

## Evaluation and Civil Society

LEHN M. BENJAMIN and DANA R.H. DOAN  
Indiana University-Purdue University, Indianapolis USA

### Key Words

- Evaluation
- Performance measurement
- Performance management
- Effectiveness
- Civil society
- Nonprofit
- Social impact
- Outcome measurement

### Definition

Evaluating civil society organizations remains a pressing issue in the sector. Civil society leaders today face a dizzying array of models and approaches but evaluating the outcomes and impacts of these organizations remains challenging. Many of the changes civil society organizations (CSOs) pursue take a long time to address and are difficult to measure. Engaging in an evaluation requires making key decisions about the criteria, standards, and evidence that will be used to judge the merit or worth of the program, organization, or network. With multiple stakeholders, each with their own ideas about what may constitute successful change, determining who makes these decisions affects what gets measured, which affects how CSOs do their work. Despite these challenges, scholars and practitioners continue to work to develop more applicable and innovative approaches to assessing the contributions of civil society organizations. After giving a brief introduction and historical background to evaluation, this essay looks at some key issues in evaluating civil society organizations and concludes with recent developments on this important topic.

### Introduction

Evaluation is a critical issue for civil society organizations (e.g., nongovernmental, nonprofit, private voluntary organizations). Policy makers, funders, activists, community members as well as civil society leaders themselves want to know whether these organizations are having an impact on the problems they seek to address. These concerns are not new and since the 1990s, we have seen an evolution of evaluative approaches in response to these concerns: from using logic models and theories of change to using social return on investment analysis and social impact bonds. What has remained consistent during this time is the demand for evidence about CSOs' effectiveness, a demand that undoubtedly will continue as long as CSOs remain important partners for both governments and philanthropic donors in addressing social problems.

But evaluating the impact of civil society organizations is not easy. Many of the changes CSOs seek take a long time to address and are difficult to measure. At the same time, CSOs' contribution to addressing these problems is partial and difficult to isolate given the variety of environmental factors that can also affect the desired outcome. For those CSOs who advocate for policy change, provide opportunities for civic engagement or organize communities, determining outcomes is even less tangible and more difficult to evaluate. Despite these challenges, those who study the field are working to develop more applicable and innovative approaches to assessing the contributions of civil society organizations to solving social

---

This is the author's manuscript of the book chapter published in final edited form as:

Benjamin, L. M., & Doan, D. R. H. (2020). Evaluation and Civil Society. In R. A. List, H. K. Anheier, & S. Toepler (Eds.), *International Encyclopedia of Civil Society* (pp. 1–7). Springer International Publishing. [https://doi.org/10.1007/978-3-319-99675-2\\_59-1](https://doi.org/10.1007/978-3-319-99675-2_59-1)

problems (e.g., Ebrahim 2019; Gugerty and Karlan 2018). After giving a brief introduction to evaluation, this essay looks at some key issues in evaluating civil society organizations and concludes with recent developments on this important topic.

## **Historical Background**

Patton (1997) defines evaluation as the systematic collection of information to make judgments, improve effectiveness, and/or inform decision making about a program, initiative, or organization. The emphasis on the systematic collection of information sets evaluation apart from the informal evaluative judgments we make every day (Benjamin and Campbell 2020: 199). The emphasis on making judgments about the merit or worth of something requires making explicit the values guiding such judgements and is what distinguishes evaluation from basic research (Shadish, Cook and Leviton 1991). Improving effectiveness and informing decision making tells us that these judgments are made in service to a larger goal: to address some issue, improve performance, and to ensure everyone is able to thrive.

Although evaluation practice has been around a long time, the professional field of evaluation emerged in the 1960s during the War on Poverty (Scriven 1996:365). In those early years, evaluators focused on developing methodological approaches that could determine whether these new programs caused observable outcomes, using experimental and quasi-experimental designs, with the hope of informing future policy making (Weiss 1998). Over the years, those contributing to the field have developed theories pertaining to the core practice of evaluation: theories about knowledge construction, social interventions, values and valuing, as well theories about use and influence (Shadish Cook and Leviton 1991). Today the field includes many approaches and models that address one or more of these theories. For example, theory-driven evaluation emerged to help stakeholders address the core assumptions underlying their program and its intended outcomes whereas utilization evaluation is premised on the idea that evaluation should be judged on its usefulness to intended users (Patton 1997).

Regardless of approach or epistemological orientation evaluation requires making certain key decisions: What criteria will be used to judge the merit or worth of the program, organization, or network (e.g., goal achievement, quality of implementation, satisfaction of beneficiaries, or financial strength)? What standards will be applied to these criteria (e.g., X percentage of those trained get full-time jobs, a majority of the residents believe the organization is effective in addressing blight)? And what evidence will be needed to understand whether such standards have been met (e.g., interviews or survey of trainees one or five years after they complete the program)? Of course, many other questions need to be addressed along the way, including: Who determines these criteria, standards, and evidence? Is there agreement among stakeholders over what constitutes success? If not, how is this resolved? Can sufficient evidence be gathered to understand the merit or worth of the program? Are the costs of collecting data feasible? What is the purpose of the evaluation and how will the results be used?

## **Key Issues**

Early writings on evaluation and civil society focused on identifying the distinguishing characteristics of these organizations, compared to public agencies and for-profit firms. This work then considered the implications of these characteristics for evaluating CSO performance (e.g., Kanter & Summers 1987). Later researchers drew on organizational sociology to consider three approaches to assessing the effectiveness of CSOs: goal attainment, resource acquisition, and multiple constituencies views of the organization (e.g., Leczy et al. 2012). By the late 1990s and early 2000s researchers shifted their attention to how CSOs were responding to increasing demands for accountability and impact. At the time,

policymakers and funders were growing increasingly concerned about the effectiveness of CSOs, their limited capacity, and lack of accountability to communities. Consequently, they started to demand evidence of measurable outcomes (Edwards and Hulme 1996). Not surprisingly, researchers turned their attention to examining how CSOs were responding to these demands, including examining the prevalence of evaluation practices among CSOs and the challenges they faced in trying to measure the difference they were making in the lives of those they served (e.g. Benjamin 2008; Ebrahim 2003).

Given limited space, we organize this literature under three broad but inter-related questions<sup>1</sup>: 1) Are CSOs evaluating what they do? 2) What challenges do CSOs face as they try to evaluate their work? and 3) How is evaluation influencing CSOs? Before addressing these questions, it is worth distinguishing between evaluation and three closely related concepts: accountability, performance measurement, and effectiveness. In brief, accountability is one possible purpose for conducting an evaluation but by no means the only purpose. Other evaluation purposes include improving programs or processes and generating knowledge. Performance measurement involves the ongoing collection of evaluative data to inform organizational decision making and is one approach to evaluation. As noted above, evaluation includes a broad range of approaches. Effectiveness is one important criterion used to evaluate a program or an organization and is often defined as some level of goal achievement, but other criteria include efficiency, beneficiary satisfaction, responsiveness, and so on.

#### *Are CSOs evaluating what they do?*

The short answer is yes, if we take the results of several surveys in different parts of the world as an indicator. These surveys have, among other things, sought to understand the prevalence of different kinds of evaluative practices among CSOs, the types of data CSOs collect and who in the organization is responsible for collecting this data and carrying out the evaluation (e.g., Campos et al. 2011; Carman 2007; Hall et al. 2003; Marshall & Suarez 2014). All of these studies suggest that the majority of CSOs, greater than 50% of those surveyed, evaluate their work. Results also show that nonprofits use a variety of evaluation methods, including surveys, focus groups, interviews and administrative data. But the results also suggest a bifurcation in evaluation capacity among CSOs. For smaller CSOs, i.e., those not able to hire an evaluator, research suggests the quality of their evaluative efforts is low as they rely on program staff, with little time and limited expertise in evaluation, to design instruments and collect data or draw on existing administrative data (Hofer 2000). Meanwhile, larger CSOs, with the ability to hire an internal or external evaluator, use sophisticated evaluation designs to assess impact (Campos et al. 2011).

However, using these studies as an indicator of whether CSOs are evaluating their work may be limited. Despite their usefulness, all the studies are one-time studies and employ unique survey instruments. Thus, we cannot compare across context or time. Moreover, the reported prevalence of evaluation practice among CSOs surveyed may be over-stated, as CSOs with limited capacity are less likely to complete such surveys. We also cannot be sure how evaluation terminology translates into different cultural contexts (Dahler-Larsen et al. 2017). All of these issues make it difficult at this time to have a firm grasp of the extent to which CSOs are evaluating their work. Such an assessment would require cross-country collaborations to develop simple instruments, that have common culturally appropriate questions, and that could be used consistently over time.

---

<sup>1</sup> This summary is adapted from an expanded review. See Benjamin, Lehn M. & David A. Campbell. 2020. Evaluation and Performance Measurement. In Helmut Anheier and Stefan Toepler (Eds.), *Routledge Companion to Nonprofit Management*.

### *What challenges do CSOs face as they try to evaluate their work?*

This question has received extensive attention by CSO scholars. This research suggests that CSOs face at least four challenges in evaluating their work. First, as noted above, CSOs have limited capacity to evaluate their work effectively. Consequently, the data they collect is low quality, often relying on internally developed survey instruments that have not been tested (e.g., Hoefer 2000). Second, the data they do collect is often in response to funders' requests (e.g., Ebrahim 2003). Consequently, CSOs use the data primarily to report rather than improve their own practices. Moreover, the reporting relationship and the desire to show positive results to funders discourages collecting and using data in ways that would encourage organizational learning (Ebrahim 2003). And for CSOs reporting to multiple funders, all of which require distinct measures, this data collection and reporting can sap organizational time that might be used for other types of data-driven decision making (e.g., Benjamin, Volda and Bopp 2018).

Third, the recent interest by public and private funders in using experimental studies to evaluate CSO initiatives—whether through impact measurement, the use of evidence-based practice, or social impact bonds—has created isomorphic pressure for CSOs to follow suit, regardless of the kinds of programs and services provided, and often with disappointing results (see Silverstein & Maher 2008). Fourth, scholars have also suggested that existing program evaluation models—which assume an instrumental project logic—do not adequately reflect the emergent nature of outcomes, where outcomes are defined as nonprofits' work in partnership with the communities they serve (e.g., Benjamin 2008).

### *How is evaluation influencing CSOs?*

Mirroring some of the themes above, the research suggests that evaluation practice by civil society organizations has both intended and unintended consequences. The intended consequences mirror the two dominant evaluative purposes: using evaluation to inform practice and improve performance and using evaluation to demonstrate accountability to important stakeholders. As already noted, many CSOs use evaluation to demonstrate accountability, primarily to funders (Lee & Clerkin 2017). However, CSOs have also taken steps to ensure they are accountable to beneficiaries, including using feedback from beneficiaries to make actionable changes (see Twersky et al. 2013). Studies also find that CSOs use evaluation to inform practice and improve performance. For example, CSOs use data to make decisions about programs and services as well as to support organizational learning more broadly, particularly when they are able to link data collection and evaluation with strategic planning (e.g., Lee & Clerkin 2017). Still, reports suggest that CSOs underutilize the data they collect. One report noted that although 75% of nonprofits collect data, only six percent report that they feel they are using data effectively (Janus 2018).

In terms of unintended consequences, a number of studies identify such consequences, including inhibiting organizational learning, undermining participation by clients, underinvestment in relationship building with individuals and communities, to name a few. These consequences seem to be the result of the failure to fully account for how social change happens in these organizations and the dominant emphasis on using evaluation for accountability (e.g., Benjamin 2008; Ebrahim 2003). With that said, more recent developments, both inside and outside the evaluation field, offer promising approaches to address some of these concerns including: developmental evaluation, which assumes the identification of outcomes is emergent (Patton 2010); evaluation approaches that use systems theory to recognize complex and contingent change (e.g., Williams & Hummelbrunner

2011); and evaluation approaches that better account for systemic causes of social problems (Thomas et al. 2018).

### **Future Directions**

As the effort to evaluate the impact of CSOs continues, four recent developments, some of which we mentioned already, are worth noting. First, we see a growing emphasis on two types of evidence: evidence from randomized controlled trials (RCTs) and evidence that pairs social outcomes with financial and sometimes environmental outcomes. The trend towards RCTs is evident in large foundations and government aid agencies that fund international development. It is also evident in the adoption of evidence-based policy by governments that partner with human service organizations applying evidence-based practices proven with accumulated studies (e.g., Gugerty and Karlan 2018). Pairing social outcomes with financial outcomes is evident among private donors, particularly those coming from the corporate world, and some foundations. Interest in measuring impact through double and triple bottom line metrics, although not new, seems to appeal to these donors because they appear to offer a level of certainty and some control in achieving measurable impact. But both the elevation of randomized controlled trials and the coupling of social outcomes with financial outcomes has generated significant debate (e.g., Cooney & Lynch-Cerullo 2014).

Second, we see continued concern with issues of race, equity, and social justice in CSO evaluative efforts. Evaluation scholars have given significant attention to issues of equity and justice in evaluation (e.g., Donaldson & Picciotto 2016). Yet, in practice, evaluations can easily reinforce inequities by failing to involve all stakeholders, question dominant assumptions, or make value claims explicit (e.g., Sturges 2015; Thomas et al. 2018). In response a number of principles and approaches have been introduced to ensure that issues of power, privilege, and systemic inequity, are acknowledged and addressed in the process as well as the context of an evaluation (Dean-Coffey et al. 2014). Third, we see greater emphasis on collective and network-based approaches to evaluating impact, recognizing that nonprofits alone can rarely address community level problems. This requires evaluative approaches that mirror these efforts. Finally, with social media and advanced analytical technologies, the possibilities of big data to both increase effectiveness and advance understanding of effectiveness is getting attention in the CSO sector (Stiffman & Tumolillo 2019).

### **Cross-References**

- Accountability
- Effectiveness and efficiency of nonprofit organizations
- Impact bonds
- Multiple bottom lines
- Performance and impact
- Social accounting for social economy organizations
- Social audits
- Stakeholders
- Values

### **References**

- Benjamin, L. M. (2008). Bearing More Risk for Results: Performance Accountability and Nonprofit Relational Work. *Administration & Society*, 39(8), 959–983.
- Benjamin, L. M., & Campbell, D. A. (2020). Evaluation and performance measurement. In H. K. Anheier & S. Toepler (Eds.), *The Routledge Companion to Nonprofit*

- Management* (1st ed., p. Chapter 14). Routledge.
- Benjamin, L. M., Volda, A., & Bopp, C. (2018). Policy fields, data systems, and the performance of nonprofit human service organizations. *Human Service Organizations: Management, Leadership & Governance*, 42(2), 185–204.
- Campos, L., Andion, C., Serva, M., Rossetto, A., & Assumpção, J. (2011). Performance Evaluation in Non-Governmental Organizations (NGOs): An Analysis of Evaluation Models and their Applications in Brazil. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 22(2), 238–258.
- Carman, J. G. (2007). Evaluation Practice among Community-Based Organizations: Research into the Reality. *American Journal of Evaluation*, 28(1), 60–75.
- Cooney, K., & Lynch-Cerullo, K. (2014). Measuring the Social Returns of Nonprofits and Social Enterprises: The Promise and Perils of the SROI in: Nonprofit Policy Forum Volume 5 Issue 2 (2014). *Nonprofit Policy Forum*, 5(2), 367–393.
- Dahler-Larsen, P., Abma, T., Bustelo, M., Irimia, R., Kosunen, S., Kravchuk, I., Minina, E., Segerholm, C., Shiroma, E., Stame, N., & Tshali, C. K. (2017). Evaluation, Language, and Untranslatables. *American Journal of Evaluation*, 38(1), 114–125.
- Dean-Coffey, J., Casey, J., & Caldwell, L. D. (2014). Raising the Bar - Integrating Cultural Competence and Equity: Equitable Evaluation. *Foundation Review*, 6(2), 80–94.
- Donaldson, S.L. & Picciotto, R. 2016. *Evaluation for an Equitable Society*. Charlotte, NC: Information Age Publishers.
- Ebrahim, A. (2003). *NGOs and Organizational Change: Discourse, Reporting, and Learning*. Cambridge University Press.
- Ebrahim, A. (2019). *Evaluating Social Change*. Stanford University Press.
- Edwards, M., & Hulme, D. (Eds.). (1996). *Beyond the Magic Bullet NGO Performance and Accountability in the Post Cold War World*. A Kumarian Press Book.
- Gugerty, M. K., & Karlan, D. (2018). *The Goldilocks Challenge: Right-Fit Evidence for the Social Sector*. Oxford University Press.
- Hall, M. H., Phillips, S. D., Meillat, C., & Pickering, D. (2003). Assessing performance: Evaluation practices & perspectives in Canada's voluntary sector. Sector Source. Retrieved July 4, 2020, from <http://sectorsource.ca/resource/file/assessing-performance-evaluation-practices-perspectives-canadas-voluntary-sector>
- Hoefler, R. (2000). Accountability in Action? Program evaluation in nonprofit human service agencies. *Nonprofit Management & Leadership*, 11(2), 167.
- Janus, K. K. (2018). Creating a Data Culture: How nonprofit organizations can do a better job with their data. *Stanford Social Innovation Review* (March 2, 2018).
- Kanter, R. M., & Summers, D. V. (1987). Doing well while doing good: Dilemmas of performance measurement in nonprofit organizations and the need for a multiple-constituency approach. In W.W. Powell (Ed.) *The nonprofit sector: A research handbook* (1st ed.). Yale University Press.
- Lecy, J. D., Schmitz, H. P., & Swedlund, H. (2012). Non-Governmental and Not-for-Profit Organizational Effectiveness: A Modern Synthesis. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organization*, 23(2), 434–457.
- Lee, C., & Clerkin, R. M. (2017). Exploring the Use of Outcome Measures in Human Service Nonprofits: Combining Agency, Institutional, and Organizational Capacity Perspectives. *Public Performance & Management Review*, 40(3), 601–624.
- Marshall, J. H., & Suárez, D. (2014). The Flow of Management Practices: An Analysis of NGO Monitoring and Evaluation Dynamics. *Nonprofit & Voluntary Sector Quarterly*, 43(6), 1033–1051.
- Patton, M. Q. (1997). *Utilization-Focused Evaluation: The New Century Text*. SAGE Publications, Inc.: Thousand Oaks, CA.

- Patton, M. Q. (2010). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. Guilford Press: New York, NY.
- Scriven, M. (2004). Reflections. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (1 edition, pp. 183–195). SAGE Publications, Inc.
- Shadish, W. R., Leviton, L. C., & Cook, T. D. (1991). *Foundations of program evaluation: Theories and practice*. SAGE Publications.
- Silverstein, L., & Maher, E. J. (2008, Winter). Evaluation Blues: How accountability requirements hurt small, innovative programs the most. *Stanford Social Innovation Review*, 6(1), 23.
- Stiffman, E., & Tumolillo, M. A. (2019). Hippocratic Oath for a Digital World: Insights From SSIR's 2019 Data on Purpose Conference. *Stanford Social Innovation Review* (February 22, 2019).
- Sturges, K.M. (2015). Complicity Revisited: Balancing Stakeholder Input and Roles in Evaluation Use. *American Journal of Evaluation*, 36(4), 461-469.
- Thomas, V. G., Madison, A., Rockcliffe, F., DeLaine, K., & Lowe, S. M. (2018). Racism, Social Programming, and Evaluation: Where Do We Go From Here? *American Journal of Evaluation*, 39(4), 514–526.
- Twersky, F., Buchanan, P., & Threlfall, V. (2013). Listening to Those Who Matter Most, the Beneficiaries. *Stanford Social Innovation Review*, 11(2), 40–45.
- Williams, B., & Hummelbrunner, R. (2011). *Systems Concepts in Action: A Practitioner's Toolkit*. Stanford Business Books.

### Further Reading

- ActionAid (2006). ALPS - Accountability, Learning and Planning System. ActionAid International. <https://actionaid.org/publications/2011/alps-accountability-learning-and-planning-system>
- Alaimo, S. P. (2008). Nonprofits and Evaluation: Managing Expectations from the Leader's Perspective. *New Directions for Evaluation*, 119, 73–92.
- Alkin, M. C. (2013). *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*, Sage Publications.
- Benjamin, L. M. (2008). Account Space: How Accountability Requirements Shape Nonprofit Practice. *Nonprofit & Voluntary Sector Quarterly*, 37(2), 201–223.
- Benjamin, L. M. (2021). If not the program then what? Evaluating social change in nonprofit and voluntary organizations. In P. Dahler-Larsen (Ed.), *A research agenda for evaluation*. Routledge.
- Benjamin, L. M., & Campbell, D. (2015). Nonprofit Performance: Accounting for the Agency of Clients. *Nonprofit and Voluntary Sector Quarterly*, 44(5), 988–1006.
- Benjamin, L. M., & Campbell, D. C. (2014, Spring). Programs Aren't Everything. *Stanford Social Innovation Review*, 12(2), 42.
- Bornstein, L. (2006). Systems of Accountability, Webs of Deceit? Monitoring and Evaluation in South African NGOs. *Development*, 49(2), 52–61.
- Cairns, B., Harris, M., Hutchison, R., & Tricker, M. (2005). Improving performance? The adoption and implementation of quality systems in U.K. nonprofits. *Nonprofit Management & Leadership*, 16(2), 135–151.
- Carman, J. G., Fredericks, K. A., & Introcaso, D. (2008). Government and accountability: Paving the way for nonprofits and evaluation. *New Directions for Evaluation*, 2008(119), 5.
- Chaplowe, S. G., & Engo-Tjéga, R. B. (2007). Civil Society Organizations and Evaluation. *Evaluation*, 13(2), 257–274.
- Cutt, J., & Murray, V. V. (2000). *Accountability and Effectiveness Evaluation in Nonprofit*

- Organizations* (Issue Vol. 2). Routledge.
- Dahler-Larsen, P. (2012). *Evaluation Society*. Stanford Business Books: Stanford, CA.
- Fine, A. H., Thayer, C. E., & Coghlan, A. T. (2000). Program Evaluation Practice in the Nonprofit Sector. *Nonprofit Management and Leadership*, 10(3), 331–340.
- Flynn, P., & Hodgkinson, V. A. (Eds.). (2001). *Measuring the Impact of the Nonprofit Sector* (1st ed.). Springer US.
- Forbes, D. P. (1998). Measuring the Unmeasurable: Empirical Studies of Nonprofit Organization Effectiveness from 1977 to 1997. *Nonprofit and Voluntary Sector Quarterly*, 27(2), 183–202.
- Fowler, A. (1997). *Striking a Balance: A Guide to Enhancing the Effectiveness of Non-Governmental Organisations in International Development* (1st ed.). Routledge.
- Hall, J. N. (2019). The Other Side of Inequality: Using Standpoint Theories to Examine the Privilege of the Evaluation Profession and Individual Evaluators. *American Journal of Evaluation*, 41(1), 20–33.
- Hargreaves, M., Cole, R., Coffee-Borden, B., Paulsell, D., & Boller, K. (2013). Evaluating Infrastructure Development in Complex Home Visiting Systems. *American Journal of Evaluation*, 34(2), 147–169.
- Herman, R. D. (2017). Methodological Issues in Studying the Effectiveness of Nongovernmental and Nonprofit Organizations. *Nonprofit and Voluntary Sector Quarterly*, 19(3), 293–306.
- Herman, R. D., & Renz, D. O. (1997). Multiple Constituencies and the Social Construction of Nonprofit Organization Effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 26(2), 185–206.
- Herman, R., & Renz, D. (1999). Theses on Nonprofit Organizational Effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 28(2), 107–126.
- Hilgendorf, A. E., Moore, T. R., Wells, A., & Stanley, J. (2020). Positioning Health Equity Within a Systems Thinking Framework to Evaluate Coalitions and Collaborative Initiatives. *New Directions for Evaluation*, 2020(165), 91.
- Hood, S. (2004). A Journey to Understand the Role of Culture in Program Evaluation: Snapshots and Personal Reflections of One African American Evaluator. *New Directions for Evaluation*, 102, 21–37.
- Hwang, H., & Powell, W. W. (2009). The Rationalization of Charity: The Influences of Professionalism in the Nonprofit Sector. *Administrative Science Quarterly*, 54(2), 268–298.
- Innonet. (2016). State of Evaluation: Evaluation Practice and Capacity in the Nonprofit Sector. Innovation Network, Inc. [https://www.innonet.org/media/2016-State\\_of\\_Evaluation.pdf](https://www.innonet.org/media/2016-State_of_Evaluation.pdf)
- Kabeer, N. (2019). Randomized Control Trials and Qualitative Evaluations of a Multifaceted Programme for Women in Extreme Poverty: Empirical Findings and Methodological Reflections. *Journal of Human Development and Capabilities*, 20(2), 197–217.
- Kang, J., Anderson, S. G., & Finnegan, D. (2012). The evaluation practices of US international NGOs. *Development in Practice*, 22(3), 317–333.
- King, J. A., & Alkin, M. C. (2019). The Centrality of Use: Theories of Evaluation Use and Influence and Thoughts on the First 50 Years of Use Research. *American Journal of Evaluation*, 40(3), 431–458. <https://doi.org/10.1177/1098214018796328>
- LeRoux, K., & Wright, N. S. (2010). Does Performance Measurement Improve Strategic Decision Making? Findings From a National Survey of Nonprofit Social Service Agencies. *Nonprofit & Voluntary Sector Quarterly*, 39(4), 571–587.
- Maier, F., Schober, C., Simsa, R., & Millner, R. (2015). SROI as a Method for Evaluation Research: Understanding Merits and Limitations. *VOLUNTAS: International Journal*



- of Voluntary and Nonprofit Organization*, 26(5), 1805–1830.
- Mook, L., Maiorano, J., Ryan, S., Armstrong, A., & Quarter, J. (2015). Turning Social Return on Investment on Its Head. *Nonprofit Management & Leadership*, 26(2), 229–246.
- Newcomer, K., Baradei, L. E., & Garcia, S. (2013). Expectations and capacity of performance measurement in NGOs in the development context. *Public Administration and Development*, 33(1), 62–79.
- Reddy, S. G. (2019, October 22). Esther Duflo and Abhijit Banerjee’s Nobel Prize Is a Cautionary Tale for Economics. *Foreign Policy*. Accessed on June 5, 2020 at: <https://foreignpolicy.com/2019/10/22/economics-development-rcts-esther-duflo-abhijit-banerjee-michael-kremer-nobel/>
- Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), 29–48.
- Schorr, L. B. (2012). Broader Evidence for Bigger Impact. *Stanford Social Innovation Review*, 10(4), 50–55.
- Schwandt, T. A. (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Stanford University Press.
- Smith, S. R. (2010). Nonprofits and Public Administration: Reconciling Performance Management and Citizen Engagement. *American Review of Public Administration*, 40(2), 129–152.
- Smyth, K. F., & Schorr, L. B. (2009). A Lot to Lose: A Call to Rethink What Constitutes “Evidence” in Finding Social Interventions that Work (Malcolm Wiener Center for Social Policy) [Working Paper Series]. Harvard Kennedy School.
- Sowa, J. E., Selden, S. C., & Sandfort, J. R. (2004). No Longer Unmeasurable? A Multidimensional Integrated Model of Nonprofit Organizational Effectiveness. *Nonprofit & Voluntary Sector Quarterly*, 33(4), 711–728.
- Weiss, C.H. 1998. *Evaluation*. Prentice Hall.
- Wolfe, S. M., Long, P. D., & Brown, K. K. (2020). Using a Principles-Focused Evaluation Approach to Evaluate Coalitions and Collaboratives Working Toward Equity and Social Justice. *New Directions for Evaluation*, 2020(165), 45.