniques including participatory rural appraisal (PRA), focus group discussions (FGDs), and key informant interviews (KIIs) to examine the local environment, identify significant challenges, and understand local language and cultural practices [13, 59]. This qualitative information then guides the creation or adjustment of the quantitative tools, thus guaranteeing that survey questions are culturally appropriate, evaluate locally relevant indicators, and accurately represent the complexities of the intended population [15, 38].

The Midline stage focuses on process evaluation, performance monitoring, and detecting early signs of change. The Convergent (Parallel) Design (QUAN + QUAL) or the Embedded Design (QUAN(qual)) are frequently employed here. The convergent approach allows for simultaneous collection of quantitative performance data (e.g., project monitoring data, administrative records, short surveys) and qualitative process data (e.g., observation of training sessions, interviews with

frontline staff and beneficiaries) [11, 64]. Quantitative data tracks outputs and early outcomes, while qualitative data provides insight into implementation fidelity, bottlenecks, and the mechanisms of change [55, 66].

The Endline stage is dedicated to assessing outcomes, impact, and sustainability. Explanatory Sequential Design (QUAN → QUAL) or a Longitudinal Mixed Methods Design that links all three stages is most common. The quantitative

phase, typically a large-scale survey or experimental/quasi-experimental design, measures the final impact on key indicators [3, 61, 62]. The subsequent qualitative phase is then strategically designed to explain the quantitative findings. This is essential for contribution analysis, where the qualitative data provides the necessary evidence to trace the causal chain and support the claim that the intervention contributed to the observed changes [57, 65].

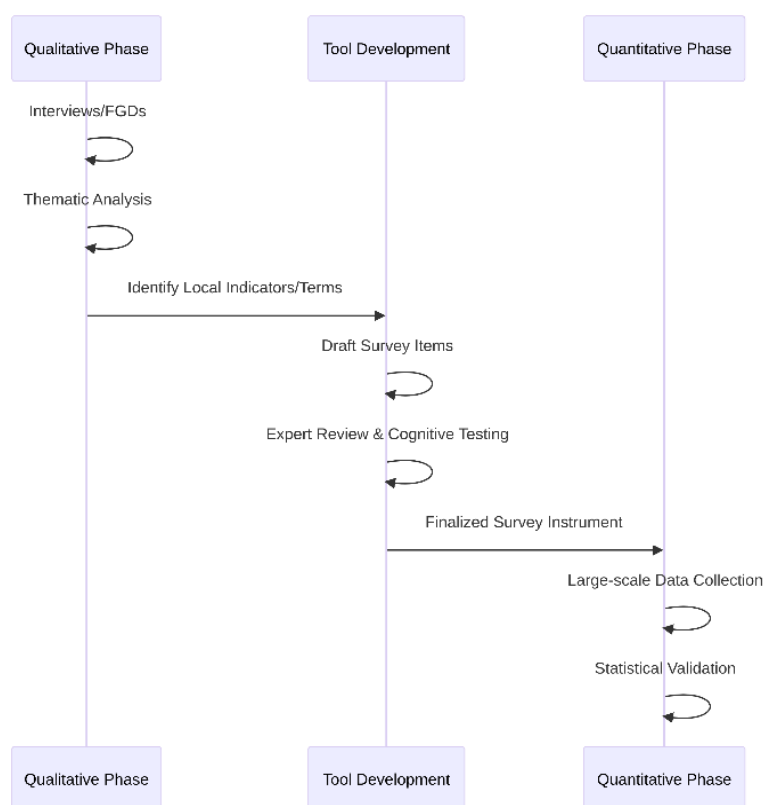


Figure 5. Tool Development Process illustrates how qualitative findings inform the creation and validation of quantitative instruments.

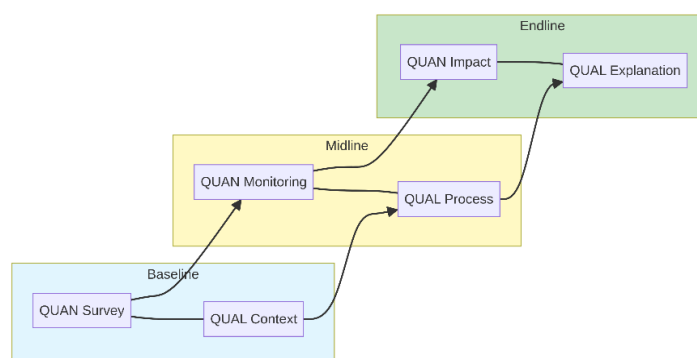


Figure 6. Longitudinal Mixed Methods Evaluation Pathway showing the tracking of quantitative and qualitative data across B-M-E stages.

Analysis of Mixed Methods at Baseline, Midline, and Endline Stages

The assessment of mixed methods data throughout the evaluation process is a complex undertaking, moving from stage-

specific synthesis to longitudinal meta-inference. This structured approach allows the evaluation to reflect the changing nature of social transformation while maintaining methodological rigor [30, 35]. During the Baseline Analysis phase, the

focus is on contextual mapping and instrument validation. Qualitative analysis, such as thematic coding of focus group discussions, helps identify key concepts and local terminology; these are then used to validate the quantitative survey items through cognitive testing and pilot studies [15, 38]. The integrated analysis at baseline establishes the counterfactual framework, where quantitative descriptive statistics are contextualized by qualitative narratives, thus providing a comprehensive, multi-dimensional basis for the evaluation [12, 19].

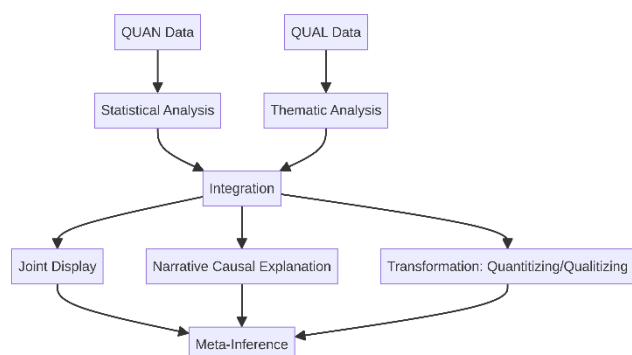


Figure 7. B-M-E Analysis Framework illustrating the specific analytical tasks at each stage and the overarching longitudinal integration.

Midline Analysis prioritizes the examination of process tracking and thematic development. Evaluators employ cross-case analysis to compare implementation experiences across diverse project locations, thereby identifying recurring patterns of both success and failure [27, 55]. Quantitative monitoring data is integrated with qualitative process data to establish adaptive feedback loops; this enables project managers to ascertain not only whether targets are achieved, but also the reasons behind the varying effectiveness of specific activities [11, 64]. This phase frequently necessitates the quantification of qualitative data, such as converting thematic frequencies into numerical scores, to facilitate statistical correlation with project outcomes [4, 10].

The Endline Analysis serves as the final phase of the mixed methods approach, concentrating on impact attribution and sustainability evaluation. The primary analytical objective is the formulation of meta-inferences—integrated conclusions derived from the systematic comparison of the ultimate quantitative and qualitative results [21, 24, 25]. This process encompasses seven distinct steps: (1) the identification of the central evaluation questions, (2) a summary of the QUAN findings, (3) a summary of the QUAL findings, (4) a comparison of the findings to ascertain convergence or divergence, (5) an exploration of the contextual factors underlying any divergence, (6) the synthesis of the integrated claims, and (7) the formulation of the final meta-inference [18, 21].

Longitudinal Integration across all three stages requires trend analysis of quantitative indicators paired with narrative

threading of qualitative themes [30, 35]. Evaluators use longitudinal joint displays to track how specific beneficiary groups evolve over time, linking their baseline characteristics to their midline experiences and endline outcomes [14, 17]. This integrated approach allows for a more thorough contribution analysis. The qualitative narrative acts as the "causal glue," connecting the quantitative data and resolving competing explanations. This results in a strong, evidence-based account of the observed changes [57, 65].

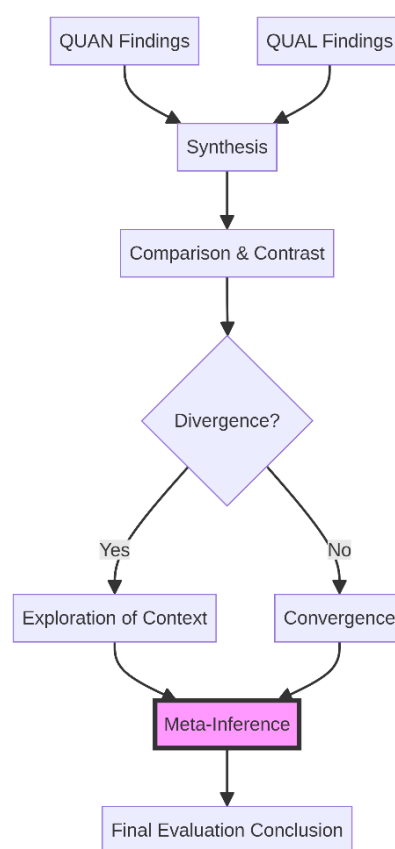


Figure 8. Meta-Inference Generation Process showing the path from separate findings to integrated evaluation conclusions.

4. Synthesizing and Reporting Findings Across the Evaluation Lifecycle

Writing the evaluation report for a mixed methods study requires a departure from traditional siloed reporting. Evaluators must follow a synthesis-first approach, where the report is structured around the evaluation questions rather than the data collection methods [7]. While the baseline, midline, and endline reports share a commitment to integration, they differ significantly in their focus, structure, and intended utility [11, 36].

The Baseline Report is primarily descriptive and exploratory. Its similarity to later reports lies in its use of both QUAN and QUAL data to establish a starting point. However, its unique requirement is the validation of the Theory of Change.

A high-quality baseline report uses qualitative narratives to "ground-truth" the quantitative indicators, ensuring that the evaluation framework is culturally and contextually relevant [15, 38]. For example, a baseline report for a nutrition project

might present quantitative stunting rates alongside qualitative data on local feeding practices, providing a multi-dimensional view of the problem [19, 58].

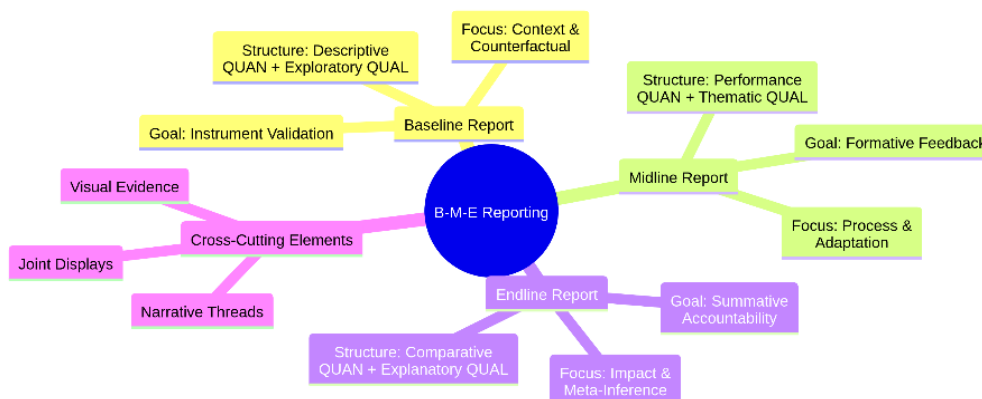


Figure 9. BL-ML-EL Reporting Structure highlighting the shifting focus and goals of reports across the lifecycle.

The Midline Report is formative and process oriented. Unlike the baseline, it focuses on change over time and implementation fidelity. The reporting should highlight the "how" and "why" of project progress, using joint displays to link performance metrics with beneficiary feedback [17, 14]. A best-practice example from the literature is the use of "success stories" and "failure cases" presented alongside quantitative progress bars, allowing donors to see the human face of the data and understand the barriers to success [11, 64]. The midline report's primary goal is to provide actionable recommendations for project adaptation [55, 66].

The Endline Report is summative and explanatory. It differs from the midline by its focus on impact attribution and sustainability. The report must culminate in a meta-inference section, where the final quantitative impact estimates are systematically integrated with qualitative evidence of causal mechanisms [21, 24, 45]. Evaluators should follow the "weaving" technique, where quantitative results and qualitative quotes are integrated within the same paragraph to provide a seamless evidence-based narrative [7, 37]. For instance, an endline report might state: "While the project achieved a 20% increase in crop yields (QUAN), qualitative interviews revealed that this benefit was primarily captured by male-headed households due to existing land tenure constraints (QUAL)" [29, 61]. This level of integrated reporting provides the most robust evidence for future policy and programming decisions [1, 63].

5. Integration and Analysis of Mixed Methods Across the Evaluation Lifecycle

The true value of mixed methods evaluation in development

goes beyond just collecting two types of data. It lies in how these data are systematically combined to create a complete understanding of a project's performance and impact [14, 28]. This integration is the most challenging part of MME. It requires careful planning and advanced analytical methods to connect the findings from the beginning, middle, and end of the evaluation into a unified evaluation strategy.

Maintaining consistent indicators is crucial for integrating data over time. While the quantitative indicators must stay the same throughout the evaluation to allow for statistical comparisons of change, the qualitative part needs a balance between consistency and flexibility. Evaluators should monitor significant qualitative themes throughout the evaluation, employing a uniform coding system established at the outset. Conversely, the qualitative protocol requires flexibility to incorporate the emergence of new, unanticipated themes at midline and endline, which often indicate important unintended consequences or shifts in context [25, 27, 55].

Sampling strategies within MME demand meticulous coordination between the quantitative and qualitative elements. Probability sampling is employed for the quantitative component to ensure generalizability, whereas purposive sampling is utilized for the qualitative component to guarantee depth and information-richness [26, 34, 35]. Integration is achieved through nested sampling (where the qualitative sample constitutes a subset of the quantitative sample), parallel sampling (where the two samples are independent but drawn from the same population), or sequential sampling (where the outcomes of the initial phase inform the selection of the subsequent phase) [8, 31].

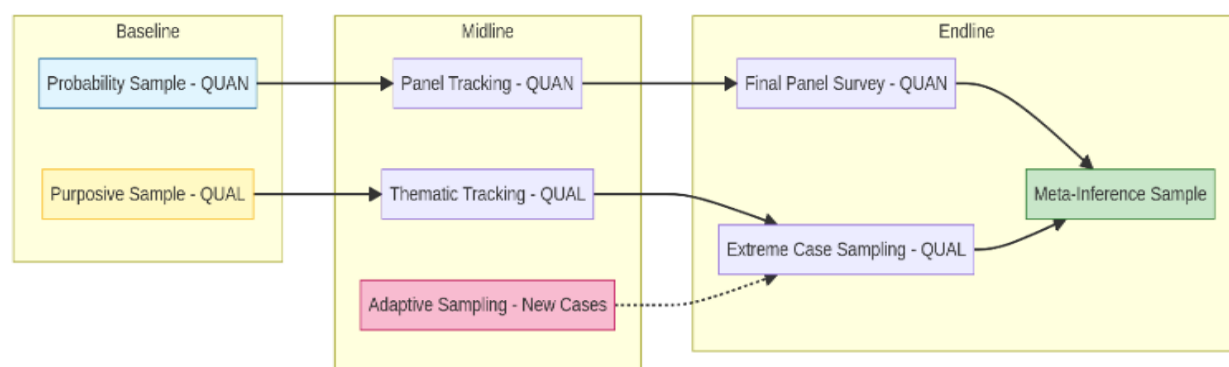


Figure 10. Mixed Methods Sampling Strategy Typology showing the integration of probability and purposive sampling.

Analytical integration techniques are the tools by which the two datasets are formally mixed. Beyond the stage-specific applications, the cross-lifecycle integration requires techniques that synthesize findings across time [4, 14]. A particularly effective method involves the employment of Joint Displays, which are visual matrices that integrate quantitative data, such as means and p-values, with associated qualitative data, including quotes and themes, thereby enabling cross-method inference [17, 14]. Furthermore, the formulation of Narrative Causal Explanations is a crucial technique; this approach utilizes qualitative data to construct a comprehensive, contextually informed narrative that elucidates the statistical relationships observed in the quantitative data [57, 65].

6. Practical Challenges, Ethical Considerations, and Mitigation Strategies

Mixed methods evaluation, despite its methodological benefits, encounters distinct practical and ethical hurdles within the evolving and financially limited context of development projects, necessitating proactive attention from evaluators [15, 52]. A prevalent practical challenge stems from time and budget constraints. MME, by its nature, demands more resources compared to mono-method evaluations, necessitating increased personnel, extended fieldwork durations, and specialized analytical skills [51, 67]. To mitigate this, evaluators should adopt a pragmatic approach to design selection, fre-

quently favoring embedded or sequential designs that facilitate resource allocation management across different phases, rather than a comprehensive convergent design [12, 52].

Data quality issues and attrition are significant concerns, particularly in longitudinal studies. High attrition rates in quantitative panels can compromise statistical power, while the loss of key informants in qualitative tracking can break the narrative continuity [35, 61]. Mitigation Strategy: For attrition, evaluators should employ rigorous tracking protocols, including collecting multiple contact points and using community-based field teams.

Power dynamics and imbalances between the quantitative and qualitative components pose a methodological challenge. Often, the quantitative component, driven by donor requirements for impact numbers, is given methodological priority, leading to the qualitative data being treated as secondary or merely illustrative. Mitigation Strategy: This requires careful team composition and leadership. The evaluation team should be genuinely interdisciplinary, with equal authority for both quantitative and qualitative leads in design and interpretation [15, 33].

Ethical considerations are especially important in MME, particularly when working with vulnerable groups. Collecting both quantitative and highly personal qualitative data from the same individuals increases the risk of identification and potential harm [22, 16]. To reduce these risks, ethical protocols must be strong. This includes getting separate, informed consent for each data collection phase (B, M, E), ensuring data is anonymized and securely stored, and having clear guidelines for reporting sensitive information [64, 39].

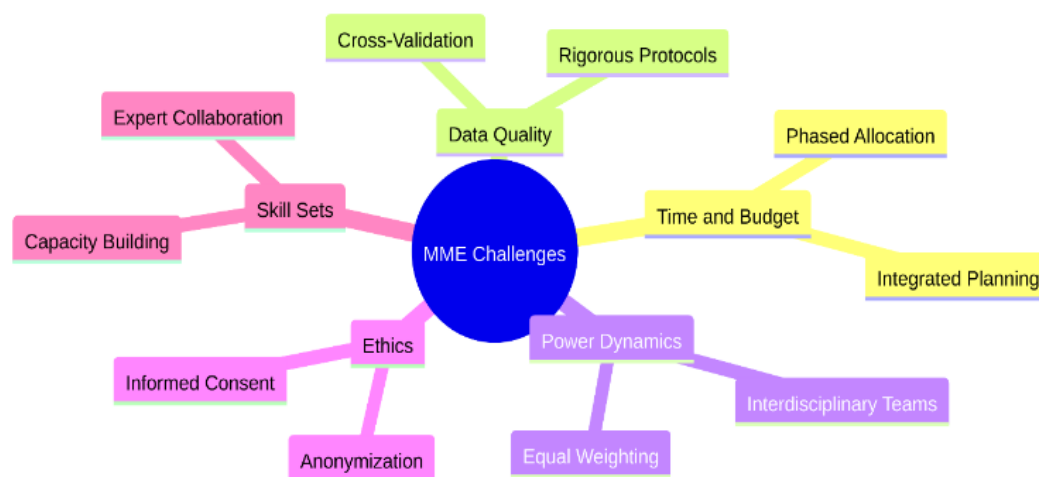


Figure 11. Practical Challenges and Mitigation Strategies in Mixed Methods Evaluation.

7. Implications for Evaluators, Donors, and Future Research

The consistent use of mixed methods designs throughout the baseline, midline, and endline phases of development projects significantly impacts learning, accountability, and decision-making in the development sector [1, 63]. For evaluators, the shift to MME requires a change in professional practice, moving from a separate approach to a collaborative, interdisciplinary model. MME enhances the evaluator's ability to create useful evidence. This practical, evaluator-focused approach is especially important when working in areas with limited data or in unstable situations, where traditional experimental designs are often not possible or ethical [51, 53].

For Donors and Commissioning Agencies, the implication is a need to evolve commissioning and management practices. Donors must recognize that MME is an investment in better evidence, not just a cost-additive exercise [62, 67]. Donors should explicitly request and incentivize integrated findings, moving away from reports that simply present two separate studies [63, 44].

Looking toward Future Research, the field of mixed methods in development evaluation has several methodological directions. There is a need for more empirical research on the effectiveness of different integration techniques, particularly in the context of longitudinal tracking [4, 14]. Specifically, future research should focus on:

- 1). Developing standardized, yet flexible, protocols for quantitating qualitative data and unalizing quantitative data across multiple time points [4, 10].
- 2). Exploring the application of transformative multilevel mixed methods designs to better address issues of power, equity, and social justice [33, 22].
- 3). Investigating the use of technology, such as machine learning for thematic coding and geospatial analysis for

contextualizing qualitative data, to enhance the efficiency and rigor of integrated analysis [59, 61, 56].

The complex nature of contemporary development issues necessitates an evaluation methodology of comparable sophistication. A mixed methods design, when systematically implemented throughout the baseline, midline, and endline phases, offers the essential structure for generating rigorous, pertinent, and comprehensive evidence. Through the application of integration principles, the resolution of practical and ethical considerations, and the promotion of interdisciplinary collaboration, evaluators and donors can ensure that development evaluations transcend mere change measurement, thereby facilitating a genuine comprehension of the underlying transformation processes.

8. Leveraging the Baseline: An Integrated Mixed Methods Approach to Robust Impact Evaluation

Baseline data availability is crucial for bolstering the methodological soundness of impact evaluations, facilitating more robust quasi-experimental designs like Difference-in-Differences (DiD), which account for unobserved confounders that remain constant over time. Within this rigorous quantitative structure, a mixed methods approach is employed not to address data deficiencies but to enrich and corroborate causal interpretations. Qualitative elements integrated from the beginning can systematically investigate the program's Theory of Change, yielding vital information on implementation fidelity, contextual disruptions, and participant experiences that are beyond the scope of the baseline survey.

This integration enables evaluators to leverage the baseline not merely as a statistical control, but as a comprehensive, contextual reference point. As an illustration, qualitative pro-

cess tracing, employing a subset of respondents and commencing at baseline, then continuing through midline and endline assessments, can directly assess the proposed causal pathways connecting the intervention to the observed outcomes, thus elucidating the mechanisms underlying the statistically measured impact. Moreover, the convergence of baseline-informed quantitative data and strategically implemented qualitative investigation substantially enhances both internal and external validity.

A sequential explanatory design offers significant advantages; following a Difference-in-Differences (DiD) analysis that quantifies the impact's magnitude, focused qualitative

inquiries with both beneficiaries and implementers can elucidate the underlying causes of varied effects observed across sub-groups, as determined by the baseline data, or clarify the factors contributing to unanticipated outcomes. This methodological interplay elevates the baseline from a static depiction of attributes to an evolving component of the explanatory framework. Consequently, the evaluation transcends the mere reporting of an average treatment effect, instead furnishing actionable insights regarding the prerequisites for success, the fidelity of implementation, and the subjective experience of change—insights that are essential for the program's scalability, adaptability, and validation of its impact.

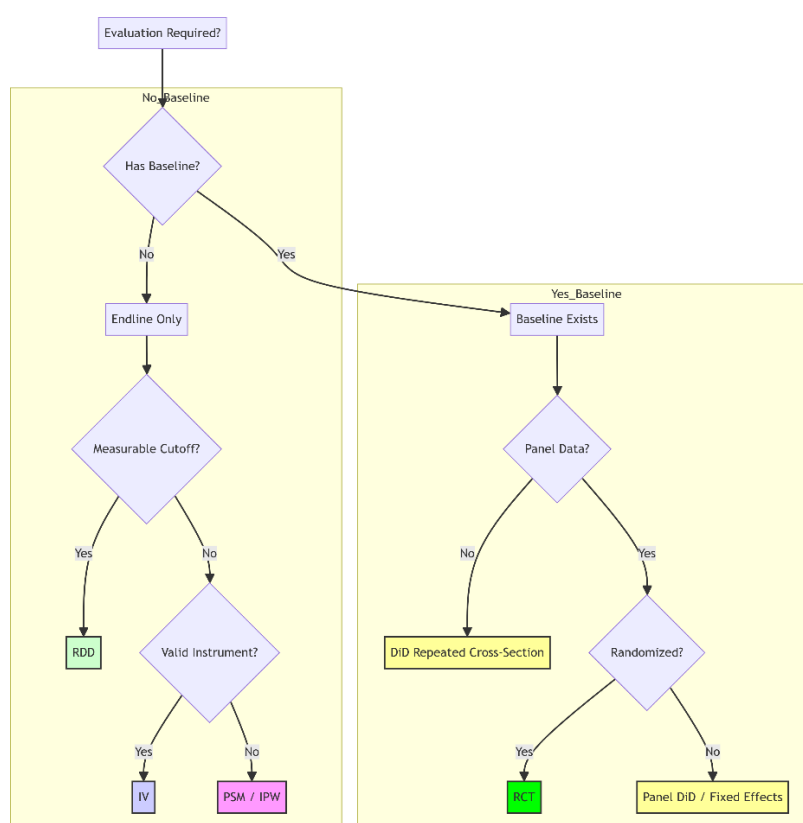


Figure 12. Impact Evaluation Design Decision Tree.

9. Conclusion: Strategic Mixed-methods Sequencing for Longitudinal Evaluation

The effective evaluation of complex development interventions necessitates a longitudinal mixed-methods approach where the sequencing of quantitative (QUAN) and qualitative (QUAL) data collection is strategically tailored to the informational needs of each project phase. This design moves beyond simple data collection to a deliberate integration strategy that maximizes validity, utility, and efficiency. At Baseline, an

Exploratory Sequential (QUAL → QUAN) design is optimal. This initial qualitative phase serves a critical function by providing deep contextual understanding, which is then used to inform and refine the quantitative tools, ensuring they are culturally relevant and of high quality. This foundational step establishes a robust and contextually sensitive measurement framework for the entire evaluation.

The subsequent phases leverage this foundation for both efficiency and explanatory power. The Midline phase is best served by a Convergent (QUAN + QUAL) design, allowing for the simultaneous collection and validation of progress data, which saves time and cost. Crucially, this phase maintains methodological flexibility, allowing for a shift back to an Exploratory Sequential (QUAL → QUAN) approach if initial

quantitative data is insufficient or if new indicators need to be developed and finalized. Finally, the Endline evaluation should employ an Explanatory Sequential (QUAN→QUAL) design. By first establishing the breadth of impact through quantitative data, the subsequent qualitative deep dive can strategically explore causal mechanisms and contextual factors. This sequence provides rich, actionable explanations for how and why the observed impacts occurred, leading to a more profound and comprehensive understanding of the intervention's success and sustainability.

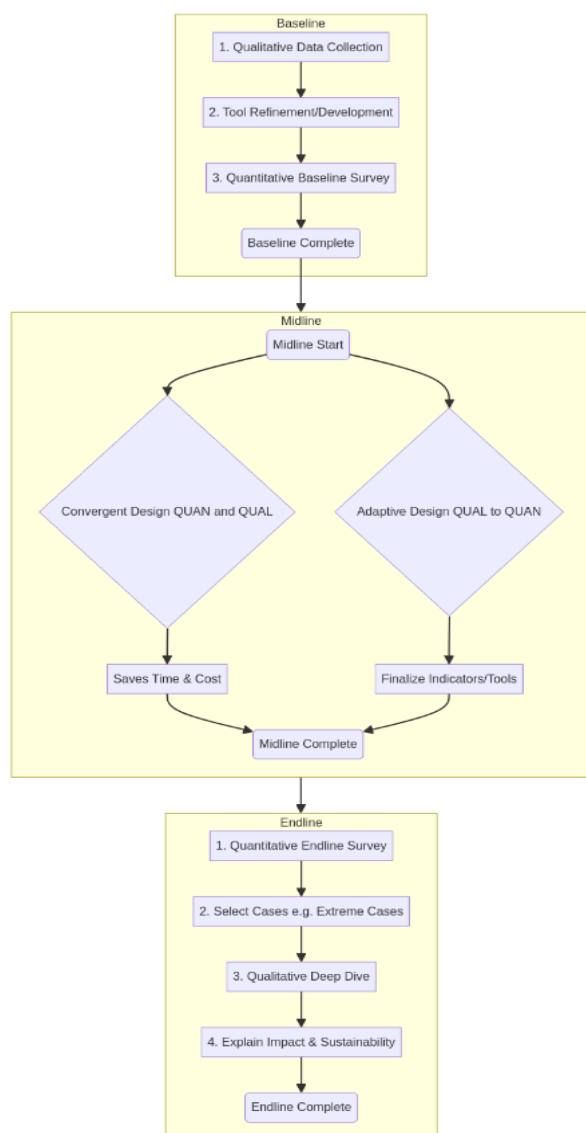


Figure 13. Strategic Mixed-Methods Sequencing for Longitudinal Evaluation.

Abbreviations

BL	Baseline
DiD	Difference-in-Differences
EL	Endline

FGD	Fixed Group Discussion
KII	Key Informant Interviews
ML	Midline
MME	Mix Methods Evaluation
MMR	Mixed Methods Research
RCTs	Randomized Controlled Trials
ToC	Theory of Change
QUAL	Qualitative
QUAN	Quantitative

Author Contributions

Peshal Kumar Puri is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Woolcock, M. (2019). Reasons for Using Mixed Methods in the Evaluation of Development Projects. Harvard University. <https://dash.harvard.edu/bitstreams/b9b0f999-1c82-4bae-81e5-3858beac6268/download>
- [2] Copestake, J. (2024). Mixed-methods impact evaluation in international development practice. Bath SDR. <https://bathcdr.org/wp-content/uploads/2024/05/Mixed-methods-impact-evaluation-in-international-development-practice-distinguishing-between-quant-led-and-qual-led-models.pdf>
- [3] White, H. (2009). Theory-based impact evaluation: Principles and practice. *Journal of Development Effectiveness*, 1(3), 229–249. <https://doi.org/10.1080/19439340903114628>
- [4] Ravallion, M. (2009). Evaluation in the practice of development. *The World Bank Research Observer*, 24(1), 29–53. <https://doi.org/10.1093/wbro/lkn007>
- [5] Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1), 97–113. <https://doi.org/10.1177/1468794106058877>
- [6] Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE Publications. <https://us.sagepub.com/en-us/nam/designing-and-conducting-mixed-methods-research/book246106>
- [7] Feters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving Integration in Mixed Methods Designs—Principles and Practices. *Health Services Research*, 48(6pt2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>
- [8] Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112–133. <https://doi.org/10.1177/1558689806298224>

- [9] Patton, M. Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. Guilford Press.
<https://www.guilford.com/books/Developmental-Evaluation/Patton/9781609180024>
- [10] Tashakkori, A., & Teddlie, C. (2010). *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications. <https://doi.org/10.4135/9781506335193>
- [11] Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255–274.
<https://doi.org/10.3102/01623737011003255>
- [12] Desalos, C. (2021). *Mixed methods in monitoring and evaluation of international development projects: A practical guide*. Wageningen University & Research.
<https://edepot.wur.nl/539427>
- [13] Farquhar, M., et al. (2013). Mixed Methods Research in the Development and Evaluation of Complex Interventions. *Palliative Medicine*. <https://doi.org/10.1177/0269216313486824>
- [14] Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating Quantitative and Qualitative Results in Health Science Mixed Methods Research Through Joint Displays. *Annals of Family Medicine*, 13(6), 554–561.
<https://doi.org/10.1370/afm.1865>
- [15] Grand-Guillaume-Perrenoud, J. A., et al. (2023). Mixed Research as a Tool for Developing Quantitative Instruments. *Journal of Mixed Methods Research*.
<https://doi.org/10.1177/15586898221146414>
- [16] Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research*, 8(3), 362–376.
<https://doi.org/10.1177/104973239800800307>
- [17] McCrudden, M. T., & McTigue, E. M. (2021). Joint displays for mixed methods research in psychology. *Methods in Psychology*, 4, 100051. <https://doi.org/10.1016/j.metip.2021.100051>
- [18] Fetters, M. D., & Guetterman, T. C. (2023). Generating meta-inferences in mixed methods research: A seven-step process. *Journal of Mixed Methods Research*.
<https://doi.org/10.1177/20597991231188121>
- [19] ILO. (2018). *A step-by-step guide to impact evaluation*. International Labour Organization.
https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@ed_emp/documents/instructional-material/wcms_627314.pdf
- [20] Sale, J. E. M., et al. (2002). Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research. *Quality & Quantity*.
<https://doi.org/10.1023/A:1014301607592>
- [21] Schoonenboom, J. (2022). Developing the meta-inference in mixed methods research through successive integration of claims. *The Routledge Handbook for Advancing Integration in Mixed Methods Research*.
<https://doi.org/10.4324/9780429432828-6>
- [22] Mertens, D. M. (2009). *Transformative Research and Evaluation*. Guilford Press.
<https://www.guilford.com/books/Transformative-Research-and-Evaluation/Mertens/9781606233718>
- [23] Mertens, D. M. (2017). *Mixed Methods Design in Evaluation*. SAGE Publications.
<https://us.sagepub.com/en-us/nam/mixed-methods-design-in-evaluation/book245787>
- [24] Younas, A., et al. (2025). Framework for types of meta-inferences in mixed methods research. *BMC Medical Research Methodology*. <https://doi.org/10.1186/s12874-025-02475-8>
- [25] Pluye, P., & Hong, Q. N. (2014). Combining the Power of Qualitative and Quantitative Research: A Methodological Study of Mixed Methods Reviews. *Annual Review of Public Health*, 35, 29–45. <https://doi.org/10.1146/annurev-publhealth-032013-182440>
- [26] Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, 12(2), 281–316.
<https://doi.org/10.46743/2160-3715/2007.1638>
- [27] Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*. <https://doi.org/10.1177/1356389007084674>
- [28] Bazeley, P. (2018). Integrating data analyses in mixed methods research. *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, 437–456.
<https://doi.org/10.4135/9781483398033.n20>
- [29] Quisumbing, A. R., et al. (2024). A synthesis of mixed methods impact evaluations from the agricultural development sector. *Journal of Rural Studies*.
<https://doi.org/10.1016/j.jrurstud.2024.103234>
- [30] Kamei, T., et al. (2022). Prospective fully longitudinal mixed methods evaluation of health literacy changes. *Journal of Mixed Methods Research*.
<https://doi.org/10.1177/1558689822111111>
- [31] Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, 23(3), 246–255.
[https://doi.org/10.1002/1098-240x\(200006\)23:3<246::aid-nur9>3.0.co;2-h](https://doi.org/10.1002/1098-240x(200006)23:3<246::aid-nur9>3.0.co;2-h)
- [32] Teddlie, C., & Tashakkori, A. (2009). *Foundations of Mixed Methods Research*. SAGE Publications.
<https://us.sagepub.com/en-us/nam/foundations-of-mixed-methods-research/book228642>
- [33] Taylor, L. (2025). Transformative multilevel mixed methods design: A worked example in health services research. *Methods in Psychology*, 12, 100107.
<https://doi.org/10.1016/j.metip.2025.100107>
- [34] Teddlie, C., & Yu, F. (2007). Mixed Methods Sampling: A Typology With Examples. *Journal of Mixed Methods Research*, 1(1), 77–100. <https://doi.org/10.1177/1558689806292430>

- [35] Watson, D. P., et al. (2020). A longitudinal mixed method approach for assessing implementation context and process. *BMC Medical Research Methodology*, 20(1), 294. <https://doi.org/10.1186/s12874-020-01164-x>
- [36] O'Cathain, A., et al. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research & Policy*. <https://doi.org/10.1258/jhsrp.2007.007074>
- [37] Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. *Handbook of Mixed Methods in Social & Behavioral Research*. <https://doi.org/10.4135/9781506335193.n14>
- [38] Zhou, Y. (2019). A mixed methods model of scale development and validation analysis. *Measurement: Interdisciplinary Research and Perspectives*. <https://doi.org/10.1080/15366367.2018.1479088>
- [39] International Rescue Committee. (2024). Monitoring and Evaluation Toolkit. <https://www.rescue.org/resource/monitoring-and-evaluation-toolkit>
- [40] Oxfam. (2024). Oxfam's Guide to Monitoring and Evaluation. <https://policy-practice.oxfam.org/resources/oxfams-guide-to-monitoring-and-evaluation-620446/>
- [41] Creswell, J. W. (2015). *A Concise Introduction to Mixed Methods Research*. SAGE Publications. <https://lcn.loc.gov/2021011444>
- [42] Ivankova, N. V., et al. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods*. <https://doi.org/10.1177/1525822X05282260>
- [43] Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*. <https://doi.org/10.1097/00006199-199103000-00014>
- [44] Save the Children. (2024). Monitoring, Evaluation, Accountability and Learning (MEAL) Handbook. <https://resourcecentre.savethechildren.net/library/monitoring-evaluation-accountability-and-learning-meal-handbook>
- [45] Plano Clark, V. L., & Ivankova, N. V. (2016). *Mixed Methods Research: A Guide to the Field*. SAGE Publications. <https://doi.org/10.4135/9781483398341>
- [46] Yin, R. K. (2006). Mixed Methods Research: Are the Methods Genuinely Mixed or Merely Run in Parallel? *Research in the Schools*. <https://www.researchgate.net/publication/232535555>
- [47] Schoonenboom, J., & Johnson, R. B. (2017). How to Construct a Mixed Methods Research Design. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. <https://doi.org/10.1007/s11577-017-0454-1>
- [48] Schumacher, K. L., et al. (2021). Methodological considerations for the design and implementation of a fully longitudinal mixed methods study. *Research in Nursing & Health*. <https://doi.org/10.1002/nur.22133>
- [49] Wittink, M. N., et al. (2006). How to use mixed methods in health services research. *Health Services Research*. <https://doi.org/10.1111/j.1475-6773.2006.00604.x>
- [50] Albright, K., et al. (2013). Importance of Mixed Methods in Pragmatic Trials and Dissemination and Implementation Research. <https://doi.org/10.1016/j.acap.2013.06.010>
- [51] Bamberger, M. (2012). Introduction to Mixed Methods in Impact Evaluation. *InterAction*. <https://www.interaction.org/wp-content/uploads/2019/03/Mixed-Methods-in-Impact-Evaluation-English.pdf>
- [52] Bamberger, M., et al. (2010). *RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints*. SAGE Publications. <https://us.sagepub.com/en-us/nam/realworld-evaluation/book233614>
- [53] IPA. (2024). Mixed Methods Research. <https://www.poverty-action.org/mixed-methods-research>
- [54] White, H., & Sabarwal, S. (2014). *Quasi-experimental Design and Methods*. UNICEF Office of Research. https://www.unicef-irc.org/publications/pdf/brief_8_quasi-experimental%20design_eng.pdf
- [55] Funnell, S. C., & Rogers, P. J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. Jossey-Bass. <https://www.wiley.com/en-us/Purposeful+Program+Theory%3A+Effective+Use+of+Theories+of+Change+and+Logic+Models-p-9780470447635>
- [56] Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. SAGE Publications. <https://www.scirp.org/reference/referencespapers?referenceid=2782576>
- [57] Mayne, J. (2008). Contribution Analysis: An Approach to Exploring Cause and Effect. *ILAC Brief No. 16*. https://www.cgiar-ilac.org/files/ILAC_Brief16_Contribution_Analysis.pdf
- [58] BetterEvaluation. (2024). Mixed Methods. <https://www.betterevaluation.org/methods-approaches/approaches/mixed-methods>
- [59] 3ie. (2024). Mixed Methods Impact Evaluation. <https://www.3ieimpact.org/evidence-hub/publications/working-papers/mixed-methods-impact-evaluation>
- [60] World Bank. (2024). Impact Evaluation in Practice. <https://openknowledge.worldbank.org/handle/10986/25030>
- [61] USAID. (2024). Mixed Methods in Evaluation. <https://www.usaid.gov/evaluation/mixed-methods>
- [62] OECD. (2024). Evaluating Development Co-operation. <https://www.oecd.org/development/evaluation/>
- [63] UNDP. (2024). Handbook on Planning, Monitoring and Evaluating for Development Results. <http://web.undp.org/evaluation/handbook/>
- [64] UNICEF. (2024). Methodological Briefs on Impact Evaluation. <https://www.unicef-irc.org/knowledge-pages/Methodological-Briefs-on-Impact-Evaluation/>

- [65] GIZ. (2024). Guidelines on Monitoring and Evaluation.
<https://www.giz.de/en/workingwithgiz/11163.html>
- [66] DFID. (2024). Evaluation Policy.
<https://www.gov.uk/government/publications/dfid-evaluation-policy-2013>
- [67] J-PAL. (2024). Mixed Methods in Randomized Evaluations.
<https://www.povertyactionlab.org/resource/mixed-methods-randomized-evaluations>

Biography



Peshal Kumar Puri is an MPhil scholar in Development Studies at Kathmandu University, Nepal. He holds a master's degree in management and a Bachelor's degree in Law from Tribhuvan University, giving him a strong academic foundation across development, management, and legal studies. He has more than ten years of professional experience with national and international development organizations, primarily in research, monitoring, and evaluation. His work focuses on data management, evaluation design, and applied research. Mr. Puri has led and contributed to over 40 evaluations and research studies across Nepal, as well as in several countries in Asia and Africa. He has extensive experience conducting complex mixed-methods and impact evaluations using both experimental and non-experimental approaches, including randomized controlled trials and quasi-experimental designs such as Difference-in-Differences (DiD), Regression Discontinuity (RDD), Propensity Score Matching (PSM), Instrumental Variables (IV), and Interrupted Time Series (ITS).