

---

# *A Literature Review of using Mixed Methods in Development Impact Evaluations*

Danielle Angers

---

*In recent decades, the mixed methods approach has gained traction in many research fields, including impact evaluation. Impact evaluation is used in many sectors, but this paper focuses on its use in international development. This paper explores what is known and unknown about using mixed methods as a possible solution for the many challenges facing impact evaluation. To provide a foundation for this claim and clarify the language surrounding mixed methods and impact evaluations, the paper first defines the terms impact evaluation, qualitative, quantitative, and methodology. It also discusses the benefits and limitations to qualitative, quantitative, and mixed methods approaches as they apply to impact evaluation. The paper then conceptualizes how rigor is applied to both qualitative and quantitative work. Finally, it outlines current best practices and gives recommendations going forward.*

<https://doi.org/10.4079/pp.v28i0.1>

**DANIELLE ANGERS** is a first-year MPP candidate at the Trachtenberg School, with an interest in program measurement and evaluation. She graduated from Virginia Tech in 2016 with a B.A. in Political Science and a minor in Applied Economics. Following graduation, she served in the US Peace Corps as a community economic development volunteer in Burkina Faso. Danielle now works as a research assistant at CNA, a nonprofit research and analysis organization supporting the Navy and Marine Corps.

## INTRODUCTION

While considerable literature on the topics of mixed methods and development impact evaluations exist separately, there is little existing literature on the topic of mixed methods in development impact evaluations. This paper brings together these topics to explain how mixed methods can improve development impact evaluations.

The mixed methods approach integrates at least two distinct methods to collect both qualitative and quantitative data. This approach differs from the more common multiple method approach in which at least two distinct methods collect only one type of data. The idea behind mixed methods is that the complementary strengths and weaknesses of qualitative and quantitative data reinforce one another. This improves the quality of a researcher's findings and increases confidence of attributing those findings to a specific intervention.

As this paper will discuss, the ability to attribute an observed change to the intervention in question is the main objective of a development impact evaluation. The government and non-governmental organizations (NGOs) funding these development programs must ensure the programs generate value equal to their cost.

## HISTORY OF IMPACT EVALUATION IN DEVELOPMENT

Prior to 1996, health science dominated impact evaluations in the development field (Cameron, Mishra, and Brown 2016). Only after 2004 did non-health publications constitute more than 40 percent of published impact evaluations. Specifically, the health, education, agriculture, and rural development sectors saw a significant rise in evaluation work from 2000 to 2012 (Cameron, Mishra, and Brown 2016). This explosion was due to several factors, including the creation of institutions such as the Abdul Latif Jameel Poverty Action Lab (J-PAL), the World Bank's Development Impact Evaluation Initiative (DIME) and Strategic Impact Evaluation Fund (SIEF), and the United Nations' Millennium Development Goals (Cameron, Mishra, and Brown 2016; White 2010). The work of these institutions, as well as the evaluations they sponsor, concentrates on low- to middle-income countries, including those in South and Southeast Asia, East Africa, and Latin America (Cameron, Mishra, and Brown 2016).

An increase in demand for impact evaluations from government agencies, banks, journals, and other organizations stems from an emphasis on measuring a program's ability to affect change in relation to its cost. Previously, there had been a "lack of evidence about what works and what doesn't—and at what cost" (White 2010, 155). Measuring program effectiveness had focused on outputs, which are the products of the program. The measurement of interest then shifted to program outcomes (Bamberger, Rao, and Woolcock 2009). Outcomes are defined as the value added by the program (Parsons, Gokey, and Thornton 2013). For example, the output of an intervention may be the number of yellow fever vaccinations administered in a community. The corresponding outcome would be the percent decrease of yellow fever cases in that community. The institutions sponsoring impact evaluations are looking to see if a program "works" in the sense that it is affecting positive change.

## WHAT IS IMPACT EVALUATION?

The main objective of an impact evaluation is deciding whether a positive change can be attributed to the program in question. The Organization for Economic Co-operation and Development's (OECD) defines impact as “positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended” (*Glossary of Key Terms in Evaluation and Results Based Management* 2002). Following this definition of impact, any evaluation looking at “directly or indirectly, intended or unintended” effects is an impact evaluation. White (2010) argues that while the OECD definition is a valid description of impact, it does not address the crucial point of attribution. One definition White proposes is “the difference in the indicator of interest (Y) with the intervention ( $Y_1$ ) and without the intervention ( $Y_0$ ). That is,  $\text{impact} = Y_1 - Y_0 \dots$  An impact evaluation is a study that tackles the issue of attribution by identifying the counterfactual value” (White 2010, 154). White prefers this definition because it explicitly links the outcomes to the intervention. White’s idea seems to have traction. In 2016, Cameron, Mishra, and Brown defined development impact evaluation as a “counterfactual-based program evaluation that attempts to attribute specific outcomes to programmatic activities by dealing with the problem of selection bias”.

## CHALLENGES UNIQUE TO IMPACT EVALUATION IN DEVELOPMENT

Several factors unique to development impact evaluations do not lend themselves well to conventional research design. First, the programs being evaluated are often complex and multifaceted with different clientele and stakeholders. Different methods are needed to address the different facets of the program. Additionally, given that the results of an impact evaluation are often delivered to clientele, stakeholders, and constituencies with different interests, mixed methods can address the information needs of these varying groups (Bamberger, Rao, and Woolcock 2009; Trochim et al. 2008). For example, the researcher may need to brief findings to both the economist directing a microfinance program and the politician advocating for the program’s government funding. While the economist may prefer to have quantitative data to inform decisions, statistics may not be as meaningful to the politician. The politician may prefer to have qualitative data, which is easier to translate to those who are not statistically literate but whose understanding of the program’s outcomes is essential to its future funding. Having multiple types of data helps target each client’s information needs and communicate the results to different audiences (Bamberger, Rao, and Woolcock 2009).

A second challenge is that strong statistical designs are often neither feasible nor practical in the low- to middle-income countries where most development impact evaluations take place. Often the data needed to construct a baseline do not exist or are very limited (Woolcock 2001). The lack of baseline data makes it nearly impossible to construct a counterfactual or comparison group. The essential question in evaluations assessing attribution is “how can we estimate what would have been the condition of the project population if our project had not taken place?” (Bamberger, Rao, and Woolcock 2009, 25). Absent a counterfactual or control group, there is no way of knowing whether the intervention is having any effect.

A third challenge is adapting to real world constraints. When research design is taught as a concept, it presents the right way to do things. In reality, conditions are rapidly changing and the research design must be able to change with it (Bamberger, Rao, and Woolcock 2009; White 2008). Researchers may not have the optimal time, budget, and support to do the research “the right way.”

Political and organizational pressures can influence who works on the project, its design, how the results are presented, and other important aspects. When it comes to evaluations funded by governments or nongovernmental organizations (NGOs), budgets and deadlines are often tight (Bamberger, Rao, and Woolcock 2009). Government agencies publish evaluations relatively quickly, with an average one-year turnaround from the end of data collection to publication. For comparison, journal articles typically take over six years to publish (Cameron, Mishra, and Brown 2016). In addition, unreliable transportation, lack of infrastructure, cultural dissonance, and security concerns in the field can cause delays to an already tight schedule (Bamberger, Rao, and Woolcock 2009). All three of these challenges require a creative research design that mixed methods can help create.

## HOW MIXED METHODS CAN ADDRESS CHALLENGES TO IMPACT EVALUATION

The definition of a mixed methods study must be established in order to discuss its benefits and limitations. Adapting White's (2008) definition, this paper defines mixed methods as a study intentionally mixing two or more distinct methods collecting both qualitative and quantitative data for the purpose of confirming, refuting, enriching, or merging results. This improves the quality of a researcher's results and increases confidence of attributing those findings to a specific intervention. For this reason, mixing methods can help solve many challenges facing impact evaluation.

First, using multiple methods can solve "problems that stem from studies relying upon a single theory, a single method, a single set of data. . . and from a single investigator" (Mikkelsen 2005, 96). The researcher may miss something they are not specifically looking for when using only one method or collecting one type of data. Mixing methods leaves enough flexibility to identify any factors the researcher did not think to account for but that may affect the program in question. (Rao and Woolcock 2003; Bamberger 2015). Looking at the problem from multiple points of view also helps to correct what Madey (1982) calls the "holistic fallacy" (231). The holistic fallacy is a researcher's tendency to perceive all observations as one part of a whole entity. In reality, life is disjointed, not a neatly divided pie. One piece of the phenomenon being studied may not be relevant to another piece, but using only one method or one type of data may make it appear so.

Second, using mixed methods can address threats to validity common in using only one method. Purely quantitative studies struggle with internal validity. Internal validity establishes causality between intended outcomes and the program's outcomes (Leeuw and Vaessen 2009). While quantitative data can describe the outcomes, qualitative data helps provide a sense of context that can attribute causality to those outcomes and the observed change. Similarly, purely qualitative studies struggle with external validity. External validity allows results of the impact evaluation to be generalized outside of the program in question (Leeuw and Vaessen 2009). Mixing methods provides the quantitative data that helps establish generalizability (Rao and Woolcock 2003).

Third, mixing methods helps neutralize researchers' biases (Madey 1982; White 2008). For example, qualitative interviews can inform the creation of a quantitative survey to make sure the target population's voice is being heard in the survey rather than only what the researcher believes should be included. A mixed approach can also help to explain discrepancies and outliers often

found in quantitative data with no explanation (Bamberger 2015). The researcher is left to come up with their own explanation as to why the discrepancies and outliers are occurring. Qualitative data can be used to investigate. As Rao and Woolcock (2003) put it, “having tea with an outlier can be very effective in understanding *why* they are an outlier” (18).

Finally, mixing methods helps with real-time feedback and replicating results. With mixed methods, researchers understand both the process and the outcomes, helping policymakers both adjust the program being studied in real time and create new successful programs (Rao and Woolcock 2003; Bamberger, Rao, and Woolcock 2009). Having researchers on the ground evaluating the process as it is happening can also save money in the end by allowing for real-time adjustments that can preempt costly mistakes (Madey 1982). It also prevents the need to fund in-depth follow-up studies that would have to retrace past events.

Methods are not inherently qualitative or quantitative, but merely different techniques a researcher can use to collect one or both types of data. Mixing methods requires researchers to have a firm understanding of the benefits and limitations of both qualitative and quantitative data. The next section discusses the use of quantitative and qualitative data in mixed methods.

## USE OF QUANTITATIVE AND QUALITATIVE DATA IN MIXED METHODS

The idea behind mixed methods is that the complementary strengths and weaknesses of qualitative and quantitative data reinforce one another. Qualitative and quantitative are types of data. Qualitative data are associated with observations, descriptions, and quotations. Quantitative data generally take the form of numbers. Until the 1970s, controlled experimental design was the standard for impact evaluation—with an exception for quasi-experimental designs when necessary. Madey (1982) claims “qualitative designs were shunned as too imprecise to produce either information of value to decision-makers or information to establish program accountability” (224). In the 1980s, researchers’ dedication to “the design and implementation of evaluations that will provide adequate information for policymakers’ needs” opened the door to collecting both quantitative and qualitative data in impact evaluation design (Madey 1982, 225).

Quantitative data is generally valued as providing rigorous, impartial, and objective information (Rao and Woolcock 2003). Although not immune to external validity problems, quantitative data are more generalizable to populations outside of the study population, thus making the study results useful beyond the scope of the research. While quantitative data are good for broad policy recommendations and generalizations, they are not as helpful for providing specific, contextual recommendations to local politicians and policymakers (Rao and Woolcock 2003). Qualitative data are deeply contextual and often subject to the researcher’s interpretation and observation. This specificity is part of qualitative data’s strength, but is also a weakness. It is impossible to completely separate qualitative data from their environment or the researcher’s interpretations, making it difficult to generalize outside the population being studied (White 2008).

Quantitative data are useful for analysis of large samples, as quantitative analysis is less resource- and labor-intensive than qualitative analysis. Thus, quantitative data lends itself well to evaluations with strict budget and time restrictions, as is often the case. However, analysis of large datasets places emphasis on the big picture and averages, which homogenizes results. Homogenization overlooks outliers, marginal groups, and special cases that are often of interest in social science research. These populations are difficult to access by outsiders. The best way to reach these communities is through on-the-ground, hands-on research to collect the target

population's perspectives (Rao and Woolcock 2003). These samples are typically small and specific, making collecting in-depth qualitative data both resource- and time-intensive.

Quantitative data is useful for summarizing outputs, but less effective for describing processes. With limited resources to put toward interventions, program funders and implementers want to know if the program design and implementation are effective (Rao and Woolcock 2003). Funders often do not have the option, time- or resource-wise, to scrap a program and start over if it is not effective. The program implementers need to be able to pinpoint specific parts of the program design that are not working so they can make adjustments (Jimenez et al. 2018). Qualitative data better describes strengths and weaknesses in process and implementation (Bamberger 2015).

A large part of determining the effectiveness of a program is attributing causality. Qualitative data does not tell us much about causality. One can observe what is taking place and a population's attitudes about it, but that does not tell us *why* that specific event is taking place (Rao and Woolcock 2003). Quantitative data can help with the problem of attribution. Attribution is not dichotomous. In other words, many factors contribute to a measured change, and the intervention can lead to a portion of that change—it is not required to be the sole attributor to be an attributor. White (2010) argues that a good study should be able to put a number on what percentage of change is attributed to the intervention being studied. A quantifiable attribution can help policymakers assess whether an intervention is cost effective or whether the intervention would have to be more productive in order to remain operational (Trochim et al. 2008). It also relieves the program from having to be 100 percent successful. If the intervention is not the only reason for change, it may be a significant reason for change, which is a lighter burden to bear.

Although qualitative and quantitative data can balance out one another, the two remain heavily polarized in the research world for a number of reasons. First, researchers are generally recruited, trained, and socialized in one discipline (Rao and Woolcock 2003). This leads them to be partial to their discipline and have trouble effectively employing the other. Splitting attention and resources between two disciplines runs the risk of doing both poorly (Rao and Woolcock 2003). A mixed methods team needs both the expertise and resources to execute a complicated design successfully, which is made more difficult due to the lack of evidence on the best way to combine methods under different circumstances (Rao and Woolcock 2003). As a relatively new concept, the kinks are still being worked out. The industry has yet to agree on a set of best practices for mixed method studies, which is discussed in the next section.

## ESTABLISHING BEST PRACTICES

In order for mixed methods to be more widely employed by researchers, the industry must come to a consensus on best practices. The failure to do so up to this point is partly due to the relatively short amount of time mixed methods has been recognized by the mainstream and partly due to the polarization of the qualitative and quantitative camps. The next three subsections cover what exists in the literature regarding rigor as it applies to qualitative and quantitative approaches, the various frameworks for mixed methods, and how the quality of a mixed methods evaluation is assessed.

## *RIGOR*

In their research on the concept of rigor, Davies and Dodd (2002) could find no explicitly agreed upon definition of the term. They could only conclude that rigor is associated with “detachment, objectivity, replication, reliability, validity, exactitude, measurability, containment, standardization, and rule” (Davies and Dodd 2002, 280). The connotations of these words create an inherent partiality to quantitative data. However, this is only one way to think about rigor. As “the authoritative evaluation of good research and the unspoken standard by which all research is measured,” rigor must be re-conceptualized in a way that can assess the validity of both qualitative and quantitative data (Davies and Dodd 2002, 280).

Perceived lack of rigor is one of the reasons researchers historically have had a problem incorporating qualitative data into studies. However, the perception that quantitative data always provides a stronger foundation on which to base policy recommendations may be misplaced (White 2002). Objectivity and truth are not synonymous. Objectivity often requires the researcher to distance themselves from the data, which can obscure their view on the process taking place and cause them to miss key issues (Davies and Dodd 2002). The objectivity associated with quantitative data gives a false sense of security that the results will be valid and true. Just as qualitative data requires interpretation, quantitative data can be manipulated to back up a specific claim (White 2002; White 2008). As R.H. Coase (1994) puts it, “if you torture the data enough, they will confess” (27). Replication and subjectivity are also mistaken as a solely qualitative problem. The casual observer might assume that the researcher chose to employ the proper mathematical techniques, which is not always the case. Even properly chosen techniques can yield inaccurate results when used improperly.

Both qualitative and quantitative studies collect observations from the world and present them in a framework. In this sense, research results are not necessarily impartial or objective. The important part of the process is the relationship between the data and the researcher (Davies and Dodd 2002). Rigor should be defined as a sense of responsibility and accountability of the researcher to present the findings as candidly as possible. The true measure of rigor is the researcher’s ability to apply proper technique and the integrity to not distort the data (Davies and Dodd 2002). This requires the researcher to think reflexively and ethically about their actions, be open to making adjustments as the study unfolds, and clearly outline the process in their write-ups to provide complete transparency regardless of how chaotic it becomes (Davies and Dodd 2002). If rigor is thought of in terms of process versus results, it becomes relevant to both quantitative and qualitative data. One step to thinking critically about the study process and being able to clearly explain it to the reader is creating a strong conceptual framework.

## *CONCEPTUAL FRAMEWORK FOR MIXED METHODS IN IMPACT EVALUATION*

Four imperative questions a researcher must answer in designing a study are: (1) What is the research question? (2) What is the type of inquiry (experimental versus natural)? (3) Which methods should be used to analyze the data? (4) What kind of data will be collected? (Lynch 1983). The order of the questions matters. White (2010) argues impact evaluations “should be issues-led not methods-led...having determined the evaluation questions, the best available method should then be used to answer them” (162). After identifying the question, the researcher must decide which methods to choose, articulate why they chose to mix methods, and determine how to integrate methods in the research design (Trochim et al. 2008). A strong conceptual framework improves the rigor of research and makes findings more meaningful and valid (Adom, Hussein,



and Agyem 2018; Trochim et al. 2008). There is more to an impact evaluation's framework than its purpose of mixing methods and points of integration, but these are essential starting points.

Greene, Caracelli, and Graham (1989) broke down the purpose of mixing methods into four categories. Many of these purposes target reasons previously discussed for why this approach is beneficial to impact evaluations, including strengthening validity, finding contradictions, increasing scope, and studying multiple facets of complex questions:

1. Complementary: multiple methods look at different facets of the inquiry in which the researcher hopes to obtain similar results to reinforce one another .
2. Development: one method used to inform the development or results of another. An example would be using qualitative data collected in focus groups to inform the creation of a quantitative survey.
3. Initiation: increases the scope of the results by looking for contradictions and discrepancies. For example, qualitative interviews can be conducted to flesh out a contradiction or discrepancy discovered after analyzing quantitative data.
4. Expansion: increases the scope of the inquiry by using different methods to study multiple components of the question. For example, the researcher can use a method gathering qualitative data to assess a program's implementation process and a method gathering quantitative data to assess the program's outputs.

After the researcher has decided on the purpose of mixing methods, they must decide to what extent the methods mix. Greene, Caracelli, and Graham define seven characteristics of mixed method studies:

1. Methods: the degree to which methods "are similar to or different from one another in form, assumptions, strengths, and limitations or biases".
2. Phenomena: describes whether the different methods assess the same or different facets of the problem.
3. Paradigm: the methods collecting either or both types of data can be used under the same or different paradigms.
4. Status: qualitative and quantitative data can be valued equally in a study or one can take precedence over the other.
5. Implementation independence: the degree to which methods are interactive or independent of one another.
6. Implementation timing: if methods are concurrent or sequential ). Greene et al. sees timing as a dichotomy, but it can also be categorized as iterative.
7. Categorical designation: the research either is composed of one study or more than one study.

Thinking critically about the purpose and the extent of mixing methods will give the researcher a solid foundation on which to begin the evaluation. After the researcher has thought critically about the question, purpose, and integration of his or her study and executed the design with integrity and focus, there must be a way for a third party to assess the study's quality.

### *ASSESSMENT OF MIXED METHODS STUDIES*

In an attempt to provide structure and consistency to the assessment of mixed method studies, Jimenez et al. (2018) developed a preliminary assessment tool. A copy of their tool can be

found in Appendix A, but what stand out are the characteristics Jimenez et al. find common in mixed method studies.

First, the studies that perform well begin with an interdisciplinary study team. Diversity provides expertise from many disciplines but only works when the members share a conceptual framework (Jimenez et al. 2018; Trochim et al. 2008). Sharing a framework helps minimize coordination issues discussed in the section describing benefits and limitations to mixed methods. The conceptual framework should also be clearly presented in the write-up (Trochim et al. 2008). The framework informs every decision made by the researcher, so the researcher should inform the reader of this foundation early in the paper (Adom, Hussein, and Agyem 2018). Along with presenting the framework clearly, the researcher should spell out the reason for using mixed methods (Jimenez et al. 2018).

Second, quality studies are highly integrated. These studies take full advantage of the benefits of mixed methods (Jimenez et al. 2018). Studies that lean toward multiple methods rather than mixed methods have the veneer of added benefit without actually providing any. Multiple method studies are those that employ more than one method but may only collect one type of data, and the data they collect are likely gathered and analyzed independently from one another. The more integrated the methods and analyses are, the more a study realizes the benefits of mixed methods. The studies were also fully transparent in their explanations of data collection and analysis as well as any supplemental materials.

Finally, high-quality studies make an earnest attempt to acknowledge limitations and discrepancies (Jimenez et al. 2018). This acknowledgment not only allows the reader to see where future research on the topic is needed, but also helps policymakers transfer the findings to new applications. By recognizing how the study environment and design affect the study, the results, and to what extent the results are tied to their context, readers can decide which pieces of the study are useful in a new setting.

None of these characteristics are surprising. All the concepts Jimenez et al. identifies as essential for quality mixed methods studies, and therefore impact evaluations using mixed methods, have been discussed in this paper. Moving forward, the industry needs to normalize and expand upon these concepts.

## CONCLUSION AND RECOMMENDATIONS GOING FORWARD

The literature up until this point has shown that impact evaluations have unique opportunities to benefit from mixed methods. A lot of work has been done so far to establish what these might be and some best practices to achieve them, but there is still a long way to go. Going forward, it is important to first establish a consensus on the definition of impact evaluations, how one thinks of qualitative versus quantitative data, and other terms relevant to mixed methods impact evaluation (Jimenez et al. 2018). The more respect mixed methods has, the more likely that governments and other funding agencies will devote the resources necessary to producing a quality study. Establishing a standard for quality, rigorous mixed methods studies will help gain the required esteem (Jimenez et al. 2018; Trochim et al. 2008). Further work needs to be done on how to best mix methods under different conditions and develop innovations that can be used to strengthen results (Jimenez et al. 2018). Researchers must ensure they are knowledgeable on the context of the study, are physically in the field often, and are intentionally building both quantitative and qualitative data into the design (White 2008; Trochim et al. 2008). Lastly, it would

be beneficial to have clearer guidelines on reporting and a more comprehensive repository of evaluations that successfully integrate methods from which researchers can model their own studies (Jimenez et al. 2018; Trochim et al. 2008). The full benefits to be gained from using mixed methods in impact evaluations have not yet been seen.

## APPENDIX A: MIXED METHODS IMPACT EVALUATION APPRAISAL TOOL

Section	Types of questions	Scoring
Preliminary information	Coder name Reference to IE report Sector focus Donor Classification of impact evaluation design (RCT/Quasi/Non-experiment)	N/A
Section A: Quantitative Aspect	A1. Clear description of the study's methods and procedures A2. Rigour of the quantitative study in conduct A3. Selection bias and confounding (if an RCT) A4. Selection bias and confounding (if a natural experiment/quasi experiment/non-experiment) A5. Post-intervention biases (motivation of participants) A6. Post-intervention biases (analysis and reporting) A7. Threats to construct and external validity A8. Reportage of statistical power	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> Not Applicable
Section B: Qualitative Aspect	B1. Clear description of the study's methods and procedures B2. Rigour of the qualitative study in conduct B3. Reportage of assumptions, values, biases B4. Attempts to address biases B5. Transferability of results B6. (Unscored) Data situated within political, institutional, cultural or social context	<input type="radio"/> Yes <input type="radio"/> No
Section C: Integration of Mixed Methods	C1. Integration theory of change/programme or logic model explored through mixed methods C2. Integration of methods to inform study design C3. Integration of methods to inform the interpretation of findings C4. Limitations of integration C5. (Unscored) Stage(s) at which qualitative evidence is incorporated into the study. C6. (Unscored) Categorisation of the type of mixed methods study into sequential exploratory, sequential explanatory and/or convergent design.	<input type="radio"/> Yes <input type="radio"/> No

Source: Jimenez, Emmanuel, Hugh Waddington, Neeta Goel, Audrey Prost, Andrew Pullin, Howard White, Shaon Lahiri, and Anmol Narainwith. 2018. "Mixing and Matching: Using Qualitative Methods to Improve Quantitative Impact Evaluations (Ies) and Systematic Reviews (Srs) of Development Outcomes." *Journal of Development Effectiveness* 10, no. 4: 400-21. DOI: 10.1080/19439342.2018.1534875.

## REFERENCES

- Adom, Dickson, Emad K. Hussein, and Joe A. Agyem. 2018. "Theoretical and Conceptual Framework: Mandatory Ingredients of a Quality Research." *International Journal of Scientific Research* 7, no. 1: 438-41. [https://www.worldwidejournals.com/international-journal-of-scientific-research-\(IJSR\)/fileview.php?val=January\\_2018\\_1514812002\\_\\_202.pdf](https://www.worldwidejournals.com/international-journal-of-scientific-research-(IJSR)/fileview.php?val=January_2018_1514812002__202.pdf).
- Bamberger, Michael. 2015. "Innovations in the Use of Mixed Methods in Real-World Evaluation." *Journal of Development Effectiveness* 7, no. 3: 317-26. <https://doi.org/10.1080/19439342.2015.1068832>.
- Bamberger, Michael, Vijayendra Rao, and Michael Woolcock. 2009. *Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development*. The World Bank. 1-27.
- Cameron, Drew B., Anjini Mishra, and Annette N. Brown. 2016. "The Growth of Impact Evaluation for International Development: How Much Have We Learned?". *Journal of Development Effectiveness* 8, no. 1: 1-21. DOI: 10.1080/19439342.2015.1034156.
- Coase, R.H. 1994. *Essays on Economics and Economists*. Chicago: University of Chicago Press.
- Davies, Deirdre, and Jennifer Dodd. 2002. "Qualitative Research and the Question of Rigor." *Qualitative Health Research* 12, no. 2: 279-89. <https://doi.org/10.1177/104973202129119793>.
- Evaluation, Development Assistance Committee (DAC) Working Party on Aid. 2002. Glossary of Key Terms in Evaluation and Results Based Management. Organisation for Economic Co-operation and Development. <https://www.oecd.org/dac/evaluation/2754804.pdf>.
- Greene, Jennifer C., Valerie J. Caracelli, and Wendy F. Graham. 1989. "Toward a Conceptual Framework for Mixed-Method Evaluation Designs." *Educational Evaluation and Policy Analysis* 11, no. 3: 255-74. <https://www.jstor.org/stable/pdf/1163620.pdf?refreqid=excelsior%3A70d4860f9383090894c2ea39548d6907>.
- Jimenez, Emmanuel, Hugh Waddington, Neeta Goel, Audrey Prost, Andrew Pullin, Howard White, Shaon Lahiri, and Anmol Narainwith. 2018. "Mixing and Matching: Using Qualitative Methods to Improve Quantitative Impact Evaluations (Ies) and Systematic Reviews (Srs) of Development Outcomes." *Journal of Development Effectiveness* 10, no. 4: 400-21. <https://doi.org/10.1080/19439342.2018.1534875>. DOI: 10.1080/19439342.2018.1534875.
- Leeuw, Frans, and Jos Vaessen. 2009. *Impact Evaluations and Development*. The Network of Networks on Impact Evaluation, Washington, DC. <https://gsdrc.org/document-library/impact-evaluations-and-development-nonie-guidance-on-impact-evaluation/>.
- Lynch, Kathleen B. 1983. "Qualitative and Quantitative Evaluation: Two Terms in Search of Meaning." *American Educational Research Association* 5, no. 4: 461-64. <https://www.jstor.org/stable/1164050>.

- Madey, Doren L. 1982. "Some Benefits of Integrating Qualitative and Quantitative Methods in Program Evaluation with Illustrations." *Educational Evaluation and Policy Analysis* 4, no. 2: 223-36. <https://www.jstor.org/stable/1164015>.
- Mikkelsen, Britha. 2005. *Methods for Development Work and Research: A Guide for New Practitioners*. Second ed. Thousand Oaks: Sage Publications.
- Parsons, Jim, Caitlin Gokey, and Monica Thornton. 2013. *Indicators of Inputs, Activities, Outputs, Outcomes and Impacts in Security and Justice Programming*. UKAID Department for International Development. 1-29.
- Rao, Vijayendra, and Michael Woolcock. 2003. "Chapter 8: Integrating Qualitative and Quantitative Approaches in Program Evaluation." In *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*, edited by Francois Bourguignon and Luiz A. Pereira da Silva, 165-90. Washington, DC: Oxford University Press.
- Stenner, A. J. 1974. *An Overview of Information Based Evaluation: A Design Procedure*. Information Based Evaluation Series, Book 1. Durham, NC: IBEX.
- Trochim, William, Stephen Marcus, Louise C. Masse, Richard P. Moser, and Patrick C. Weld. 2008. "The Evaluation of Large Research Initiatives: A Participatory Integrative Mixed-Methods Approach." *American Journal of Evaluation* 29, no. 1: 8-28. <https://doi.org/10.1177/1098214007309280>.
- White, Howard. 2002. "Combining Quantitative and Qualitative Approaches in Poverty Analysis." *World Development* 30, no. 3: 511-22. [https://doi.org/10.1016/S0305-750X\(01\)00114-0](https://doi.org/10.1016/S0305-750X(01)00114-0).
- White, Howard. 2010. "A Contribution to Current Debates in Impact Evaluation." *Evaluation* 16, no. 3: 153-64. <https://doi.org/10.1177/1356389010361562>.
- White, Howard. 2008. "Of Probits and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation." *Network of Networks on Impact Evaluation*: 3-18. DOI: 10.1111/j.1759-5436.2008.tb00436.x.
- Woolcock, Michael. 2001. *Social Assessments and Program Evaluation with Limited Formal Data: Thinking Quantitatively, Acting Qualitatively*. World Bank Social Development Note No. 68.