



UNIVERSITY OF THE PHILIPPINES
CENTER FOR
INTEGRATIVE AND
DEVELOPMENT
STUDIES

DISCUSSION PAPER SERIES 2026-22

Impact Evaluation Methodologies and the Challenge of Complexity

Towards a Theory-Driven and Pathway-Based Evaluation

Edylinda Annette Balaoing-Pelkmans  and *Jane Lynn D. Capacio* 

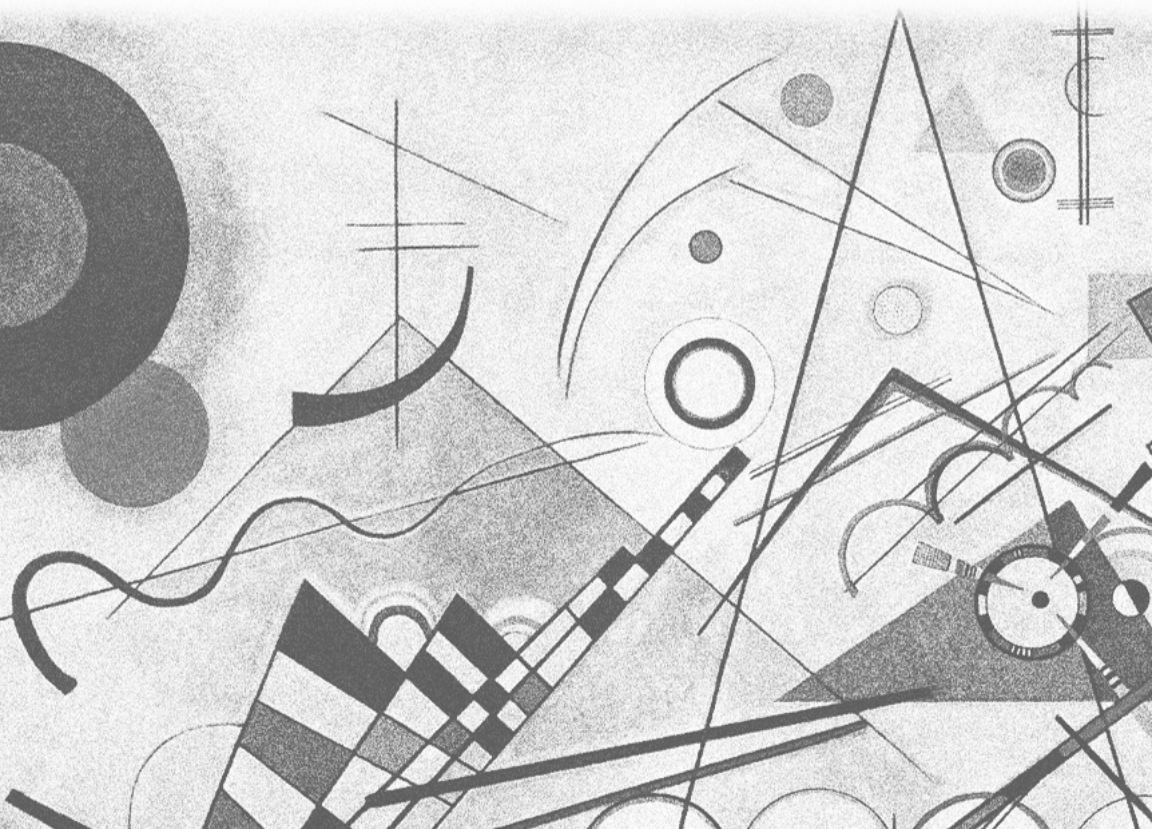


Program on Escaping the Middle - Income Trap: Chains for Change

Impact Evaluation Methodologies and the Challenge of Complexity

Towards a Theory-Driven and
Pathway-Based Evaluation

Edylinda Annette Balaoing-Pelkmans and Jane Lynn D. Capacio



UP CIDS DISCUSSION PAPER SERIES

2026-22

UP CIDS Discussion Paper Series is published by the

**University of the Philippines
Center for Integrative and Development Studies**

Lower Ground Floor, Ang Bahay ng Alumni
Magsaysay Avenue, University of the Philippines
Diliman, Quezon City 1101

Telephone: (02) 8981-8500 loc. 4266 to 4268 / (02) 8426-0955

Email: cidspublications@up.edu.ph

Website: cids.up.edu.ph



cids.up.edu.ph

**Copyright © 2026
by the UP Center for Integrative and Development Studies**

The views and opinions expressed in this discussion paper are those of the author/s and neither reflect nor represent those of the University of the Philippines or the UP Center for Integrative and Development Studies. No copies can be made in part or in whole without prior written permission from the authors/editors and the publisher.

**ISSN 2619-7448 (Print)
ISSN 2619-7456 (Online)**

Cover Image Credit

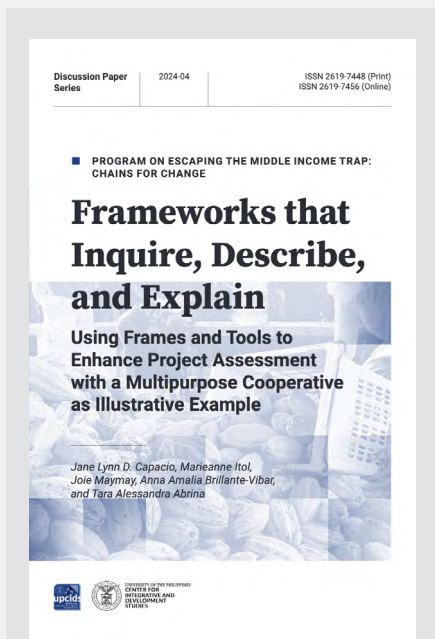
"Wassily Kandinsky's "Composition VIII" (July 1923)."

Table of Contents

Key Highlights	1
Introduction	2
Impact Assessment and the Problem of Complexity	3
Evaluation Methodologies	5
Experimental and Quasi-Experimental Impact Evaluation	5
Theory-Based Impact Evaluation	6
Formalizing Theory: Mechanisms, Pathways, and Intermediate Outcomes	7
Mixed Methods Impact Evaluation and the Challenge of Integration	8
Contribution Analysis and Realist Evaluation	9
Action-Oriented Inquiry and Its Implications for Impact Assessment	11
Environmental and Social Impact Assessment as Parallel Traditions	12
Persistent Methodological Tensions Across Impact Assessment Approaches	13
Impact Assessment, Learning, and the Question of Use	14
Research Impact and Program Impact: Distinct but Related Evaluative Challenges	15
Designing Impact Assessments Under Real-World Constraints	16
Implications for Integrative Evaluation Frameworks	17
Toward Theory-Driven, Pathway-Based Evaluation	17
Conclusion	19
Recommendations	19
References	20

Download related policy papers for FREE

cids.up.edu.ph/collection-database



DISCUSSION PAPER

Frameworks that Inquire, Describe, and Explain: Using Frames and Tools to Enhance Project Assessment with a Multipurpose Cooperative as Illustrative Example



POLICY BRIEF

Unlocking Government Data for Evidenced-Based and Agile Policymaking

Impact Evaluation Methodologies and the Challenge of Complexity

Towards a Theory-Driven and
Pathway-Based Evaluation

Edylinda Annette Balaoing-Pelkmans¹ and Jane Lynn D. Capacio²

1 Convenor of the Escaping the Middle-Income Trap: Chains-for-Change (EMIT C4C) Program of UP CIDS and Assistant Professorial Fellow of the UP School of Economics. Email: eobalaoingpelkmans@up.edu.ph

2 Research Fellow of EMIT C4C and Assistant Professor of the UP Diliman Extension Program in Pampanga and Olongapo (UPDEPPO). Email: jdcapacio@up.edu.ph

Key Highlights

- Given the importance of impact evaluation in development policy, program implementation and applied research, a wide range of impact evaluation methodologies has emerged including experimental and quasi-experimental designs, theory-based approaches, mixed methods, and contribution analysis.
- The methodological diversification reflects a growing recognition that development interventions are complex, multi-actor, and embedded in institutional processes rather than discrete, isolated treatments.
- The plethora of methodologies highlighted persistent tensions in impact evaluation: attribution and contribution, linear and systematic causality, assessment and learning, as well as the relationship between methods and processes.
- This paper reviews and analyzes major impact assessment methodologies by focusing on their suitability for complex development interventions. The paper argues that the challenge of impact evaluation lies not in selecting a single methodological paradigm but in integrating theory construction, causal pathways, and empirical analysis within a coherent evaluative framework.
- Rather than advocating a new method, this paper combines existing approaches to support rigor, learning, and policy relevance in complex development contexts. It proposes a theory-driven, pathway-based evaluation as an integrative approach for assessing the impacts of complex interventions.

Introduction

Impact assessment occupies a central place in contemporary development practice. Governments, donors, civil society organizations, and research institutions increasingly require evidence that programs and policies generate measurable and meaningful change. This demand has intensified over the past two decades alongside the rise of evidence-based policymaking, results-based management, and performance accountability frameworks (Gertler et al. 2016; Rossi, Lipsey, and Freeman 2004). In response, the field of impact evaluation has expanded rapidly, producing a diverse set of methodological approaches and an extensive body of literature on evaluation design and implementation.

Simultaneously, the nature of development interventions has evolved in ways that complicate evaluation. Many contemporary interventions no longer consist of discrete, time-bound activities delivered to clearly defined beneficiaries. Instead, they operate through networks of actors, span multiple institutional levels, and adapt over time in response to learning and contextual change. Agricultural value-chain programs, industrial upgrading initiatives, social protection systems, and governance reforms typically combine economic, social, and institutional objectives and rely on coordination among public agencies, private firms, civil society organizations, and community institutions. Thus, outcomes emerge through interaction, sequencing, and feedback rather than through linear cause–effect relationships (Rogers 2008; Bamberger, Rugh, and Mabry 2012).

This evolution has exposed growing tensions between the assumptions underpinning dominant impact assessment methodologies and the realities of complex development practice. Approaches that prioritize causal identification often rely on assumptions of stable treatments, limited interaction among units, and linear causality (Shadish, Cook, and Campbell 2002). While these assumptions may be defensible in narrowly defined interventions, they become increasingly difficult to sustain in settings characterized by institutional complexity, adaptive implementation, and systemic interdependence. Conversely, approaches that emphasize context, process, and participation may struggle to maintain analytical discipline, comparability, and credibility, particularly in policy environments accustomed to quantitative estimates of impact (White 2009; Bamberger 2012).

These tensions are now widely recognized in impact assessment literature. They are reflected in debates over the limits of randomized controlled trials, the role of theory in evaluation, the promises and pitfalls of mixed methods, and the

appropriate balance between attribution and contribution (Weiss 1997; Pawson and Tilley 1997; Mayne 2011). Despite extensive discussion, however, there is no consensus on how best to evaluate complex development interventions in ways that are both analytically rigorous and substantively meaningful.

This paper engages with these debates by reviewing major impact assessment methodologies and examining how they conceptualize causality, handle complexity, and support learning. Rather than advocating for the dominance of any single approach, it argues that methodological progress lies in integration. Specifically, it suggests that theory-driven, pathway-based evaluation provides a structured way to combine insights from multiple traditions while remaining attentive to complexity, institutional dynamics, and policy relevance.

Impact Assessment and the Problem of Complexity

The challenge of evaluating complex development interventions is not merely technical; it is also conceptual. At its core lies the problem of understanding causality in settings where change is produced by multiple interacting factors rather than by isolated treatments. Traditional impact evaluation approaches often rest on an implicit model of causality wherein interventions act as exogenous shocks that generate measurable changes in outcomes. While this model has proven powerful in certain contexts, it is increasingly misaligned with the nature of many contemporary interventions (Shadish, Cook, and Campbell 2002; White 2009).

Complexity in development interventions manifests in several interrelated ways. First, interventions are frequently multi-component. Programs combine training, financing, institutional support, and policy engagement in ways that are difficult to disentangle analytically (Rossi, Lipsey, and Freeman 2004). Second, interventions are multi-actor. Outcomes depend on the behavior and interaction of implementing agencies, beneficiaries, intermediaries, and external stakeholders, often operating across different levels of governance (Rogers 2008). Third, interventions are dynamic. Implementation strategies evolve over time as actors learn, adapt, and respond to changing conditions, rendering static evaluation designs insufficient (Bamberger, Rugh, and Mabry 2012). Fourth, interventions are embedded in broader social, economic, and institutional systems. They interact with markets, regulatory frameworks, social norms, and political incentives, producing intended and unintended

effects. As environmental and social impact assessment traditions have long emphasized, such interactions can generate cumulative, indirect, and delayed impacts that are not captured by narrow outcome measures (Glasson, Therivel, and Chadwick 2012; Vanclay 2003).

These features complicate core evaluative tasks. Defining the treatment becomes problematic when program components change over time or vary across contexts. Constructing credible counterfactuals becomes difficult when spillovers and interaction effects are central to program logic. Interpreting observed outcomes becomes challenging when multiple causal pathways operate simultaneously and feedback effects are present (White 2009; Mayne 2011).

Impact assessment literature has responded to these challenges in different ways. Some approaches seek to extend existing methods through more sophisticated designs, such as clustered randomization or longitudinal analysis (Gertler et al. 2016). Others advocate for a shift in emphasis toward theory-based reasoning, mixed methods, or contribution analysis (Weiss 1997; Bamberger 2012; Pawson and Tilley 1997). Still others draw on parallel traditions, such as environmental and social impact assessment, that have long grappled with cumulative effects, significance, and systemic change (Thompson 1990; Sadler 1996).

Despite these advances, a recurring concern is methodological debates often proceeding in isolation from one another. Experimental, theory-based, mixed methods, and contribution-focused approaches are frequently presented as alternatives rather than as complementary tools. As a result, evaluations may become method-driven rather than question-driven, with methodological choices shaped more by disciplinary norms than by the nature of the intervention or the evaluative purpose (Better Evaluation 2023).

This fragmentation points to a deeper issue: the absence of a coherent framework for integrating different approaches to impact assessment. Without such a framework, evaluators face a trade-off between rigor and relevance, attribution and explanation, simplicity and realism. This paper argues that this trade-off is not inevitable. By grounding evaluation in explicit theories of change and formalized causal pathways, it is possible to integrate multiple methodological traditions in ways that enhance both analytical rigor and substantive insight.

Evaluation Methodologies

Experimental and Quasi-Experimental Impact Evaluation

Experimental and quasi-experimental approaches have come to dominate contemporary impact evaluation, particularly in development economics and policy analysis. These approaches are grounded in a counterfactual conception of causality, where the impact of an intervention is defined as the difference between observed outcomes for treated units and the outcomes that would have occurred in the absence of the intervention (Shadish, Cook, and Campbell 2002). Randomized controlled trials, difference-in-differences designs, regression discontinuity, matching methods, and instrumental variable approaches are all variations on this core logic.

The appeal of experimental and quasi-experimental approaches lies primarily in their emphasis on internal validity. By carefully constructing counterfactuals and controlling for confounding factors, these designs aim to isolate causal effects with a high degree of credibility. Practitioner-oriented guides such as *Impact Evaluation in Practice* have played a significant role in institutionalizing this paradigm within development agencies, framing rigorous impact evaluation as synonymous with credible counterfactual estimation (Gertler et al. 2016).

This emphasis has yielded important benefits. Experimental approaches have raised standards of transparency and methodological discipline, encouraged the use of pre-analysis plans, and reduced reliance on post hoc causal claims. In contexts with narrowly-defined interventions, standardized implementation, and minimized spillovers, these methods have generated persuasive evidence that has informed policy debates and resource allocation decisions.

At the same time, a growing body of literature has highlighted the limitations of experimental and quasi-experimental approaches applied to complex development interventions. One fundamental challenge concerns the definition of treatment. Many development programs consist of multiple, interacting components that evolve over time. Training may be combined with financing, institutional support, and policy engagement, with different elements emphasized at different stages of implementation. In such settings, the assumption of a stable, well-defined treatment becomes difficult to sustain (Rossi, Lipsey, and Freeman 2004).

A second challenge concerns interaction and spillovers. Experimental designs typically rely on assumptions of limited interference among units, often formalized as the stable unit treatment value assumption. Yet many development interventions operate precisely through interaction, coordination, and network effects. Agricultural value-chain initiatives, for example, may generate spillovers through market integration, information diffusion, or changes in bargaining power. Institutional reforms may alter incentives and behaviors beyond the immediate scope of the intervention. These spillovers are not incidental; they are often central to program logic, yet they complicate counterfactual estimation (Shadish, Cook, and Campbell 2002; White 2009).

A third limitation relates to explanation and learning. While experimental estimates can establish whether an intervention had an effect on average, they provide limited insight into how change occurred, why impacts varied across contexts, or which mechanisms were most important. As a result, their usefulness for informing program adaptation and policy design may be constrained. Rossi, Lipsey, and Freeman (2004) note that evaluations focused narrowly on impact estimation risk neglecting implementation processes, institutional constraints, and unintended effects.

These limitations have prompted efforts to extend experimental approaches through more sophisticated designs, including clustered randomization, longitudinal follow-up, and the incorporation of mediating variables. While such extensions can address some concerns, they do not fully resolve the underlying tension between attribution-focused evaluation and the complexity of many development interventions.

Theory-Based Impact Evaluation

Theory-based impact evaluation emerged partly in response to dissatisfaction with what came to be known as “black-box” evaluation. Rather than treating interventions as exogenous treatments, theory-based approaches emphasize the causal mechanisms, assumptions, and contextual conditions through which interventions are expected to generate change (Weiss 1997; White 2009). Central to this tradition is the articulation of a Theory of Change that links activities, intermediate outcomes, and final impacts.

Theory-based evaluation offers several important advantages in the context of complex interventions. By making assumptions explicit, it enhances transparency and provides a framework for interpreting observed changes. It

encourages evaluators to ask not only whether an intervention worked, but how and why it produced particular outcomes. These features make theory-based approaches particularly attractive for interventions that operate through multiple interacting pathways and involve diverse actors (Rogers 2008).

In practice, theories of change are often developed through participatory processes involving implementers, stakeholders, and beneficiaries. This can facilitate shared understanding and buy-in, but it also introduces challenges. Theories of change may reflect negotiated consensus rather than analytical scrutiny, with competing explanations glossed over in the interest of coherence. Moreover, theories of change are frequently presented as linear diagrams obscuring feedback effects, non-linear dynamics, and institutional interaction (White 2009). A recurring critique in the literature is that theory-based evaluations often lack formal analytical structures. While theories of change articulate plausible causal narratives, they are rarely translated into testable propositions or subjected to systematic empirical examination. As a result, theory-based evaluation risks remaining at the level of plausibility rather than explanation. White (2009) emphasizes that theory-based approaches must go beyond description and engage explicitly with causal inference if they are to contribute meaningfully to evaluation practice.

This critique points to a central challenge of operationalizing theory in ways that discipline empirical analysis rather than merely accompanying it. Without formalization, theory risks becoming an add-on instead of being a core component of evaluation design.

Formalizing Theory: Mechanisms, Pathways, and Intermediate Outcomes

One response to the limitations of both experimental and theory-based approaches has been to focus more explicitly on mechanisms and intermediate outcomes. Rather than treating impact as a single end-point, this perspective emphasizes the processes through which interventions influence behavior, institutions, and outcomes over time (Rogers 2008; Weiss 1997).

From this perspective, theories of change are most useful when they specify not only sequences of activities and outcomes, but also the mechanisms linking them. Mechanisms refer to the underlying processes where interventions produce change, such as learning, incentive shifts, coordination, or norm change. Making mechanisms explicit allows evaluators to examine whether

interventions operated as intended and to identify points of failure or adaptation.

However, specifying mechanisms alone is insufficient. For theory-based evaluation to support rigorous analysis, mechanisms must be embedded within structured causal pathways that can be examined empirically. This requires identifying intermediate outcomes that are observable, measurable, and theoretically meaningful. Intermediate outcomes serve as bridges between activities and final impacts, allowing evaluators to trace how change unfolds over time.

The challenge lies in balancing analytical tractability with realism. Overly detailed theories risk becoming unwieldy, while overly simplified theories may obscure important dynamics. Moreover, intermediate outcomes are often interdependent rather than sequential. Changes in income, empowerment, and resilience, for example, may reinforce one another rather than occur in a fixed order. Capturing such interdependencies requires moving beyond linear causal chains toward pathway-based representations.

These considerations underscore the need for evaluation frameworks that can integrate theory, empirical analysis, and complexity. While theory-based approaches provide a foundation, they must be complemented by methods that can examine interdependent pathways and feedback effects. This insight sets the stage for the discussion of mixed methods, contribution analysis, and pathway-based evaluation in the next sections.

Mixed Methods Impact Evaluation and the Challenge of Integration

Mixed methods impact evaluation has gained increasing prominence as dissatisfaction with single-method approaches has grown. At its core, mixed methods research combines quantitative and qualitative forms of evidence to address different dimensions of impact, most commonly magnitude and meaning, outcomes and process, or attribution and explanation (Greene 2007; Creswell and Plano Clark 2017). In development evaluation, mixed methods are frequently promoted as a pragmatic response to the limitations of experimental and quasi-experimental designs, particularly in complex or data-constrained settings.

The appeal of mixed methods lies in their promise of complementarity. Quantitative methods provide statistical inference, comparability, and a degree of generalizability. Qualitative methods offer insight into institutional dynamics, behavioral responses, sequencing, and contextual conditions that are difficult to capture numerically. When effectively combined, mixed methods can strengthen both causal inference and policy relevance.

In practice, however, mixed methods literature consistently emphasizes that integration is the central challenge. Many mixed methods evaluations consist of parallel quantitative and qualitative components that are weakly connected. Qualitative findings are often used to illustrate or contextualize quantitative results, while quantitative analysis proceeds independently of qualitative insight. This form of additive combination falls short of the deeper integration envisioned in the mixed methods literature (Tashakkori and Teddlie 2010).

Bamberger (2012) argues that effective mixed methods impact evaluation requires integration at the level of research design rather than merely at the reporting stage. Qualitative inquiry should inform the formulation of research questions, the selection of variables, and the interpretation of causal relationships. Without such integration, mixed methods risk reproducing the weaknesses of single-method approaches while adding complexity and cost.

A further challenge concerns epistemological coherence. Quantitative and qualitative methods are often grounded in different assumptions about causality, validity, and generalization. Combining them mechanically does not resolve these differences. Instead, it can obscure underlying tensions and lead to ambiguous interpretations. This issue is particularly salient in impact assessment, where causal claims carry significant policy weight.

These critiques point to a central insight: mixed methods are most effective when anchored in explicit theory. Theory provides a common frame of reference that can guide method selection, integration, and interpretation. Without theory, mixed methods risk becoming a technical fix rather than a coherent evaluative strategy.

Contribution Analysis and Realist Evaluation

Contribution analysis and realist evaluation represent a further shift away from exclusive attribution as the defining criterion of rigorous impact assessment. Instead of asking whether an intervention caused an outcome, these approaches

ask whether it plausibly contributed to observed change, given alternative explanations and contextual conditions (Pawson and Tilley 1997; Mayne 2011).

Contribution analysis explicitly acknowledges that development outcomes are often co-produced by multiple actors and influences. It seeks to build a credible contribution story by articulating causal mechanisms, examining evidence along the theory of change, and assessing the plausibility of alternative explanations. The emphasis is on triangulation and theory-based reasoning rather than on statistical identification alone (Mayne 2011).

Realist evaluation shares this orientation, but places greater emphasis on context–mechanism–outcome configurations. It asks not only whether an intervention contributed to change, but also under what conditions and through which mechanisms (Pawson and Tilley 1997). This focus aligns closely with the realities of complex development interventions, where the same program may produce different outcomes in different contexts.

These approaches offer important conceptual advances. They foreground mechanisms, acknowledge institutional interaction, and emphasize learning rather than binary judgments of success or failure. They are particularly well suited to policy interventions, institutional reforms, and programs involving multiple stakeholders.

At the same time, the literature also identifies significant challenges. One concern is the potential loss of analytical discipline. Contribution analysis relies heavily on narrative reasoning and qualitative evidence. Without structured analytical tools, contribution claims can become difficult to compare, aggregate, or generalize. Critics argue that this can weaken credibility, particularly in policy environments accustomed to quantitative estimates of impact.

Another concern relates to boundary definition. In open systems, determining which factors to include in a contribution story is inherently subjective. Without clear analytical frameworks, evaluations risk becoming overly expansive or unfocused. These challenges do not invalidate contribution analysis or realist evaluation, but instead underscore the need for methods that can formalize contribution pathways without reverting to reductionist causal models.

Action-Oriented Inquiry and Its Implications for Impact Assessment

Action-oriented inquiry, including action research, brings these methodological tensions into sharp relief. In action research settings, researchers engage with practitioners, policymakers, and other stakeholders not only to observe change, but to shape it. Knowledge production and intervention are intertwined, and research questions evolve through engagement and learning.

From the perspective of conventional impact evaluation, action-oriented inquiry poses significant challenges. Treatments are not fixed, implementation is adaptive, and researcher involvement is itself part of the causal process. These features violate key assumptions underpinning experimental and quasi-experimental designs, including stable treatments and researcher neutrality (Shadish, Cook, and Campbell 2002).

At the same time, action-oriented inquiry reflects the realities of many development interventions. Programs aimed at institutional reform, capacity building, or structural transformation are inherently adaptive and participatory. Evaluating such interventions using rigid designs risks mischaracterizing both the nature of the intervention and the sources of observed change.

The methodological challenge, therefore, is not whether action-oriented inquiry is compatible with impact assessment, but how evaluation approaches can be adapted to accommodate iterative learning and institutional engagement without sacrificing rigor. Contribution analysis and realist evaluation offer partial responses, but they require complementary analytical tools to support systematic examination of pathways and outcomes.

This challenge reinforces the argument that impact assessment methodologies must be capable of integrating theory, process, and empirical analysis. Mixed methods, contribution-focused approaches, and action-oriented inquiry each highlight different dimensions of complexity. Bringing these dimensions together requires frameworks that can formalize causal pathways while remaining attentive to context and learning.

The next section turns to environmental and social impact assessment traditions, which have long grappled with cumulative effects, indirect impacts, and significance, offering additional insights for evaluating complex development interventions.

Environmental and Social Impact Assessment as Parallel Traditions

Environmental and social impact assessment traditions developed largely outside the mainstream of development impact evaluation, yet they offer valuable insights for assessing complex interventions. Environmental impact assessment (EIA) emerged in response to large-scale infrastructure and industrial projects whose effects were cumulative, indirect, and often irreversible. From its inception, EIA emphasized the need to anticipate impacts before they materialize, to examine interactions among activities and environmental systems, and to assess the significance of change rather than merely its existence (Glasson, Therivel, and Chadwick 2012; Noble 2015).

Early methodological reviews in environmental impact assessment document a wide range of tools designed explicitly to address complexity. These include interaction matrices, checklists, scoring and weighting systems, and approaches to cumulative impact assessment (Carlin et al. 1977; Thompson 1990). While these tools were not designed to establish causal attribution in the econometric sense, they reflect a recognition that impacts arise from interactions among activities, environmental systems, and social processes. The analytical task in EIA is therefore one of tracing pathways, identifying thresholds, and judging significance rather than isolating net effects.

The concept of significance occupies a central place in environmental impact assessment. Significance is not determined solely by statistical magnitude but by judgments concerning reversibility, distributional effects, uncertainty, and long-term consequences (Thompson 1990; Sadler 1996). This broader conception of impact contrasts with development impact evaluation, where statistical significance is often privileged over substantive importance. The EIA literature thus raise fundamental questions about what counts as impact and how it should be assessed.

Social impact assessment (SIA) extends this perspective by expounding on social relations, power dynamics, and institutional change. Vanclay (2003) emphasizes that social impacts are not merely outcomes to be measured but processes that unfold over time, shaped by perceptions, expectations, and interactions among stakeholders. Participation is therefore not simply a data collection strategy but an integral component of impact assessment.

For development impact evaluation, these traditions offer two important lessons. First, they demonstrate that complexity can be addressed through structured analytical frameworks that do not rely exclusively on counterfactual estimation. Second, they show that impact assessment can incorporate process, participation, and judgment without abandoning rigor. The relative marginalization of EIA and SIA insights within development evaluation reflects disciplinary boundaries rather than methodological inadequacy.

Persistent Methodological Tensions Across Impact Assessment Approaches

Across the impact assessment literature reviewed in this paper, several persistent methodological tensions emerge. One such tension concerns attribution versus contribution. Experimental and quasi-experimental approaches prioritize attribution, seeking to isolate the causal effect of an intervention. Theory-based, contribution-focused, and realist approaches emphasize contribution, recognizing that outcomes are often co-produced by multiple actors and influences (White 2009; Mayne 2011).

A second tension concerns linear versus systemic conceptions of causality. Many evaluation designs implicitly assume linear causal chains linking interventions to outcomes. Yet complex development interventions operate through interdependent pathways, feedback loops, and emergent dynamics. Environmental and social impact assessment traditions have long emphasized cumulative and indirect effects, highlighting the limits of linear causal models (Glasson, Therivel, and Chadwick 2012; Vanclay 2003).

A third tension concerns method versus process. Some approaches treat evaluation as a technical exercise conducted after implementation, while others view it as an ongoing process of learning and adaptation. Action-oriented inquiry and participatory approaches challenge conventional boundaries between research and intervention, raising questions about the role of evaluators and the purposes of evaluation (Bamberger, Rugh, and Mabry 2012).

These tensions are not anomalies; they reflect fundamental differences in how impact, causality, and evidence are conceptualized. Attempts to resolve them by privileging a single methodological paradigm have proven unsatisfactory. Instead, they point to the need for integrative frameworks that can accommodate multiple perspectives while maintaining analytical coherence.

Impact Assessment, Learning, and the Question of Use

A recurring but often under-theorized issue in impact assessment literature concerns the purpose of evaluation and, relatedly, the use of evaluation findings. While impact evaluation is frequently framed as a tool for accountability, particularly in donor-funded development contexts, there is growing recognition that evaluation also serves learning, adaptation, and strategic decision-making functions (Rossi, Lipsey, and Freeman 2004; Bamberger, Rugh, and Mabry 2012).

The distinction between evaluation for accountability and evaluation for learning is not merely rhetorical. It has direct implications for methodological choice. Accountability-oriented evaluations often prioritize clear attribution, standardized metrics, and comparability across programs. Learning-oriented evaluations place greater emphasis on explanation, context, and process, seeking to understand why interventions work or fail and under what conditions. Tensions arise when a single evaluation is expected to serve both purposes simultaneously.

Experimental and quasi-experimental approaches align closely with accountability-oriented evaluation. By producing estimates of average treatment effects, they offer apparently clear answers to questions of effectiveness. However, as the literature has repeatedly noted, such estimates often provide limited guidance for program adaptation or scaling, particularly in heterogeneous contexts (Gertler et al. 2016; White 2009). When evaluation findings are reduced to binary judgments of success or failure, opportunities for learning may be lost.

Theory-based, mixed methods, and contribution-focused approaches are more explicitly oriented toward learning. By foregrounding mechanisms, contextual conditions, and implementation processes, these approaches generate insights that are more readily actionable for practitioners. At the same time, they may be perceived as less credible or less decisive in accountability-focused environments, particularly when they do not produce single summary measures of impact.

This tension is not accidental. It reflects deeper institutional incentives shaping evaluation practice. Donor agencies and implementing organizations often operate under pressure to demonstrate results within short timeframes, encouraging the use of methods that produce clear, quantifiable outputs.

Learning-oriented evaluation, by contrast, often requires longer engagement, iterative analysis, and tolerance for ambiguity (Bamberger, Rugh, and Mabry 2012).

Recognizing these dynamics is essential to understand why certain impact assessment methodologies dominate in practice, despite well-documented limitations. It also highlights the importance of designing evaluation frameworks that can accommodate multiple uses without collapsing into methodological incoherence. Pathway-based evaluation, by explicitly linking mechanisms, intermediate outcomes, and final impacts, offers one way to bridge accountability and learning objectives within a single evaluative structure.

Research Impact and Program Impact: Distinct but Related Evaluative Challenges

A second issue that merits explicit attention is the distinction between program impact and research impact. Most impact evaluation literature focus on assessing the effects of development interventions, such as policies, programs, or projects. However, in applied policy research contexts, evaluators are increasingly asked to assess the impact of research itself, including its influence on policy, practice, and institutional behavior.

The methodological challenges associated with research impact assessment differ in important respects from those associated with program impact evaluation. Research outputs rarely function as discrete treatments. Instead, they influence decision-making through complex pathways involving dissemination, interpretation, negotiation, and adaptation. Attribution is particularly difficult, as research competes with other sources of evidence, political considerations, and institutional constraints (Reed et al. 2021).

Recent frameworks for evaluating research impact emphasize contribution rather than attribution, focusing on impact pathways that link research activities to intermediate outcomes such as changes in knowledge, attitudes, relationships, or capacities, and ultimately to policy or practice change (Reed et al. 2021). These frameworks draw heavily on theory-based and contribution-focused approaches, highlighting the importance of articulating and testing causal assumptions.

The distinction between research impact and program impact underscores a broader point: impact assessment methodologies must be tailored to the nature

of what is being evaluated. Applying attribution-focused designs developed for discrete interventions to research impact assessment risks producing misleading conclusions. Conversely, insights from research impact evaluation can inform the evaluation of complex development programs, particularly those that rely on learning, coordination, and institutional change.

Better Evaluation (2023) emphasizes that evaluative questions should drive method selection, rather than the reverse. Whether the object of evaluation is a program, a policy, or a body of research, the choice of methods should reflect the underlying causal logic and the intended use of findings. This principle reinforces the argument for integrative frameworks that can accommodate diverse evaluative purposes.

Designing Impact Assessments Under Real-World Constraints

A further theme running through impact assessment literature concerns the constraints under which evaluations are conducted. Ideal evaluation designs are often infeasible due to budgetary limitations, data gaps, political sensitivities, or timing constraints. Bamberger, Rugh, and Mabry (2012) emphasize that “real-world evaluation” requires methodological flexibility and pragmatic decision-making rather than strict adherence to idealized standards.

These constraints interact with the complexity of interventions in ways that further complicate evaluation design. For example, baseline data may be unavailable or unreliable, precluding certain quasi-experimental designs. Randomization may be politically unacceptable or ethically inappropriate. Program timelines may not align with evaluation cycles, limiting opportunities for longitudinal analysis.

Under such conditions, evaluators must make trade-offs. The key question is not whether an evaluation meets an abstract standard of rigor, but whether it produces credible and useful evidence given contextual constraints. This perspective challenges method-driven approaches and underscores the importance of evaluative judgment.

Pathway-based evaluation can be understood as a pragmatic response to these realities. By focusing on articulating and testing causal pathways, it allows evaluators to make systematic use of available data while remaining explicit about assumptions and limitations. Quantitative analysis can be applied where

data permit, while qualitative evidence fills gaps and supports interpretation. Importantly, this approach does not require full control over implementation or perfect data conditions to generate meaningful insights.

This pragmatism does not imply methodological laxity. On the contrary, it requires careful design, transparency, and reflexivity. Evaluators must be explicit about what can and cannot be inferred, how evidence supports particular claims, and where uncertainty remains. Such transparency is essential for maintaining credibility, particularly in policy-relevant evaluations.

Implications for Integrative Evaluation Frameworks

This section reinforces the paper's central argument that impact assessment methodologies must be capable of integrating rigor, learning, and use in complex settings. Accountability-oriented and learning-oriented evaluations need not be mutually exclusive, but integrating them requires frameworks that go beyond single-method designs.

By linking theory construction, pathway specification, and empirical analysis, pathway-based evaluation provides a structure for such integration. It accommodates multiple forms of evidence, supports explanation and learning, and remains compatible with accountability requirements when carefully designed. Moreover, it aligns with emerging approaches to research impact assessment, suggesting broader applicability beyond program evaluation alone.

These considerations set the stage for the synthesis developed in the following section. The task is not to resolve every methodological tension, but to design evaluation frameworks that make these tensions explicit and productive rather than obscured or ignored.

Toward Theory-Driven, Pathway-Based Evaluation

Against this background, theory-driven, pathway-based evaluation can be understood as an attempt to integrate insights from multiple impact assessment traditions. Instead of privileging exclusive attribution or abandoning rigor, this approach seeks to formalize contribution pathways articulated through theory-based and action-oriented analysis.

At the core of a pathway-based approach is explicit theory construction. Rather than treating theory as a preliminary planning exercise, theory is used as a disciplining device that specifies how change is expected to occur, through which mechanisms, and under what conditions (Weiss 1997; Rogers 2008). Theories of change are articulated not as linear chains, but as systems of interrelated pathways linking actions, intermediate outcomes, and final impacts.

These pathways are then formalized analytically. Intermediate outcomes are treated as endogenous and interdependent rather than as isolated endpoints. Relationships among outcomes are specified explicitly, allowing for feedback effects and mutual reinforcement. This formalization creates a bridge between theory-based reasoning and empirical analysis, addressing a key weakness identified in the theory-based evaluation literature (White 2009).

Quantitative methods play an important role within this framework, but their function shifts. Rather than estimating a single net effect, quantitative analysis examines the strength and direction of relationships along specified pathways. This allows evaluators to assess which mechanisms matter most, how outcomes interact, and where complementarities or bottlenecks arise. Such analysis does not claim exclusive attribution, but provides disciplined evidence on contribution.

Qualitative methods are integrated through process tracing and interpretive analysis. Rather than serving as illustrative add-ons, qualitative evidence is used to examine whether estimated pathways align with observed practices, institutional behaviors, and stakeholder experiences. This form of integration responds directly to critiques raised in the mixed methods literature regarding superficial combinations of methods (Bamberger 2012; Tashakkori and Teddlie 2010).

The impact assessment design developed for an entrepreneurship program involving smallholder farmers provides an illustration of how such an approach can be operationalized (ongoing project). In that design, income, resilience, and empowerment are treated as interdependent outcomes contributing jointly to quality of life. Institutional actors enter the analysis through the mechanisms they influence rather than as binary treatments. Quantitative modeling examines interdependencies among outcomes, while qualitative inquiry assesses whether these relationships are enacted in practice. The value of this design lies not in its novelty, but in its coherence with the nature of the intervention and the evaluative questions posed.

Conclusion

Impact evaluation methodologies have diversified in response to growing recognition that development interventions are complex, multi-actor, and embedded in institutional processes. Yet this diversification has not resolved enduring methodological tensions concerning attribution and contribution, linear and systemic causality, and method and process. This paper reviewed major impact assessment traditions, examining their strengths and limitations in relation to these challenges.

The analysis suggests that the limitations of existing approaches stem less from methodological weakness than from isolated application. Experimental, theory-based, mixed methods, contribution-focused, and environmental and social impact assessment traditions each illuminate important aspects of impact, but none are sufficient on their own. Integrative approaches that combine theory construction, pathway specification, and empirical analysis offer a way forward.

Theory-driven, pathway-based evaluation represents one such integrative response. By formalizing causal pathways articulated through theory and engagement, and by combining quantitative and qualitative evidence within a coherent framework, it allows impact assessment to remain analytically rigorous while engaging seriously with complexity. Such an approach does not eliminate uncertainty or judgment, but it makes both explicit and accountable.

As development interventions continue to evolve toward greater complexity, the demands placed on impact assessment will only intensify. Meeting these demands requires moving beyond methodological silos toward evaluative frameworks capable of capturing interaction, learning, and institutional change. The challenge is not to choose between rigor and relevance, but to design impact assessments that can sustain both.

Recommendations

Program and project managers, as well as evaluators (who include academics) could consider these recommendations. First, treat impact evaluations as part of program or project design by jointly constructing explicit, pathway-based theories of change that reflect multi-component, multi-actor, and adaptive implementation realities rather than linear input-output chains. In this way, evaluations are more grounded to how development programs and projects have been implemented. Second, closely look into intermediate outcomes

(e.g., behavior change, institutional practices, deepening relationships between institutional partners) that sit along key pathways, so evaluation also tackles how change happens, not just whether final targets were met. Lastly, treat evaluations as learning processes by building iterative reflection, adaptation, and feedback loops with partners and communities, especially where interventions evolve over long periods of time.

References

- Bamberger, Michael. 2012. "Introduction to Mixed Methods in Impact Evaluation." *InterAction Impact Evaluation Notes*, No. 3. <https://www.interaction.org/wp-content/uploads/2019/03/Mixed-Methods-in-Impact-Evaluation-English.pdf>.
- Bamberger, Michael, Jim Rugh, and Linda Mabry. 2012. *RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints*. Thousand Oaks, CA: SAGE Publications.
- BetterEvaluation. (2023). "Impact evaluation: Methods and approaches." Better Evaluation, 2023. <https://www.betterevaluation.org/methods-approaches/themes/impact-evaluation>.
- Carlin, A., Ian, et al. Maurice Warner, and Edward Preston. (1977). *A Review of Environmental Impact Assessment Methodologies*. Washington, DC: U.S. Environmental Protection Agency. <https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=9101BOX1.TXT>.
- Creswell, John W., and Vicki L. Plano Clark. 2017. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: SAGE Publications.
- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M.J. Vermeersch. 2016. *Impact Evaluation in Practice*. Washington, DC: World Bank. <https://www.worldbank.org/en/programs/sief-trust-fund/publication/impact-evaluation-in-practice>.
- Glasson, John, Riki Therivel, and Andrew Chadwick. 2012. *Introduction to Environmental Impact Assessment*. London: Routledge.
- Mayne, John. 2011. "Contribution analysis: Addressing cause and effect." In *Evaluating the Complex: Attribution, Contribution, and Beyond*, edited by R. Schwartz, K. Forss, and M. Marra, 53–96. New Brunswick, NJ: Transaction Publishers.

- Noble, Bram F. 2015. *Introduction to Environmental Impact Assessment: A Guide to Principles and Practice*. Don Mills, ON: Oxford University Press.
- Pawson, Ray, and Nicholas Tilley. 1997. *Realistic Evaluation*. London: SAGE Publications.
- Reed, Mark S., Marie Ferré, Julia Martin-Ortega, Rachel Blanche, R. Lawford-Rolfe, Martin Dallimer, and Joseph Holden. 2021. "Evaluating impact from research: A methodological framework." *Research Policy*, 50 (4): 104–120.
- Rogers, Patricia J. 2008. "Using programme theory to evaluate complicated and complex aspects of interventions." *Evaluation* 14 (1): 29–48. DOI: 10.1177/1356389007084674.
- Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman. 2004. *Evaluation: A Systematic Approach* (7th ed.). Thousand Oaks, CA: SAGE Publications.
- Sadler, Barry. 1996. *International Study of the Effectiveness of Environmental Assessment*. Ottawa: Canadian Environmental Assessment Agency.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin.
- Tashakkori, Abbas, and Charles Teddlie. 2010. *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oaks, CA: SAGE Publications.
- Thompson, Mark A. 1990. "Determining impact significance in environmental impact assessment: A review of 24 methodologies." *Journal of Environmental Management* 30 (3): 235–250. <https://www.sciencedirect.com/science/article/pii/030147979090004G>.
- Vanclay, Frank. 2003. "International principles for social impact assessment." *Impact Assessment and Project Appraisal* 21 (1): 5–11. <https://doi.org/10.3152/147154603781766491>.
- Weiss, Carol H. 1997. "Theory-based evaluation: Past, present, and future." *New Directions for Evaluation* 76 (1): 41–55. <https://doi.org/10.1002/ev.1086>.
- White, Howard. 2009. "Theory-based impact evaluation: Principles and practice." *Journal of Development Effectiveness* 1 (3): 271–284. <https://doi.org/10.1080/19439340903114628>.

THE UP CIDS DISCUSSION PAPER SERIES

The UP CIDS Discussion Paper Series features preliminary researches that may be subject to further revisions and is circulated to elicit comments and suggestions for enrichment and refinement. They contain findings on issues that are aligned with the core agenda of the research programs under the University of the Philippines Center for Integrative and Development Studies (UP CIDS).

Center for Integrative and Development Studies

Established in 1985 by University of the Philippines (UP) President Edgardo J. Angara, the UP Center for Integrative and Development Studies (UP CIDS) is the policy research unit of the University that connects disciplines and scholars across the several units of the UP System. It is mandated to encourage collaborative and rigorous research addressing issues of national significance by supporting scholars and securing funding, enabling them to produce outputs and recommendations for public policy.

The UP CIDS currently has twelve research programs that are clustered under the areas of education and capacity building, development, and social, political, and cultural studies. It publishes policy briefs, monographs, webinar/conference/forum proceedings, and the Philippine Journal for Public Policy, all of which can be downloaded free from the UP CIDS website.

The Program

The Program on Escaping the Middle-Income Trap: Chains for Change (EMIT C4C) examines inclusiveness and competitiveness in its research agenda and its efforts on working with its partners to address social development challenges. It undertakes three projects: (1) big data analytics on manufacturing, labor, and exports; (2) action research on agriculture and fisheries value chains; and (3) action research on business-for-peace in Mindanao. EMIT C4C provides learning and partnering platforms for the academe, government, civil society, and the private sector to share lessons and act on recommendations.

Editorial Board

Rosalie Arcala Hall
Editor-in-Chief

Honeylet L. Alerta
Deputy Editor-in-Chief

Program Editors

Education and Capacity Building Cluster

Dina S. Ocampo
Lorina Y. Calingasan
Education Research Program

Rosalie B. Arcala Hall
Program on Higher Education Research and Policy Reform

Romylyn Metila
Marlene Ferido
Assessment, Curriculum, and Technology Research Program

Ebinezer R. Florano
Program on Data Science for Public Policy

Social, Political, and Cultural Studies Cluster

Rogelio Alicor L. Panao
Program on Social and Political Change

Darwin J. Absari
Islamic Studies Program

Herman Joseph S. Kraft
Francis Rico C. Domingo
Strategic Studies Program

Marie Aubrey J. Villaceran
Frances Antoinette C. Cruz
Decolonial Studies Program

Editorial Staff

Jheimeel P. Valencia
Bryan Patrick Garcia
Copyeditors

Development Cluster

Annette O. Balaoing-Pelkmans
Program on Escaping the Middle-Income Trap: Chains for Change

Antoinette R. Raquiza
Julius Lustro
Political Economy Program

Eduardo C. Tadem
Maria Dulce Natividad
Program on Alternative Development

Iris Thiele Isip-Tan
Program on Health Systems Development

New Programs

Maria Angeles O. Catelo
Food Security Program

Weena S. Gera
Urban Studies Program

Benjamin M. Vallejo, Jr.
Conservation and Biodiversity

Rosalie B. Arcala Hall
Local and Regional Studies Network

Alexa Samantha R. Hernandez
Editorial Assistant

Mikaela Anna Cheska D. Orlino
Layout Artist

Get your policy papers published. Download open-access articles.

The Philippine Journal of Public Policy: Interdisciplinary Development Perspectives (PJPP), the annual peer-reviewed journal of the UP Center for Integrative and Development Studies (UP CIDS), welcomes submissions in the form of full-length policy-oriented manuscripts, book reviews, essays, and commentaries. The PJPP provides a multidisciplinary forum for examining contemporary social, cultural, economic, and political issues in the Philippines and elsewhere. Submissions are welcome year-round.

For more information, visit cids.up.edu.ph.
All issues/articles of the PJPP can be downloaded for free.

Get news and the latest publications.

Join our mailing list to get our publications delivered straight to your inbox! Also, you'll receive news of upcoming webinars and other updates.

bit.ly/signup_cids

We need your feedback.

Have our publications been useful? Tell us what you think.

bit.ly/dearcids



**UNIVERSITY OF THE PHILIPPINES
CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES**

Lower Ground Floor, Ang Bahay ng Alumni, Magsaysay Avenue
University of the Philippines Diliman, Quezon City 1101

Telephone (02) 8981-8500 loc. 4266 to 4268
(02) 8426-0955

Email cids@up.edu.ph
cidspublications@up.edu.ph

Website cids.up.edu.ph