



Unique Methods in Multi-Stakeholder Advocacy Evaluation

Anne Gienapp

ORS Impact

Jewlya Lynn

Spark Policy Institute

Rebecca Ochtera

Spark Policy Institute

Jared Raynor

TCC Group

Introduction

The field of advocacy evaluation has grown up, transitioning from an emergent area to an established field of practice. However, evaluation of *multi-stakeholder advocacy* is still in its infancy with evaluators, funders, and advocates working to untangle the complexity of how multiple organizations come together to influence policy change, build cohesive fields of practice, and accomplish more than any one group can do.

This brief is a follow-up to a seminal resource in the advocacy evaluation field, “Unique methods in advocacy evaluation” (Coffman & Reed, 2010). This work increased evaluators’ awareness of methods uniquely positioned to evaluate advocacy. This second brief adds to that toolbox, this time focusing on methods that have been successfully used to understand multi-stakeholder advocacy.

Of course, no different from any other field of evaluation, advocacy and multi-stakeholder advocacy evaluations can draw on the many different methods commonly deployed, from surveys and interviews to focus groups and observations. The goal here is not to suggest these more traditional methods are not important, but rather, to add to the toolbox. The unique methods presented here build on these techniques, adding frameworks and specific types of analysis to strengthen a traditional interview or survey process. The three methods are: Assessing Mature and Emergent Advocacy Fields, Machine Learning, and Dashboard Tracking.

Each of the methods presented here have been tested and refined in real-life settings. They have been found to generate information in a reasonable timeframe without needing excessive resources (time or money) to complete. Perhaps most important, the results of evaluations using these tools have been used by advocates and funders to inform their strategic decisions, evidence that the methods have generated actionable, meaningful information.

Table of Contents

Assessing Mature and Emergent Advocacy Fields.....	4
Machine Learning.....	8
Dashboard Tracking	11

Three Unique Methods in Multi-Stakeholder Advocacy Evaluation

Method (Developed by)	Synopsis
<p>1 Assessing Mature and Emergent Advocacy Fields</p>	<p>Assessment methods differ depending on how mature a field is, but can use the same five dimensions as a measurement framework. More mature fields are appropriate for social network analysis and quantitative methods, along with qualitative, while more emergent fields primarily benefit from qualitative assessments.</p>
<p>2 Machine Learning</p>	<p>An analytic tool that uses natural language processing and text analytics, mixed with manual annotations to categorize large amounts of unstructured data. Based on annotations, patterns in language can be detected, creating a model that can be run to categorize a much larger set of data than evaluators could efficiently code manually.</p>
<p>3 Dashboard Tracking (TCC Group)</p>	<p>This method addresses one of the largest logistical issues facing multi-stakeholder advocacy evaluations: bringing data together across multiple stakeholders. This requires agreement on content, process, and use. Engagement is obtained by anchoring the method in dashboard creation, while the process creates the imperative for shared monitoring and outcomes reporting in a shared reporting template.</p>

Conducted By	Type of Data Returned	Use This Method When
<p>An outside evaluator, but benefits from active involvement of advocates in design and interpretation of results</p>	<p>Qualitative interview data or Qualitative and quantitative survey data, including network data</p>	<p>The intent is to understand how a field of advocacy organizations can collectively influence a specific policy domain and identify ways to strengthen the field.</p>
<p>An outside evaluator, in partnership with an expert in machine learning.</p>	<p>Raw output file that includes language content, relevance scores, and key descriptors (e.g., author, date, source, other pertinent information)</p>	<p>The intent is to surface key themes contained in very large, unstructured data sets.</p>
<p>Group leadership and managers; may benefit from external consultant to facilitate process</p>	<p>Dashboard and corresponding indicators (gathered through various types of data collection)</p>	<p>The intent is to facilitate agreement on measures and set up the tracking system to collect relevant data. It leads to the greatest efficiencies when partners are acting fairly autonomously or covering a wide range of activities.</p>

Assessing Mature and Emergent Advocacy Fields

By Jewlya Lynn and Rebecca Ochtera, Spark Policy Institute (contact Rebecca@sparkpolicy.com)

Recent work by Spark Policy Institute, the Center for Evaluation Innovation and others to develop a framework for understanding and evaluating advocacy fields has made it easier to conceptually untangle the concept of an advocacy field. In a framework for evaluating advocacy fields, the five dimensions (see box to the right) are defined, relevant evaluation questions explored, and specific methods identified.¹ Evaluators are using a variety of methods for measuring the dimensions, from interactive activities with advocates to qualitative style interviews asking questions about each dimension.

Below are outlines of two distinct measurement approaches, each of which explores all five dimensions, including seeking to untangle the relationships between them. One approach is for emergent fields, the second for mature fields, defined as:

The Five Dimensions of Advocacy Fields:

- Field Frame
- Skills & Resources
- Adaptive Capacity
- Connectivity
- Composition

- 1. MATURE ADVOCACY FIELD** – a field with a variety of organizations familiar with each other’s work, many working under the same field frame, and with ample connectivity between different types of groups.
- 2. EMERGENT ADVOCACY FIELD** – a field whose issue, geographic reach, and/or organizational composition is relatively new, where many organizations are not familiar with each other’s work and not actively working together.

	Emergent Field	Mature Field
Type of Data Collection	Key informant interviews, primarily qualitative data.	Mixed methods survey with a social network analysis component.
Sample	Interviewees representing the range of perspectives in the field, each well placed to talk about the field as a whole.	Respondents in leadership positions representing the majority of organizations in the field, each well placed to describe their own organization and talk about the broader field.

¹ Lynn, J. 2014. Assessing and evaluating change in advocacy fields. Washington, D.C.: The Center for Evaluation Innovation. Available at: http://www.pointk.org/resources/files/Spark-Evaluating_Change_In_Advocacy_Fields.pdf

Emergent Field

Capacity questions (field)

- We'd like to better understand the types of advocacy skills and capacities that are deployed related to this issue area. I'm going to name different types of capacity and I'm interested in understanding whether there is enough capacity to do this type of advocacy, as well as how effectively it is deployed.

Capacity questions (field)

- Overall, when you think about organizations working on this issue, to what extent do these organizations act together to advance policy issues? [Behaviors of the field].
- Are there certain coalitions, planning groups, or other collaborative spaces where advocates come together to make decisions about policies to pursue or strategies? [Infrastructure supporting the behaviors].

Adaptive capacity questions (field)

- In general, are advocates working on these issues good at assessing changes in the policy and political environment? How effectively do advocates adapt strategies and priorities in response to those changes?

Mature Field

Capacity questions (individual organization)

- Which of the terms below best describe the types of work your organization regularly undertakes related to <issue area>? (List of advocacy skills).

Capacity questions (field)

- When you think of the entire landscape of organizations working to improve <issue area> in <geographic area>, are there gaps that have affected progress? In other words, are there some skills, tactics, roles that you wish the advocacy community were better at using or used more frequently?

Capacity questions (individual organization)

- When you think about the work your organization does on <issue area>, who are your key organizational partners - that is, organizations with which your organization shares information, engages in shared activities, or even shares resources, in order to achieve your organization's goals related to <issue area>?

Capacity questions (field)

- In general, in what ways are organizations in <geographic area> that are advocating for <issue area> good at working collaboratively to identify goals to pursue?

Adaptive capacity question (individual organization)

- How, if at all, has your organization changed its advocacy approach to [issue area] due to shifts in the policy, political and funding environment over the last two years?

Adaptive capacity questions (field)

- In general, in what ways are organizations in <geographic area> that are advocating for <issue area> good at assessing changes in the policy and political environment? How effectively do advocates adapt strategies and priorities in response to those changes?

Example Questions

Analysis Techniques

Visualization Techniques

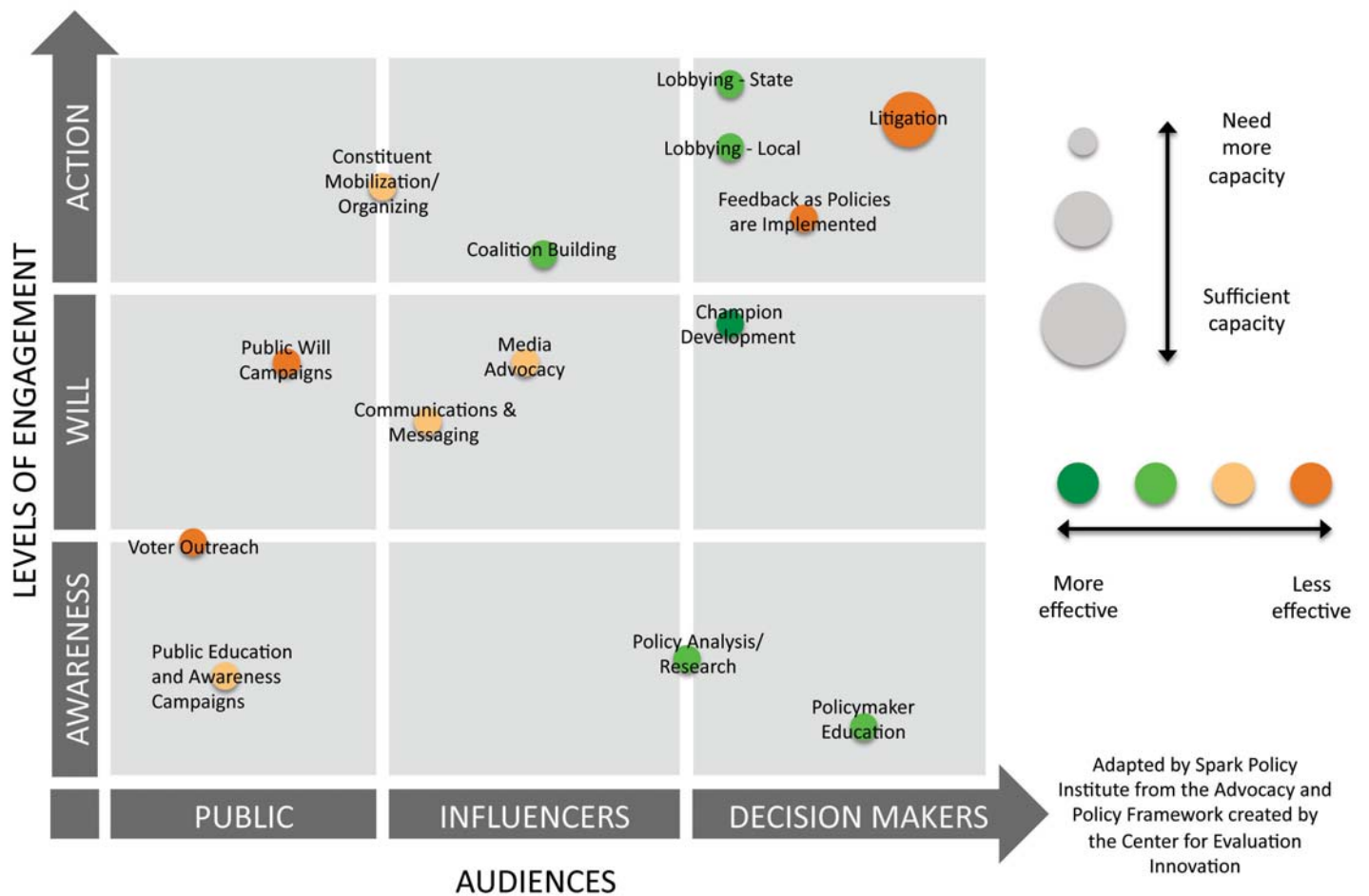
- Qualitative theming within each dimension.
- Use of a data display to compare themes across dimensions and surface cross-cutting findings.

- Use of the Advocacy & Policy Framework to share capacity findings (highlighting availability of capacity and effectiveness) (see Image 1).
- Use of heat-map in bar formats to create quick visual of the strength of key elements of the field across different dimensions.

- Qualitative theming within each dimension.
- Each question analyzed quantitatively, as appropriate to the question.
- Multivariate analysis to bring multiple dimensions together, allowing for things like cluster analysis to develop groupings and descriptions of similar organizations.
- Social network analysis, not weighted, directional, and analyzed with attributes that come from the other dimensions.
- Data display to compare qualitative themes and quantitative findings across dimensions and surface cross-cutting findings.

- Use of the Advocacy & Policy Framework to share capacity findings (highlighting availability of capacity and effectiveness) (see Image 1).
- Use of heat-map in bar formats to create quick visual of the strength of key elements of the field across different dimensions.
- Use of social network analysis visualizations to show relationships between different types of organizations, as identified by constituents they represent, issue areas, capacities, etc.

Image 1: Example of an emergent field visualization based on Advocacy & Policy Framework



Perceptions of field-wide advocacy capacity in an emerging climate adaptation field. This analysis indicated that advocates saw too few organizations and too few capacities available to act on the opportunities in the policy arena. It further highlighted that litigation capacity, a core strategy of many of the field actors, was not effective in this issue area.

Limitations

EMERGENT

- Accuracy of findings heavily influenced by distribution of respondents and whether they can speak at a “field” level.
- A need for accuracy makes selection of the “right” respondents critical.

MATURE

- Accuracy of organizational level findings on specific dimensions (e.g., capacity and connectivity) heavily influenced by response rate.
- Accuracy of field level findings influenced by whether the respondents can speak to the full range of their organization's work and still have an understanding of the field.

Benefits

EMERGENT

- Can surface critical strengths and disconnects of the emerging field.
- Can identify points where a funder or leading advocate can intervene to strengthen the field as it matures.

MATURE

- Can clarify why the mature field is continuing to have challenges or failing to achieve critical policy goals.
- Identifies points where advocates can experiment with different ways of acting in order to strengthen on one or more dimensions.

Example

A field assessment was completed of a mature health coverage field (via a survey of over 60 organizations) and an emergent health equity field with some overlap (discovered via the health coverage survey, followed by key informant interviews with over 20 leading stakeholders). Key findings included:

- The isolation in the network of equity organizations prioritizing the needs of specific communities of color;
- The existence of broker organizations that had an equity frame and both reached into the densely connected network of health coverage advocates and the more loosely connected network of health equity advocates;
- Lack of policymaker influence capacities among health equity advocates;
- Negative past experiences, and consequently biased current perceptions, between health equity advocates and health coverage advocates;
- A strong desire to diversity staff and advocate for equity by health coverage advocates; and
- A strong desire to build capacity and be better connected to the health coverage advocacy field by health equity advocates.

The assessment was completed in two months and used immediately by a funder and advocates to plan a field building strategy for the emerging health equity field. The strategy brought together health coverage and health equity advocates developing a more robust and shared field of health equity, inclusive of, but not limited to, the priorities often held by organizations historically focused on coverage and healthcare. The strategy prioritized taking collective actions, fostering increased connections, and engaging diverse communities and affected populations.²

² Krause, E.M.S. and Armijo, C. (2014). Views from the field: Public policy and the equity agenda. Washington, D.C.: Grantmakers in Health. Available at: http://www.gih.org/files/FileDownloads/Views_Policy_Health_Equity_TCT%26CTHealth_January_2014.pdf

Machine Learning

By Anne Gienapp, ORS Impact (contact agianapp@orsimpact.com)

Multi-stakeholder advocacy is aimed at advancing shared goals via aligned or well-coordinated efforts. As with other types of advocacy efforts, stakeholders' efforts may be focused on specific policy-related goals, but shared goals may also reflect what stakeholders believe to be the necessary conditions for favorable policy decisions and implementation. In other words, via coordinated action, stakeholders are interested in achieving changes in awareness, dialogue, opinion or action; changes in the depth of support for a particular issue or decision, or changes in systems that facilitate or monitor policy implementation. While it is possible to explore these types of changes via traditional quantitative or qualitative evaluation methods, examination of publicly available data such as media stories, social media feeds, or organizational websites can add depth and rigor to evaluation efforts and enhance the utility of findings. However, the retrieval and efficient coding and analysis of such unstructured data may be a daunting prospect for evaluators.

Machine learning is a data analytic approach that facilitates the retrieval of unstructured text-based data and involves the use of algorithms, informed by sample inputs, which are applied to large data sets in order to generate meaningful patterns or predictions. Machine learning uses natural language processing and text analytics, mixed with manual annotations to categorize large amounts of data. Based on manual annotations, computational systems “learn” how to recognize

relevant desired information and/or perform a desired task based on information extracted from the data. Models can then be developed that, when run, allow for categorization and filtering of data. Machine learning can be useful for extracting findings from large amounts of unstructured data—websites, social media outlets, SMS, emails, or other text data—to generate insights about notable trends or themes, sentiment, or the relationship between sentiment and actions. Machine learning is a specialized sub-field, and its application requires evaluators to partner with trained computer scientists.

Within multi-stakeholder advocacy evaluations that are aiming to create the conditions for policy decisions or implementation, the types of questions that might be addressed through machine learning include:

- What are perceptions or sentiment regarding a particular topic or issue and how has sentiment changed over time?
- What are differences in perceptions or sentiment based on key descriptors?
- What is the volume of discussion about a particular issue? Among whom is the volume of discussion strongest? How has the volume of discussion changed over time?
- How do documented actions or trends reflect support for policy decisions or policy implementation?

The steps involved in machine learning are detailed below.

Steps	Notes	
1	<p>Determine parameters of the data set to be examined, e.g., US newspapers (online editions) for the period January-December 2015; all federal appeals court decisions handed down between 2000 and 2010; tweets made by a sample of individuals during a critical time period.</p>	<p>The appropriate data set will be determined based on the evaluation questions to be addressed and resources available. Machine learning experts are able to facilitate retrieval of these types of data in order to generate a useable output file, though there are costs involved in obtaining certain types of data.</p>
2	<p>Multiple raters annotate data by identifying words or combinations of words that are relevant to the evaluation questions being addressed.</p>	<p>Once the data is obtained, annotation of data establishes the relevance model, which can then be applied to the entire data set to generate the content that is most relevant to evaluation questions. Annotation is likely to occur through an interface created by the machine learning expert.</p>
3	<p>Ensure inter-rater reliability among those annotating data for relevance.</p>	<p>Establishing a strong relevance model requires raters to maintain an acceptable level of agreement, e.g., at least 75%.</p>
4	<p>Use annotations to inform categories for analysis by the machine model.</p>	<p>Once the machine knows what relevant data are, modeling generates primary categories or themes emerging from the relevant data. Evaluators can help refine the categories or themes based on evaluation questions, and identify those that are highest priority.</p>
5	<p>Examine the full data set against established categories.</p>	<p>Once the priority categories/themes are established, the machine can examine the full data set to generate output that addresses questions of interest.</p>

Limitations

- Machine learning has inherent costs. Besides the cost of partnering with experts, evaluators' time is required for sufficient annotation of data. In addition, since machine learning delivers raw data rather than a "product," there may be additional costs associated with creation of graphs, charts, visuals or other useful evaluation products.
- It can be difficult to create strong relevance models when concepts contained in the data are nuanced – as may frequently be the case in an advocacy-related context.
- Like other inquiry methods, the strength of findings generated via machine learning is conditional on the precision and quality of the data set.

Benefits

- Machine learning makes it possible to answer questions about issue sentiment, discourse and the effectiveness of collective action that would otherwise be very difficult for evaluators to answer.

Example

A foundation invested in multiple stakeholders to conduct aligned advocacy and communications efforts regarding a particular set of education-related topics. An assessment was conducted to determine whether changes in dialogue or sentiment regarding a particular set of topics had occurred within the education field as a result of the advocacy and communications efforts undertaken by multiple grantees. Evaluators partnered with machine learning experts to retrieve Twitter data from 215 selected handles reflecting individuals and organizations active in the education field. Evaluators sought Twitter data from two time periods – one before stakeholders' efforts had begun, and the second, a year later, after stakeholders' efforts had been underway for several months.

Once evaluators annotated the Twitter data to determine relevance, an algorithm was created. As the algorithm was applied to the full data set, an output file was generated that identified certain categories that emerged from the content of tweets as well as descriptor variables. Evaluators were then able to run analyses that illuminated changes in the volume of dialogue and sentiment regarding the topics of interest over the time period when stakeholders were implementing aligned communications. Through analysis, evaluators were able to see how the volume of dialogue and sentiment had changed overall, as well as among key groups – e.g., thought leaders in education. The results of the evaluation shed light on where stakeholders' efforts had made a difference and provided insights about where there were opportunities to target efforts in the future.

Dashboard Tracking

By Jared Raynor, TCC Group (contact jraynor@tccgrp.com)

This method, developed by TCC Group, addresses one of the largest logistical issues facing multi-stakeholder advocacy evaluations: bringing data together across multiple stakeholders. This requires agreement on content, process and use. Engagement is obtained by anchoring the method in dashboard creation, while the process creates the imperative for shared monitoring and outcomes reporting in a shared reporting template.

Dashboard tracking is most effective early on in multi-stakeholder engagements, as it facilitates agreement on measures and sets up the tracking system to collect relevant data. Dashboard tracking leads to the greatest efficiencies when partners are acting fairly autonomously or covering a wide range of activities.

There are three distinct steps in the process:

1. ALIGNMENT

Having the group come to agreement on what is most important to track, from both an outcomes and process perspective

2. DESIGN

Setting up how data will be displayed in the most effective manner

3. DATA ENTRY

Ensuring clear data entry processes and follow-through on getting/giving data

Alignment

During the alignment process, stakeholders work toward agreement on a common dashboard. This can be done using a tool like logic models or more organically through group conversation. **During alignment**, the group is directed to come to agreement on key indicators in three areas: group processes, group activities, and group outcomes.

Indicator Category	Definition	Advised # of Indicators
Group Process	How they want to work together—levels of trust, agreed upon decision-making approaches, accountability for follow-through, etc.	3 - 5
Group Activities	Identify shared categories for defining the work that the group will do as a whole or as individual organizations on behalf of the group.	As many categories as is deemed relevant to capture the group's activities.
Group Outcomes	Relate to the outcome targets that the group sets at an aggregate level—what their work collectively is meant to address. See Table 1 for samples.	3 - 5

Again, we have found that three to five indicators is a good target number for group outcomes indicators, with **two important caveats**. First, at least one indicator should rely on data that can be independently verified from a source outside the group. Second, the outcome indicators should be focused on macro goals. However, given the different roles stakeholders may play, the alignment on outcomes should include one open-ended space that asks each organization to describe the outcomes that they feel they are individually contributing. This is the crucial space that allows organizations to agree to the macro outcomes while still having a place to share the value of their own work in a more nuanced way.

Design

During design, the organizations create a reporting tool that works best to aggregate the agreed upon indicators. It is best if one organization serves as the backbone for the reporting, a role that can also be effectively played by an outside evaluator. The platform can be high tech, such as an interactive interface using a survey mechanism; medium tech, such as a shared cloud-based document (e.g., Google doc); or low tech, such as phone calls or paper forms that get submitted and aggregated. The goal of the design phase is to make the data submission process as easy as possible.

Data Entry

During data entry, organizations input the data into the reporting tool. Given the three types of indicators, data collection should happen at distinct points in time.

- **Group process indicators** are collected at inflection points. Inflection points can occur, a couple of weeks after intense periods of work (e.g., budget session).
- **Group activities indicators** should be collected as close to the point of action as possible (real time) in an ongoing manner. See Table 1 for sample groupings.
- **Group outcomes indicators** should be collected at pre-set time periods when it is most relevant. This could include after intense periods of work or in anticipation of reporting deadlines for external stakeholders such as funders.

These three steps are followed by more traditional analysis and interpretation processes that focus on data use such as longitudinal tracking, group analysis, and decision-making.

Table 1: Sample Activity Groupings for Tracking

Category	Activity	Category	Activity
Media	<ul style="list-style-type: none"> • Op-Eds • Social Media 	Grasstops	<ul style="list-style-type: none"> • Grasstops Organization Sign Up • Grasstops Organizations Action Taken • Grasstops Individual Sign Up • Grasstops Individual Actions Taken
	<ul style="list-style-type: none"> • Focus Groups • Trainings on Quality and Cost • Provider Engagement Breakfasts • Interview with Parents 		Research <ul style="list-style-type: none"> • Target Question Analysis Completed
Mobilizations	<ul style="list-style-type: none"> • Providing Campaign Communications to Providers • One-on-One Meetings with Providers • Recruiting Peer Mentors • Parents • List Building Event • Engaging School Districts • Relationship Building 	Legislative Engagement	<ul style="list-style-type: none"> • Adopt-A-Legislator Events • Legislative Breakfasts • Training Constituents

Limitations

- May feel overly process-heavy for some group members.
- May be difficult to align on what is most important to the group as a whole (which doesn't delegitimize value at the individual organizational level).
- Sensitive or proprietary data might limit willingness to share some data.
- The data is derived from information from within the group (not external).
- Gathering data from all the group's organizations can be challenging.

Benefits

- Clear alignment of what is important to the group as a whole.
- Enhanced communication ability to outside stakeholders (e.g., funders) on what the group is doing as a whole.
- Improved cohesion of the group.
- Quick snapshot of how the group is doing to inform decision-making—keeps the group on track.
- Examines both process and outcomes indicators.

Example

A campaign of 10 organizations was formed to promote access to high-quality pre-kindergarten programs for all families. Each organization had a distinct role, but also collaborated on common activities. The campaign aligned around a set of indicators that were then tracked by an outside evaluator using a mix of Google Docs and an online survey that would lead organizations through the reporting requirements. The data was used regularly to assess the health of the campaign as well as progress toward agreed-upon outcomes.

About ORS Impact

ORS Impact serves as a thought partner to social impact leaders, supporting their work to accomplish their missions. Our team has earned a global reputation for innovation and the application of approaches that deliver the insights our clients need to move confidently from ideas to action to impact. Since 1989, we've worked collaboratively with our clients to pursue the change they seek in their communities' health, wellbeing, and prospects to flourish. We support clients engaged in advocacy and policy change, systems change and social change, bringing our distinctive expertise in planning, measurement and evaluation to their most vexing challenges.

About Spark Policy Institute

Spark Policy Institute partners with policymakers, communities, foundations and non-profits to develop actionable strategies for solving society's complex problems. We combine community and stakeholder engagement with innovative, adaptive, data-driven solutions that bridge sectors, issues, beliefs and values to achieve meaningful, measurable outcomes. Spark is also the creator of the the Tools for Social Innovators site, a community toolbox contributed to by experienced and cutting-edge social innovators. Each of the 12 tools, including the Advocate's Evaluation Toolkit, provides practical hands on information, templates, checklists and more to help funders, nonprofits, business and community leaders enhance their ability to ignite social change.

About TCC Group

TCC Group is a mission-driven consulting firm that collaborates with leaders to solve complex social problems. TCC Group has more than 35 years of experience working in the social impact field with companies, philanthropies, and nonprofit organizations. Our approach is data-driven and outcomes-based, our approach also draws from the knowledge of in-house program management and evaluation teams. We work to ensure that our clients develop actionable and measurable strategic goals to communicate effectively with their stakeholders.



ORSIMPACT
Missions Accomplished



SPARK POLICY INSTITUTE
igniting public policy and community change

tccGROUP
solutions for social impact