WHO USED OUR FINDINGS?
FRAMING COLLABORATIONS BETWEEN AMERICAN FOUNDATIONS AND THINK TANKS
THROUGH PRACTICES IN THE EVALUATION OF POLICY INFLUENCE

A Thesis
submitted to the
Faculty of the Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
Master of Arts in Communication, Culture and Technology

By

Jaime A. Gonzalez-Capitel Martorell, B.A.

Washington, DC
April 22, 2016
American foundations measure returns on investments through evaluations that often gauge program outcomes and impact, but evaluation research has seldom been conducted on public policy research centers. This study exposes theoretical reasons for this disconnect, proposes the advocacy coalition framework (ACF) to frame collaborations between foundations and think tanks, contrasts indicators of think tank influence with evaluation policies of prominent foundations and existing methods and techniques for the appraisal of policy influence in the fields of advocacy and international development, and identifies recommendations and best practices. Original data were collected through 22 interviews to foundation staffers, think tank staffers and evaluation experts. The primary hypothesis that foundations drive the evaluation of think tanks is supported with several case studies in which think tanks are building their capacity or conducting self-evaluation research.
To Francesca,
Whose love makes everything happen.
Acknowledgements

My thesis would not have been possible without the gracious support of many experts and professionals that generously offered their time and helped inform my research. The critical commentaries of Kent Weaver, Diane Stone and Andy Richman were instrumental in shaping my research question at different stages, some of which felt close to a dead end. Carrie Kolasky at the Urban Institute and Meredith Hanley at the New America Foundation made time in their busy days to talk to me and identified fantastic participants for my interviews. My thesis advisor, Dr. Diana Owen, made an extraordinary effort to coordinate a cohort of thesis writers and succeeded in keeping all of us in track; and my second reader, Margaret O’Bryon, set it all in motion when she introduced me to the fascinating world of American foundations. But I would be oblivious if I failed to mention Marisol Pastor, my first boss in Spain and the person that introduced me to the use of social sciences in a workplace environment.
# Table of Contents

Chapter 1: Introduction.................................................................................................................. 1

Chapter 2: Literature Review........................................................................................................ 10
  2.1 Theories of research use........................................................................................................... 10
    2.1.1 Carol H. Weiss on knowledge use and policy influence.............................................. 13
  2.2 Placing think tanks in the knowledge utilization debate....................................................... 17
  2.3 Foundations and evaluation................................................................................................... 23
    2.3.1 What a traditional program evaluation looks like.......................................................... 25
    2.3.2 The expansion of evaluation practices............................................................................. 27
  2.4 Evaluation as outcome assessment of public policy research .............................................. 30
    2.4.1 Basic concepts in outcome evaluation............................................................................. 32
    2.4.2 Integrating processes........................................................................................................ 35
    2.4.3 Evaluation and advocacy.................................................................................................. 35
    2.4.4 Annotated list of methodologies and techniques............................................................ 37
    2.4.5 Some conclusions on evaluating policy research.............................................................. 41

Chapter 3: Methodology.................................................................................................................. 43

Chapter 4: Metrics of Influence and Evaluation Practices............................................................. 48
  4.1 Overview............................................................................................................................... 48
  4.2 Measuring the influence of think tanks.................................................................................. 48
    4.2.1 Metrics of influence and performance in selected think tanks................................. 48
    4.2.2 Rankings and indexes...................................................................................................... 55
    4.2.3 Case study 1: Center for Global Development............................................................... 58
    4.2.4 Key takeaways................................................................................................................ 62
  4.3 Evaluation policies in selected foundations.......................................................................... 64
    4.3.1 Overview of evaluation brands....................................................................................... 64
    4.3.2 Analysis and comparison................................................................................................. 65
    4.3.3 Case study 2: The W. T. Grant Foundation.................................................................... 68
    4.3.4 Key takeaways................................................................................................................ 69
Chapter 5: Interview Analysis

5.1 Think tanks: case studies on evaluation and influence assessment

5.1.1 Brookings Institution: the conversation on quantitative and qualitative metrics

5.1.2 Urban Institute: building evidence use with a dedicated unit

5.1.3 Internal evaluation unit

5.1.4 Building capacity with an earmarked grant

5.1.5 Building capacity with no dedicated funding

5.1.6 IDRC and the Think Tank Initiative: combining unrestricted funding with specific objectives

5.2 Role of public policy research for foundations

5.2.1 Can policy research be evaluated?

5.2.2 Research, advocacy, and collaboration between nonprofits

5.3 Foundations and their evaluation practices

5.3.1 Uses of evaluation by funders

5.3.2 The importance of general operational support

5.4 Measuring the influence of policy research

5.4.1 Interviews to key informers

5.4.2 Stories of success

5.4.3 Other methods for evaluating policy research: Social Network Analysis

5.5 Other factors of success

5.5.1 The value of dissemination

5.5.2 The value of combined skillsets

5.6 Conclusions

Chapter 6: Recommendations and Best Practices

6.1 Recommendations

6.2 Best practices

6.2.1 For funders

6.2.2 For evaluators and commissioners of evaluation studies
6.2.3 For think tanks................................................................. 110

Chapter 7: Conclusion............................................................. 111
  7.1 Limitations.......................................................................... 113
  7.2 Further research................................................................. 114

Appendices...................................................................................... 116
  Appendix A: Operational definitions of measures of influence........ 117
  Appendix B: Stories of impact....................................................... 118
  Appendix C: Evaluation policies at selected foundations............... 126
  Appendix D: Interview protocols................................................ 133
  Appendix E: List of respondents.................................................. 136

Bibliography....................................................................................... 138
List of Figures

Figure 1: Decisions evaluation is called upon to support .................................. 26
Figure 2: Simplified model of a traditional program evaluation .............................. 26
Figure 3: Why the simplified model doesn’t work .............................................. 26
Figure 4: Relative influence along the results chain ............................................ 34
Figure 5: Five-fold screening of issues and advocacy efforts .................................. 36
Figure 6: CGD, Core questions guiding the strategic review ................................. 60
Figure 7: CGD, Decision-making processes leading to new projects ..................... 61
Figure 8: CGD, (E)xpected (R)eturn Worksheet Sample ..................................... 62
List of Tables

Table 1: Percentage of foundation funding in DC-based think tanks…………… 2
Table 2: Organizations and acronyms……………………………………………….. 53
Table 3: Number of indicators used by each organization………………….. 53
Table 4: Influence Indicators used by selected think tanks………………….. 54
Table 5: Summary of the influence indicators used by selected think tanks…….. 55
Table 6: Absolute scores and ranking for US think tanks…………………… 56
Table 7: Budget adjusted scores and ranking for US think tanks……………… 57
Chapter 1: Introduction

This study asks two different but interrelated sets of questions. The first set is located in the academic field of policy studies that Carol H. Weiss, among others, named the sociology of knowledge application, and it inquires about the influence of public policy research. How can public policy research, and more generally social scientific knowledge, affect the decisions of public officers, policy makers and other relevant actors, such as advocates, non-profits, or the intangible general public? How is research used? And what methodological tools are available to gauge, track or measure utilization?

The second set belongs to the realm of knowledge-driven organizations, and asks whether American foundations are promoting the evaluation of think tanks, their research and their programs. These two quintessentially American types of organizations can be examined in the donor-grantee relationship, in which the former provides funding, whereas the latter obtains support to conduct independent research on policy issues. The grantmaking blueprint allows for conversations and interactions beyond monetary support, as foundations and think tanks do not go in opposite directions once the check has been cashed. Instead, the service or value provided by the grantee becomes a subject of formalized conversation through reporting requirements. The question is, then, whether those requirements foresee evaluative research, and of what nature.

When the two questions are combined, the academic about the influence of policy research is situated at the intersection of two organizational environments marked by distinct policies, practices, expectations and priorities, and embodied in specialized forms of expertise. At the same time, the question about non-financial interactions can be asked again: How do foundations assess the value of public policy research developed by the think tanks they fund? How do they measure, track and gauge their impact, and what methodological tools are available to them?

In 2012 there were in the United States 86,192 registered foundations, of which 91% are independent. Their aggregated giving for that year totaled $52 billion, spearheaded by the $3.2 billion donated by the Bill and
A group of transparency advocates recently found that just 21 prominent American think tanks added to over a billion-dollar in expenditures in 2013. In spite of the fact that foundations are more homogeneously distributed across the US than think tanks, which are still a heavily DC-centric phenomenon, these two types of organizations work consistently together. While think tanks are also supported by corporations and individual donors, in many cases foundation money makes up a very significant portion of their financial blood:

<table>
<thead>
<tr>
<th>% of foundation grants over total revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipartisan Policy Center</td>
</tr>
<tr>
<td>New America Foundation</td>
</tr>
<tr>
<td>American Enterprise Institute</td>
</tr>
<tr>
<td>Center for Global Development</td>
</tr>
<tr>
<td>Brookings Institution</td>
</tr>
<tr>
<td>Urban Institute</td>
</tr>
<tr>
<td>World Resources Institute</td>
</tr>
<tr>
<td>Heritage Foundation</td>
</tr>
<tr>
<td>Peterson Institute</td>
</tr>
<tr>
<td>Cato Institute</td>
</tr>
</tbody>
</table>

**Table 1: Percentage of foundation funding in DC-based think tanks**

From the selected list in Table 1, the Bipartisan Policy Center stands out with a 70 percent of foundation grants in 2014, followed by the New America Foundation, with 60% in the same period. The percentage represents a still significant 31% at Brookings in fiscal year 2015, being the first source in grants and second only to the sum of contributions and the endowment. At the bottom, the Peterson Institute for International Economics, the Cato Institution and the Heritage Foundation all have percentages below 15%, but still above 5%. These data have been retrieved directly from the organizations’ websites, and should not be taken at face value as being representative of the industry as a whole –although the range of values certainly is.
Despite the many connections between foundations and think tanks, there have been very few attempts at capturing their interactions in a systematic way. James Allen Smith, for example, did touch on the role of some pivotal philanthropic endeavors in his pioneering history of think tanks, describing the Russell Sage Foundation as a “think tank prototype”, or underscoring how important think tanks like the Urban Institute or the Carnegie Endowment for International Peace, among others, received in their beginnings substantial support of the Ford Foundation and the Carnegie Corporation. Nevertheless, the goal of Smith’s account was to explain the historic formation of the elite of policy experts and analysts that populate the Washington, DC scene, not to elaborate on whether the collaborations between foundations and think tanks are shaped by the assessment of grantee effectiveness.

This gap in the literature is far from surprising. In fact, a systematic examination of the intersection is a rather hard exercise for at least two reasons. Firstly, foundations function with a great degree of freedom and have no legal need to justify their choices and priorities to anybody.

By its very nature, the world of philanthropy is idiosyncratic and decentralized. Or, as is commonly heard among observers in the field, “if you’ve seen one foundation, you’ve seen one foundation!” (Prager 2003, 7)

A strange exception to capitalist economies in that they are not naturally regulated by any markets, the core activity of grantmaking is thus particularly devoid of accountability, which means that there are fewer regularities and common rules than in other fields. This rather arbitrary nature, which manifests itself in a wild diversity of organizational sizes, missions and values, is a natural obstacle for sharing lessons learned that can be especially cumbersome in the area of evaluation. As Peter Frumkin notes,

the results of this research have rarely been circulated widely or incorporated into the decision making of the field as a whole. Without meaningful channels through which to share knowledge, too much evaluation research remains proprietary information, consumed only by those closest to the subject matter. (Frumkin 2008, 339)

Secondly, the literature on organized philanthropy offers abundant case studies on collaborations with nonprofits involved in service delivery in education, health, immigration, and other major areas of social reform; in more recent years, there have also been important advances in evaluating advocacy campaigns
with specific goals in policy change. Neither of these two activities dwell specifically on the relationship with organizations that only do research. In other words, although the term nonprofit includes all or most think tanks, think tanks are an unusual form of nonprofit. As we will see in the following chapters, the particular nature of the question about the impact and relevance of social science may have contributed to absconding this organizational realm from the attention of experts in philanthropy and added layers of invisibility to the possibilities for shared knowledge.

Frumkin signals Rockefeller’s evaluation of the Minority Female Single Parent program (1982-1988) as an example of constructive failure in which an expensive but well-designed evaluation on four different sites allowed external actors to learn from what didn’t work and avoid repeated mistakes (Frumkin 2008, 69). However, the reliance of the largest independent foundations on effective grantmaking supported through evaluation has not yet consolidated in a community of practice with agreed upon notions of success and failure. Very often there is not a clear sense of how to extract value from the results of evaluation beyond the common sense rule of measuring to tell apart what work from what doesn’t. Every difficulty advancing scholarship in this empirical field, where generalizations are a hazardous material, underscores the need for efforts that contribute with the addition of knowledge from new angles. Asking the harder question about the impact of policy research may help identify initiatives that, given the lack of attention to the intersection, would otherwise remain siloed, and stimulate conversations about what kind of assessments are worth a try. Indeed, evaluations may offer an opportunity for the cross-examination of institutional alignment between the two organizational types, which often display different languages and sets of motivations. While the predominant discourse at foundations includes the overarching goal of advancing the organizational mission to promote their values while leveraging limited resources, and include nonprofits in their strategies as the tactic means to meet those ends, think tanks work towards building relevance in the world of public policy making by maintaining their credibility as independent actors that partially
overlap but remain distinct from the spheres of media, government, academics and business (Medvetz 2012).

Another reason for pursuing this avenue of research connects the problem of institutional alignment discussed above with the already canonic distinction between think tanks that function as universities without students and those that perform an advocacy role (Weaver 1989), which implies different models of reputation and institutional credibility. While universities without students are much more likely to define themselves as non-partisan, conduct all their programming independently from the agendas of their donors and construct a reputation for credible expertise based on unbiased, high-quality academic work, advocacy tanks tend to base their credibility in their identification with an ideological agenda that is pursued with the tools of modern marketing and the agility required to stay relevant in the high-speed context of current events. Although this “hunger for impact” has expanded, especially in the digital arena, by all policy research organizations (Rich 2004), the model of credibility is still significantly divided between centers modeled after academic standards and centers seeking to shape the political agenda in a much more direct manner.

Why does this matter for a discussion on the institutional alignment between think tanks and their funders? If think tanks, like other nonprofits, are considered as means to the end of furthering mission and values, the alternatives between funding an academic think tank or an advocacy tank splits quite dramatically regarding the nature of the programs that are funded and, subsequently, the options available to evaluate them. In other words, while the impact assessment of academic think tanks will still bump against the hard question about the utilization of policy research, advocacy think tanks would seem to fall under the scope of the methods that are already being used to evaluate the impact of other nonprofits actively engaged in policy change and advocacy strategies. xiii

Lastly, investigating practices and uses of evaluation methodologies to assess the impact of think tanks can contribute to framing the judgment of program officers at foundations. Frumkin’s thesis on the contradiction
between the professional program officer and the expressive philanthropist will be indirectly scrutinized by the question about the declared purpose and utilization of evaluations (Frumkin 2008, 62). If the mandate of program measurement were to offer a mere appearance of rationality, this study would contribute to building up on the skepticism over the idea of effective philanthropy. By identifying practical uses of evaluation research, the empirical component of this study has been able to document counter-examples to the skeptical position. Whether they remain exceptions or samples of trends that are gaining traction is out of the scope of this study.

**Research hypotheses**

These considerations lead into the three working hypotheses of this work:

- **H1:** Think tanks are more likely to evaluate their work when they receive foundational funding.
- **H2:** Think tanks with an orientation to advocacy are more likely to evaluate their work than nonpartisan think tanks.
- **H3:** Think tanks invested in international development are more likely to evaluate their work than think tanks invested in domestic policy.

Collected data support the first hypothesis. In spite of the sampling bias, the percentage of foundational funding is an important factor for the think tanks that reported existing evaluation practices or were embarked in building such capacity. While the findings are not conclusive for the second and third hypotheses, they are compatible with them. Firstly, integration in advocacy efforts appeared in many respondent’s view as the main frame under which public policy research can be evaluated. Secondly, the number of examples from initiatives in international development seem to indicate that this remains a driving sector for the evaluation of policy-oriented research. However, the realization that often the body of research that is subject to evaluation has been developed by academics casts shadows on the credibility of think tanks as knowledge producers, or at least on the overall weight of their contributions.
Beyond confirming or refuting the hypothesis, the results of the qualitative study offer answers to secondary questions that may be most interesting for practitioners in the field, including: recommendations and best practices, some of the most widely used methodologies, institutional rationales for conducting evaluation, and professional perspectives on the collaborations between foundations and think tanks.

**Chapter 2** places this investigation in the context of the existing literature. A first section examines the scholarship on knowledge utilization. The core of the chapter is dedicated to the scholarly work on the utilization of social science research, starting with a much needed overview of Carol H. Weiss’ theories, which straddle from evaluation theory to the role of intermediary organizations. More recent contributions to the research on evidence-based policymaking are reviewed. The second section surveys the scholarship on think tanks, their functions, roles, and how scholarship has examined the question about think tank influence through the quantified notions of media visibility and political reputation. R. Kent Weaver, James McGann, Donald Abelson, Andy Rich and Diane Stone are the main authors in this dialog. The third section offers a brief review of the adoption of evaluation by foundations, its functions and limitations, with a special attention to the expansion overtime in scope, research subjects and core organizational competencies. The efforts to build shared knowledge and methods in advocacy evaluation stand out as the principal body of practice that can be drawn into the question of public policy research evaluation, while the question about the use of evaluation in the context of philanthropy allows to apply the key takeovers on knowledge use, thus circulating back to the first section with the situated notions of the rest of the chapter. Finally, the fourth and last section examines the notion of evaluation, defined as the outcome assessment of public policy research programs. Contributions in international development and advocacy and policy change are analyzed and a number of techniques adopted in both fields is reported.

**Chapter 3** describes the methodologies used to collect empirical data and analyze it. Two methodologies were combined: collection of public information from organizational websites, and a set of semi-structured interviews to three professional profiles: staffers at foundations, staffers at think tanks and independent
evaluators. One questionnaire was designed for each profile so as to capture cultural and professional differences in terms of expectations, needs and practices. Respondents were identified by a combination of direct web searches, author extraction, and snowballing techniques.

In chapter 4 the annual reports of a dozen of prominent think tanks based in Washington, DC are analyzed to show how these organizations make the outward-facing argument about their influence; while in the case of foundations the focus was on the evaluation policies of four large independent foundations, which were analyzed to track intended uses, contexts for evaluation, and whether they conduct evaluation at the organization level and/or of public policy research. Each section is completed with a case study that serves as a contrast to the general patterns identified in the core analysis: the Center for Global Development as a think tank that has driven two major efforts to evaluate the overall organization influence; and the W.T. Grant Foundation, which is pushing forward the research agenda on knowledge utilization under a grant portfolio which may produce tested strategies to increase research utilization.

Chapter 5 reports the results of the interviews through an open, qualitative analysis. A total of 22 interviews were conducted to 12 foundation staffers, 6 think tank staffers and 4 independent evaluators. The main themes covered are methodologies, research evaluation, advocacy evaluation, evaluation utilization, good practices and recommendations. Since the sample shows a strong bias towards think tanks with an existing interesting in evaluation, the results cannot be considered representative of the whole field. However, the six think tanks offer a diverse muster of case studies that supports the main hypothesis linking foundational funding to evaluation adoption. While one of them has a working evaluation unit, another two are currently building their monitoring and evaluation systems. The other three cases include the Canadian think tank IDRC, analyzed in its role as managing intermediary of the Think Tank Initiative (TTI); the Urban Institute, which has created an internal unit that promotes research uptake among policymakers and is funded primarily by foundations; and the Brookings Institution, where conversations about influence are ongoing but have not crystallized thus far in evaluation studies.
Chapter 6 reports recommendations and best practices identified in the analysis of the interviews.

Chapter 7 summarizes the findings of the thesis, describes limitation of the study and points at its use to inform future research.

---


13 As Andrew Rich has pointed out, the distinction may be mirrored in the approach to funding that foundations take depending on their ideology. In an article published in 2005, an interviewee reported the following perception: “If you’re on the left, you have to go to the foundations and say you’re neutral, unbiased – not politicized. You’re certainly not liberal. If you’re ideological, they don’t want to support you. It’s frustrating – because, by contrast, if you’re on the right, the foundations will only fund you if you toe the ideological line, if you want to do battle for the conservative cause”. Andrew Rich. 2005. “War of Ideas”, Stanford Social Innovation Review. Spring. Last retrieved on January 22, 2016 from [http://ssir.org/articles/entry/war_of_ideas](http://ssir.org/articles/entry/war_of_ideas).
Chapter 2: Literature Review

2.1 Theories of research use

The concern for the relevance of social scientific enterprises for the public policy and decision making communities is not new, although it has been focused traditionally on the value of use of publicly funded research, and not so much on the role of independent policy research organizations, or think tanks. However, an overview of the discussion about the former builds the frame in which think tanks and, more generally foundation-driven policy research, can influence public discourse and decision making.

A rationalistic inheritance from the Enlightenment supporting the Great Society effort on social programs (Bowling, 2005: 68), in the early 70s it was still a broadly shared view among government officials in the United States that publicly funded research in the social sciences should aim at producing knowledge that would be useful in supporting decisions (Lynn 1978, 9). In this time, experts were compared to doctors healing a social disease, social planning engineers or, in the conception of President Hoover, to the technocratic remedy to partisan politics (Smith, 1993).

However, as evidence against those expectations soon accumulated, some studies started investigating why was the linkage of research to decision makers an “uncertain connection”, and how research was actually being used (Riecken and Boruch 1974; Cohen and Garet 1975; Floden and Weiner 1978; or Lindblom and Cohen 1979). Caplan et al. confronted both the lack of evidence on research utilization and the lack of recommendations on the issue with a pioneering study in 1975, which found 575 instances of use in different contexts, but identified an important limitation of predominant utilization theories:

> only rarely is policy formulation determined by a concrete, point-by-point reliance on empirically grounded data[…].[O]ur data suggest that there is widespread use of soft information and that its impact on policy, although often indirect, may be great or even greater than the impact of hard information. (Caplan 1975, 47)

The following years gave way to some attempts at explaining knowledge utilization as happening neither naturally nor in a vacuum, but in “epistemic communities” (Holzner and Marx 1979, 103:111), as
“subjectively mediated social processes” (Dunn, 1980, 516), and with a “political nature” (Weiss, 1977). The most common approach to understanding the issue was the image of the two communities of scientists and policy makers, first explicitly formulated by Caplan himself. He identified an existing trend in the literature to explain the lack of utilization through the argument that “social scientists and policy makers live in separate worlds with different and often conflicting values, different reward systems, and different languages” (Caplan 1979, 459). In this conception, mistrust reigns between the two communities:

to oversimplify a complex story, there have been two principal sources of controversy: legislators distrustful of social engineers who promote rigid ideas or pursue irrelevant academic interests, and social scientists worried that dependence on government might compromise their objectivity. (Lynn 1978, 2)

The image of the two communities proliferated, as it was very productive heuristically in uncovering factors, barriers and constraints based in social differences and in the act of transfer. With the pedagogic aim of generating realistic expectations for knowledge utilization, Carol H. Weiss and Michael Bucuvalas scanned contemporary contributions to the debate and were able to list 17 obstacles based on the nature of research, 10 based on the nature of the policy making environment, and 8 due to defective communication (Weiss and Bucuvalas 1980). Examples of this obstacles include mismatch in problem formulation, lack of general theories in social science that can be applied to specific contexts, or policymaker’s lack of time and inability to express their knowledge needs. However, when Lester and Wilds constructed a conceptual framework to consolidate this subfield of research, they noticed that “there is no dominant model or set of variables in the literature to predict use by decision makers” (Lester and Wilds 1990, 314). In their proposed model, the independent variables are of two types: 1) contextual factors, like the nature of the problem and its political feasibility, and 2) technical factors, like the timing and size of the study. In this framework, bureaucratic factors like the decision-maker’s style and participation are the intervening variables (Lester and Wilds 1990, 316). Their framework allowed to generate a conceptual model oriented towards prediction in which the two communities lose explanatory power. In fact, the two communities model had rapidly been uncovered as a metaphor rather than a theory, with very little predictive power due to the amount of
undiscussed assumptions and the compatibility with different models of interaction (Dunn 1980; Wingens 1990).

Set against this historical background, the metaphor appears as a useful analogy for confronting the challenge, both political and epistemic, that the lack of utilization of publicly funded R&D posed in the public debate about the role of the behavioral sciences in a democratic and knowledge-based society. However, the lessons that this meta-scientific agenda uncovered had very little use for understanding the role, function and influence of independent organizations involved in producing and communicating information aimed at the public policy community.

After little progress in the debate since the literature boom from the mid-1970s, new contributions on evidence-based policymaking have appeared in recent years. Thanks to these voices, the problem of research use is beginning to transcend the simplistic model of knowledge producers and knowledge consumers towards an eco-systemic understanding of networks of actors and organizations. Bogenscheider and Corbett have pushed beyond community-based theories to uncover the notions of professional and institutional cultures, defined as the “underlying values that inform the way [an organization] does business, [which], in turn, shape the way the agency functions and makes decisions” (Bogenschneider and Corbett 2010, 83). The notion of institutional culture is supported by that of core technology, borrowed from James Thompson’s 1967 classic *Organizations in Action*, which refers to “the core activities or functions performed by an agency or program” (ibid.). In this sense, while some agencies’ core technology is to make laws, the foundation’s core technology is to make grants, and that of think tanks is, among other roles, to generate ideas and connect them to the right personalities (see section 2 of this chapter).

Others have devoted their attention to the role of intermediary organizations, especially in the promotion of research for political objectives (Scott et al 2014). In particular, foundations have been recognized as hubs that can determine “which organizations are best positioned to produce, disseminate, and promote research for policymakers and the broader public”. (Scott and Jabbar 2014, 254). When foundations support
both implementation and evidence base through their grantees, they can create research “echo chambers” (Scott and Jabbar 2014; Lubienski 2009).

Although evaluation is one of the main technologies with which foundations support the decision making processes underlying their giving, until now the strong tie between the role of foundations and the function of evaluation in determining the influence of foundation-driven intermediary organizations has escaped the attention of academics. But before screening the scholarly debate on evaluation and on independent policy research organizations, it is useful at this stage to examine more closely the concept of knowledge use by policy makers.

It is important to notice that the use of research for decision making in the realm of public policy is not exclusively embodied in the figures of public officers. In a very strict sense, evaluations commissioned or conducted by foundations also offer data captured with methodologies drawn from the social sciences, and are intended to inform decisions that often have a direct bearing on public policy. For that reason, the following paragraphs on the notion of use will resonate in the later discussion on evaluation and its uses by foundations.

2.1.1 Highlight: Carol Weiss on knowledge use and policy influence

The work of Carol H. Weiss spanned from the very early debate on knowledge utilization to the exploration of the non-academic channels by which scientific information influences decision makers. Both a sociologist and an evaluation specialist, Weiss did not draw a clear line between evaluation, social science and policy analysis (Shadish et al. 1991). Therefore, her legacy deserves special attention because it can be applied in two separate sections of this study. Firstly, in studying how the social sciences can influence public policy, Weiss paved the path to understanding the modes of information transfer between the two communities. In the following pages it will become clearer how think tanks are part of those intermediaries, and how Weiss contributed to the theoretical understanding of think tank influence. Secondly, her theory
of evaluation as an activity that is political in nature can be applied to the relations between foundations and think tanks despite the fact that she never addressed them explicitly.

In 1980, Weiss and Bucuvalas published *Social Science Research and Decision Making*, an empirical study that presented summaries of reports to Public Health officers and prompted them to identify which documents would be more useful to them. The study offers two important contributions. Firstly, it identified the main dimensions that explain research relevance: the factors that contribute to the credibility of a report, its Truth Test, are both its quality and its conformation to the expectations of the user; whereas the Utility Test is composed by the orientation to action and the challenge to the status quo — the only unexpected factor. Although the representativeness of the study is limited, these four factors are a useful guide for ensuring relevance by packaging findings of high quality research in action-oriented, non-academic formats adapted to specific audiences.

Secondly, and perhaps more importantly, Weiss and Bucuvalas advanced a broader definition of use that escapes the narrow metaphor still underlying in the notion of *knowledge application*, where the evidence is represented as an external tool that can be used to shape reality. What Weiss called a “policy nugget”, or Nutley and Webb have called “decisionism” (Nutley and Webb 2000), represents an instrumental use that happens only under extraordinary circumstances:

Research directly relevant to an issue up for decision, available before the time of decision, that addresses the issue within the parameters of feasible action, that comes out with clear and unambiguous results, that is known to decision makers, who understand its concepts and findings and are willing to listen, that does not run athwart of entrenched interests or powerful blocs, and that is implementable within the limits of existing resources (Weiss and Bucuvalas 1980, 10).

However unlikely, *policy nuggets* sometimes occur and are sought. It has become an often-told, pivoting anecdote in the history of think tanks that the decision of the American Enterprise Institute to withhold a report on supersonic transport for fear of affecting a vote inspired Edwind Feulner to found the Heritage Foundation in the hope of effecting precisely that kind of timely, decisive influence (Smith 1993, 198).
But instead of elaborating on the factors that could maximize the political success of research in the unfrequent instrumental model, Weiss offered an important and much cited typology of five models of knowledge use (see Davies et al. 2005, or the simplified version reported by Pestieau 2003). Since Weiss referred to this typology several times throughout her career with slight variations, the following can be considered a good consolidated version:

- **Instrumental uses** occur when research knowledge is directly applied to decision making to address particular problems.
- **Conceptual uses** occur when research influences or informs how policy makers and practitioners think about issues, problems, or potential solutions.
- **Tactical uses** involve strategic and symbolic actions, such as calling on research evidence to support or challenge a specific idea or program, such as a legislative proposal or a reform effort.
- **Imposed uses** (which is perhaps a variant on instrumental uses) describe mandates to apply research knowledge, such as a requirement that government budgeting be based on whether agencies have adopted programs backed by evidence.
- **[...] Symbolic or ritualistic use** in the organizational practice of collecting information with no real intent to take it seriously, except to persuade others of a predetermined position or even to delay action (Leviton and Hughes, 1981; Shulha and Cousins, 1997). (Prewitt et al 2012, 38)

Classifications of use for specific applied knowledge practices can be mapped into Weiss’s typology. The main political uses of policy analysis, like the legitimation of political choices or the containment of a program are tactical; while the control of bureaucratic agents can be classified as an imposed use (Jenkins-Smith, 1990, 47). Evaluation studies, on the other hand, have three main intended uses: rendering judgments, facilitating improvements, and generating knowledge (Patton 1997, 65; Chelimski & Shadish 1997). While the generation of knowledge is oriented towards academic principles (conceptual use), judgments can be considered instrumental applications of the evaluation findings on the merit of a program, when there is a readiness for decision; but also, depending on the circumstances, tactical or imposed uses; whereas improvements are likely to be facilitated by an accumulation of evidence in a conceptual use.
That typology served as a baseline for Weiss’ research agenda “to understand the conditions and processes of research use”\(^1\). The model to which Weiss devoted more attention is the conceptual use of research by means of a process that she called Enlightenment (Weiss 1987). Central to her argument is that, contrary to the decisionist model, no single decision maker has the power to shape policy. Instead, public officers work in a complex stream of inherited, interdependent processes that constrain possible action so that decisions happen by *accretion*.

Small choices over time, each one barely noticed and made with little rational review, leave few options open; they lead with seeming inevitability to “the only thing we could do under the circumstances”. (Weiss 1976, 226)

It is in the context of these limitations that social scientific knowledge gradually diffuses and informs, but also loses its identifiers. It has been anonymized, re-elaborated, appropriated.

Bits of information seep into his mind, uncatalogued, without citation. He finds it very difficult to retrieve the reference to any single bit of knowledge. If we ask him about the effect of social research on his decisions, he usually will not be able to give an accurate account - or even be aware that he derived his ideas from the social sciences. (Weiss 1977b, 534)

The theory of Enlightenment points at the major obstacle for any individual, organization or campaign that seeks to track message acceptance in the policy community, and why it is an accepted practice to frame impact only in terms of *contribution*, never of *attribution*. The greatest influence on policy is not obtained by single research outputs, but by the accumulation of findings from different sources that make their way to common knowledge, (Lindblom and Cohen 1979). Weiss, in fact, stresses the importance of non-authoritative sources to inform and enlighten policy makers, who don’t have neither the incentives nor the time to read academic journals. That communicative function happens through “micro-processes” enacted by different types of intermediaries, like prominent newspapers or interest groups. Mass media are an important carrier of information because the main outlets have the power to determine what the people

---

\(^1\) This vision was endorsed in 2012 by the National Research Council’s proposal of a research framework for the future that includes investigations into the “psychological processes influencing [the] acceptance” of policy arguments, cognitive operations, “as well as institutional logics, practices, and cultural assumptions”. Op. cit., p. 54
know and react to, creating a “common language” in terms of the types of issues and problems that constituencies care about:

But from the policymakers’ perspective, media reporting of a study or commentary on a topic in their domain is a significant event. Policymakers prick up their ears, not only because Time or the Washington Post reaches them with a brief and simple version of social science, but even more important, because they know that the same story reaches all the other players in the policy game. (Weiss 1986, 279)

In other words, assuming the image of the two communities for the sake of the argument, media outlets are one of the primary channels for transferring knowledge from the scientific community to the policymaking community. What information is reported in newspapers is likely to be shared, commented orally and recollected at a later point, gaining authority through this socialized process. The second micro-process relates to the social and political nature of policymakers: interest groups are relevant because they already speak in political terms that are easy to interpret and related to possible courses of action. Additionally, through professional lobbyists and all sorts of professional connections, they cultivate personal relationships that are allow to present an argument succinctly and in oral form, a strategy that is more likely to persuade busy politicians with no time to read in-depth investigations (Weiss 1989). Interest groups representatives are, however, only a part of “the issue network”, a term that Weiss borrows from Hugh Heclo (1978), and that includes think tank experts.

Weiss described think tank experts as members of an “analytic community” with a “mediating function” (Weiss 1992), a categorization that hardly captures the complexity and flexibility of this organizational model. The following section explores the notion of “think tank” and ties it to the problem of knowledge utilization.

2.2 Placing think tanks in the knowledge utilization debate

While there is no agreement on how to define what a think tank is, a classic classification distinguishes between universities without students, like the Brookings Institution and the American Enterprise Institute; contract research organizations, like the Rand Corporation and the Urban Institute, which perform most of
their research through contracts with government agencies; and advocacy tanks, like the conservative Heritage Foundation, or the liberal Center for American Progress, which “combine a strong policy, partisan or ideological bent with aggressive salesmanship and an effort to influence current policy debates” (Weaver 1989). What all these organizations have in common is that they are formally independent and conduct public policy research. But even limiting the discussion to the United States, similarities start to dissolve there. A difference that is important for this study is that, while universities without students and advocacy tanks receive primarily private funding, contract research organizations usually work for government agencies.

It is also not clear what think tanks do: “The briefest answer is that think tanks may do several different things, but not all think tanks do the same thing” (Weaver and McGann 2000, 5). Weaver and MacGann signal five roles: 1) basic research, 2) advice on issues of current concern to policymakers, 3) evaluation of government programs, 4) facilitators of issue networks, and 5) suppliers of personnel for government. Contract research organizations concentrate their activities in the third role, but the remaining are practiced by organizations across the spectrum. In other words, think tanks have at their core two main activities: researching and networking. It is in this duality that the question about the influence of think tanks should be asked. Since there is not a one-to-one logical correspondence between the typology of knowledge utilization and the five roles listed above, which can be fulfilled in uncountable ways, the use of think tank’s knowledge production remains an empirical question. The purpose of the analytical chapters in this study (Chapters 4 and 5) is to survey whether foundations and think tanks are responding to that question through evaluation studies.

Think tanks span the divide between the two communities through brokerage: “Brokering involves filtering, synthesizing, summarizing, and disseminating research findings in user-friendly packages.” (Prewitt et al. 2012, 45). Think tanks have “boundary transcending qualities” (Krause Hansen et al. 2002), and work as “boundary organizations”:
Boundary organizations are institutions that straddle the apparent politics/science boundary and, in doing so, internalize the provisional and ambiguous character of that boundary. Negotiating these elusive qualities becomes the daily work of the boundary organization (Guston 2000, 30).

Their boundaries, however, are multi-dimensional, straddling a precarious equilibrium between the fields of media, academics, politics and business (Medvetz 2012). In this multi-dimensional middle ground, it is particularly difficult to tell apart facts from values, scientific knowledge from political judgments (Jasanoff 1990, 229). In exercising their expertise and defending their role as “neutral transmission belts between science and policy”, think tanks can also act as gatekeepers and draw additional boundaries with voices that tend to remain unheard by policy makers, including the layman, the non-expert advocate (Stone 2007, 276) and the liberal academic (Medvetz 2012; Smith 1993).

In the contribution paradigm, the unit of analysis for the application of social science in public policy contexts should be dislodged from the isolated report and the single organization to a more macro-level approach that can consider how contributions from different sources and actors coalesce around an issue. This realization has dramatic consequences for any institution seeking to make the case for their influence controlled in separation from the environment in which they act.

An approach that has the virtue of offering “policy subsystems” as a proper unit for explaining the role of knowledge in policy change is the Advocacy Coalition Framework (Sabatier 1988), in which subsystems are simply defined as “the set of actors who are involved in dealing with a policy problem such as air pollution control, mental health, or energy” (Sabatier 1988, 138). ACF thus transcends descriptions based on institutions and on individual actors, and proposes ‘advocacy coalitions’ within policy subsystems as an aggregate of actors with common goals. While the empirical exploration of the role of think tanks in such coalitions is still underdeveloped, some analysis have pointed at their possible function as policy fora for cross-coalition learning (Stone 2004, 102; Smith 1993, 212), an important function of policy oriented learning predicted in ACF that seems to validate the secondary mediation role of the analytic community (Weiss 1992).
That the role played by organizations labeled as think tanks cannot be described in absence of their situated context is clear. However, the ACF offers interesting explanatory powers that are highly compatible with the general characteristics of think tanks: in particular, the independence of each institute to set its own research agenda, and the freedom of scholars with low institutional allegiance in environments that usually lack rigid hierarchies speak of an organizational type that is particularly apt for tapping into the processes by which knowledge and learning are used in public policy issues.

Different think tanks have different models for internalizing boundaries, very often in dependence of the particular “ideas, values or philosophies that serve as their organizing principles” (McGann 1995: 64), and of the degree of aggressiveness towards policy change. The institutional diversity translates in an array of professional profiles for which there is not a consolidated typology. While Smith put forward a distinction between scholar statesmen, policy specialists, policy consultants, government experts, media professors, and policy entrepreneurs (Smith 1993, 225), Medvetz’s theory of an interstitial field unfolds in four ideal types, one for each type of mother-field: wonks (professors-academia), hacks (lobbyists-politics), quotemeisters (pundits-media), and policy entrepreneurs (fundraisers-business). These typologies allow to distinguish think tanks according to their preferred notions of expertise and credibility, which vary from a rationalistic, post-Enlightenment model of relevance sustained on academic quality to the Heritage Foundation’s belief that “people are policy” (Smith 1993, 202). As a consequence of the institutional model, different think tanks are more likely to engage at different phases of the policy process: issue articulation, policy formulation and policy implementation (Abelson 2010, 56). The engagement with the policy process through boundary work suggests the importance of networks for influence, as both the material driving factor and the unit of analysis for gathering evidence about these organizations and their members across time. However, researchers should be wary of the warnings that longitudinal network analysis of policy subsystems may be too costly (Sabatier and Pelkey 1987).
On the other hand, these typologies may contribute to losing from sight the fact that most think tanks host policy analysts, and that cost-benefit analysis is perhaps the most common evaluative methodology for informing decision making in public policy.

Even more instructive are the possibilities for deploying policy analysis in confrontations within policy subsystems. The strategic alternatives for disputing a prevailing policy proposal are fueled by “political and analytical debates”:

Those who feel themselves aggrieved by the proposed policy and have the resources to do something have a number of options. They can:

1. Challenge the validity of the data concerning the seriousness of the problem
2. Challenge the validity of causal assumptions concerning:
   a) Technical aspects, e.g. the links between emissions, ambient air quality, and health effects
   b) Institutional arrangements which will provide the necessary changes in behavior (Sabatier 1988, 152)
3. Attempt to mobilize political opposition to the proposal by pointing to costs to themselves and others, i.e. by creating/enlarging their coalition. (Jenkins Smith 1990, 93)

The styles of analyst participation in the debates can also be inflected, if not in the skill deployed, in the degree of freedom and institutional obedience: the objective technician, the issue advocate and the client’s advocate (Jenkins-Smith 1990, 109:116). While the client’s advocate working for a think tank would be but a lobbyist in disguise, the objective technician has also been widely criticized, leaving the issue advocate as the better candidate for participation in the debates. A similar idea is supported by Bogenschneider and Corbett, who defend the notion of advocacy with a lower case “a” as “bringing the needs of a particular group to policymaker’s attention, but without lobbying for a particular policy option” (Bogenschneider and Corbett 2010, 228).

That think tanks contribute to numerous debates within policy subsystems is quite obvious, but the exact footprint of their efforts is very hard to measure. Some findings suggest their involvement intensifies precisely when the debate has become highly politicized and entrenched, facilitating tactical use of their outputs but offering little room for enlightenment (Rich 2004, 26). However, few generalizations can be
made on the modes of engagement. The existing scholarship on think tanks has combined still limited case studies to track the influence on selected debates with a more systematic attempt to understand the holistic influence of think tanks as institutions based on some of the metrics used by think tanks for tracking outputs (see Chapter 4). While the standard metric of academic impact is the scholarly citation, media citations, especially in prominent newspapers that are likely to be noticed in Washington, DC circles, like the Washington Post, the New York Times, and the Wall Street Journal, are the first indicator of visibility of a think tank (Rich 2004). Whether it is an Op-ed, a commentary on breaking events, or coverage of a report, media citations are “standardized packages” (Guston 2000) and necessary points of passage with different meanings for different actors. For the expert, they are a public endorsement, a recognition of their authority; for the journalist, a selective procedure for placing items in the agenda; for the policymaker, they flag issues that might become relevant. The link to the public opinion, vague as it might be, is also, significantly, one of the most powerful avenues for democratic life. Other useful indicators are congressional testimonies and appearances before committees and subcommittees, as well as consultations (Rich 2004; Abelson 2010). However, indicators have an important margin of error. Both Rich and Abelson recognize that visibility and influence are different concepts, and that therefore the most visible institutions are not necessarily the most influential. In fact, Rich’s findings suggest that an excessive orientation to the marketing of ideas can be harmful to the credibility of a think tank (Rich 2004, 83).

Abelson is quite skeptical of the effort to gauge the impact of policy research:

Documenting how think tanks attempt to influence policy is a reasonably straightforward exercise, but determining how to assess their impact at various stages of policy-making is not. (Abelson 2010, 15)

Instead, his book complements the visibility metrics with case studies that analyze some of the different strategies by which think tanks have sought to intervene in the policy process: preparing the arrival of new administrations in the US, and the preliminary conferences that led to the Charlottetown Accord in Canada. Abelson selected his case studies for the exceptional notoriety of the think tank involvement and the
simplicity of identifying the relevant institutes. The implication is that this exercise is impractical in most real scenarios.

In conclusion, while advocacy coalition frameworks are not apt for analyzing think tanks at the institutional level, they might be proper lenses for understanding the contribution of particular programs to a debate and for examining interactions with other members of policy subsystems. This is particularly true with regards to the formal and informal connections between funders, analysts and other advocates. In contrast, the practiced profiles of institutional visibility have protected the view of think tanks as actors with intangible value proposals and unmeasurable impact, the influence of which can only be gauged in terms of media coverage and congressional testimonies.

The following section reviews the adoption of evaluation research by foundations. The increasing importance of evaluative research to support strategic considerations is commensurate with ACF and the multi-actor perspective on sub-policy systems. The fourth and last section of the chapter offers an overview of the techniques for measuring policy outcomes that can be adopted by evaluators, including the most widely adopted tools in the fields of advocacy and policy change, and international development.

2.3 Foundations and evaluation

The adoption of evaluation by foundations is fraught in an undercurrent of tensions and paradoxes. Aimed at maximizing efficiency and providing a process of institutional accountability, more than four decades after the first foundation pioneers started practicing the applied social science methods that had been embraced during the Great Society, there is not a clear sense of the extent to which evaluation studies support the philanthropic mission and inform decisions – yet the scale of adoption has not stopped growing. Foundations started practicing program evaluation in the 1970s, precisely in the same years when the accumulation of social scientific research left unused was being marked as problematic. As Congress was debating how to limit the undue influence of big fortunes through what would become the 1969 Reform Tax Acts, in 1968 John D. Rockefeller III convened the Commission on Foundations and Private
Philanthropy, also known as the Peterson Commission because it was chaired by renowned industrialist Peter G. Peterson\textsuperscript{i}. The recommendations of the final report published by the University of Chicago two years later included a section on “evaluation of programs, government and not government, and impartial appraisals”. Although this section proposed foundations to sponsor “impartial appraisals of government policies and programs” and other established institutions, it also mentioned the benefits for philanthropic self-evaluation in terms of public accountability:

\begin{quote}
The foundations, with no electorate or competitors and no need to raise money, are subjected to fewer pressures to measure and improve the performance than are most other institutions. Precisely for this reason, some periodic review of foundation performance would help reduce some of the mystery from foundation activities and convey a sensitivity to the public that has not always been present. (Commission on Foundations and Private Philanthropy 1970, 130:131)
\end{quote}

Although the immediate influence of the report was limited, thanks to the leadership of Orville G. Brim Jr., recommendations were taken up by the Russell Sage Foundation (Hall 2004, 32), which had funded important contributions to evaluative research. Brim, who had been President of the Foundation until 1972, published the essay “Do we Know What We’re Doing?” in which he criticized the unexamined conventional wisdom at the base of foundation’s decisions and recommended the adoption of scientific evaluation to gain knowledge at five levels: “the ‘who did what to whom’ type of inquiry” at the project level, comparisons between similar projects, between programs, between foundations and between the goals and activities of different foundations (Brim 1973, 220-221). The commitment was later joined by other pioneering organizations, like the Ford Foundation or the Robert Wood Johnson Foundation (Hall 2004, 35).

Much like in governmental settings, unrealistic expectations on the usefulness of research were frustrated from the outset. Instead of providing answers, evaluations seem to unearth more problems and new questions. Hall’s historic review shows that, over and over, the evidence on how foundations use evaluation has pointed at many problems in their managing practices. (Hall 2004). For David Rogers, President of the
Robert W. Johnson Foundation until 1987, the benefits of evaluation were forging agreement between multiple staffers with different backgrounds and value systems:

Rather than producing objective measurements of the impact of foundations interventions, its primary value was to reduce uncertainty and disagreement within grantmaking organizations (Hall 2004, 39).

As Hall notes, this had already been perceived by Weiss: “evaluation is not meant to inform grantmaking as much as to help the organization” (Weiss 1973). Practitioners and experts still perceive reporting may as a compliance function (Patton et al. 2004, 83), an oftentimes merely ritualistic procedure:

Foundation reporting often reverts to an empty ritual in which grantees report to evaluators, evaluators report to program officers, and program officers report to trustees, each going through motions without using the information in any meaningful way (Crutchfield et al. 2011, 169).

Evaluation can also serve a number of covert purposes like postponing decisions, ducking responsibilities, advertising an initiative or merely fulfilling grant requirements (Weiss 1972, 11). Tactical and symbolic uses of evaluation are not an exclusive attribute of foundations – much rather they follow program evaluation wherever it’s practiced.

2.3.1 What a traditional program evaluation looks like

Figures 1 to 3 are an illustration of traditional program evaluation. They quote and summarize the insights contained in Carol H. Weiss 1972 classic Evaluation Research. Figure 1 offers a list of some of the most common decisions that can be informed by evaluation findings. These are decisions concerning support to the program, program improvement, adoption or rejection of strategies and techniques, and program adoption in new locations. Figure 2 illustrates what are the basic steps in designing and conducting an evaluation, from determining the goal to designing indicators, capturing data and comparing it. Figure 3 examines some of the complexities that evaluators confront and offers an account of why the results of evaluation tend to provide new insights, and do not provide a base for binary decisions. To this regard, it’s important that Weiss underlined the use of evaluation for learning: “In practice, evaluation is most often
called on to help with decisions about improving programs. Go/no-go, live-or-die decisions are relatively rare.” (Weiss 1972, 16).

- To continue or discontinue a program
- To improve its practices and procedures
- To add or drop specific program strategies and techniques
- To institute similar programs elsewhere
- To allocate resources among competing programs
- To accept and reject a program approach or a theory

(Weiss 1972, 16)

Figure 1: Decisions evaluation is called upon to support

- Find out the program goals
- Translate the goals into measureable indicators of goal achievement
- Collect data on the indicators for those who participated in the program (and for an equivalent control group who did not)
- Compare data on participants (and controls) with the goal criteria

(Weiss 1972, 24-25)

Figure 2: Simplified model of a traditional program evaluation

- Programs often have multiple goals or ambiguous goals.
- Programs have unintended consequences: they accomplish other things, sometimes in addition and sometimes instead.
- There is no easy way to separate the essential from the accessory in the array of people, structures and activities that make up the program in order to inform decision makers about transferability. What features make the program work?
- The evaluation question asks how well the program performs, but not why it works.

Adapted from Weiss 1972, 25

Figure 3: Why the simplified model doesn’t work
2.3.2 The expansion of evaluation practices

In a sector where many grantmakers still give expressively, the principle of effectiveness is hardly ubiquitous, and many of the smaller foundations can hardly afford to conduct rigorous serious evaluation (Hall 2004, 48). But despite the early warnings, the rhetoric about the benefits of evaluations on foundational performance is pervasive. Evaluation has surged in the world of organized philanthropy due to a combination of factors, including the increasing bureaucratization of larger organizations and the professionalization of nonprofit management (Frumkin 2008), the business orientation and the focus on results of many philanthropists, the appearance of new tools, the emergence of strategic philanthropy and the role model of larger, evaluation-oriented foundations, which exercise a gravitational effect over the field as a whole (Coffman et al 2013, 37).

Studies to portray the state of the art in foundation evaluation are not a new idea. A survey conducted in 1997 to a population of 959 philanthropic foundations in Michigan found that 45% of respondents used evaluations to make grant proposal decisions, and 41 percent used post-project evaluation. Paradoxically, while “meeting project objectives” was rated as the most useful evaluation strategy, it was also the least practiced, at only 21.2 percent (Alie and Seita, 1997). The authors concluded that, given the little information about the field, “we have no idea how to improve evaluation efforts”.

A survey conducted to 31 foundations in 2012, including 26 organizations in the US, found that the commitment to evaluation had increased both in terms of investments and in staff of exclusive dedication (Coffman et al. 2013, 39:40). The survey also found the role and scope of evaluation to be expanding:

> Evaluation-unit responsibilities now go well beyond managing evaluation contracts or assessing individual grantee results. Many are now leading a range of evaluation practices that include performance management, knowledge management, organizational learning, and strategic learning. (Coffman 2013, 40)

Survey data also supported a strong thesis according to which strategy can be treated as an evaluand. Adapting Henry Mintzberg’s notion of strategic planning, Patton and Patrizi suggest that “evaluation involves, in part, comparing rhetoric about strategy with the behavioral realities of how the organization
operates strategically, thereby helping an organization separate the rhetoric from the reality of their work. (Patton and Patrizi 2010, 26). The early focus on program evaluation and service delivery has shifted towards integrated, systems-wide learning. Policy change is one of the areas of activity where evaluation has evolved most. The *Guide to Measuring Advocacy and Policy commissioned* by the Annie E. Casey Foundation recognized that, despite the “growing desire to gauge the impact of investments in this area […], there are no common approaches”, “no standards of practice, acknowledged best practices, nor commonly used methods or tools” (Reisman et al 2007, 1:2). Significant progress has been made in the last decade, including some of the tools and techniques illustrated at the end of the next section. Traditional evaluation approaches have been adapted to specific outcomes, and sophisticated interviews protocols to key informants have replaced randomized control trials as the predominant technique.

Public policy evaluation methodologies can aim at grantee performance, strategy-level success, specific strategy indicators or grantee capacity (Sherman and Peterson 2007, 215). Crutchfield et al. recommend switching to a learning culture by 1) using information prospectively and on an ongoing basis, 2) helping grantees adopt a learning culture, 3) following a “three-step learning cycle” of planning, implementation and assessment, and 4) shifting the analysis to the system level (Crutchfield et al. 2011, 182). Evaluation has also been defended as a democratizing tool that can compensate for top-down, paternalist program design by opening up to participatory processes (Greene et al. 2004).

It is this expansion beyond program assessment that supports the reasonableness of the research question that makes the subject of this study. If until this moment the question was posed at a rather abstract level, it can now be anchored in an empirical trend: Has the expansion of evaluation practices reached public policy research? Has it extended to the relationships with nonprofits that do not fall under the sphere of social programs and social movements?

The field-building effort spearheaded by the Casey Foundation offers some hints about the role of research under their framework. Many of the activities supplied by think tanks, like policy development, policy
analysis and debate, scientific research, or preparation of working papers, are examples of strategies for two of the main outcome categories: strengthened base of support and improved policies (Reisman et al. 2007, 19). But the question on how to evaluate public policy research is still underdeveloped. What is perhaps the only published attempt at framing the evaluation of public policy research under the prism of policy change and advocacy outcomes offered six propositions (Devlin-Foltz 2006):

1. Develop a theory of change in collaboration between the research organization and the funders.
2. Establish clear benchmarks to guide planning and continuous evaluation.
3. When possible, include an external evaluator to obtain objective data on how research relates to policy outcomes.
4. Determine how to measure influence in presence and in absence of windows of opportunity.
5. Adopt the standard of contribution, instead of attribution.
6. Determine who the research users are.

A second reference in the literature is Best Practices for Funding and Evaluating Think Tanks, a report commissioned by the Hewlett Foundation to inform the selection of think tanks in developing countries as grantees under the Think Tank Initiative (McGann 2006). Although the selection criteria and parameters identified by McGann are not applicable to the donor/grantee relationship in the United States, the report signals the inception of a paradoxical situation in which the pioneering work on the monitoring and evaluation of policy research organizations is being conducted in developing countries under the TTI flagship, rather than in the collaboration between foundations and think tanks in America. A decade after the debate on the evaluation of advocacy outcomes hit policy research organizations, it is time to take stock on what is actually happening on the domestic field.

From the perspective of organized philanthropy, the inclusion of knowledge management, organizational learning and strategic learning as competencies of evaluation units seems to support the idea that evaluation is evolving from an objective assessment of external performances to a complex information technology
that seeks to provide staffs and boards with the best possible information, supporting the sharing of data, experiences, findings and possibly generalizations in ways that are applicable to day-to-day tasks. The question of whether think tanks’ products and activities may be falling under this expanding scope seems more reasonable given their expertise in producing usable information in a timely manner. The next section will provide more insights on existing approaches to understand what the role of public policy research can be in larger policy change initiatives and what techniques and methodologies have been tested to make the argument about the influence and outcomes of public policy research.

2.4. Evaluation as outcome assessment of public policy research

When thinking about influence, think tanks must recognise that there is a difference between their goals and what they should be measuring. It is possible to recognise a long-term goal as a direction of travel or as a rallying call for an organisation or a network, while being perfectly comfortable measuring changes closer to home.

Enrique Mendizábal (2016)

The consensus of the main scholars on think tanks points at the difficulty of tracking influence. However, given the incentives of funders to grasp the impact of the programs that they fund, in the last decade there have been important attempts at evaluating the influence of policy research programs and using strategic assessment tools for organizational learning. These efforts have experienced a particular degree of sophistication in two different areas: international development and strategies for policy advocacy. This section introduces some of the main methodologies that have been deployed in the field.

In the context of this study, evaluation is defined as outcome evaluation of activities consisting in the conduction and dissemination of public policy research. This definition follows Pestieau’s adaptation of Patton’s criteria: “At a general level, evaluation is the systematic collection of information about the operations and the outcomes of an activity to make judgments about it, improve its effectiveness and inform decisions about its future.” (Pestieau 2003, 7). The orientation to outcomes is a main thread in the literature on evaluation. However, conducting outcome evaluations of public policy research is a relatively new trend.
that has found pioneering efforts in the work of scholars affiliated to leading developmental think tanks. Particularly salient is the work on outcome mapping developed by Sarah Earl, Fred Carden and others at the Canadian International Development Research Center (IDRC) (Earl 2001), or Ingie Hovland’s work on the RAPID Outcome Assessment (RAO) at the British Overseas Development Institute (ODI) (Hovland 2007). There have not been to date any comprehensive studies on the adoption of these or similar techniques in domestic American think tanks.

The selection of foundations and think tanks as representative models of intermediary organizations involved in such modality of evaluation does not imply the assumption that influential research happens only or mostly in this connection; much rather, it recognizes that, while this is only a relatively small component in the production of policy relevant knowledge in the United States, both organizational types do collaborate on a regular basis and have incentives to assess influence beyond the indicators reported in section 2 of this chapter. On the one hand foundations, as seen in section 3, have an increasing interest in assessing the effectiveness of their giving, combined with a diminishing pudor for funding initiatives aimed at influencing policy. On the other, since think tanks have one of their primary sources of funding in foundations, making the argument about their influence with more scientific tools is likely to be an important strategy for their long term sustainability.

Hence, this study is based on the assumption that, given the difficulties of asserting links between knowledge production and decision making, and the flexible position of think tanks as intermediary organizations, one of the main criteria for evaluating public policy research must be a pragmatic one: whether research is used, and how it is used. Assessing outcomes means as much as elevating the view above the activities and outputs that fall within the reach of the organization to track the life of those messages and their reception by targeted audiences.

The focus on outcomes sets other assessments in the background. Quality assurance and research integrity policies, for example, are essential safeguards of credibility enacted by many prominent organizations.
Many of the more academic organizations, like Brookings and AEI, have peer review processes for their books and main papers that are modeled after academic journals (Weidenbaum 2011). More examples are possible. The Peterson Institute for International Economics and the Center for Global Development, both based in Washington, DC but with very different areas of expertise, have policies that encourage the replication of their analysis by sharing their datasets and code. Such replications, rare as they may be (McCullough and McKitrick 2009), do indeed evaluate the accuracy of academic findings. All of these internal policies are important, and they are very significantly tied to the credibility and outreach of a think tank. However, the superabundance of information in the knowledge societies implies that quality research is ignored every day. To evaluate what information works, why it matters and how it is used, many other factors have to be taken into account.

The focus also leaves out evaluation in the sense of policy analysis. Many think tanks conduct evaluations of federal, state or local programs on a regular basis. Program evaluations of this nature are also conducted by internal agency evaluators or by independent evaluators with the aim of providing objective, non-partisan information about the efficiency of such programs, devise possible improvements and alternatives:

Evidence-based policy is a rigorous approach that draws on careful data collection, experimentation, and both quantitative and qualitative analysis to answer three questions: What exactly is the problem? What are the possible ways to address the problem? And what are the probable impacts and costs of each? A fourth question that figures into all public policy decisions—What political and social values do the proposed options reflect?—is largely outside the scope of evidence-based policy. (Dunworth et al. 2008: 1)

In those cases where policy research programs consist in evaluations of public policies, outcome evaluation becomes a form of meta-evaluation.

2.4.1 Basic concepts in outcome evaluation

Some basic distinctions can be made depending on the purpose and focus of an evaluation. From the point of view of the purpose, the distinction between summative and formative evaluation, coined by Michael Scriven in 1967 in an educational context, has become a standard (Scriven 1967). “Summative evaluation provides data to support a judgment about the program’s worth so that a decision can be made about the
merit of continuing the program” (Patton 1997, 68), whereas formative evaluation focuses on extracting learning from implementation with the aim to applying it to future improvement (ibid.).

It’s important to note that the two basic typologies of evaluation are not mutually exclusive. A formative evaluation, for example, can be outcome-oriented or process-oriented. Evaluation is a highly contextual activity, and its value depends on its intended use (Patton 1997), but even given a particular context, there is no perfect recipe for what an evaluation should examine. This world of possibilities that can be challenging, and usually requires a strategic integration of data collection and evaluative approaches.

Another distinction of particular relevance for this discussion affects the focus of what is measured. Process evaluations examine the implementations of activities and how they correspond to planning; whereas outcome evaluations examine the footprint of those activities on the environment. Examining an important event, process evaluation may collect how many invitations were sent and how many people attended; whereas an outcome evaluation would, for example, survey attendants to gain feedback on how the event was useful for their work. In other words, the objective of outcome evaluations is to link activities or outputs, over which the organization has control, to outcomes and impact. Although there is not an agreed upon use for these two terms, a quite extended use sees outcomes as the medium term effects that were achieved by the outputs, while impact refers to longer term goals at a systemic level.

One of the most usual tools that guide the integration of an organization in a cohesive evaluation strategy is the logic model, defined by the Kellogg Foundation as “a picture of how your organization does its work – the theory and assumptions underlying the program. A program logic model links outcomes (both short- and long-term) with program activities/processes and the theoretical assumptions/principles of the program” (Kellogg 2004, 1). Kellogg’s logic models arrange components in 5 categories: Resources or Inputs, Activities, Outputs, Outcomes, and Impact. Components are conditions that have to be met to proceed towards the fulfillment of program expectations. However, only resources and activities are sufficient to produce the planned output. In fact, Figure 4 offers a much cited visualization of the degree of
influence of an organization compared to the influence of factors within the given environment. This image is a very powerful reminder of how causal links between outputs and outcomes can almost never be ascertained causally, and the discourse on influence has to remain at the level of assessing modest contributions to change. In other words, a researcher has control over the phases of research planning and development. However, once the research has been conducted, it escapes her sphere of influence. The attention that the disseminated message will receive can never be predicted exactly, as it depends on many exogenous factors. The contribution to certain outcomes and impacts is even more difficult to assess.

![Relative Influence Along the Results Chain](image)

**Figure 4. Relative Influence Along the Results Chain**

(Smutylo 2001, figure captured from Earl et al. 2001)

Intended outcomes for policy research projects will depend on the style of engagement. Approaches characterized as non-partisan typically include changing the terms of a debate, developing a new area of research or positioning an issue on the agenda (Weaver and McGann 2000; Abelson 2010); whereas advocacy-oriented efforts can aim at influencing legislation, or improving the capacity of a target audience. However, recent approaches that characterize policy change outcomes at a higher level of abstraction may have the virtue of being applicable to different styles. Following the work of Keck and Sikkink (1998) and Steven (2007), Pasanen and Shaxson compile eight “policy change options to monitor and evaluate”: 1) Attitudes of key stakeholders to get issues on the agenda, 2) public opinion, 3) capacity and engagement of other actors, 4) change in discourse among policy actors and commentators, 5) improvements in
policymaking procedure/process, 6) change/no change in policy context, 7) behavior change for effective implementation, 8) networks and systems for supporting delivery (Pasanen and Shaxson 2016, 24).

2.4.2 Integrating processes

Different processes have to be coupled in order to evaluate organizations. The most immediate and low level is monitoring, intended as an “ongoing process to verify systematically that planned activities or processes take place as expected or that progress is being made in achieving planned goals and output/impact” (McGann 2006: 19). Monitoring captures data around indicators designed to measure the attainment of an objective. Another related activity are holistic assessments of an organization, usually “conducted as a learning process for both the donor and grantee” in order to identify areas of improvement “and guide capacity building efforts” (Ibid.). However, some experts have contested the view that an organizational assessment can be considered an evaluation. In a response to Johanna Morariu’s strategic assessment of the Post Carbon Institute, Enrique Mendizábal notes that “an organization […] does not operate in bursts of activity than then wind down to nothing. It does not start and then stop to be evaluated. Its parts may do, but not the organisation”. (Mendizabal 2011b). In Morariu’s own words, the work with the Post Carbon Institute is an evaluation framework, not an organizational evaluation per se. In this case, the independent expert guided the work of generating a theory of change for the organization, with “corresponding outcomes, indicators, and data collection tools” (Morariu 2011). Building such systems and infrastructures is a preliminary step for conducting specific evaluations, which can happen at the project, program or organization level, while making sure that they are aligned with the overarching goals that express the organization’s mission.

2.4.3 Evaluation and advocacy

In the United States, outcome evaluation of advocacy and policy change initiatives has experienced significant progress in the last decade. It has been suggested that when advocacy organizations embrace one of the many techniques designed to collect opinion data from key respondents, they should guide their
work on the idea of political will (Charney and Coffman 2015). The main assumption is that, by mobilizing the public about the prominence of a certain issue, that issue will start appearing on the agenda. To Charney and Coffman, public will has five components: 1) opinion, 2) intensity, 3) salience, 4) willingness to act, and 5) capacity to act (ibid). Each of them can be measured with polls to specific target audiences, and offers guidance on framing and messaging (see figure 5).

1. **Does the audience have an opinion on the issue?** If not, then frame it so they do.
2. **Are opinions intense?** If not, then increase their intensity by emphasizing the consequences of inaction or the promise of proposed solutions.
3. **Is the issue salient?** Show why it matters. Reveal policymakers’ positions on it.
4. **Are audiences willing to act?** Make the “ask” clear and doable. Convince the audience that doing the ask will have concrete results.
5. **Is there capacity to act?** If not, address the existing barriers or build the necessary skills.

(Charney and Coffman 2015)

**Figure 5. Five-fold screening of issues and advocacy efforts**

This concept does not match the orientation of nonpartisan think tanks towards the critical spaces of decision making. However, it can be useful for the discussion about the role about think tanks for several reasons. Firstly, the base concept of opinion corresponds with one of the most commonly stated missions of policy research organizations: to educate and inform. Secondly, the general adherence to this roadmap can be seen as what distinguishes an advocacy tank from a nonpartisan or academic think tank. This is not to say that organizations that don’t seek to galvanize partisanship are devoid of advocacy- Jenkins-Smith’s defense of the issue advocate should be a reminder of the differences between an institutional allegiance to a political direction, which many think tanks reject, and the issue advocacy individual scholars or programs can embark in when recommending policy formulations and supporting them with evidence without necessarily lobbying for them. Secondly, treating policy research as one component of broader strategies
deployed by networks of organizations that collaborate loosely around a common advocacy goal may support a better understanding of its potential uptake and effectiveness (Coffman and Beer 2015). Highlighting the necessary outputs in an overarching strategy can promote an awareness of how different nonprofits contribute differentially as tactical elements. In their Advocacy Strategy framework, Coffman and Beer organize tactics along two dimension: changes in awareness, will and action, and types of audiences (Public, Influencers, and Decision Makers). In this framework, policy analysis and research is classified as affecting changes in awareness and straddling the audiences of influencers and decision makers. Thirdly, advancing the saliency and politization of an issue may be an adequate tactic to describe the engagement of advocacy tanks, both in influencing the public opinion through mass communications and in providing advice with a partisan lens.

2.4.4 Annotated list of methodologies and techniques

Examinations conducted on particular strategies and programs can make use of a number of techniques, depending on whether their use is for diagnostic, monitoring or evaluation purposes. While many of the methods included, like questionnaires and interviews, have been traditionally part of social science research, recent innovations have adapted them for the evaluation of policy influence. Below is an annotated list of the main techniques that can be deployed for the monitoring and evaluation of public policy research, drawing from the two fields of international development and advocacy (Weyrauch et al. 2010).

Monitoring:

- **Questionnaires:** Questionnaires allow a certain degree of automation in the data collection phase, and also enable the construction of quantitative indicators in the analysis phase. With the aim of assessing policy influence, questions prompting respondents to report “the progress or results of implemented actions” should be based in the indicators and objectives of the initiative. (Weyrauch et al. 2010)
• **Focus Groups:** Focus groups are group interviews facilitated by an expert. Interviewees should have common traits (age, gender, qualification, etc.). All interactions are captured (usually video recorded) and transcribed for analysis, which is time demanding. (Weyrauch et al. 2010)

• **Reports from Field Workers:** This technique can be deployed whenever the program has a component of implementation in the field. Field workers can then record data, including their own and third party opinions, related to predefined indicators stemming from the policy influence objectives of the initiative. (Weyrauch et al. 2010)

• **Participant Observation:** A traditional data collection method in cultural anthropology, this technique stands out because the reporter is also an active participant. Used to analyze events and processes, it can be very challenging to participate and keep track of observations. (Weyrauch et al. 2010).

• **Media Tracking:** Counting the appearances of an organization in TV, Radio, Newspapers and digital news outlets. Experts recommend running content analysis to measure the extent to which media coverage is projecting a positive image of the organization. (Coffman 2009; Weyrauch et al. 2010)

• **Modular Matrix:** This mapping tool, which draws from social network analysis and seeks to escape the linearity of logic models towards networked models, plots links between outputs, impacts, and stakeholders. The results are descriptive and intended for self-assessment purposes, especially in mid-term reviews. (Davies 2005)

• **Public Policy Tracking:** This method monitors “policy proposals in the lawmaking process” and establishes at which stage in the adoption process they are. Budget information can also indicate how much funding has been mobilized around an issue. (Coffman 2009)
Evaluation:

- **Innovation Stories**: A narrative report of an innovation, developed by CGIAR. It consists in four steps: 1) Clarifying objectives and expectations of stakeholders; 2) defining the innovation; 3) Constructing innovation timelines and plotting actor network maps onto them, and 4) Writing up the story. The technique includes recommendations on sharing the stories and engaging external audiences. (Douthwaite and Ashby 2005)

- **Episode Studies**: Episode studies are retrospective analysis of the impact that research had on a selected policy change. The process that led to the policy change can be reconstructed through interviews, literature reviews, debates with stakeholders or workshops. (ODI 2009)

- **After Action Review**: Initially developed in the US Army in the 1970s, After Action Reviews (AAR) are open discussions about a project that allows participants to learn about what worked and what didn’t. The process can be documented to record tacit knowledge and support institutional memory. It is mentioned expressly as a “good tool to include in accountability reports for funders”. (Weyrauch et al. 2010)

- **Rapid Outcome Assessment**: Based on Outcome Mapping, with which it represents a series of approaches for systems mapping, ROA focuses on changes in the environment of a policy debate and has three stages: preparation, during which the context of the project and intended outcomes are identified; a workshop, in which participants identify key processes; and a follow-up, in which interviews are used to refine the conclusions and to prepare a report. (Hovland 2007)

- **Intense Debriefing Periods**: Inspired in the After Action Reviews but applied at a larger scale, this method seeks to document the experience of advocates in the field during a policy window with a high activity. It includes feedback from a number of advocates, either in a focus group or following an interview protocol. (Coffman and Reed 2009).
• **Indexes of Public Officials**: In this method, developed by the Harvard Family Research project, advocates rate key policymakers in relation to an issue according to three scales: support, influence and confidence. (Coffman 2009, 20).

**Both Monitoring and Evaluation:**

• **Interviews with Key Informants**: Interviews conducted with subject matter experts that have no ties to the organization. Their insights can be useful and more objective. Weyrauch et al. warn that informants’ feedback should be contrasted and not taken at face value. (Weyrauch et al. 2010)

• **Bellwether Methodology**: A particular interview protocol developed by the Harvard Family Review, the main objective of the Bellwether methodology is to assess the prominence of an issue in the public agenda through bellwethers, i.e. influential actors “whose positions require that they track a broad range of policy issues”. Cognitive and sampling biases are avoided by selecting half of the interviewees that are not directly connected to the issue area, and by blinding the initial questions: participants know the general area they will be discussing (i.e., education, public health, taxation), but not which is the campaign, product or organization that is seeking to determine its influence. (Coffman and Reed 2009)

• **Outcome Mapping**: The objective of Outcome Mapping is to track how the positions, behaviors and beliefs of key actors change in relation to the initiative and to the environment, and in a process that internalizes evaluation by focusing on behavioral change, boundary partners and contributions. This articulated tool, developed by the Canadian developmental think tank IDRC in 2001, establishes intended outcomes against which change is assessed, and consists in three phases: Intentional Design, Outcome and Performance Monitoring, and Evaluation Planning. (Earl et al 2001). Together with ROA, they compose a set of approaches also known as System Mapping.

• **Most Significant Change**: A participatory form of evaluation, as it invites many stakeholders to select what are the most important changes in the project. It straddles between monitoring and
evaluation, as it is typically conducted during program implementation but offers insights into the performance that can be used for evaluation purposes. The balance between the stories collected at the field level and a selection at the system level aims at transcending the anecdotal. (Davies and Dart 2005)

2.4.5 Some conclusions on evaluating policy research

Although it’s hard to link outputs to outcomes, there is a wealth of experimentation happening around the data collection tools and the processes by which reactions to programs can be integrated with organizational strategies.

By breaking down the notion of political will, it is possible to go beyond the vague reference to informing and elevating the discourse on a certain issue towards the measurement of actual changes in opinion, priorities and willingness to act. This stands in opposition to the traditional measurement of influence as media and political visibility. However, it remains an empirical question whether these existing methods are being adopted by the think tank industry. The following two chapters of this book are an attempt to start answering that question.

The first part of Chapter 3 analyzes the information published on the annual reports of think tanks to track what are the indicators of influence that are presently being used by some of the most prominent centers in the US. The second part examines the approaches of selected foundations to evaluation, and whether their policies and published practices make them likely candidates for assessing the outcomes of think tank programs, and policy research more generally.

Chapter 4 reports the results of 22 expert interviews on the evaluation of policy research and the relations between foundations and think tanks to this regard. Respondents’ feedback is examined to offer insights on how the policy influence of think tanks is falling under the scope of evaluative strategies, and how these strategies are driven by the relationships with foundations.
ii In those same years, the power of record-keeping organizations on the life of individuals was seen as both a promise for efficient, bureaucratic administration and a potential threat. The most persuasive proof of that duplicity is the 1977 report on Personal Privacy in an Information Society by the Privacy Protection Study Commission.

ii The same Peterson who founded the Peter G. Peterson Foundation and had the leading economics think tank Peterson International Institute named after him.
Chapter 3: Methodology

This study employs two empirical methods to test how public policy research is evaluated at the encounter between foundations and think tanks. The first method is a survey of grey literature; the second consists in a set of interviews to experts.

Chapter 4: Surveying the grey literature

Chapter 4 reports the results of surveying the websites of the organizations that are the subject of this study in search of two types of grey literature. Annual reports were scanned in order to identify both quantitative indicators and qualitative arguments used as evidence of influence. Evaluation policies of foundations were analyzed in order to identify and differentiate approaches to the evaluation of policy research. The results are supplemented with two case studies, which function as a recognition of the diversity that can be found in the field and stand in strong contrast with the general patterns identified in the preceding analyses.

Chapter 5: Expert interviews

a. Building the sample

Chapter 5 reports the analysis of a set of semi-structured interviews to experts in three different settings: foundation staffers, think tank staffers, and independent evaluation experts. Staffers at foundations and think tanks can include both program officers and evaluation specialists. Experts were identified with a combination of four sources:

1. Existing contacts in the field;
2. Direct search in the directories of organizations of interest;
3. Authors of prominent literature in the field of policy change evaluation. This source was especially productive for recruiting independent evaluators.
4. Snowballing technique with contacts from the above.

A number of existing contacts in the field had already been identified as potential participants during a preliminary round of informational interviews during the months of September and October of 2015.
Although it informed the formulation of the research questions, none of the feedback provided in those interviews has been included in this study.

Target organizations were selected according to the following parameters:

- **Foundation parameters:** Independent foundation with an interest in effective grantmaking, grantee evaluation, and/or organizational learning, a proven record in funding research and independent policy research organizations, and/or advocacy and policy change

- **Think tank parameters:** At least 20% of foundation funding, consensual recognition as a think tank.

The objective was to interview a total of 25 participants: 12 foundation staffers, 12 think tank staffers, and 3 evaluation experts.

**Conducting the interviews**

The final sample consists of 22 experts, of which 12 are foundation staffer, 6 are think tank staffers, and 4 are independent evaluators. The positive bias towards think tanks engaged in evaluation is discussed at the beginning of Chapter 5. Interviews were conducted per telephone and audio recorded. During the interviews, the author took notes about the respondents’ feedback.

Three different interview instruments were designed (see Appendix D). The aim of the questions is to surface first hand experiences, as well as expectations and beliefs around policy research evaluation, its value of use, and the types of judgments and decisions that are based on the knowledge that derives from them.

The decision to conduct semi-structured interviews is strategically anchored in the nature of the project. Open interviews with a fixed set of questions provide several benefits to this study. Firstly, open interviewing offers more flexible rules for collecting data, recognizing the subject’s institutional context and underlying assumptions. Connections to the object of study that might go unnoticed under a stricter protocol can be traced, explored and brought to focus. This flexibility happens both in preparation of each interview and throughout its conduction. Before each interview, the author explored organizational websites and screened the questionnaire against information regarding the organization’s focus, products and
activities. Although the essence of the questionnaire is preserved for comparison purposes, additional questions were added that best addressed the known professional experience and day to day practices of the interviewee, while other questions appear as non applicable. As opposed to a quantitative approach, which imposes operational definitions and frames a view of the world with the relations between different variables, interviews are a space where new categories, challenges, issues, barriers and attitudes can emerge at any time. In fact, in-depth interviews are recognized as the best qualitative method for obtaining information about generally known rules and status (Patton 1990). Secondly, by recognizing and discussing the particular context of each respondent, semi-structured interviews help creating a climate of trust between the researcher and the interviewee. This “safe zone” is further supported by the integrity protocol of the project, which lent participants the ability to select the degree of attribution for their feedback, ranging from complete confidentiality to full attribution. A good number of participants accepted to be listed in the respondent’s list, but declined to have their name, organizational affiliation and job title referenced alongside their comments.

The study was conducted from January 5 to May 1, 2016. Interviews were conducted from February 1st to April 1, 2016.

Analyzing the interviews

Although qualitative data has an incremental and iterative nature (Patton 2002, 436), the methodology of this study is limited by the scope of a master’s thesis with a total timeline of less than four months. Instead of different cycles of data collection, coding and analysis, the analysis consisted in a single cycle. Additionally, no full transcripts of interviews were generated. The author reviewed the live notes, coded them to identify recurring themes, and reviewed the audio recordings to reconstruct missing passages and extract precise quotations. The annotations were then used to create an outline of recurring themes supported by the topics included in the interview instruments.

Interview data presents three major combine in the analysis. Respondents’ profiles arrange data according to the three expert types; context identifies examples and situations in which practices are described with
sufficient detail; and *themes* draw patterns and relationships around and between the concepts discussed. While all dimensions are important, they cannot be represented in their totality, for the overlaps would result in presenting the same information many times over.

From a high level perspective, the main consideration for data analysis is, in the first place, identifying the examples of detailed contexts that allow to answer the primary question of how the evaluation of public policy research is practiced by think tanks and by foundations. Those descriptions represent simple case studies from which general conclusions may be extracted. Subsequently, the analysis devotes attention to the *themes*, including methods, uses of evaluation and the relationship between research and advocacy. When discussing the themes, the profile dimension can be used to distinguish important nuances, but its use has not been superimposed on each of the discussion points.

Lastly, a third moment extracts best practices and recommendations, moving from a descriptive to a normative phase that is immanent to respondent data.

**Conceptual framework**

a) **Research question, hypotheses and main variables**

Given the research question:

**RQ:** *How is public policy research evaluated at the intersection of foundations and think tanks?*

The main hypothesis is the conformation of the research question:

**H1:** *Think tanks are more likely to evaluate their work when they receive foundational funding.*

While the secondary hypothesis regard the connection between think tanks, evaluation and a third element, be it advocacy or geographic scope (domestic v. international development):

**H2:** *Think tanks with an orientation to advocacy are more likely to evaluate their work.*

**H3:** *Think tanks invested in international development are more likely to evaluate their work than think tanks invested in domestic policy.*

The dependent variable is the *presence of evaluation practices in think tanks*. The three main independent variables are ascribed to think tanks: *relative weight of foundation funding, relation towards advocacy,* and
geographical domain (domestic vs. international). Given the fact that most of the techniques for monitoring and evaluating policy research have been developed in the fields of advocacy and international development, the expectation is that think tanks with expertise in those areas will be more likely to engage in evaluation than non-advocacy oriented organizations with a domestic focus.

b) Other themes

Other major themes included in the interview protocols are:

- Intended purpose of evaluation
- Uses of evaluation
- Levels at which evaluation is conducted (project, program, organization, initiative)
- Perceptions of foundations
- Perceptions of think tanks
- Techniques and methods

Acknowledgments

The author is exclusively responsible for data collection and for all the mistakes and inaccuracies that may be contained in the analysis and reporting phases. This study would not have been possible without the exchange scholarship awarded by the Spanish Fulbright Commission, the contribution of the Spanish Fulbright Alumni Association, and the additional support of Fundación Universidad-Empresa in Madrid.
Chapter 4: Measures of Influence and Evaluation Policies

4.1 Overview

This chapter analyzes how some of the most prominent think tanks and foundations in the United States communicate their efforts through information published in organizations’ websites. The first section consists in a content analysis of annual reports of think tanks to track the metrics that support arguments on the organizations’ influence. This is supplemented with an overview of the indicators on which the main rankings of think tanks are based.

Since many foundations do not publish annual reports, and the exceptions may not include information that is relevant to the discussion about the influence of policy research programs, the second section of this chapter describes the evaluation policies of several prominent foundations which devote at least part of their portfolio to fund policy-oriented research, advocacy and policy change programs, and/or programs that include grants to one or more think tanks. As section 4 in Chapter 2 has shown, evaluation is an essential accountability mechanism that foundations build into their grantmaking planning and strategies.

Each section is completed with a case study of an organization that furnishes insights into the research question but does not fit within the strict parameters of indicators of influence and evaluation policies. The first case study covers the Center for Global Development, and how think tanks can include evaluations as a part of their strategy, including its adoption as a public relations tactic. The second case study connects the classic debate on knowledge utilization to ongoing efforts sponsored by the William T. Grant foundation, which recently completed its 2009-2015 portfolio on the use of research evidence and is launching a research agenda to promote empirical testing of strategies to increment utilization.

4.2 Measuring the influence of think tanks

4.2.1. Metrics of influence and performance in selected think tanks

Contrarily to for-profit business organizations, nonprofits are not required by law to publish an annual report. However, the annual report is broadly recognized in the popular literature on nonprofit management as a useful instrument for communicating the organization’s achievements in a single place.¹ Heavily
oriented towards dissemination, think tanks are likely to have a website as the foundation of their communication strategy, and will often have dedicated communications teams. Depending on the size and capacity of the organization, annual reports can go from simple brochures to complex documents including financial information, the identity of donors, members, staffers and the board, as well as a summary of activities developed throughout the year and arguments about the organization’s relevance and influence. This section looks at the latter: metrics of outputs and activities, and narratives that underscore the relevance of the organization.

Normative research has determined *relevance* as one of the five major dimensions for the annual reports of nonprofits, recommending a combination of program summaries, quantitative and qualitative elements. Quantitative indicators include the total program recipients and the number of countries in which operations are conducted, while qualitative narratives to render lobbying results were found in a 22.7% of the reports examined across three different fields (Gordon et al. 2010). As this section shows, narrative stories of impact are one of the most common indicators used in annual reports, and the single most common for showcasing impact. An important reminder is mandatory here. While Bettman and Weitz hypothesized for corporate annual reports that “the proportion of attributions associated with internal causes relative to external causes will be greater when the outcome is favorable rather than unfavorable”, the temptation of *egotic attribution* that ascribes adverse outcomes to external factors and favorable outcomes to internal factors (Bettman and Weitz 1983) is moderated, in the context of policy change, by the explicit assumption that any single organization can only *contribute* to change, and not attribute it as an effect of its activities.

Table 4 in page 61 contains an overview of the indicators used by a dozen of think tanks. Measures are grouped in three clusters: *Outputs*, which include *Publications* and *Events* that depend directly on the activities planned and developed by the organization; *Outreach*, including web visits and social media, traditional media, and third party awards and recognitions; and *Impact*, which includes narratives of the successful contribution to changes in legislation, regulation, litigation or other policies, as well as congressional testimonies.
As a general rule, indicators were coded as present when they were already provided in numeric form in the annual report. For instance, congressional testimonies were coded as present in the annual reports of the American Enterprise Institute because the report offers the total number during the 113th Congress: 89. Most often, the total number of events celebrated is provided as a highlighted digit in the report. Descriptions of selected events only were not coded. There are, however, some exceptions to this rule. In the case of books published, the indicator was coded as present in two cases: when the report provided the overall number, and when it provided a comprehensive list of the titles. In other cases, unique channels for the dissemination of information have been included when specifically mentioned as sources of curated information with its own identity. Some good examples are the Heritage Foundation’s Dairy Signal (http://dailysignal.com/), a news outlet of its own right; or the Aspen Institute’s Journal of Ideas, which since 2014 publishes the best 5 ideas of the day, every weekday at noon (www.aspen.us).

The other exception are the stories of impact, which are presented as narrative elements that tie changes in public policy as a result of or contributed by the organization’s efforts, presented in highlighted sections of the report. This treatment implies that some other stories of impact, which may be contained in the letters from the President, and longer, descriptive essays have not been considered for this category. Annex B includes a selection of these stories of impact and captures their graphic presentation.

Another difficulty in coding is the lack of consistency in reporting social media metrics. Some organizations report yearly increases, others report aggregated numbers on the followers of their social media channels, and still others report the number of video minutes viewed instead of the number of viewers. Additionally, since some organizations make explicit reference to video streaming, as well as podcast and webcast services, an additional category was created to code for those web-based multimedia channels. Despite the inconsistencies in measuring social media, Facebook, Twitter and YouTube appear as the most popular channels in the think tank community. A final comment on the coding procedure refers to the categories considered for the impact cluster. As a symptom of how idiosyncratic the selection of impact indicators can be, some of the annual reports contain arguments that can be considered unique and tied to the institutional know-how of a given organization: a sort of DNA of how to exert influence. The Cato Institute lists the
number of amicus briefs they submitted to the Supreme Court and how many of them fell on the side of the final sentencing —counted as “victories”; the American Enterprise Institute highlights a success story in which one of their proposals for the banking sector were implemented in the industry. Although sufficiently different to be distinguished from other types of evidence, these arguments have not been isolated in the content analysis because they would have required the creation of ad hoc categories. For that reason, both cases were coded as stories of impact.

The table shows that the most frequent indicators are the total number of events (8), Facebook likes (8), and Twitter followers (9), an intuitive result given the fact that data is very easy to capture. Additionally, Facebook pages and Twitter handles are relatively easy to maintain, moderate and include in a communications plan, as opposed to a YouTube channel, which is more labor-intensive, requires more technical means and dedicated production experts not all organizations may have. Despite that fact, as many as four think tanks (a third of the total) report their numbers, although the total number of organizations with a YouTube channel could be higher: a search on the YouTube website performed on March 22nd, 2016 identified official channels of the Heritage Foundation, the Aspen Institute or Urban Institute, some of them containing videos dated several years back.

With six mentions appear the number of books published, a fundamental driver of academic credibility and a very high score for a single type of output. In comparison, the seven organizations that were coded as reporting stories of impact (see Appendix B) refer to a much broader range of rhetorical constructs: in general, stories of impact make the case for organizational effectiveness. Usually, the narrative elements allow for a more explicit and nuanced connection between the organization’s expertise, efforts and accomplishments in what could be considered as a minimalized form of case studies. However, their variety is reflected by the fact that in the case of the Cato Institute, amicus briefs and Supreme Court “victories” were coded in this category in spite of the lack of explicit. The claim for success and impact on the highest spheres of the judicature is a strong, idiosyncratic expression of the organization’s influence that differentiates the brand from most of its competitors.
The policy brief, a publication format extensively used in the think tank world, has given way to a number of other designs for presenting summarized, easily consumable information that are represented in three of the annual reports. Among the 15 publication series listed by the Cato Institute, there are letters, bulletins, essays, briefs, studies, reviews, and several articles. CSIS counts its Commentaries, defined as “brief policy analysis by CSIS experts”, and Critical Questions, “quick answers to the key questions surrounding today’s top international issues”; while the Heritage Foundation produced 204 Issue Briefs, 116 backgronders, and 35 legal memoranda. In a related note, the Heritage Foundation also has its own online news outlet: The Daily Signal.

Annual reports offer highly selected information that is chosen as an indicator of excellence. For example, the fact that only one quarter of the think tanks reported the number of web visits does not imply that web visits are not used to monitor performance of digital communications—rather, it may imply that for the remaining 75%, whichever numbers they had were not impressive enough to make it to the final report. From that point of view, the content analysis may be very instructive in diagnosing which are the areas where each think tank considers to be performing outstandingly. The overall numbers suggest that twitter, with its open network of conversations, is currently the best social media channel for think tanks, although a good number of organizations has invested significant efforts in building a solid audiovisual presence, as reflected in YouTube channels. Regarding the impact area, it is interesting to note that only one third of the think tanks reported the number of congressional testimonies. Since testimonies before congressional committees are one of the top measures of visibility, as seen in section 2 of Chapter 2, these data may suggest that the remaining two thirds may not have obtained outstanding measures in that area. However, without closer examination other reasons cannot be excluded: the number of expert testimonies in each area surely varies from year to year, making this metric very dependent on the political environment and current agenda.

By communicating in a highly selective fashion, annual reports do not offer insights on how information about the visibility and influence of the organization is used internally. These indicators of visibility and influence blackbox discussions around what can be learned about the organization from the results of
programs and projects, what should be done differently, how and why investments may be altered and resources may be allocated differently. However, the range of indicators used is likely to represent all or most of the elements that make up the monitoring systems of think tanks. Additionally, the number of measurements used by each organization is likely to be an indicator of how important it is for each think tank to base the argument about its influence on objective data capturing its reputation and outreach. The fact that Brookings and Heritage, perhaps the two most influential think tanks in Washington, DC, top the list presented in table 2 seems to support this view.

The case study at the end of this section will examine how the Center for Global Development has de-blackboxed those conversations with two organization-level evaluations, and how the monitorization of influence has been intentionally built into a broader planning and evaluation framework.

| AEI       | American Enterprise Institute |
| AI        | Aspen Institute |
| BI        | Brookings Institution |
| BPC       | Bipartisan Policy Center |
| CI        | Cato Institute |
| CGD       | Center for Global Development |
| CFR       | Council for Foreign Relations |
| CSIS      | Center for Strategic and International Studies |
| GMF       | German Marshall Fund |
| HF        | Heritage Foundation |
| NAF       | New America Foundation |
| UI        | Urban Institute |

**Table 2: Organizations and acronyms**

<table>
<thead>
<tr>
<th>Organization</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brookings Institution</td>
<td>12</td>
</tr>
<tr>
<td>Heritage Foundation</td>
<td>11</td>
</tr>
<tr>
<td>CSIS</td>
<td>9</td>
</tr>
<tr>
<td>AEI</td>
<td>8</td>
</tr>
<tr>
<td>Cato Institute</td>
<td>7</td>
</tr>
<tr>
<td>New America</td>
<td>7</td>
</tr>
<tr>
<td>Council for Foreign Relations</td>
<td>6</td>
</tr>
<tr>
<td>German Marshall Fund</td>
<td>6</td>
</tr>
<tr>
<td>Bipartisan Policy Center</td>
<td>3</td>
</tr>
<tr>
<td>Center for Global Development</td>
<td>3</td>
</tr>
<tr>
<td>Aspen Institute</td>
<td>2</td>
</tr>
<tr>
<td>Urban Institute</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 3: Number of indicators used by each organization**
Table 4: Influence Indicators used by selected think tanks

<table>
<thead>
<tr>
<th>OUTPUTS</th>
<th>Publications</th>
<th>Books</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Journals</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>Briefs and short pieces</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>25%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>News outlets</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8%</td>
</tr>
<tr>
<td>Events</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>66%</td>
<td></td>
</tr>
<tr>
<td>VISIBILITY</td>
<td>Web &amp; Social Media</td>
<td>Web visits</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>33%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Newsletter</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Facebook</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>66%</td>
</tr>
<tr>
<td></td>
<td>Twitter</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>YouTube</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>33%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other (podcast, webcast, etc.)</td>
<td>X</td>
<td>X</td>
<td>8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Op-ed</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>25%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Citations</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X*</td>
<td>X</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interviews</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>17%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Awards</td>
<td>GETTI</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>17%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transparency</td>
<td>X</td>
<td>X</td>
<td>17%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>17%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPACT</td>
<td>Congressional testimonies</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>41%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stories of Impact</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>58%</td>
</tr>
</tbody>
</table>
### Table 5: Summary of the influence indicators used by selected think tanks

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outputs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Books</td>
<td>6</td>
<td>50%</td>
</tr>
<tr>
<td>Journals</td>
<td>1</td>
<td>17%</td>
</tr>
<tr>
<td>Briefs and other</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>News outlets</td>
<td>1</td>
<td>8%</td>
</tr>
<tr>
<td>Events</td>
<td>8</td>
<td>66%</td>
</tr>
<tr>
<td>Events</td>
<td>8</td>
<td>66%</td>
</tr>
<tr>
<td><strong>Visibility</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web &amp; Social Media</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web visits</td>
<td>4</td>
<td>33%</td>
</tr>
<tr>
<td>Newsletter</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>Facebook</td>
<td>8</td>
<td>66%</td>
</tr>
<tr>
<td>Twitter</td>
<td>9</td>
<td>75%</td>
</tr>
<tr>
<td>YouTube</td>
<td>4</td>
<td>33%</td>
</tr>
<tr>
<td>Other*</td>
<td>1</td>
<td>8%</td>
</tr>
<tr>
<td>Traditional Media</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Op-ed</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>Citations</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>Interviews</td>
<td>2</td>
<td>17%</td>
</tr>
<tr>
<td><strong>Awards</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTTTI</td>
<td>2</td>
<td>17%</td>
</tr>
<tr>
<td>Transparency</td>
<td>2</td>
<td>17%</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>17%</td>
</tr>
<tr>
<td><strong>Impact</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congress. Test.</td>
<td>5</td>
<td>41%</td>
</tr>
<tr>
<td>Stories of Impact</td>
<td>7</td>
<td>58%</td>
</tr>
</tbody>
</table>

#### 4.2.2 Rankings and indexes

The question about the influence of think tanks has also been answered with rankings. The Go to Think Tanks Index (GTTTI) is the most famous ranking of think tanks, prepared by James McGann, Director of the Think Tanks and Civil Societies Program at the University of Pennsylvania. This global index, which is published every year the last week of January, is based on a survey sent to peers around the world, who express their opinion in which are the leading think tanks across a series of categories, including Geographic Region, Area of Research, and Special Achievements. While GTTTI still is the major reference in the field and a good general indicator of overall brand recognition, it does not offer any accurate measurements of impact and has received several criticisms. Some voices have expressed concern at the lack of rigor in the selection of organizations and the general lack of insights provided by the Index (Mendizabal 2011a).
Others have pointed at inconsistencies and inaccuracies in the preparation of the index. More fundamentally, some experts have cast doubts upon the fact that any given can be well informed to express substantial judgments about the work of policy research centers across all continents and disciplines (Seiler and Wohlrabe 2009).

Some alternative rankings have appeared of late, the most notable of which is elaborated by the Center of Global Development, with a first release in 2012 (Clark and Roodman 2012) and an update in 2015 (Gelb et al. 2015). Originally designed by David Roodman and Julia Clark, the methodology takes into account a combination of digital metrics and academic visibility metrics: Social media fans, web traffic, incoming links, media citations and scholarly citations. The big difference of this ranking is that it measures the efficiency of the organizations by entering the annual operating expenses of each think tank as a factor, allowing to compare across different sizes and budgets, and also promoting the relatively small CGD as a top-tier player. The tables below report the absolute and adjusted scores for the top 10 US think tanks with 2014 data.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Change from 2013</th>
<th>Name</th>
<th>Social media fans</th>
<th>Web traffic</th>
<th>Incoming links</th>
<th>Media mentions</th>
<th>Scholarly citations</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+1</td>
<td>Brookings Institution</td>
<td>4.7</td>
<td>13.6</td>
<td>11.0</td>
<td>21.1</td>
<td>35.5</td>
<td>17.2</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>Heritage Foundation</td>
<td>37.9</td>
<td>14.0</td>
<td>9.9</td>
<td>9.2</td>
<td>1.6</td>
<td>14.5</td>
</tr>
<tr>
<td>3</td>
<td>+2</td>
<td>Pew Research Center</td>
<td>4.8</td>
<td>12.4</td>
<td>8.6</td>
<td>13.8</td>
<td>6.6</td>
<td>9.2</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>Council on Foreign Relations</td>
<td>7.3</td>
<td>9.7</td>
<td>8.0</td>
<td>5.8</td>
<td>3.5</td>
<td>6.9</td>
</tr>
<tr>
<td>5</td>
<td>-2</td>
<td>Cato Institute</td>
<td>8.1</td>
<td>6.6</td>
<td>8.4</td>
<td>4.1</td>
<td>2.4</td>
<td>5.9</td>
</tr>
<tr>
<td>6</td>
<td>+1</td>
<td>Center for American Progress</td>
<td>1.8</td>
<td>3.8</td>
<td>6.4</td>
<td>4.7</td>
<td>5.2</td>
<td>4.4</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>RAND Corporation</td>
<td>1.5</td>
<td>5.5</td>
<td>6.7</td>
<td>3.7</td>
<td>4.3</td>
<td>4.4</td>
</tr>
<tr>
<td>8</td>
<td>+1</td>
<td>American Enterprise Institute</td>
<td>2.5</td>
<td>5.2</td>
<td>4.7</td>
<td>6.7</td>
<td>2.4</td>
<td>4.3</td>
</tr>
<tr>
<td>9</td>
<td>-1</td>
<td>Center for Strategic and Int'l Studies</td>
<td>6.2</td>
<td>2.7</td>
<td>3.8</td>
<td>5.7</td>
<td>2.5</td>
<td>4.2</td>
</tr>
<tr>
<td>10</td>
<td>+1</td>
<td>Carnegie Endowment for Int'l Peace</td>
<td>5.9</td>
<td>2.3</td>
<td>3.2</td>
<td>3.0</td>
<td>6.2</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Table 6: Absolute scores and ranking for US think tanks (Gelb et al. 2015).
By leveraging new digital tools, like alexa.com and lexisnexis.com, which were not available but few years ago, the CGD measurements build on Rich’s attempt at gauging the visibility of think tanks in the media and in academic debates by generating an aggregate measure of public profile:

The rankings are best understood as an indicator of public profile. While policy impact is hard if not impossible to measure, the strength of a think tank’s public profile is likely to be a good marker of its influence and potential for impact: ideas need to be noticed to be adopted (Clark and Roodman 2013)

In this consideration, impact on public policy is a “known unknown” that can only be measured indirectly by means of the effectiveness of good communications and the quality (or popularity) of research, but never, or very rarely supported with direct evidence. The thought is almost identical to Weiss’ argument:

If they are to influence policy, a necessary (although not sufficient) first step is to see that their message is heard (Weiss 1992, viii)

Note, however, how Weiss stresses the insufficiency of this condition. Rankings can be a reliable way of assessing the brand of think tanks –their reputation, the extent to which their brand is recognized, their papers cited and their messages disseminated. However, they do not explain how the message disseminates,
or whether it resonates among the intended audiences. Seeking to answer that question, the Center for Global Development has conducted two organizational evaluations that are the subject of the following case study.

4.2.3 Case study 1: Center for Global Development

4.2.3.1 Evaluation of impact by Arabella Advisors, 2006

In 2006, five years into the creation of CGD, an independent evaluation was commissioned to Arabella Advisors by four of the largest independent foundations in the country: the Bill & Melinda Gates Foundation, the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, and the Rockefeller Foundation (Hipple et al. 2006).

The evaluation confronted core questions related to the policy niche that the research agenda sought to fill, the rigor and influence of the research product, the communications and outreach strategy and the partnerships to assess the recognition of the center among its diverse audience of policymakers, academics, and advocates. Based on 150 in-depth interviews and a survey to 1,400 respondents, the evaluators found that 86% of the US policymakers interviewed said to have discussed or used CGD research, and 89% of the academic researchers that were familiar with CGD expressed their satisfaction with the center’s research output. Advocates, while sharing many of the broader goals of CGD, expressed their concern that the center was failing to coordinate and leverage meaningful partnerships with advocacy organizations and other stakeholders.

4.2.3.1.1 Making Markets for Vaccines Initiative

The evaluation of impact conducted by Arabella Advisors contained three case studies: the Nigerian Debt Relief Initiative, the Population Dynamics & Economic Development Initiative, and the Making Markets for Vaccines Initiative, on which this subsection is focused.

The Making Markets for Vaccines Initiative was inspired in the work of Michael Kremer, who in 1999 had proposed government-driven vaccine research as a way to guarantee their low cost and availability in their
public domain. CGD embraced the issue when it was starting to resonate on the Hill. In 2003, a working
group supported by the Gates Foundation commenced to explore the feasibility of a pull mechanism for the
production of vaccines in developing countries. The group included Kramer and Ruth Levine, who then
was a CGD senior fellow and director of programs. The four key findings of the case study indicate that 1)
CGD contributed to raising the issue at the 2006 G8 meeting, 2) CGD translated the pull mechanism for
the vaccine from an academic idea to a specific proposal handed to international actors, 3) CGD made an
effective use of the working-group approach and 4) substantial financial support to the mechanism had
already begun, with a C$100 million pledge from the Canadian government which exceeded by far the total
estimated cost of the initiative in $400,000.

The evaluative case study reports the outputs that the grant application with the Gates Foundation promised
to deliver. They included a policy report, a book, an advocacy tool, briefings and presentations at
conferences, as well as 3 types of sample agreements for two different diseases. The evaluators conclude
that the implementation of the initiative delivered more than promised:

Not part of the original scope were the many consultations and briefings for policymakers and sherpas (who do the preparatory work for the political leaders) in the G8 and G7 forums in which the working group and staff engaged, and CGD’s continued involvement in the evolution of the process after the working group. In assessing this initiative, it is useful to look at not only at what was accomplished but also the way the work was done.

Balancing the outstanding judgment of the initiative outputs, the modal lens found alleged flaws in the
process. While the Policy Research Network model from which the working group had originated starts
with open-ended questions to shape the agenda, CGD’s working group adopted an existing formula.

Alternative options went unheard:

One member who left the group said dissenting views were dismissed. Greater attention to the work process could have avoided the “my idea versus your idea” impression in favor of a report that represented “our idea.”

While the case study mentions some salient third parties, like a legal team whose work was instrumental in
drafting contract protocols, it also underlines inconsistencies between the objectives of the working group
and its composition. In particular, the lack of private industry representatives in the consideration of a proposal that aimed at altering substantially the manufacturing process of vaccines was an important obstacle for their buy-in. Similarly, the absence of voices from the developing world made the working group detached from real ground.

Regarding the closure of the project, the evaluators express some difficulties in reaching any final conclusions. The role of CGD as a “think and act tank” is described as clearly overstepping classic roles as policy research suppliers “in doing good analysis and presenting a balanced report”. However, the lack of precedents in transition phases and of intended outcomes in the grant application impedes the assessment of whether the CGD played a proper role at the outset of the initiative. In any case, the question about the initiative’s outcomes is positive, with a number of them clearly identified:

CGD and the vaccines working group added value far beyond the cost of engagement and achieved real impact. They took an idea, added more analysis, and turned it into a policy package – a proposal, not just a report. They identified champions to advance the idea and tried to clarify key issues for each stakeholder group. They got it placed on the public agenda. They used a team model for the work and in doing so demonstrated that in harvesting the wisdom and managing the diversity of a group of experts, they could deliver far more than any individual or small grouping of experts could have achieved. By working in the technical, political, and international financial realms in succession, the process evolved and they achieved their impressive impact. (Hipple et al. 2006, 91)

4.2.3.2 Strategic review by Redstone Strategy Group, 2012

In 2012, the findings of a second evaluation were published. The prompt given to the evaluator were four broad, strategic questions:

- How are we doing relative to our peers?
- How are we perceived and what could we do better?
- Should we become more global, and if so, how?
- How can we better connect our work to real-world impact?

(CGD and Redstone Strategic Group, LLC, 2012: 1)

Figure 6: CGD, core questions guiding the strategic review
The evaluation was based in a series of internal and external interviews, which yielded a positive image of the CGD as the most recognized think tank in development. But the most relevant piece of the evaluation is the link between the second and the fourth question. In fact, one of the themes that emerged regarding the second question was the need for “more formal intellectual structure and documentation” that would add to the informal monitoring and evaluation procedures. At the time, CGD monitored “media mentions, attendance at CGD events, and a set of Web-traffic metrics (subscriptions to email newsletters, open and click rates, unique visits to the CGD Web site, and page views). For impact, CGD document[ed] specific success stories”. The evaluation report announced that CGD had started experimenting with expected return estimates to assess future initiatives and research and add that to the decision-making processes already in place, and consisting in 5 steps:

1. Begin with the goal of improving the policies and practices of the influential to promote shared global prosperity
2. Identify the set of development issue areas in which the influential play a crucial role
3. Hire smart, passionate, and rigorous policy entrepreneurs who can identify and are interested in high-return opportunities within that framework
4. Build on the entrepreneurs’ ideas with “crowd sourcing” from other staff
5. Use internal processes (e.g., internal presentations and discussion, peer review, and expected return estimates as described later in this paper) to assess CGD staff’s instincts before taking on a new research project or policy initiative.

(CGD and Redstone Strategy, LLC: 2012, p. 8-9)

Figure 7: CGD, Decision-making processes leading to new projects

The Expected Return estimates (see worksheet sample in Figure 7), based in the formula:

\[ \text{benefit} \times \text{likelihood of success} \times \text{contribution} / \text{cost} \]

would offer critical cost assessments to be included in the deliberative process in step 5. The report explicitly states that ERs are not designed to replace deliberation, but rather to guide strategic thinking towards setting the right goals, targets and metrics, while selecting cost-effective means.
4.2.4 Key takeaways

The comparison between the content of annual reports and the assessments in the two evaluation studies reported above yields four main takeaways:

- **Attention to process and outcomes**: Compared to the annual reports, which focused predominantly on quantitative measures of output and media visibility, the attention in these external evaluations at the organizational level is shifted to assessing processes, influence and outcomes. The 2006 report assesses whether the Center for Global Development has effectively established its presence in a policy development niche, and how its influence is perceived by its audiences; the 2012 report focuses on similar questions, but connects them to an internal process of decision-making. In the case studies of the 2006 report there is a close attention to what the outputs of the initiatives don’t tell: how the initiative originated, how the process was and what alleged flaws it showed, and which other voices should have been included. Also, in assessing
impact it described how CGD exceeded the role of a simple knowledge producer to set an academic proposal in the agenda of international development, and connected that effort to specific outcomes.

- **Attention to audiences:** Far from being based in an analysis of media visibility, influence was assessed through surveys and interviews with CGD’s main audiences: policymakers and development advocates. Listening to key audiences allows to compare the self-image of the organization, as can be selectively represented in a carefully crafted annual report, with the external image of third parties. In fact, this approach surfaced biases that may have passed unseen to the organization, like the excessive focus on the US scene and the lack of industry representatives and Southern experts in the vaccines working group.

- **Evaluation as reflection in use:** The external evaluations are tools to identify the current position, assess weaknesses and strengths and support medium and long-term planning. While this is true of both documents, it’s much more obvious in the 2012 report, which has had as its immediate consequence the adoption of a new tool to inform decisions on project development.

- **Evaluation as a Public Relations tactic:** While the reflective component is certainly an important motivation for commissioning these studies, they can also be considered a public relations tactic conducive to the goal of communicating the CGD brand as associated with the values of due diligence, transparency and openness (Weiss 1972). At the same time that they are a balanced display of the organization’s strengths and achievements, the evaluation reports, which are widely shared and accessible at the organization’s website, are designed to persuade about the excellence of CGD based on objective data. The communicative strategy is smart and balanced. In the 2006 report, the first page of the executive summary, probably the single most read page in any report, contains a summary of six key findings. While the three first findings report the main achievements (policy niche, talented staff and demonstrable impact), the last three signal deficiencies in the inclusiveness of the organization and attach the corresponding recommendations.
4.3 Evaluation policies in selected foundations

This section analyzes the evaluation policies and practices of four selected foundations: Bill and Melinda Gates, MacArthur, Robert Wood Johnson, and William and Flora Hewlett. The analysis is based on the information published on the foundations’ websites, either in the form of working papers and other documents, or in the website’s main body. The statements and policies analyzed have been summarized in Appendix C for easy reference. The excerpts have been arranged in “quick sheets” composed with the following categories: Definition, Approach, Purposes, Questions, Contexts, Utilization, Sharing, Output, Examples. Since the summaries only contain quotes and literal paraphrases of the original sources, the authored decided not to infer any statements that were not included literally in the information offered. For that reason, the information retrieved does not cover every category for each organization.

4.3.1 Overview of evaluation brands

4.3.1.1 Bill and Melinda Gates

The Bill and Melinda Gates foundation bases its evaluation policy in the notion of variation. Since not all efforts have to be evaluated, the policy is designed to assist program officers in the selection of those initiatives that better align with institutional priorities and predefined purposes. The rationale for evaluation is primarily outcome-oriented and strongly based in the search of effectiveness, as can be seen in the three predefined purposes: 1) understanding and improving program effectiveness, 2) testing the causal effects of pilot projects and 3) improving the performance of institutions or operating models.

4.3.1.2 MacArthur Foundation

On its website, the MacArthur Foundation frames the publication of evaluation reports as part of the institutional commitment to accountability. The questions that evaluations seek to answer range from the “evolving nature of the problem” and the context, to the foundation’s strategies, partners, accomplishments and impacts. This broad perspective is tackled by evaluating exclusively at the initiative level.
4.3.1.3 Robert Wood Johnson Foundation

Perhaps the proponent of the most comprehensive approach to evaluation, the grantmaking activity of the Robert Wood Johnson Foundation is defined as “a continuous cycle of research, evaluation and learning”. This is the only among the foundations analyzed that conducts an institutional assessment or internal evaluation of the foundation itself. The assessment combines interviews to grantees and policymakers to situate the recognition and reputation of the brand with indicators of communications outreach, like social media following. Additional resources include Program Results Reporting, a joint effort of the evaluation and the communications departments that seeks to communicate the work of the foundation with external audiences, including grantees, partners and policymakers. While there is not an explicit document on evaluation policy, documentation on conducted evaluations is very vast.

4.3.1.4 William and Flora Hewlett Foundation

In an internal working paper on evaluation, Twersky and Lindblom posit a series of questions that the William and Flora Hewlett Foundation can seek to answer through evaluation, ranging from context and overall strategy to implementation, outcomes and impact. The definition of evaluation focuses on achieving goals and objectives, allowing for a broader number of interpretations than definitions focused on outcomes. At the higher level, evaluation is framed by a chain of seven principles that underscore the importance of evaluation for learning, its inclusion in the strategy, and the effort to share and use findings. (Twersky and Lindblom 2012)

4.3.2 Analysis and comparison

4.3.2.1 Do the 4 foundations analyzed evaluate organizations?

Two out of the four foundations analyzed establish conditions under which organizations can be evaluated. For the Gates Foundation, the organization needs to be at a critical point of development that can benefit from an independent performance assessment; for the Hewlett Foundation, it is about informing grantee’s practices and decisions. Two elements are associated with organization-level evaluations: building the
capacity of the grantee, and ensuring involvement and interest in the process, including the willingness to listen to results. The 2006 evaluation of the CGD is a clear example of this type of assessment. The Center was a new actor in the field of development and needed to assess its influence and recognition. The main findings of the study included recommendations on how CGD should engage in a more inclusive manner with advocates and actors outside the US. The study was not a formal evaluation in the sense that it did not consisted of any randomized control tests.

Although the McArthur Foundation was one of the four funders behind the CGD 2006 evaluation, their policy is to evaluate strategies and initiatives that include multiple organizations, not to run studies of single grantees. The case of the Robert W Johnson foundation is quite unique, as the only organizational assessment published on their website is the 2013 Assessment Report of the foundation as a whole. The Assessment Report is based on three sources of data: a survey to health policy experts, with a particular focus at the state level; a grantee survey, which has been administered by the Center for Effective Philanthropy since 2004 and allows comparisons with 14 national funders; and internal and administrative data. While the first two elements share many characteristics with both the 2006 CGD evaluation, internal and administrative data includes, among other things, the types of aggregate output numbers that are typically found in the annual reports of think tank: web traffic, email subscribers, and social media following in Twitter, Facebook and YouTube.

4.3.2.2 Do the 4 foundations analyzed evaluate research?

Each of the foundation policies analyzed has a different approach to the evaluation of research grants that depends largely on the specific nature of the portfolio. The Gates Foundation’s policy does not have an explicit position on this aspect, although it clearly states that it does not evaluate basic scientific research. Additionally, only 6% of its spending portfolio goes to advocacy initiatives. However, prioritizing contexts in which outcomes are less observable, including “systems, policies, and institution-wide change” makes it
a good candidate for the evaluation of policy research. The McArthur Foundation also lacks an explicit perspective on this issue.

On the contrary, the Robert Wood Johnson Foundation evaluates research on a regular basis, conceiving its grantmaking strategy as a “continuous cycle of research, evaluation and learning”. As the sampled documents show, the foundation is attempting the use of innovative approaches to outcome evaluation, including the insights that can be gained from social network analysis. The Hewlett Foundation recognizes a specific purpose in evaluating field-building efforts, which are primarily composed of research in an emerging area.

4.3.2.3 Who will use the evaluations?

Each of the four foundations exhibits nuances in the audiences of the evaluations conducted. In the case of the Gates foundation, for example, maximizing the efficiency of the foundation is stressed as the main objective, although organizations are recognized as important partners that can benefit from the findings. The Hewlett Foundation has a stronger emphasis in dissemination to the general public, in particular in the case of field-building efforts – while the disclosure regimes of organization level assessments are decided on a case by case basis. The dissemination and the use of evaluation findings are the sixth and seventh principles of the seven-step approach.

In the case of the MacArthur Foundation, the publication of evaluations is part of a broader effort for internal accountability, and includes the publication of other information, such as tax disclosures, financial information, etc., with no particular mention of grantees or internal audiences as beneficiaries of the studies. Finally, the Robert W. Johnson Foundation targets specific external audiences, beyond the donor-grantee relationship. The streamlined collaboration between the evaluation and the communications team for the Program Results Reports suggests a clear effort of transforming evaluation findings in amenable information that is easy to use as a resource in the field, especially by policy makers and healthcare experts.
4.3.3 Case study 2: The W. T. Grant Foundation

In 2009, the William T. Grant Foundation launched an initiative to study research use by policymakers and practitioners in youth education, the organization’s traditional field of interest, and encouraged researchers to take an “ecological approach” (W.T. Grant Foundation 2016) to research use that would consider how users are embedded in larger sociotechnical systems. In doing so, it declined to ask the still linear question about the lack of research use. The cycle of the portfolio ended in 2014. As findings accumulated, the foundation published in late 2015 the essay Leveraging Knowledge: Taking Stock of the William T. Grant Foundation’s Use of Research Evidence Grants Portfolio (Du Mont 2015). Among the three major themes that emerged from the findings, the importance of professional networks and intermediaries is of particular relevance for this study. Together with vendors and professional associations, think tanks and advocacy groups are among the main intermediaries considered. In particular, Scott et al. studied intermediaries in the context of the debate about educational incentives. They found that intermediary organizations, which can perform multiple research functions sometimes within the same organization, “operate in a highly politicized landscape” and “are increasingly blurring the lines between research, policymaking, and political advocacy” (Scott et al. 2014, 70). Additionally, they note:

To date, there has not yet been sufficient scrutiny on the role of intermediaries in research acquisition, interpretation, and implementation, particularly in terms of their role in shaping educational policies. (Scott et al. 2014, 70)

This study found a clear political function in intermediary organizations: despite the uncertainty about the effectiveness of educational incentives, intermediary organizations have used research evidence to promote the reforms.

The next cycle of the Grant Foundation’s initiative on research use “aims to identify and test strategies to ensure that research reaches the hands of decision makers, answers their most pressing questions, and is used in ways that benefit youth” (W. T. Grant Foundation 2016, 2). The focus has shifted from
understanding use to understanding *how to increase it*. Where the first research portfolio resulted in mostly descriptive theories, the second encourages researchers to:

- **On the demand side**, “identify or test strategies to improve the use of existing research”, including descriptive studies and “evaluations or deliberate efforts to increase routine and beneficial uses of research in deliberations and decisions”.

- **On the supply side**, identify or test strategies for producing more useful research evidence, “examining incentives, structures, and relationships that facilitate the production of research in ways that respond to decision makers’ needs and optimize researchers’, decision makers’, and intermediaries’ joint work to benefit youth”.

- **On the side of policy outcomes**, test the assumption that decisions are improved by the use of high quality research.

Thanks to one of the first grants awarded in the new cycle, a team led by Janelle Scott, Associate Professor at the Graduate School of Education at the University of California, Berkeley will focus on the governance structures of Los Angeles and New York City to examine “whether non-partisan and independent research organizations, such as research–practice partnerships, shift the use of research evidence by local decision makers”.[iv] The study builds on previous work that examined the role of foundations as central actors in education reform, supporting and directing new yet increasingly influential intermediary organizations (Scott and Jabbar 2014).

### 4.3.4 Key takeaways

Evaluation is not an important component in the planning process of the W.T. Grant Foundation in the same sense that it is for the four foundations considered in this section. With significant differences, those organizations represent the current state of the art in grantmaking technologies, with a bureaucratic culture that measures the effectiveness of interventions as the principal accountability mechanism. Given the W.T.
Grant Foundation’s unique approach to research in the world of education, grantees are evaluated through the lens of academic rigor, not with evaluative research.

Of the four foundations, the Robert Wood Johnson has the most similar portfolio to the W.T. Grant Foundation: both specialize in funding research on a specific policy area (healthcare and youth education). For that reason, the fact that the W.T. Grant Foundation does not have an explicit approach to grant evaluation cannot be attributed exclusively to a misalignment between the nature of research products and outcome evaluation. Differences are subtler than what a superficial assessment may suggest. From an essential point of view, the entire portfolio on research use and the evaluation studies are in fact making very similar questions: Under which conditions do policy makers use research? What is the impact of that use? What can be done to further support evidence-based decision making? No significant difference can be found in the methodological assets either: formal evaluations are scientific studies that use many of the same methods, both quantitative and qualitative, adopted by the grantees of the research in use initiative. There are, however, differences in the value of use and in the formalization of these pieces. Evaluation policies examined above signal the primary purpose of informing the internal public of the organization and supporting the decision making process: if done properly, they are designed for circulation and use (Patton 1997). They are usually shared with the broader field, but it is hard to know whether they are valued by external actors. In other words, they are inward-looking evidence embedded in foundation-grantees relationships, which present many contextual coordinates that can be difficult to generalize to different situations, teams and organizations. In that sense, the learning purpose of evaluations outside the political framework that originated them is always challenging.

Instead, the research studies funded by the W.T. Grant foundation do not have an immediate instrumental use for the foundation. They are part of a long term effort of building a new body of research following the established protocols of the scientific community, with highly formalized standards regarding the recognition of research quality, publication, dissemination and contribution to existing knowledge. In
particular, with regard to assessing the influence of intermediary organizations the portfolio shows a high impact potential for several reasons. Firstly, it is a disinterested approach. In the usual grantmaking framework, both foundations and think tanks have obvious interests in making the argument about the effectiveness of a grant. Nonfindings and certain unexpected outcomes, including negative uses of evidence, may prevent sharing the studies. Secondly, the scope is much larger and ambitious. By focusing in measuring actual uses of research instead of the influence of specific outputs, researchers can analyze the contributions of many different organizations to entire policy debates, focusing on policy subsystems instead of specific actors and escaping the contribution vs. attribution debate. These studies can potentially be larger than initiative-level evaluations, such as the ones conducted by the MacArthur Foundation, which are still quite limited if compared to the system scale. Thirdly, precisely because of that scale and in spite of its limitation to state and local level policymakers in youth education, this body of research is much more likely to generate useful generalizations among experts in other fields, and ultimately to impact the perception of the value of think tanks and other intermediary knowledge producers.


Note the victory ratio hidden in the columns of the Supreme Court icon (10 victories out of 11 amicus briefs).

Sources for the content analysis:
  http://i.cfr.org/content/about/annual_report/ar_2015/AR2015_web.pdf


Chapter 5: Interview Analysis

This chapter reports the analysis of a total of 22 semi-structured interviews, conducted between February 1st and April 1st of 2016. Over half of the respondents (12) were evaluation specialists or program officers working at foundations, with the remaining split between 6 members of think tanks and 4 independent evaluation experts. As described in Chapter 3, interviewees were coded based on the live notes of the researcher and organized based on three basic criteria: case studies of think tanks, covering their organizational context and how it unfolds in evaluation practices and approaches to influence; themes, including research, advocacy and their combination, uses of evaluation research and techniques and methodologies for evaluating policy research; and professional nuances, which emerge in the discussion of specific points.

5.1 Think tanks: case studies of evaluation and influence assessment

Recruitment of think tank staffers for the semi-structured interviews proved harder than expected, with only a 50 percent of the target reached, or 6 interviews conducted out of the 12 planned. There is an important positive bias in the sample: think tanks are likely to have actively or passively rejected participation in the project under any of the following circumstances:

- Leadership does not believe in the importance of evaluations
- Evaluations have not produced the expected results
- Lack of evaluation capacity and/or the resources to build it
- Lack of dedicated evaluation specialists
- Concern that evaluation has a political aspect that may harm the organization’s reputation

The two conditions that seem to have prevailed when think tank staffers accepted to participate in the interview are 1) a strong contact with the author and 2) the involvement of the organization in some form of evaluation practices. It seems significant that the invited members of the Center for Global Development declined to participate in the project in spite of the fact that CGD is a pioneer in organizational evaluations,
as featured in the case study on Chapter 4. In other cases, the lack of dedicated evaluation specialists may have prevented the identification of staffers involved in ongoing conversations or projects around evaluation.

The political aspect of both the conversation around the collaborations between foundations and think tanks, and the critical assessment of the influence of these organizations is reflected also in the confidentiality regime selected by many participants both at foundations and at think tanks. While only one interviewee decided to remain entirely confidential, a number of the respondents listed by name, job description and affiliation opted out of having their name cited alongside their comments. It follows that findings must be reported somewhat abstractly, describing case studies, practices and uses without the contextual coordinates that may allow readers to identify an undisclosed source by deduction.

Each think tank represented in the sample reported a unique approach to evaluation. 2 out of the 6 organizations reported that they did not conduct any evaluation of their activities, but had different approaches to policy influence. While at the Brookings Institution the conversation on the assessment of influence is framed on the difference between quantitative and qualitative approaches, the Urban Institute has recently created a Policy Advisory Unit with the sole mission of connecting evidence with users. All of the remaining 4 think tanks reported their involvement in outcome evaluations, albeit in different capacities and at different stages. One of them has an evaluation unit that conducts both public policy evaluations and evaluations of internal programs. At the time when the interviews were conducted two of the organizations were building their monitoring and evaluation systems. Finally, the Canadian International Development Research Center is captured in its role as manager of the Think Tank Initiative, which includes a supervisory role both in the initiative’s sophisticated monitoring process and in the commissioning of independent evaluations at the initiative level.

In spite of the positive sampling bias and the fact that only two thirds of the think tanks interviewed had adopted some type of monitoring and evaluation system, it is reasonable to claim that the collected data
support the primary hypothesis of this project. In fact, foundational support is a clear driving force in four of the case studies, including the Policy Advisory Unit of the Urban Institute.

5.1.1 Brookings Institution: the conversation on quantitative and qualitative metrics

The position at the Brookings Institute is coherent with the findings of think tank scholarship, like Abelson (2010) and Rich (2004), and the analytical findings in chapter 4. Richard Bush III, Director of the Center for Northeast Asian Policy Studies of the Brookings Institution and a Senior Fellow of Foreign Policy reported that the question on how to assess influence is an ongoing conversation, but to his knowledge it has not crystallized in any evaluations. There is an internal debate between the proponents of quantitative metrics and those who believe that the kind of evidence captured by counting outputs does not account for truly influential interventions:

the things that can be measured quantitatively may not indicate high quality impact, they may be more measures of activity rather than impact. If somebody writes fifty op-ed articles that are placed in various journals around the country, and in some way enter the competition of ideas, how can that compare with a scholar spending fifteen minutes with the President of the United States? You know, I think the question answers itself.

(Personal interview with Richard Bush, Brookings Institution, March 8, 2016)

The latter position tends to be defended by experts whose work is more likely to have impact in direct contact with decision makers:

The people who feel that quantitative measures for doing assessment are people who to the extent that they have an impact it’s through face-to-face encounters. For example you can have twenty emails to a policy maker but no impact at all and on the other hand one phone call can have a huge impact in terms of going this way or that way.

(Personal interview with Richard Bush, Brookings Institution, March 23, 2016)

However, Brookings does have an internal performance review. Each senior fellow does an assessment of her work for the fiscal year, and presents it together with a work plan for the succeeding fiscal year, offering a baseline for managers. While there are guidelines for what should be included in the assessment and in the work plan, with an increasing emphasis on perceived impact, managers can make judgments at their discretion and may, for example, prioritize planning instead of focusing on assessments of past performance.
5.1.2 Urban Institute: building evidence use with a dedicated unit

Justin Milner, Senior Policy Analyst at the Urban Institute, was the first hiree of the new Policy Advisory Unit when it was created in early 2014. The unit will reach a staff of 25 “in a couple of months”. Since the Urban Institute as a contract researcher is devoted mostly to providing public policy evaluation and analysis to government agencies, its knowledge products are already designed for consumption by policymakers. This scenario has shifted the discussion on influence away from outcome evaluation to an approach that seeks to actively maximize the connection.

The new unit seeks to engage with the rest of the organization and connect it to the specific needs of policymakers, especially in initiatives like Pay for Success. Although the unit is still in “start-up phase”, Milner’s vision about where the effort can lead in the next five years includes three aspects: 1) establishing a consulting practice to support urban leaders and policy makers in cities around the country; 2) establishing policy issue areas of expertise for more general advice; and 3) consolidating the reputation of the Unit through signature products with a proven record in connecting evidence and practice.

External and Internal Challenges

Externally, the identification of users appears as the main challenge:

> When you talk about connecting research to policymakers, the next question is: Which policymakers? And who exactly are you talking about? Defining that audience and being super clear (sic) and cognizant that the needs of agency leaders at the federal level look radically different from the needs of a commissioner at the county level, and what they might require in terms of the set of tools and information might look really different. One of the challenges I think for our team is how do we make our work, the products that we create and the meetings that we hold, how do we make sure that they are connecting with people in a practical way at the level at which they operate.
> (Personal Interview with Justin Milner, Urban Institute, March 21, 2016)

A second external challenge lies in the means for translating findings to that audience. “In that sense we are more often taking the work of other researchers and thinking about how do we present it differently, how do we summarize it, how do we incorporate the lessons from multiple studies and do so in a paper, in a product, in a data visualization, or in a blog or series of blogs that can be more accessible” (Ibid.).
Internally, the unit encountered skepticism regarding its added value, as some researchers considered that they were already thinking about the audience; additionally, there was a “cultural component” in passing from the traditional work organization of isolated researchers to an environment that promotes team efforts.

Foundations as drivers

For the sake of this study it is very important to note that while overall support for the Urban Institute is provided in two thirds by government funding, foundational funding makes up the largest piece of support to the Policy Advisory Unit.

There’s been real interest from philanthropy in doing a better job and a more strategic job in connecting research investment with practical and policy outcomes. And I think philanthropy in general is much less interested in simply funding research products, and far more interested in thinking about how it’s going to be applied in the real world. Before I was at Urban I was at the Annie E. Casey Foundation, and that was a trend that I could see taking place at the foundation level as well.

[The foundations behind the Unit include] the Arnold Foundation, the Kresge Foundation, JP Morgan Chase, Open Philanthropy Projects, the MacArthur Foundation, and I think the Ford Foundation, too.

(Ibid.)

5.1.3 Internal evaluation unit

Exceptionally, some think tanks have an evaluation unit or an evaluation specialist. In the case of the evaluation specialist interviewed for this study, her work is divided in different areas:

- As an external evaluator, assessing the work of third party-nonprofits commissioned by funders. This role is comparable to the profile of the independent evaluators interviewed in this study. Under this role, sometimes external clients commission research that is not intended as a formal evaluation, but rather as an exercise in strategic learning that can help approach funders more effectively and allocate resources.

- As an internal evaluator:
o Evaluating service delivery options, similarly to the work conducted for the government by an evaluation firm or by research contractors such as the Urban Institute.

o Evaluating the outcomes of programs related to policy development. Out of the service portfolio, this is by far the hardest activity. Identifying impact with certainty was defined as a “dream outcome”:

We subsequently interviewed the commissioners and not only did they agree that the procedure had made a difference, but also we can actually see language that came out of our study reflected in subsequent regulatory change. That would be the dream outcome. There’s a very specific issue that’s being tackled, a very specific desired outcome, and people who are actually in a position to make the change made the change and reported that it was our institution’s intervention that helped influence. Now, we won’t say that it absolutely made the difference, we will not attribute change specifically to our institution’s influence but in a circumstance like that we have a very good evidence that we contributed to the desired outcome.
(Personal Interview, February 9, 2016)

Given the large volume of internal programs that are evaluated, the respondent estimated the dream outcome to happen once a year. This exceptional nature of positive results explains why the expert reported that managing expectations is one of the tasks of her job. Sometimes expectations are too high:

Some foundations can be unrealistic, there is a legacy and a culture of precise evaluation, and it’s sometimes very difficult to let go of that expectation of rigor. But I hope that changes in the future.
(Ibid.)

5.1.4 Building capacity with an earmarked grant

Thanks to a grant provided by a foundation, a think tank is currently developing a Monitoring, Learning and Evaluation Plan as part of the overall strategic planning for a new center. The staff is experimenting with different tools to capture two types of measurements: operational indicators, like the number of publications and their communication; and the engagement with stakeholders. In this second aspect there are two main components: keeping track of outreach and conversations between center members and external parties, and a survey to policymakers.
The strategic plan is based in the design of a logic model. Although outputs, products and conversations are easy to track, “the goals are the most difficult part”. The interviewee believes that this challenge is especially true because the organization does not engage in advocacy:

We don’t have an outcome in mind when we start our work, so the message is not as controlled as when you have an advocacy objective. […] We do policy recommendations, but we don’t lobby around them.

(Personal interview, February 29, 2016)

An external evaluation will be conducted three years after the center’s activities launch to “see how the resources and activities on the left hand of the model map out on the policy outcomes sphere”.

Although the parameters for this effort were set by the funder and the annual report will be prepared for it, the leadership in the center will adopt the system and use it for reporting to other funders.

5.1.5 Building capacity with no dedicated funding

Capacity building in monitoring and evaluation can also happen without dedicated funding, as was the case of another think tank that relies on foundations as its first source of financial support. While the organization already tracks output information at the project and program level, these streams have not been formalized yet. The organization has just started measuring the diversity of its revenue in the last year, with special attention to the longitudinal variations between restricted versus unrestricted funding. In that context, the organization is currently setting the technological foundations and databases on which standardized metrics and frameworks for reporting will draw.

A self-started initiative by the development team, it was later joined by the communications team due to their interest in tracking the volume quality and reach of media campaigns, compiling digital analytics into a dashboard that would make them more readily accessible. The vision is to streamline the technological base so that certain metrics can be attributed to programs and projects in automated reports. Discussions across programs compare the needs for different types of measurements, mostly depending on scope and scale: while larger projects seek to have more reliable evidence on their influence, smaller projects can often rely on the communications and output indicators.
However, the effort is an important burden that is difficult to address in an efficient and sustainable way due to two main reasons. Firstly, the current lack of dedicated funding implies that there are no resources allocated to the development. Secondly, staff has no incentive to dedicate time and resources to the reflection process that is required to frame the new monitoring system. With no formal buy-in from the board, the overhaul is an exceptional effort that falls outside the staff’s job descriptions.

There’s not a lot of discussions […] at the board level. Boards have a lot on their plate, they're worried about governance issues, they're worried about fundraising, they're worried about the big general 40,000 foot view: is the organization successful and impactful? But there’s not really been a focused conversation at the board level yet about how do we know if we’re doing the right thing.

(Personal Interview, February 19, 2016)

5.1.6 IDRC and the Think Tank Initiative: combining unrestricted funding with specific objectives

The case of the Think Tank Initiative is quintessential to the discussion on the power of philanthropy in supporting independent policy research organizations; however, its lessons cannot be generalized to the relationship between American foundations and American think tanks. TTI starts from the premise that the sustainability of think tanks in developing countries is challenging. In the words of Peter Taylor, TTI Program Manager at the Canadian International Development Research Center (IDRC):

TTI at its core is about institutional strengthening, helping independent policy research institutions along a path towards institutional sustainability. It does that by providing them core flexible funding, and technical support over an extended period of time. The basic premise of the program is that stronger institutions, which are more sustainable organizations who do high quality research and engage effectively in policy processes, are able to make a contribution through their evidence, analysis and data to public policy making.

(Personal interview with Peter Taylor, IDRC, February 2, 2016)

Grant agreements are the basis for a results framework that allows to capture change “at the ground level” through a number of tools, including an annual questionnaire, peer-reviewed research outputs, and a policy community survey. Data are captured in the TTI database and build up the monitoring system, in which progress at each organization can be tracked against individual objectives.
On the other hand, evaluation takes place at the initiative level, becoming an examination of the assumptions on which the program is based, and the level of achievement of its declared goals. Evaluation does not inform judgments of grantee performance, but is conceived as a learning and accountability mechanism. Each of the two phases in which the ten-year long initiative is split is subject to an external evaluation. Sarah Lucas noted that a major difference was the timing between evaluation and monitoring, an important improvement in the second phase:

Since the first evaluation was initiated fairly far along into Phase I, the TTI monitoring systems were not built with an eye towards answering evaluation questions. The timing wasn’t aligned, so the evaluation provided a look back but was not as fully data informed as one would like, and it was not as able to capture change overtime because of the incomplete access to usable data and the late start. […] As an executive committee we decided that for Phase II we really wanted the evaluation to start roughly at the same time. The evaluators would then have the capacity to help inform the results framework and the collection of monitoring data […] to inform the answers for the evaluation questions. Also, the evaluation team could get to know the think tanks and understand their changes over time more deeply. (Personal interview with Sarah Lucas, February 8, 2016)

5.2 Role of public policy research for foundations

One of the more general themes that emerged from interviews to foundation staffers refers to the role of policy research. Broadly speaking, the generation of new knowledge tends to assume either an instrumental or a knowledge-driven role. Instrumental roles can consist in policy development around an issue of interest, in policy analysis and policy evaluation for evidence-based decision making, or in the integration into a hybrid model where research organizations serve the factual evidence that is used in advocacy campaigns that are usually tied to specific policy outcomes; whereas knowledge-driven situations are typical when foundations identify gaps in the literature and provide grants to develop new bodies of evidence.

Transitioning to the collaboration between institutions, foundation staffers reported different roles and perceptions in the collaboration with think tanks. A common reaction to the question was to indicate that relations are no different to other grants. In foundations specializing in research grants, it was noted how think tank researchers can perform a similar role than traditional academic researchers provided that their
organizations have a proven record of nonpartisanship. In other cases, think tanks do appear as a special type of organization, more a partner than a grantee:

We have lots of relationships with think tanks. We work with the Aspen Institute, Brookings… I’m thinking of individual people but I can’t think of where they work. When we work with think tanks we usually work with them to either help inform us of something or to help us with the convening. We don’t evaluate think tanks. We work with them as partners.

(Personal interview, March 4, 2016)

Jackie Kaye, Director of Research and Evaluation at Wellspring Advisors, a firm that coordinates grantmaking programs for donors, held the view that think tanks often receive support because of their anchoring function in a certain niche:

In my view, support to think tanks is less related to specific research, and more related to its role within a field. In addition to thinking what type of work you need this field to be doing, you can ask yourself: is it a field that needs anchor institutions, institutions that have strong reputation, that are well known to the people that make decisions and to the people who influence those who make decisions? And so the usefulness of the think tank is assessed more at the organizational level and less in terms of a specific project.

(Personal interview with Jackie Kaye, Wellspring Advisors, February 22, 2016)

This perspective seems close to the strong collaborations that exist between the McArthur Foundation and the think tank world:

Think tanks are key partners in almost all our areas of work. MacArthur has always had a very strong research component, it’s part of our DNA. We see them as partners in both providing knowledge for the field, or for the world, and for disseminating that knowledge and helping translate that into action. So it’s a critical leverage point for us.

(Personal interview with Chantell Johnson, MacArthur Foundation, April 1, 2016)

Another axis is whether think tanks are perceived as academic actors that just publish their findings, or whether they try to connect research to action in a more active way. Some funders expressed their preference for action-oriented “think tanks that don’t want to remain in the think tank sphere” and package their proposals for easy consumption, while others expressed diffidence in the “usual think tanks”, as opposed to those that can be considered credible academic players, such as the Urban Institute and the RAND
Corporation. Sarah Lucas also indicated her preference for organizations that push beyond the cycle of research in innovative ways.

5.2.1 Can policy research be evaluated?

Many respondents conceded that the evaluation of policy research is either too troublesome or simply outside their scope. For Tom Kelly, policy research and evaluation don’t go hand in hand:

Policy research centers at national and state levels produce products that are less informative to a nonprofit operator than they would be to a policymaker. So there are rare cases in which evaluations can draw on that.

(Personal interview with Tom Kelly, Hawai‘i Community Foundation, March 21, 2016)

This mismatch is consistent with the recognition, quoted above, that some foundations do not evaluate think tanks, especially if the former are involved in direct delivery service and traditional social programs. However, other interviewees disagreed with this view, indicating that “everything can be measured”, although proxies and indirect work might be needed. Some experts indicated that policy research might not be evaluated at the level of specific reports, but at a higher level perspective that allows funders to identify changing patterns over time in how the public and specific audiences are informed around a certain issue. Foundations practicing the evaluation of policy research have not adopted particular methods or techniques as a standard practice. Methodological adoption depends on a number of contextual elements such as field, role and subject of research, integration in larger strategies with other partners and networks, etc. However, two methods for qualitative data collection stand out as more congenial to the task: interviews of key informers, usually aimed at gauging research uptake by policy makers, and narrative reports for identifying change. The adoption of these methods is covered in more detail in section 5.4.

5.2.2 Research, advocacy, and collaboration between nonprofits

The views on the integration, combination or antagonism between research and advocacy show a high degree of complexity among respondents that speaks to how these activities are framed through different organizational lenses. For that reason, this section reports the perception of each respondent type in a separate subsection.
5.2.2.1 Foundations

Some voices in the foundation sector expressed the need to integrate research and advocacy. In particular, a program officer in the field of education expressed her satisfaction with the supply of policy development products for the consideration of policymakers by what she named “action tanks”, like the Center for American Progress, and with other organizations like the New America Foundation and the American Enterprise Institute, as they integrate their research with communications strategies.

In the field of higher education, policy development, policy design, and policy analysis, think tanks have sufficiently generated enough supply to the market of ideas for policymakers to consider. [The success in policy development is met when] ideas are as backed up by quantitative analysis as possible, and [there is an] awareness among policymakers of those ideas.

(Personal interview, February 23, 2016)

An evaluation specialist, on the other hand, expressed the view that programs integrating research and advocacy components can eliminate barriers in the assessment of uptake. An interviewee indicated that, given the visibility of legislative outcomes, evaluation research does not focus on determining whether new bills or regulations were passed:

What’s more interesting is what our organization’s contribution was to the work. And part of the reason for evaluating contribution is that our board and executive leadership want to make sure that we are putting our money in the right places. So in particular, if we put X million dollars on a program over a certain number of years, did it matter that our organization was there, or would that stuff have gotten done regardless of our participation? And if it did matter, what was it in the strategy that made a difference?

(Personal interview, February 19, 2016)

Chera Reid, Director of Strategic Learning, Research and Evaluation at the Kresge Foundation, perceived the mentioned integration to be already in place in virtue of the collaborations that some think tanks can activate. They have connections to advocates that promote their findings, so the goals that these organizations have around policy influence are then included in Kresge’s strategy:

It’s a workflow, a chain of actors. In and of ourselves, as funders, we are not going to be the people who are directly informing policy, but we support leading organizations that are going to do that. The goals are included in our strategy at a different level. Our focus is expanding opportunities for low income people and cities. Given that centrality of people in our work, when
we’re thinking about policy we’re not only thinking about federal policy, we’re often thinking about local policy.
(Personal interview with Chera Reid, Kresge Foundation, March 16, 2016)

Theories of change and logic models are perceived as too rigid to promote this type of workflow. Foundations can encourage collaboration in a field and attract organizations to their network by focusing on a limited set of design principles to which organizations can contribute in a plastic manner. In the words of Reid:

One of the hallmarks of our work is that we are working on complex issues that are also changing, they’re dynamic, and we also need to hold multiples hypothesis in our work. Often the theory of change and the logic model, unless they are open to revising and adapting that as we go, could become a limiting feature. Superimposing that frame onto a network of organizations would be a difficult thing to do. We often think in terms of design principles. It’s not that each partner in the network is doing the exact same thing, but we may have three or four principles for the work that you get in any case, in any organization.
(Ibid.)

An evaluation expert at another foundation offered a similar view:

We are a field building foundation, we think of frameworks for the larger field and have a portfolio looking at specific domains, and as a national organization we might work with others to come up with a framework that is not just for a single program and works in the domain. We put that framework out, and do some competitive processes where we invite others that may be interested.
(Personal interview, March 4, 2016)

By contrast, a foundation staffer has used in the past financial stimulus to incentivize collaboration. With the aim of reducing the number of different messages, grantees were requested to create a conglomerated platform with a unified message that could be recognized as an important actor, bigger than the sum of its parts. “These folks didn’t want to work with each other, and I said that wasn’t going to work” (personal interview, February 23, 2016). The program officer promoted a structure that allowed to explore areas of convergence, work through challenges and disagreements, and collaborate on specific points. The interviewee reported that the platform, which included a single-issue think tank, had a longer life than what was provided by the incentives, and was able to appear as a much stronger force in the policy debate.
5.2.2.2 Think Tanks

The sample of think tanks is not large enough to be considered representative of the possible attitudes towards the integration between research and advocacy. In particular, none of the centers included is considered an advocacy tank. However, the three positions represented in the classic distinction between academic think tanks, advocacy tanks and research contractors (Weaver 1989) resonated in some of the interviews. Representing the classic position of academic neutrality, an interviewee noted:

> The message is not as controlled as when you have an advocacy objective, it has a life of its own. That makes the use of tools that are adequate for advocacy purposes, such as position tracking, much more difficult.
> (Personal interview, February 29, 2016)

Other centers did not have the same type of reluctance to speak about specific changes in policy, and admitted that their portfolio includes some projects more strongly oriented towards change. Since these tend to be larger projects, the pressure to show impact is also higher. Finally, the discourse at the Urban Institute is simply one of providing the evidence base that is required by the work of decision makers, in a value-free zone that does not focus on making explicit recommendations.

5.2.2.3 Independent Experts

According to most of the independent experts interviewed, think tanks’ usual focus on academic rigor has not motivated them to the conduction outcome assessments. Jane Reisman indicated that “think tanks are not likely to commission us evaluations, as they are more interested in pure research, quality and objectivity”. However, all the independent experts interviewed shared the view that advocacy campaigns need to be based in persuasive evidence, so that it’s often part of a solid strategy to provide advocacy groups with knowledge resources. According to Johanna Morariu, “in the work that we do often policy research is taken as an input into an advocacy effort, and most of the evaluation focus is on the advocacy effort itself”. In this context, Julia Coffman explained how think tanks provide that information but do not engage in understanding the aftermath of a research product:
I don’t think that think tanks are very good at this [measuring how a message resonates with its audience]. There aren’t really good examples. Think tanks are better now about thinking of audiences first hand and getting the questions right. They have a good design at the front end, but then it kind of stops.
(Personal interview with Julia Coffman, February 2, 2016)

That difficulty is anchored on the nature of work. As indicated by Jackie Kaye, advocacy and policy research have contrary barriers for evaluation: in the case of advocacy work, the intended outcomes are very clear and include passing new legislation or other forms of policy change, while the strategies are more difficult to identify. On the contrary, research has a clear methodology but no obvious outcomes.

5.3 Foundations and their evaluation practices

Although some foundations conduct evaluations internally or commission them to external consultants on a regular basis, most of the respondents employed at foundations clearly stated that evaluations are never conducted for the sake of it, as they are an important investment, both financially and in terms of time management. In other words, no organization, neither among foundations nor among grantees, has the resources to evaluate every grant. The most general view is that evaluations are thus to be conducted selectively, although the criteria for selection may vary significantly from one institution to the next. Some principles are quite commonsensical, like the alignment of the activity with the foundation’s priorities, or the significance of the program—a major effort is more likely to be subject to methodical scrutiny that a small, isolated grant to an unknown organization.

An important trend in the selection principles is the evaluation at the initiative or cluster level, which are seen as better assessing the effectiveness of the foundation’s work by measuring the cumulative contributions of a number of grantees and other actors in the same space, instead of trying to determine the results of individual efforts that might be isolated and difficult to track. In particular, when measuring impact, the MacArthur Foundation looks for trends around defined variables:

We fully recognize that as a grantmaking foundation, in the issues that we tackle we won’t be the sole player nor will any impact be attributable to us. That doesn’t mean that we don’t define our aspirational goals and our rationale for what we do and what we want our impact to be for a population or an issue. We define a number of variables so that we can observe movement along trends
Respondents reported this level of analysis with different intensities. The MacArthur Foundation only evaluates at the initiative level. The Robert Wood Johnson Foundation is moving towards the same scenario, but is still conducting evaluations at the project and program level, which have been very frequent in the past. In the words of Brian Wilson, Assistant Vice President of Research-Evaluation Learning, “the conclusions that we reached when looking at particular programs were much more difficult to generalize than in these broader studies” (personal interview with Brian Quinn, Robert W. Johnson Foundation, February 16, 2016). Another respondent indicated that the benefit of such cluster or initiative-level approaches is that they can include different pieces and contributions by grantees with different core activities, which may include research, outreach, direct service or advocacy.

In absence of the principle to evaluate at the higher level, the selection process is not streamlined and may depend from discrete decisions. Very often the green light is given when an opportunity for learning is perceived either internally or in grantee organizations. Methodologies and objectives change from time to time depending on what is being evaluated. Respondents indicated that an important driver for evaluating an organization is the express request of the grantee’s leadership to do so, as was seen in the case study of the Center for Global Development in Chapter 4.

Respondent feedback also supports the notion that, when evaluation capacity and systems exist internally at grantee organizations, evaluation is likely to happen internally to the grantee and in a more distributed manner. For example, an evaluation specialist at a think tank explained that his organization evaluates many of their programs, including events bringing together experts for facilitated and informed dialogue, either in specific issues or in broader areas likely to become the subject of discussion. Requests for evaluation can also come from an internal client at the foundation:

one of my colleagues asked me to conduct an evaluation of one of their grantees, and I saw, well, your manual is not specific and you don’t have a logic
model, so this is not ready for evaluation. With my staff I helped with the manual, and with the logic model we hired an external researcher to go in there and help them develop a logic model. Without that, they wouldn’t be ready to talk to an evaluator. We try to build off their capacity so then they can do self-evaluation, and if they need a logic model we’ll do that.

(Personal interview, March 4, 2016)

Often times, the goal of the evaluator is to elevate the grantee to a stage at which it can conduct self-assessments. Johanna Morariu reports about her work with the Post-Carbon Institute in similar terms. In this case no foundations were involved.

Our role was to help them with strategic planning and theory of change development. They had gone through an executive director transition and had a number of programs running and research initiatives, but they weren’t aligned. Our work was to review the work that they had done, do a number of interviews with different programs, and develop a theory of change to guide and support their work, and then we handed them a number of data collection tools because they wanted to evaluate their work themselves. We developed those tools so that they could conduct their own evaluation.

(Personal interview with Johanna Morariu, February 3, 2016)

5.3.1 Uses of evaluation by funders

There was significant variation in the answers to the question about the use of evaluators among foundation experts, although some general patterns could be identified. Firstly, evaluation is generally considered a tool that serves strategic planning, never an end in itself. Secondly, most respondents affirmed that their organizations conduct formative evaluations. In any case, there were significant exceptions to this trend, with some respondents indicating that evaluations are conducted to decide which checks to write. The boundary between the two may not be definitive, as the adoption of a new or overhauled strategy can have the consequence of not extending grants to organizations that were formerly considered instrumental to the organization’s mission. In any case, the statement that “evaluations inform decision making”, while being the most common response, needs further qualification to discern between summative and formative uses. As an example, Chantell Johnson explained the different ways in which evaluations are shared and discussed at the MacArthur Foundation:

Each program team discusses our evaluation activities as a team and writes a management response (both of which we make public). We also bring all evaluations to our president for dialog with the program team, and we share our
evaluations, findings and results, with our board. Sharing with our board has taken two forms – formal and informal. We have shared our evaluations, including measures of outcomes and impacts, in formal board sessions where decisions are being made but also in learning roundtable sessions where the focus is on learning about the evaluation, the measures and from our work. (Personal interview with Chantell Johnson, Macarthur Foundation, April 1, 2016)

For Tom Kelly, Vice President of Knowledge, Evaluation and Learning at the Hawai’I Community Foundation, the most common and easiest use of an evaluation is applied to the subsequent grantmaking cycle. Kelly suggests that funders should fit findings into existing strategies and make sure not to expect evaluation reports to trigger a new agenda. More ambitious uses for organizational learning are limited by the usually narrow focus of evaluation questions, and the fact that most reports are “extremely context-bound”:

> Evaluations can uncover issues or changes or levers within a system that could occur. But unless the evaluation was designed to assess the system’s change over time, I think you get small pieces of information that are not very influential. And you require multiple versions of a system evaluation to go deeper on those items, and particularly program evaluations don’t give enough [insight] at the systems level. (Personal interview with Tom Kelly, March 21, 2016)

In recognition of these limitations, some funders have added different streams of information. An evaluation specialist at a foundation made abundantly clear that evaluations are only one of the elements contributing to a decision, and highlighted that it might be seen as more important for the evaluation unit to produce a study that is worth sharing with the field in order to multiply the learning factor. The evaluation team can be only one among several units working in the production, management and distribution of information and knowledge. This is how complex systems can be at a data-driven foundation:

> I am in the research and evaluation group, but we also have a data development group, data management, performance measures, and knowledge management. And all of those groups are about helping our foundation in different ways, so for example performance measurement may sit down with a programmer and their grantee and work on developing performance measures so they can do their own self-evaluations. Knowledge management works for policy a lot and aims at identifying internal and external resources, […] moving them and making sure that they get to the right people. And then in data development we work with census group and different integrated data systems throughout the country, and try to convince policy makers at different levels that integrated
data systems are a good thing. There’s a lot of ways in which we are informing, and the research and evaluation [unit] is just one of them. (Personal interview, March 4, 2016)

As a learning component, respondents considered logic models to be helpful in identifying gaps and untested assumptions. When supporting advocacy efforts, the evidence obtained from surveying audiences can help redefine tactics addressing the issue from different angles. Two evaluation experts at different foundations stressed the importance of understanding what the decision points of a program are in order to define what information will be needed at a specific time and address it with evaluation findings:

There is a number of reasons why [program officers] would want to have input in particular at that point in time, so we try to target the evaluations there to coincide with a key decision or set of decisions that need to be made. (Personal interview, March 4, 2016)

5.3.1.1 Evaluations as integrated know-how

A number of respondents at foundations indicated that their organizations had formalized different information streams fed by evaluation findings. At the MacArthur Foundation, there are four information needs: landscape, process, outcomes, and impact. Landscape questions were described as follows:

We have landscape questions. What are the things that we need to be paying attention to in a given space? May be political or legislative, community based, or it may be a shift in the space where we’re working. We outline a series of landscape variables we want to pay attention to. Our primary questions are: What is the context, what are the social and environmental conditions, and how are they changing over time and space? Which external variables pose a threat and to which must we be sensitive? (Personal interview with Chantell Johnson, MacArthur Foundation, April 1, 2016)

Process information is closely tied to programmatic work, and focused on reflection and feedback loops by means of which the foundation hears directly from partners and grantees. “Primary questions are: what are the inputs, activities and outputs associated with the program or project? What is working? What is not working and why? What could be improved?” Outcomes evaluation often entails considerable field work to measure progress towards specified goals. Finally, impact evaluation aims at better understanding contributions. “Our primary questions are: To what extent have the goals been achieved? What contributions has the program made? With what intended and unintended results?”
An evaluation specialist at a foundation that seeks to combine policy research and an orientation to advocacy offered a particularly sophisticated account of how her organization uses evaluation.

1) **summative evaluations** to wrap up initiatives after their closure and diagnose whether objectives had been met.

2) **strategic evaluations**, typically conducted at mid-term as a point of reflection and to allow for strategic shifts during the implementation of the program. They combine a summative aspect with learning goals that are reflected in planning the remaining duration of the program.

3) **tactical evaluations**, conducted on request of the project team in order to answer questions deemed necessary to move the project forward. They are not shared with the board and can be especially valuable for programs “that are exploratory”

These programs may not have a policy goal, it may be a research program that’s trying to inform or make policymakers use information in their decision making. We’re not trying to tell them how to use the information, we just want them to be more informed about things they may [be unaware of]. A great example is our work on fiscal federalism. It’s about connecting policymakers at the federal level, and at the state and local level. A lot of times, particularly federal decision makers don’t understand how their decisions affect policymakers at other levels, and this has a particular relevance with decisions that have financial implications. So this whole program is designed around providing information to federal policymakers to help them recognize, understand and just think more about the potential implications of their financial decisions for state and local [policy makers]. We did convenings and provided other face to face environments so they have opportunities to understand not just clinically or from a written report, but from talking to real people —what does it look like in practice. That may end up down the road in something that has more policy goals, but initially we did a tactical evaluation to understand whether the program’s approach was actually working, whether they were reaching the targeted audience, did folks find that information useful, how they were using it, and how can we do a better job. (Personal interview, February 19, 2016)

4) and **retrospective evaluations** or reviews, conducted 5 or 6 years after the foundation decided to step off a certain issue or area with the purpose of testing the assumptions on which the exit strategy was based.

In a similar tone, an evaluation specialist at a foundation underscored the importance of timing the evaluation in accordance to the knowledge needs of the corresponding project. Different projects may
benefit from incorporating feedback at the outset or at midterm, instead on focusing of post-completion assessments.

5.3.1.2 Evaluating bodies of research

When funding research, the most common use of evaluation assesses the effectiveness in building a new field of knowledge, and the influence of that new body outside of academic circles. This type of effort happens at a regular basis at the Robert W. Johnson Foundation. The Hewlett Foundation’s evaluation policies analyzed in Chapter 4 also list this as one of the main contexts that are prone to assessment. For example, the Population and Poverty Research Initiative, which sought to “develop evidence on the economic impact of population dynamics and reproductive health while reinvigorating interest in economic demography”, was evaluated by the RAND Corporation after the for-profit think tank won the Request For Proposals.

Helena Choi, Program Officer in the Global Development and Population Program at Hewlett, described that the findings on the policy communication were frustrating, as few respondents could point to particular studies, but that some secondary components of the initiative, like the support to junior researchers, gained weight:

One of the most rewarding things was supporting PhD dissertation fellowships as a way to bring junior scholars into the field. We thought of them anecdotally because we would see them on an annual basis at these conferences, it was a second line contribution, but the evaluation really helped underscore the importance of that as a strength in the field-building work. You can just support the more established researchers, but bringing new blood is important. (Personal interview with Helena Choi, Hewlett Foundation, February 26, 2016)

5.3.1.3 Evaluating organizations, evaluating think tanks

Finally, foundations also evaluate organizations with the objective of supporting their growth. Although it only happens exceptionally, these efforts can also encompass think tanks. As seen in Chapter 4, the Hewlett Foundation has publicly documented organization level evaluation, usually when an organization is in a pivotal moment, as was the case of the Center for Global Development in 2006. The evaluation that was
conducted then had the support of four major foundations. Unfortunately, the members of the CGD that were invited to participate in the study did not respond to the invitation or declined it altogether. However, there may be other reasons for evaluating think tanks, including the goal of understanding their role as actors in the issue networks that populate the policy arenas, especially at the federal level in the Washington, DC circles. During the interview with Julia Coffman it surfaced that the Hewlett Foundation has recently commissioned her external evaluations of the Brookings Institution, Pew Research and the Bipartisan Policy Center with the overarching objective of better understanding how Congress works.

I’m about to do for the Hewlett Foundation an assessment of a couple of large think tanks, the Pew Research Center, the Bipartisan Policy Center, and Brookings. Since they’re huge, we’re going to do some work around looking at particular projects that they do. They [the Hewlett Foundation] are interested in the functioning of Congress. As they think about their future strategy they want to understand better the roles that their funding of think tanks plays in their strategy.

(Personal interview with Julia Coffman, February 2, 2016)

5.3.1.3 Great expectations

Some of the data collected pointed at a gap between expectations of use and the types of use that evidence can actually support. Sarah Lucas, for example, was aware of the dissonance and distinguished what TTI evaluations can really provide from what she would ideally like to learn from them:

And then the question that you really cannot answer but I wish we could, which is the degree to which TTI support and engagement has had an attributable and measurable difference in think tank’s capacity to engage and influence in the policy environment of their country. Without a solid counterfactual we can’t explore that fully. There may be some anecdotal ways to explore it, and it’s something where we push the evaluation team quite hard. Ultimately, we’d like to be able to answer that question.

(Personal interview with Sarah Lucas, Hewlett Foundation, February 8, 2016)

Other respondents expressed that dissonance as a hope for future developments. In particular, Jackie Kaye explained that the philanthropic sector has until recently failed to see the importance of the issue of how to get research to users, but expected to see answers in the next two to three years as findings begin to accumulate. Yet another source expressed her hope that evaluation results could become an important tool both for program officers, for the general portfolio of a foundation, and as an essential part of the advocacy
toolkit. Her expectation is to be able to see a moment in the future where knowledge about what has worked in past situations is fed into a sequenced mapping that could be used to create more accurate strategies by composing the strengths of different grantees in different moments.

Being able to map type of partner (I define type as “think tank” or, “mobilizer”, or “grassroots” or “grasstops”), the sequencing of where that particular issue might be in the policy cycle, to the activities that work well or not in moving the particular policy issue further down the field. Having that information would be very important at the front end. And I also think that if we had that information and it could be deliberately built into the grantmaking process and the investment decision process, it would be extremely useful.

(Personal interview, February 23, 2016)

5.3.2 The importance of general operational support

The question of whether the output of a grantee can or should be evaluated also depends largely on whether a philanthropic organization awards grants that are earmarked for specific activities, or provides general operational support. The difficulty of securing unrestricted funding is a recurring topic in conversations in the field, with many intermediate organizations recommending it as the best way to make nonprofits sustainable on the long term. Although this issue was not part of the questionnaire, it emerged at several of the interviews. In the context of a discussion about the different cultures of grantmaking across the ideological spectrum, a respondent noted how project-based funding is more likely to be evaluated:

I share the perception that conservative foundations are more likely to provide core funding to organizations whose overall mission they support, rather than tying their support to specific projects, which I think is more often the case from liberal-leaning foundations. The difference does not lie so much in that liberal foundations require rigorous evaluation — it’s more that the liberal foundations are more likely to fund on a project by project basis, which is often tied up with evaluation.

(Personal interview, February 9, 2016)

Grantees often perceive that project-related grants can result in artificial time allocation and the inability to cover expenses that are not tied to project activities, like “buying computers for everybody”. On the other hand, to Jackie Kaye the consideration of think tanks as anchoring institutions that can elevate the discourse in a certain field is closely linked to general operating support:

This is an example of an organization where I would hope funders would provide general operating support. […] If a funder identifies a think tank as a leading
anchor organization for a field, then the funder is more interested in the evaluation of its organizational contribution of that think tank. Partly the way that we think of the evaluation depends on our rationale for the funding of a think tank.

(Personal interview with Jackie Kaye, Wellspring Advisors, February 22, 2016)

As depicted in case study of the Think Tank Initiative above, the rationale for supporting think tanks unfolds in the consideration of think tanks as important pieces in the policymaking ecosystems of developing countries. The unusual combination of unrestricted funding and a complex monitoring and evaluation apparatus deserves additional attention.

5.3.2.1 The Think Tank Initiative and the case for General Operating Support

The example of the Think Tank Initiative (TTI) shows how, although there is a tension between offering general operating support and including an evaluation component, they do not compose a dichotomy. Peter Taylor explained the challenge in combining unrestricted funding with accountability measurements:

What is a bit unusual about TTI is that each institution supported has its individual sets of objectives, laying out its aspirations related to the support. There’s a challenge in understanding how providing core funding and technical support can actually translate into an organization being able to achieve certain things.

(Personal interview with Peter Taylor, February 2, 2016)

That challenge was addressed by aligning institutional objectives with the initiative’s three main areas of emphasis: research quality; policy engagement; and organizational performance.

These objectives, a final version of which is arrived at through negotiation between the TT and TTI, gives us the basis for a grant agreement which affords some accountability and monitoring of results and outcomes, without being overly restrictive on what the TT can work towards.

(Personal interview with Peter Taylor, personal message to the author, March 22, 2016)

External evaluations at the program level seek to gauge the overall performance of the initiative and possible improvements. But Sarah Lucas also noted that evaluators were pressed to answer the differentiated impact between different types of support, comparing the effects of core flexible funding and other items in the TTI toolkit, like capacity building or opportunities for collaboration. Thanks to
unrestricted funding, TTI grantees have been able to depend less on consulting work and efforts around scattered revenue sources.

The more I work with think tanks, the more convinced I am about the importance of core flexible support. Especially in the absence of a country footprint, we have no way to know what the policy priorities in those contexts are.

(Personal interview with Sarah Lucas, February 8, 2016)

The chasm between grantmaking in international development and in the domestic scene was reflected in Sarah Lucas’ opinion that general operational support for US think tanks would be harder to justify, as there is a weaker alignment between think tank output and the program’s interests.

5.4 Measuring the influence of policy research

The question on how to measure the influence of policy oriented research produced significant agreement among respondents, with admitted differences in the nuance given to some aspects. It is a unanimous view on the issue that while it is simple to count research outputs to build quantitative indicators, those indicators have a weak relation with whether and how messages are heard and used in the community of policy makers. Many respondents indicated the need to complement output numbers with qualitative considerations, although it might be hard to find a term of comparison, as Richard Bush noted (see case study A in this chapter). Another recurrent consideration is the impossibility of attributing change to research: independently of what sophisticated studies may accomplish, their objective shall never be to establish a causal relationship between a report and a decision. The agreement in these areas does not come not as a surprise, being solidly consistent with the major takeaways of the research on knowledge utilization and advocacy evaluation reported in the literature review of this study (see Chapter 2).

The ‘influence by contribution’ paradigm was referenced among some practitioners at foundations from a slightly different angle. To these experts, the objective is not to understand the influence of a book or an organization, but to assess the cumulative effect of a body of research. The challenge of the assessment has different aspects. Firstly, impact can take a long time. Secondly, users may remember having read a study, but they may not remember the title, the name of the author or who sponsored it. Thirdly, while the visibility
of isolated studies or experts concentrating the attention of the public is relatively easy to identify, the overall influence of the body of research is surely larger but also less observable. This difficulty connects with the recognition (see paragraph 1 in this chapter) that, especially when compared to advocacy, the outcomes of research are much more difficult to establish and have a very weak observational component. The general methods reported more often are interviews to key informers and the narrative stories of influence. While the first extracts new evidence from targeted audiences and may provide new information, the second is constructed from the experience of program officers and participants with the aim of extracting a meaningful lesson.

For Jackie Kaye, research has a very clear process and methodological design, but it’s much more difficult to establish specific outcomes and to operationalize the objective of identifying research use in the policy making process. However, Julia Coffman indicated that it’s always possible to survey key audiences to measure message reception through proxies in attitude, language and knowledge.

In measuring research influence and uptake, the McArthur foundation combines the attention to public sentiment with discourse analysis in visible settings and traces in the legislative process:

> Often times we look at more than just policy. We look at public sentiment, public discourse around a topic. We’ll be doing public polling, media scan work. We’ll be looking at what happens and what is said in political environments that are transparent and visible to us, like a political debate, a summit, or a bilateral environment. […] And then of course we would track legislative aspects, but that are one step below legislation: proposals, things that showed up in bills or were debated, part of other legislative components or administrative regulations. It’s not always about a federal change. (Personal interview with Chantell Johnson, MacArthur Foundation, April 1, 2016)

Regarding the application to policy research efforts of different templates or frameworks that have been shared by prominent foundations in the field of advocacy like the Kellogg Foundation or the Casey Foundation, Jackie Kaye downplayed the importance of particular tools:

> What it comes down to is: Are we asking the right questions? In terms of policy research I would probably ask the same core evaluation questions, and I do believe that they could be equally effective. There is nothing in public policy research that makes you start with something else. And those questions are:
What do we want to accomplish? How do we get there? How do we know that we get there? What do we want to learn? What should we put in place to make sure that we learn that, so we can build on it later? And for policy research: Who do we want to reach? Did we reach them effectively? And that guides your evaluation plan.

(Personal interview with Jackie Kaye, Wellspring Advisors, February 22, 2016)

Tom Kelly defended a similar position from his perspective of learning about systems change:

The methods should be defined by the questions, time frame and resources – the challenge here is defining good systems questions to evaluate. I am less worried about a method that I am about a clear description about a system, a clear understanding of effect, and relationships to the net system. Systems mapping and other tools are there, but what drives the success of the evaluation is that the intention and the focus around what parts of the system you’re interested in are clear.

(Personal interview with Tom Kelly, Hawai’i Community Foundation, March 21, 2016)

5.4.1 Interviews with key informers

For Julia Coffman, Director of the Center for Evaluation Innovation, it is possible to measure the influence of policy research, but the key is not to look at decisions and policy change, and instead focus on proxies related to the reactions, beliefs, attitudes and knowledge of target audiences. Coffman stated that the bellwether methodology, which has found a widespread acceptance in the advocacy world, can also apply to policy research. Bellwether interview protocols prevent priming by blinding the focus of the interview, and ensure that there are no sampling biases with diverse respondent lists. “There is a number of things that you can capture, including what they think about an issue, how they talk about it, how they are framing it themselves, who they know is out there speaking on the issue or who they see as influential voices, and whether the messages that they are relaying back to you really line up with information you’ve put out there” (personal interview with Julia Coffman, February 2, 2016)). Respondents from some of the foundations interviewed signaled the use of similar techniques. In particular, the same attention to bias and priming that defines bellwether interviews was reported at another foundation in which evaluators are required not only to keep the name of the organization whose influence they are tracing undisclosed, but to avoid sampling biases by including three different categories of interviewees: respondents with ties to the
organization, neutral individuals (usually scientists or journalists) and respondents on the other side of the issue.

According to one of the respondents, in the case of debates and events with high level personalities attendants are the surveyed audience in what may be called “anecdotal evaluations”. In these cases, interviews or surveys ask attendants what they have learned, and what they would do differently. Ideally, a follow-up is administered one year after the event to obtain feedback on whether the event effectively shaped their behavior. In any case, a very strong positive bias was recognized by the specialist, as respondents are likely to provide favorable feedback as a token of courtesy or loyalty.

5.4.2 Stories of success

Both Peter Taylor, Program Manager of the Think Tank Initiative at IDRC, and Sarah Lucas, Program Officer at the Hewlett Foundation, pointed at the Stories of Influence as the major tool to document the influence of the policy research institutes funded under TTI. Ruth Levine, Global Development and Population Program Director, made reference to the “most significant change”, a similar tool by which grantees select the innovation that better describes the effect of support to recipient organizations. Levine expressed confidence that program officers could use the distance between their expectations and those changes, often around capacity building and organizational growth, to base summative judgments.

It’s important to note how these stories of success and significant changes are similar in their narrative nature to the stories of impact that many think tanks use to make the case for institutional influence in annual reports, as seen in Chapter 4.

5.4.3 Other methods for evaluating policy research: Social Network Analysis

Brian Quinn explained how Social Network Analysis is proving to be an increasingly important tool at the Robert Wood Johnson Foundation, where it has been used to measure the social networks of the researchers involved and to help advocates identify gaps in their outreach strategies. In particular, the deployment of social network analysis seems useful in identifying intermediate outcomes that can help teams assess the
situation in a nuanced way, avoiding the binary assessments that a strict focus in policy change may carry along.

Through SNA we identified one challenge in the program [the Consumer Voices for Coverage]. We were funding a consortia of consumer advocates in each state, and when we did the social network analysis we saw that in many states they were not reaching out to minority groups as effectively as they possibly could. That piece of feedback was really important to us. We said well, part of what we’re doing here is expand your reach, when we see the analysis we see that there are some holes, so this is an opportunity to start thinking of how to expand the program or shift gears. (Personal interview with Brian Quinn, Robert W. Johnson Foundation, February 16, 2016)

5.5 Other factors of success

5.5.1 The value of dissemination

In spite of the difficulties in generalizing findings, evaluations are actively shared both internally and externally. Some foundations underscore the importance of public dissemination to foster the advancement of a field that otherwise acts in isolated sockets. Ruth Levine pointed that, when evaluations are conducted at the grantee level or report information about specific grants, the conversation on what can be published and what should remain confidential can be hard, as grantees tend to prefer the redaction of anything that may affect their reputation.

Internally, information flows differ enormously depending on the organization. At a foundation which has only adopted advocacy strategies recently, there is only an informal level of internal coordination and peer learning driven by simply comparing notes among colleagues. Other foundations have units named “evaluation and learning” that have competencies regarding staff training, and conducting retreats that can be at the team level or organization-wide.

5.5.2 The value of combined skillsets

Some funders have very large evaluation teams. Brian Quinn, for example, indicated that the evaluation team at the Robert Wood Johnson Foundation has between 15 and 20 evaluation experts, all of them PhD holders. An evaluation specialist at a foundation formulated the ideal composition of a team that combines
research and advocacy efforts with the integration of four different skillsets: researchers, campaigners, communicators and government relations experts. While the communicators can help package research findings in a way that resonates according to the different audiences, government relations experts can identify who are the policymakers interested in the issue, what is blocking them from collaborating or what motivates them. Campaigners, on the other hand, have a very accurate idea of the process moving a policy issue forward, the different pieces that have to be combined and in which order they should be positioned. At the same time, the inclusion of this strategic mentality helps shape projects in a way that makes them almost naturally aligned with outcome evaluations.

The combination of skillsets was also an important conclusion for Helena Choi and the Population and Poverty research Initiative. As a result of the evaluation, funding was directed towards scholars with recognized public profiles and a demonstrated interest in divulging their academic activity among other audiences. This idea connects with the notion of the public intellectual that is so well represented by think tank staffers.

Other funders expressed the view that, often times, a research project is funded because of the prominence of a given scholar – not because of particular expectations of use. This attitude can be compared to Jackie Kaye’s theory of the anchor institutions.

5.6 Conclusions

The six case studies at the beginning of this chapter show that the attitudes of think tanks towards the assessment of policy influence are as diverse as the industry itself. In their relations with philanthropic funders, think tanks conduct a number of different functions, including instrumental as well as knowledge-driven roles. A strong positive selection bias towards those organizations that have decided to adopt some sort of monitoring and evaluation can be deducted from the scarce number of think tank staffers among the list of respondents, supporting the general diagnosis that evaluation has not been adopted as a standard practice by managers. Ongoing debates at the Brookings Institution regarding the comparison between
quantitative and qualitative data to base assessments of influence are a good indicator of the general tone of internal discussions in the industry, as witnessed in the think tanks scholarship discussed in Chapter 2. The annual reports analyzed in Chapter 4 also support this idea, as most of them sought to strike a balance between output figures that account for the volume of intellectual production and media visibility, and narrative stories of success about the influence and uptake of specific projects, which tend to underline exceptional achievements.

Against that background landscape, the main hypothesis of this study is supported by the collection of four case studies in which foundations can be attributed as a driving force in the adoption and practice of monitoring and evaluation tools and systems. In two cases, think tank staffers were currently involved in developing a monitoring and evaluation framework that was seen as instrumental for fostering relationships with funders and getting as close as possible to demonstrating influence. At another think tank, evaluations are conducted both internally and on demand of third party funders, usually foundations. The Urban Institute, a research contractor that defends the paradigm of evidence-based policymaking, belongs to that background landscape in that it does not evaluate the influence of its products. However, the fact that its Policy Advisory Unit is funded mainly by foundations suggests that organized philanthropy is behind innovations designed for bridging the gap between knowledge producers and knowledge users. While evaluation builds an information loop with research uptake with the aim of improving and adapting strategies, the new Unit at the Urban Institute is an action-oriented response to the same overarching needs: improving and demonstrating influence.

The last case study was devoted to the Think Tank Initiative. Although TTI is certainly an exception that falls outside the otherwise American scope of the study, its global recognition and scope justify its presence. TTI’s success may play an important role in the industry-wide adoption of monitoring and evaluation tools for the promotion of high-quality, sustainable independent policy research with a double eye on the principles of accountability and organizational learning. To this regard, TTI’s defense of unrestricted core
funding as the best way to safeguard institutional responsiveness to the political environment is deeply ingrained with the combination of a monitoring framework that adapts program principles to the goals of each institution and independent evaluations conducted at the higher program level. While the adoption of a similar program domestically is out of discussion, domestic funders may extract important lessons from the program. These lessons may not be very visible in the short term. In fact, none of the three interviewees at the Hewlett Foundation were able to point clearly at how the experiences of TTI and other funding in international development may affect domestic relationships with think tanks.

Related or not to the TTI experience, the Hewlett Foundation has recently commissioned independent evaluations of sampled programs pertaining to three major research organizations (the Brookings Institution, the Bipartisan Policy Center and the Pew Research Center) with the objective of understanding “how Congress works” and how think tanks mediate in the policy arena. The organizations selected stand out for their different approaches to influence: while Brookings is the most renowned nonpartisan think tank in the world, the Bipartisan Policy Center has a model that actively combines party players in order to draft practicable regulation. Finally, the Pew Research Center brands itself as a “fact tank” that conducts opinion polling and refrains from any particular policy positions.

The sample of think tank staffers was too small to defend or reject the secondary hypotheses regarding the increased likelihood of adopting evaluation in centers that are advocacy oriented or engaged in the field of international development. Although data are limited, they point at a more complex, nonlinear relation between research, evaluation and advocacy: while programs with advocacy outcomes are much easier to evaluate, Case Study D depicts a nonpartisan center working on an evaluation, monitoring and learning framework. This supports the idea that the financial incentives for adopting similar frameworks can be very persuasive to leadership. Similarly, foundations with a strong focus on advocacy are likely to measure research uptake towards certain policy goals, but the example of the MacArthur Foundation shows that strict advocacy goals are not a necessary condition for the evaluation of initiatives with a strong research
component. Intermediary outcomes in public sentiment and prior to legislative action are considered good proxies of how the trends around a certain issue evolve over time. Where grant portfolios consist mainly in service delivery and evaluation is established at the program level, policy research and think tanks tend to be seen as established in a different stream of knowledge production, oriented towards policymakers and incompatible with the objectives of evaluation research.

In accordance with the fuzziness of the organizational typology, data provide no uniform image of what a credible think tank is or what it should do. While some foundations prefer actors that conduct strictly academic work, others prefer centers that go beyond the activity of pure research towards a heavier orientation to action. In any case, think tanks are important partners for many funders in multiple roles. They are seen as strong academic players, anchor institutions, partners, conveners, and well-connected actors.

Regarding the methods and techniques for measuring public policy influence polling, interviews to key respondents, and narrative case studies stand out as the most important ways for measuring and reporting influence. Another innovative technique that seems to be taking force is social network analysis, which allows to visualize influence streams, and identify both power actors and gaps in outreach.

The treatment of internal evaluation use at foundations points at quite complex organizational structures in which evaluation activities provide one among many sources of information and knowledge in combination with knowledge management systems and internal research units, for example. The wild range of variation from organization to organization impedes a general description of how evaluation is used specifically and how it informs grantmaking decisions and strategic planning. In any case, the existing cycles of research, learning, and evaluation speak of an interactive model of research use: “the use of research is only one part of a complicated process that also uses experience, political insight, pressure, social technologies, and judgment” (Weiss 1979, 429).
Sarah Lucas also noted that, in spite of the alignment between the two layers in the second phase, the results of the evaluation were not as eloquent as expected (Sarah Lucas, e-mail message to the author, April 15).

Chantell Johnson, e-mail message to the author, April 8, 2016.

Ibid.
Chapter 6: Recommendations and Best Practices

6.1 Recommendations

Recommendation 1: Learn from the field of strategic communications

One of the most recurring pieces in valorizing the contributions of think tanks was the consideration that evidence ought to be packaged strategically to stay influential. Some of the interviewees, like Justin Milner at Urban Institute, mentioned the production of visualizations, videos and blog posts.

Jackie Kaye highlighted the benefits of connecting policy research with the lessons of strategic messaging and communications from a more experimental perspective:

> If you have policy research findings, wouldn’t it be great to do randomized control mini-experiments, where half of the people you want to influence get a two-page summary of the findings, and half of the people get a video and a two-page summary of the findings? I think that the work to evaluate policy research needs to get a little more creative.

(Personal interview with Jackie Kaye, Wellspring Advisors, February 22, 2016)

Similarly, Julia Coffman saw that the “wave of the future” may lie in the power of digital analytics for capturing “audience exposure to information and electronic consumption”. Even without advancing that far, some of the tools of market research that gauge behaviors, attitudes and beliefs of target audiences could be adopted more often by public policy research organizations.

Recommendation 2: Think about audiences

A recurring recommendation linked to the above is to focus on the users of information. This recommendation dilates into multiple dimensions, as the information consumed can consist in evaluation findings, or in the policy research project; while intended users can be internal to the organization, members of a partnering organization, or decision makers. For that reason, the recommendation to understand the work culture and information needs of users can take many forms. Some organizations focus on understanding the needs of policymakers at the national and local level; others have focused on fostering the understanding of how policies at the federal level constrain the possible decisions of other policy
makers; while still others have in mind what can be done to make evaluation findings more useful beyond the context that generated them.

As a design principle, the focus on users can be adopted in the earliest phases of research design by addressing the information needs of specific audiences and concentrating dissemination on them. Additionally, that intensive focus decreases the barriers for measuring uptake in terms of opinions, behaviors, attitudes and use.

**Recommendation 3: Map the inclusion of public policy research in larger initiatives**

Some of the respondents expressed the hope that the findings of public policy research will be integrated in a purposeful and intentional manner in larger initiatives including different tactics, like grassroots mobilizations, awareness campaigns, etcetera. In this context, evaluation is seen as a test that can help identify what is the most effective placement of new information in a planning sequence.

6.2 **Best practices**

6.2.1 **For funders**

**Best Practice 1: Build anchor institutions by aggregation**

Offer financial incentives for collaborations between grantees working in the same area with the objective of working through their disagreements and finding a common message that can be disseminated through a collaborative platform.

**Best Practice 2: include unnatural allies in the coalition**

Including communities and voices that are usually on the other side of an issue, like a hunter and an environmentalist, can decrease the ideological perception of a message and strengthen its credibility.

6.2.2. **For evaluators and commissioners of evaluation studies**

**Best Practice 3: Avoid superimposing strict models**

Although logic models and theories of change can be useful tools for identifying and testing assumptions and assigning intended outcomes to a project, many funders avoid superimposing their own model, and
focus instead on offering support to organizations with similar goals, including support for capacity building in evaluation and planning.

**Best Practice 4: Establish intermediate outcomes**

Outcome evaluation of policy research cannot be tracked with legislative change only. Other intermediate outcomes that signal changes in attitudes, behaviors, discourse and opinions can be measured. For example, public polling, discourse analysis of public political debates and bill proposals that were not passed, etc.

**Best Practice 5: Use evaluation to better understand the world**

In the realm of public policy research, the assessment of influence is usually not a summative look on whether a program delivered what it promised. Much rather, it is an examination of how the surrounding environment interacts with the objectives of a program to understand whether the message has to be repeated, translated or shifted.

**Best Practice 6: Identify and address key decision points**

Evaluation is more likely to be informative and useful when it has been designed to address program-related information needs and is delivered on time for a decision. This often implies designing the evaluation and the program as interdependent processes.

**Best Practice 7: Make Social Network Analysis part of the influence assessment toolkit:**

The Robert Wood Johnson Foundation has adopted the use of SNA as a way to measure the social outreach of academics, but also to provide insights to advocacy organizations into which groups and communities are underexposed to a certain message. The references to the utilization of SNA for assessing policy influence are augmenting: see Borgenschneider and Corbett (2010, 300) and the guide published by Network Impact and the Center for Evaluation Innovation (2014).
6.2.3 For think tanks

Best Practice 8: Package information differently

Push beyond the traditional publication of information in books and pdf reports towards formats that summarize information for easy consumption: videos, infographics, visualizations, stories or shorter texts.

Best Practice 9: Target specific audiences

Although federal legislation is usually targeted due to its high profile, less known policymakers at the state and local level, as well as mid-rank employees in government agencies have an important weight in policy implementation. Targeting them at the right time with information that is relevant and connected to their sphere of influence is hard, but can make a difference.

Best Practice 10: Connect the dots

Policymakers are rarely aware of the consequences of their actions for the work of other agencies. Relevant information about unintended consequences in other policymaking communities can promote more nuanced, self-aware and responsible decisions.

Best Practice 11: Create multidisciplinary teams

Varied skillsets can effectively triangulate the lifecycle of a project between the phases of planning, research, communication, and evaluation.

Best Practice 12: Collaborate with advocates

When there are clear legislative goals, collaborate with advocates that can make the message heard in an active way. The inclusion of advocates can also facilitate the assessment of which groups and communities have been informed by a project or are using evidence to support their argument.
Chapter 7: Conclusion

The study set out to examine the collaborations between foundations and think tanks in order to identify if the former are driving the evaluation of the latter, and to determine whether and how the influence of public policy research is being assessed in that intersection. The motivation for the study was grounded in a paradox: while many collaborations between foundations and think tanks are longstanding and well known, there has been to date no attempt at framing how they are affected by the increasing reliance of organized philanthropy on impact and outcome assessment.

The overview of the classic academic debate on the application of social scientific knowledge has proved the difficulty of delivering public policy research to decision makers, and the exceptional character of linear, instrumental uses in which research findings are applied to form a decision. The scholarship on think tanks and their influence built on that recognition and constructed arguments about the relevance and influence of these organizations based on their media profiles, congressional testimonies and their intervention at different stages of the policy process. While these notions allow to compare the reputation, visibility and credibility of different centers, they have contributed to the idea that the influence of think tanks cannot be measured. In opposition, a number of tools and techniques for outcome evaluation have been designed and adopted by numerous organizations working in the spheres of advocacy and international development. Polling, focus groups, and specific techniques like bellwether interview protocols all aim at measuring beliefs, opinions, attitudes and discourses of key audiences against the message and framing of the evaluand. Methods for mapping and tracking change in complex systems including networks of actors have been used to help link research outputs and outcomes. The recognition of these two pioneering areas has served as the basis for the elaboration of the main hypothesis of the study into two secondary hypothesis: think tanks are more likely to evaluate their work when they are invested in advocacy efforts or in international development.
The information published on the websites of selected think tanks and foundations confirms the general disconnect between the prevalent arguments on influence of the former and the priority of learning from experience to maximize impact of the latter. In general, think tanks make the case about their influence with a combination of output indicators, such as the number of books, reports, op-eds and blog posts published, as well as social media following, and short narrative elements, often anecdotal or lacking a clear rationale, which reconstruct the success of particular projects. Against that, some prominent foundations do measure research uptake and influence (Hewlett, Robert W. Johnson, MacArthur), or conduct evaluations at the organizational level to support grantees at a particular point in time (Gates, Hewlett). The case study of the Center for Global Development musters how evaluations at the organizational level can measure influence, recognition and research use in a much more objective and nuanced way, while at the same time serving as a public relations tactic promoting, in this case, the association of the think tank with the values of transparency and openness. On the other hand, the case study of the W.T. Grant Foundation shows that the role of intermediary organizations, such as think tanks, can also be understood at a more systemic level by commissioning academic studies on research utilization that build on the knowledge use debate and contribute to it with a specific lens on local policymakers. Indeed, the portfolio on testing strategies to increment the use of research is very closely connected to the activity of the Policy Advisory Unit at the Urban Institute, which is actively working on connecting research findings to users, particularly decision makers in large cities across the country. These two examples show that outcome evaluation of policy research is not the only option to foster influence and research use: action oriented strategies that embed a two-way communication process with final users seem to be taking the lead.

In contrast with that general landscape, the main hypothesis of the study has been supported with the identification of four case studies in which foundations can clearly be attributed as the driving force for the construction of monitoring and evaluation capacities and the conduction of evaluation research. When think tanks build such systems, they are significantly improving their ability to make a persuasive argument about
their influence and contribution to the policymaking environment; and this argument is an important part of their fundraising strategy. At the same time, the Think Tank Initiative and other examples of think tank evaluation may have opened the field to studies on the impact of prominent think tanks, like the Brookings Institution, the Bipartisan Policy Center, and the Pew Research Center, that have recently been commissioned by the Hewlett Foundation.

For most foundations engaged in advocacy efforts, policy research findings are usually included as the factual base supporting a policy position. The persuasiveness and outreach of that base can be measured with the methodologies indicated, allowing to establish intermediary outcomes that complement the binary vision of policy change, which is often insufficient for nuanced assessments of complex initiatives. The common recognition that no single organization or study can claim attribution over change stands in harmony with the notion that, either formally or informally, intermediary organizations like foundations and think tanks enter into advocacy coalitions around their issues of interest. The study embraces the claim that a notion of advocacy “with a small a”, or an “issue advocacy” (Borgenschneider and Corbett 2010) lends the correct perspective for properly framing the influence of scientific evidence for policy arguments, and enhancing more transparent collaborations between advocates and researchers in the promotion of change and progress.

7.1 Limitations

This study is only a modest step towards a better understanding of how evaluation research can be adopted to promote the influence of policy. Given the scope of the fieldwork, the case studies and best practices collected are examples of excellence that cannot advance a general diagnostic of the field, nor answer the question of whether think tanks will benefit from the adoption of evaluation practices.

Since the sample of interviewees contained no advocacy tanks and only one think tank involved in international development (IDRC, managing organization for the Think Tank Initiative), the secondary hypotheses have neither been endorsed nor rejected. In other words, it remains unclear whether the
In the study of philanthropy and think tanks, it is observed that the proportion of advocacy tanks and developmental think tanks that is adopting evaluation is higher than among nonpartisan and domestic policy centers. Further research could focus on surveying the state of the art in the assessment of policy change outcomes among the many advocacy tanks that populate the industry, and in the reception of lessons learned from initiatives and centers like TTI, IDRC and ODI across the developmental field.

Another important bias that affects both the foundations and the think tanks discussed is ideological: none of the organizations represented is situated in the conservative side of the spectrum, and therefore Andy Rich’s important thesis that conservative foundations tend to support their anchor institutions with unrestricted funding, and are hence less inclined to measure the effectiveness of their investments with evaluation procedures, has not been tested. The results of this study represent only a small number of high profile institutions, most of which are clearly espoused with liberal and progressive agendas. Under this lens, the treatment of advocacy efforts and their combination with research is usually associated with policy positions that defend the interests of minorities, low income communities and communities that struggle to find an equal stance on national, state and local American politics. It is stained with the nuance of defense, equality and civil rights. But once the term “advocacy” absorbs its complete meaning and pushes all around the political spectrum, including the aggressive tactics of pro-business and free enterprise politics, it falls prey to the infamous notion of special interest and corporate lobbying that is often associated with elitist theories of power. In other words: by selecting only the gentler side of both advocacy and philanthropy, this study has largely ignored the problem of power and how it undermines the credibility of both American think tanks and foundations.

### 7.2 Further research

The effectiveness of philanthropy in shaping policy will without doubt emerge as an important topic in the academic literature. This study seeks to better understand the role of public policy research institutions by supporting the factual view that think tanks can be and are, occasionally, evaluated. But the true extent of
these practices is out of the present scope: are think tank evaluations isolated exceptions, or is there a building momentum towards the application of evaluative research to understanding the influence of think tanks and public policy research?

There are multiple possibilities to extend the direction of this study towards more reliable and generalizable results. Other efforts may expand the sample including advocacy tanks, developmental think tanks and conservative organizations. More evidence is also needed in order to validate the professional stories collected in Chapter 5. Experiences and perceptions of different professional profiles within the same organization could be compared and contrasted, including the role of communications officers, with important responsibilities in both packaging information and measuring its outreach, without neglecting the important differences between program officers and evaluation specialists. In a similar line, boards’ perceptions of the utility of evaluation research is another fundamental empirical factor that falls entirely outside the scope of this study. By shedding light on dissonances and divergences, such efforts can provide interesting evidence regarding the internal function of evaluations. Empirical work on the consequences of adopting monitoring and evaluation systems for think tanks is another important direction for future researchers, especially in the comparison of external benefits (secured funding, credibility) against internal benefits (streamlined management practices, or improved knowledge management, for example). As many funders concentrate their efforts in state-level policies, attention to state-based think tanks and their funders may offer an important direction to expand the question of whether evaluation and monitoring is becoming an important part of the conversations around the role and influence of public policy research institutes.

On a more general note, this study was guided by a number of more vital, less technical questions: Does research matter? Can it be used to improve our lives, to ameliorate our social coexistence? How can political and social leaders understand scientific evidence? None of these questions have a definitive answer. Precisely for that reason they are worth asking, once and again. Hopefully, my study will be a helpful orientation to thinkers that want to keep looking for queues to these questions in their own ways.
Appendix A: Operational definitions of measures of influence

Books: Total number of books published within the year, or annotated list of the titles.

Journals: Periodical publication of academic nature is featured as a distinct product.

Briefs and other: Number of outputs of named briefer formats that reflect the editorial brand of the organization.

News Outlets: Publication or broadcast program that offers news to a more general public.

Events: Similarly to the books indicator, events can be tallied with the overall number celebrated or with a comprehensive summary of each of them.

Web visits: Overall visits to the website, as well as related metrics such as unique visitors, yearly increases of number of pages viewed.

Newsletters: Number of subscriptions to the main newsletter of the organization.

Facebook: Number of likes and/or followers on Facebook, and/or yearly increase

Twitter: Number of twitter followers, and/or yearly increase

YouTube: Number of YouTube views, as well as related metrics such as minutes viewed, and/or yearly increases.

Other (multimedia): Other multimedia formats, including podcasts, event webcasts or streaming.

Op-ed: Number of Op-eds placed or outlets where they are placed.

Citations: Number of fellows and/or reports cited by the media.

Interviews: Number of interviews conducted to fellows, or a highlighted mention to such an appearance.

GTTTI: Rank assigned in the Go To Think Tanks Index.

Transparency: Transparency rating assigned by an independent group, like Transparify or Charity Navigator.

Other (awards): Other awards mentioned

Congressional testimonies: Number of times fellows testified before a congressional committee, or a highlighted mention to such an appearance.

Stories of impact: Explicit connections between an activity or product of the organization and a visible outcome, usually in narrative form (see appendix B).
Appendix B: Stories of impact

American Enterprise Institute

Wealth Building Home Loan: In 2014, Edward Pinto and Stephen Oliner created the Wealth Building Home Loan (WBHL), a new approach for banks to provide homebuyers a more reliable and effective means of building wealth while maintaining buying power similar to a 30-year loan. The Neighborhood Assistance Corporation of America (NACA), in conjunction with the Bank of America, has decided to incorporate the WBHL concept into its low-income lending program in NACA’s 40-plus offices nationwide.

Trade Promotion Authority: Derek Scissors was a go-to expert on trade promotion authority, being interviewed for background information by members of Congress. The House Ways and Means Committee circulated Scissors’s April AEIdeas blog post, “Trade Promotion Authority, finally,” as a prime example of the conservative argument in support of free trade, and Scissors was frequently called to Capitol Hill to brief congressional leaders.

Higher Education Reform: Andrew Kelly has been recognized as the intellectual architect for several promising new policy ideas for higher education reform put forward by major political figures. In his contribution to the YG Network’s Room to Grow volume last year, Kelly outlined a proposal to reform the college accreditation system. Slate’s Jordan Weissmann noted that these ideas were carried through to legislation introduced by Sen. Mike Lee (R-UT).

Acquisition Reform: Many of AEI’s ideas on acquisition reform, especially those of former visiting fellow William Greenwalt, can be found in Senate Armed Services Committee Chairman John McCain’s (R-AZ) defense authorization bill being considered in conference. In June, Mackenzie Eaglen published a chart of Greenwalt’s work on acquisition reform to show AEI’s progress in affecting actual legislation and areas for continued improvement.
AEI AS INFLUENCER

AEI has made a concerted effort to disseminate its scholarship to those best placed to act, thus expanding its outreach to policymakers and their staffs on both sides of the political aisle.

AEI scholars testify before congressional committees more often than any peer organization, with a total of 99 testimonies before the 113th Congress and 99 testimonies through mid-November before the 114th Congress.

- **New Member Orientations**
  AEI played a crucial educational role in the wake of the 2014 midterm elections, helping to lead three different policy orientations for new members of Congress.

- **Trade Promotion Authority**
  Derek Scissors was a go-to expert on trade promotion authority, being interviewed for background information by members of Congress. The House Ways and Means Committee circulated Scissors’s April AEIdeas blog post, “Trade Promotion Authority, finally,” as a prime example of the conservative argument in support of free trade, and Scissors was frequently called to Capitol Hill by brief congressional leaders.

- **The Elementary and Secondary Education Act Reauthorization**
  Frederick Hess has played a leading role since 2004 in critiquing the No Child Left Behind (NCLB) Act and offering reforms that would decrease bureaucracy while increasing transparency and innovation. The 2015 Student Success Act and the Every Child Achieves Act, put forward by the House and Senate, respectively, both closely track Hess’s vision for NCLB reauthorization. On the day the House’s bill was brought to the floor, AEI scholars published two op-eds in National Review Online. One was tweeted by House Speaker John Boehner (R-OH) and emailed to the entire House GOP conference by Cathy McMorris Rodgers (R-WA), and the other was cited by Chairman John Kline (R-MN) in a meeting he convened with outside organizations.

- **Higher Education Reform**
  Andrew Kelly has been recognized as the intellectual architect for several promising new policy ideas for higher education reform put forward by major political figures. In his contribution to the YG Network’s Room to Grow volume last year, Kelly outlined a proposal to reform the college accreditation system. State Senator Weissmann noted that these ideas were carried through to legislation introduced by Sen. Mike Lee (R-UT).

- **Acquisition Reform**
  Many of AEI’s ideas on acquisition reform, especially those of former visiting fellow William Greenwall, can be found in Senate Armed Services Committee Chairman John McCain’s (R-AZ) defense authorization bill being considered in conference. In June, Mackenzie Eaglen published a chart of Greenwall’s work on acquisition reform to show AEI’s progress in affecting actual legislation and areas for continued improvement.

Bipartisan Policy Center

The Bipartisan Policy Center published the influence of its ‘lobbying arm’, BPCAN:

- **BPC, through its Governors’ Council, helped shape the Workforce Investment and Opportunity Act to increase state flexibility and encourage innovation. The act was signed into law in July 2014.**
- **BPC’s housing recommendations were incorporated into the Senate Banking Committee’s bipartisan government-sponsored enterprise reform bill that passed out of committee.**
- **BPC’s health care cost containment recommendations were reflected in several pieces of new legislation, including the Tri-Committee health care bill, which seeks to resolve the chronic problem of funding physicians under the Medicare sustainable growth rate.**
  [...]
- **The Insurance Capital Standards Clarification Act was enacted into law. The bill was championed by BPC’s Financial Regulatory Reform Initiative to clarify that the Federal Reserve has the authority to tailor bank capital rules for insurance companies under the Dodd-Frank Act.**

Brookings Institution

Making an Impact on Marijuana Reform

When Oregon, Alaska, and the District of Columbia passed measures to legalize recreational marijuana in 2014, they joined Colorado and Washington in defying federal drug laws to institute a new regulatory regime that promised greater justice for minority populations and fiscal benefits from reduced policing costs and new tax revenues. Adding these jurisdictions to the 19 others that have legalized medical cannabis, it becomes clear that voters in a large swath of the United States are reconsidering the value and effectiveness of the “War on Drugs” and its impact on vulnerable communities. Fellows John Hudak and Philip Wallach have written extensively on the issues that have been raised by various levels of legalization and the tensions between federal, state, and local laws. As governments grapple with how to regulate growing, selling, and use of marijuana, Wallach and Hudak have provided analysis and recommendations through the 420 Series of the widely read FixGov blog. Marijuana policy will be a closely watched issue in the 2016 presidential election and in referenda in a number of states, making their research timely and important in educating the electorate and policymakers.

Cato Institute

The Cato Institute is the only think tank analyzed that reports its engagement with Supreme Court cases. In doing so, it not only adopts an agonistic language (“victories”), but also implicitly attributes an effect of the briefs on the Court’s rulings, with what is an impressive success ratio of 10 out of 11 or 91% displayed in the hidden infographic function of the icon, in which each column represents a case.

Center for Global Development

Examples of solutions (think-and-do-tank)

- **Advance Market Commitments for vaccines**: CGD’s proposals led five countries and the Bill and Melinda Gates Foundation to pledge $1.5 billion as an incentive for the development of vaccines for diseases that kill children in poor countries.

- **Cash on Delivery Aid**: Our proposals to make aid more effective by paying for outcomes continue to set the trend in development. The United Kingdom funded a $10 million pilot in Ethiopia, followed by another in Rwanda.

[...]

- **Haiti immigration policy**: Two years after the devastating earthquake in Haiti, CGD helped bring about a change in US visa rules to allow Haitians to come to the United States as temporary workers, increasing their earnings tenfold and helping their families back home.

- **Commitment to Development Index**: Every year CGD calculates and publishes a ranking of rich countries according to how development-friendly their policies are across a range of areas, including aid, trade, finance, and migration. Currently 27 countries are included, allowing us to compare how the policies of the world’s richest economies affect the poor.

New America Foundation

Letter from the President

[...]

Everyone in our business talks about impact. This year I am proud to say that New America’s impact was profound. Our Open Technology Institute played a central role in engineering the surprise victory of net neutrality through the passage of the 2015 Open Internet Order. OTI logged more visits to educate the Federal Communications Commission than any other public interest group. The FCC’s order captured every key recommendation and cited OTI 19 times.

Urban Institute

A few examples of how nonprofits and city agencies have used Urban’s research:

- When public school officials were deciding about closing elementary schools and redefining enrollment boundaries, Urban helped develop and apply criteria to inform those decisions and analyzed potential effects on travel distances for students and enrollment by school.
- Urban scholars are projecting the effect of DC’s new minimum wage on incomes and on eligibility for district- and federal-level programs, in a study commissioned by Mayor Vincent C. Gray.
- Urban’s online resource Neighborhood Info DC provides data about the District at the ward and neighborhood level. Urban’s analysis of DC foreclosure data helped nonprofits and government programs assist troubled homeowners.
- Acting as the data analyst and local evaluator, Urban is partnering with the DC Promise Neighborhood Initiative, a collaborative effort committed to improving academic and life success for children in targeted neighborhoods.
- Since 2009, Urban’s District of Columbia Crime Policy Institute has produced timely research to help practitioners develop and implement evidence-based crime and justice policy.

Appendix C: Evaluation policies at selected foundations

Quicksheet 1: Melinda and Bill Gates Foundation

Definition of Evaluation:

“Evaluation is the systematic, objective assessment of an ongoing or completed intervention, project, policy, program, or partnership. Evaluation is best used to answer questions about what actions work best to achieve outcomes, how and why they are or are not achieved, what the unintended consequences have been, and what needs to be adjusted to improve execution.”

Approach: Variation

The foundation has an evaluation policy intended to help staff and grantees “align their expectations in determining why, when, and how to use evaluation” in a selective way. The approach is largely formative, conceiving “evaluation to be a collaborative learning tool that provides us and our partners with feedback so we can learn, adjust, and decide how best to achieve outcomes.”

Since not all programs are evaluated, the policy prioritizes those contexts in which “outcomes are difficult to observe and knowledge is lacking about how best to achieve results”. Examples are the improvement of service delivery, behavioral change, innovative models, and systems, policies, or institution-wide change. On the contrary, projects consisting in basic scientific research or with results that are easily observable have a low priority.

Purposes:

- Evaluations to understand and strengthen program effectiveness
- Evaluations to test the causal effects of pilot projects, innovations, or delivery models
- Evaluations to improve the performance of institutions or operating models

In the latter case, there is a strong focus on a close collaboration with the organization involved, which should be at “a critical juncture where additional, detailed information on a specific area of operation can inform next steps, strengthen collaboration, and depersonalize decision making”.

126
**Contexts:**

“When evidence is needed to fill a knowledge gap or evaluate a significant policy decision.

When we and our partners need a better understanding of how a cluster of important investments or a specific program or project is performing.

When an organization, intermediary, or consortium that we work with is at a critical stage of development and can benefit from an independent performance assessment.

When a program team needs to assess the progress of a new operational model or approach.”

**Utilization:**

Evaluations are demand-driven and arise from conversations between program officers and grantees, who “jointly determine the evaluation’s intended users, the general evaluation design, and resources needed to ensure that the purpose, timing, and use of the evaluation findings are well matched. All the studies are compiled in a central registry, which is instrumental in “track[ing] evaluation spending and findings, and ensure[ing] continuity and consistency regardless of any foundation or partner staff turnover.

**Sharing:**

Evaluation plans are shared by program teams with grantees and other partners. The plan identifies existing evidence and the critical gaps that we and our partners need to fill to inform decision making and build knowledge.

“The foundation’s central Strategy, Measurement & Evaluation team […] works with all program teams at the foundation to find opportunities to invest in and share evaluations that have crossprogram relevance and to advance innovation in evaluation methods.”

**Quicksheet 2: MacArthur Foundation**

**Approach to Evaluation:**

Evaluation of initiatives for learning purposes.

**Questions:**

- What is the evolving nature of the problem or opportunity in which we aim to have an impact?
- What is the context and environment in which we are working?
- What is the Foundation’s role and strategy for making a difference?
- Who are the best partners and grantees to effect change?
- What are we accomplishing together?
- Who are we impacting and in what way can we broaden or deepen our impact?
- In what ways could we better leverage our resources?

**Sharing:**

The MacArthur Foundation operates with the presumption that as a private trust for the public good the information that it generates through its grantmaking activities, as well as information about its operations, should be publicly available in a timely and understandable fashion.

The Foundation hopes that this presumption of openness will inform the work of others, encourage communication that provides new perspectives on the Foundation’s work, enrich public debate, and foster a better understanding of issues for interested persons.

**Output:**

Foundation-funded evaluations and assessments of grantmaking strategies.

Quicksheet 3: Robert Wood Johnson Foundation

Approach:
The Robert Wood Johnson Foundation’s portfolio is specialized in research and advocacy in public health and health care, conceiving of its grantmaking activity as “a continuous cycle of research, evaluation and learning”.

Sharing:
One of the best examples of how the foundation shares the results of its projects and evaluation findings is the Program Results Reporting, “a joint effort of the Foundation’s Research & Evaluation and Communications Departments” with four purposes:

- To create a historical record of research, applied interventions and lessons learned for program staff, grantees and partners.
- To document and describe the findings and lessons of our grantees’ work and encourage reflection on them.
- To communicate the work of the Foundation openly to the public.
- To assemble a comprehensive body of work serving many purposes, including education, guidance, research, evaluation and analysis.

Output:
Several types of documents publish the findings of evaluations:

- **Assessment Report of the Robert W Johnson Foundation**: The Assessment Report (2013) contains many of the same metrics considered by think tanks, as reported in section 1 of this Chapter. Social media: Youtube, Facebook, Twitter. The assessment report also includes an extensive survey to health experts, state legislators and grantees. RWJF is particularly interested in the attitudes and behaviors of state legislators because they are an important constituency for the Foundation’s work and have always differed from the other types of health policy experts.

- **Program Results Reports** (see above)

- **Evaluations commissioned to external evaluators** on specific methods and approaches. For example:

This study looked at the usefulness of Social Network Analysis for evaluation purposes, specifically in combination with advocacy strategies and either for summative or formative purposes. It also identified challenges associated with the use of this methodology, including methodological limitations, “implications presented by the size and scope of the projects, and the need to develop the technical expertise and capacity to manage the projects and use the data”.

**Example(s):**

Policies for Action, a program founded in 2015 to fund “research identifying policies, laws, and other regulatory tools in the public and private sectors that can support RWJF’s mission to build a Culture of Health”, is fueled by the collaboration with the Urban Institute, New York University and the University of Michigan.

**Sources:**


Last accessed on March 16, 2016.


[http://www.rwjf.org/content/dam/farm/toolkits/toolkits/2011/rwjf71517](http://www.rwjf.org/content/dam/farm/toolkits/toolkits/2011/rwjf71517)
**Quicksheet 4: William and Flora Hewlett Foundation**

**Definitions:**

*Evaluation* is an independent, systematic investigation into how, why, and to what extent objectives or goals are achieved. It can help the Foundation answer key questions about grants, clusters of grants, components, initiatives, or strategy.

*Grant or portfolio monitoring* is a process of tracking milestones and progress against expectations, for purposes of compliance and adjustment. Evaluation will often draw on grant monitoring data but will typically include other methods and data sources.

**Approach:**

7 Principles of Evaluation:
1) We lead with purpose
2) Evaluation is a learning process
3) Evaluation is an explicit and key part of strategy development
4) We strategically choose what to evaluate
5) We choose methods that maximize rigor without compromising relevance
6) We share our findings with appropriate audiences
7) We use the data!

**Purposes:**

1. To inform Foundation practices and decisions on
   - funding or adapting an overall strategy, component, or initiative;
   - setting new priorities;
   - or setting new targets for results.
2. To inform grantees’ practices and decisions.
3. To inform a field.

**Questions:**

**Implementation:** “How well did we and our grantees execute on our respective responsibilities? What factors contributed to the quality of implementation?”

**Outcomes:** “What changes have occurred? How do they compare with what we expected? To what extent and why are some people and places exhibiting more or less change? What is the relationship between implementation and outcomes?”

**Impact:** “What are the long-term sustainable changes? To what can we attribute them?”

**Context:** “How is the landscape changing? Have changes in the world around us played an enabling or inhibiting role in our ability to affect change?”

**Overall Strategy and Theory of Change:** “Did our basic assumptions turn out to be true, and is change happening in the way we expected?”
Utilization:
Uses depend on purposes. When the purpose is to inform practices and decisions, evaluations are designed to test assumptions so findings can be used to make decisions on funding or adapting initiatives, establishing new priorities and targets. When grantees are the main users, obtaining their buy-in and building on technical capacity “to conduct evaluations and use the findings” are an important piece of utilization strategies. Finally, field-building efforts have a much broader audience that may be attracted towards the field, and evaluations are designed to be disseminated widely.

Sharing:
Case-by-case-basis supported by activities such as convening, internal briefings and discussions, and dissemination of public versions of the studies.

Output:
Field building efforts have the most visible outputs, like the Evaluation of the Population and Poverty Research Initiative (Pop Pov), conducted in 2013 by the Rand Corporation (DaVanzo et al. 2014).

Example(s):
The evaluation of the Think Tank Initiative addresses multiple questions that are driven by the funders’ different perspectives:

- **Initiative design**: How and to what extent did our theory of change play out in practice?
- **Program implementation**: How successfully did IDRC provide the needed support to its grantees to ensure their success?
- **Think tank impact**: In what ways and to what extent did the funded think tanks improve their performance and influence over the course of the grant period?

Appendix D: Interview protocols

SET 1: Foundation staffers

1- Could you please describe briefly what is the role that evaluation plays in [NAME OF FOUNDATION] and at what level of programming is it conducted?
   a. Are any evaluation requirements included by default in all grants?

2- Although most evaluation methodologies are aimed at measuring the effectiveness of service delivery, efforts towards policy change can also be evaluated. What are in your opinion the differences between assessing policy research activities and other programs related to policy change and advocacy?

3- Has your organization recently assessed the impact of a policy research or social science program, or otherwise evaluated it?
   a. Yes: What are the names of some think tanks you have worked with in this capacity?

4- Does your organization negotiate logic models and theories of change with the organizations they fund?
   a. What type of goals and benchmarks are usually set for think tanks?
   b. Could you offer any recent examples?

5- How are the evaluations you conduct used?
   a. And specifically when evaluating policy research?
   b. Can you describe a grant that your foundation made to a think tank in which evaluation was especially helpful?

6- When funding a think tank, what are usually your expectations in terms of research utilization?
   a. By policymakers / By other advocates/nonprofits/ by the foundation /general public/ specific constituencies
   b. Could you provide any examples?

7- What type of conversations, either internal or external, are spurred by evaluation?
   a. Could you give an example of organizational learning?
   b. Do you ever have discussions after one program is concluded and before the next grant is approved/extended on how to improve certain aspects, look at new angles, etc.
      i. If yes: Do those conversations ever serve as the ground for the design of new research?

8- What are the four most important factors when making the call to fund a think tank?
   a. Does evaluation inform these conversations?

9- Have you ever decided to discontinue funding for a program or grantee based on the results of evaluation?
   a. Have you heard of a similar situation?

10- Have changes in leadership affected the relationships with think tanks?
    a. In what ways?

11- Could you suggest members of those organizations with which you have worked closely and who may be willing to contribute to this study?
SET 2: Think tank staffers

1- Does a specific staff member or department manage evaluation in your organization?
   a.    If not, how is it handled?
   b.    Do you use an external evaluator?
2- Do you conduct formal evaluation of any of your programs?
   a.    If yes: Could you describe how you plan for evaluation in the research programs conducted in your area?
          i.  Put together a team
          ii. Benchmarks
   b.    If yes: Does your organization use a standard model for evaluation? Which one?
          i.  For example, the logic models promoted by Kellogg, or the template circulated by the Casey Foundation
   c.    If no: Do you assess influence and research uptake in any way?
3- Could you describe the importance of logic models and theories of change in your work?
   a.    Are those logic models negotiated with foundations?
   b.    What kinds of benchmarks are established?
4- When Foundations supporting [NAME OF TT] demand formal evaluation in the grant proposals, what are their terms and expectations?
   a.    What role does outcome evaluation play in your relationship with funders, specifically with foundations?
5- How do you think the knowledge created with evaluations is used by foundation officers?
6- Can you give an example of a situation in which the evaluation of a program required by a foundation was particularly helpful?
   a.    For example, did it help advance the objectives of the program or the relation with funders?
   b.    Did it allow to identify new directions for research? Did it drive new conversations?
   c.    Did it allow to assess the reasons why the research had been taken up/disregarded?
7- Do you ever have discussions after one program is concluded and before the next grant is approved on how to improve certain aspects, look at new angles, etc.
   a.    Does evaluation inform these conversations?
8- When considering research design and the potential utilization of the research you and your team conduct, do you ever assess the informational needs of other nonprofits that have received funding for the same foundational strategies?
9- Does [NAME OF TT] try to track the utilization of its research among policy makers? How?
10- Could you suggest members of those organizations with which you have worked closely and who may be willing to contribute to this study?
SET 3: Independent evaluation experts

1- It is usually said that evaluation is important in order to distinguish what works from what doesn’t. What are in your opinion the differences between evaluating policy research and advocacy efforts? (evaluating the organization / specific programs)

2- Is evaluation a useful tool to determine the value of policy research? Why?

3- Have you ever worked in assessing the impact of a think tank? Which one?
   ○ What did you do?

4- How do the foundation officers you have worked with use the results of evaluation?
   ○ Can you give any examples?

5- Do foundations use evaluation results to take decisions? To discontinue funding?

6- Could you give an example in which evaluating the work of a think tank was particularly useful?

7- In your opinion, which think tanks take evaluation more seriously?
   ○ Why? What have they learned?

8- Which foundations do a better job at assessing the impact of policy research?

9- Do you see any particular trends in the evaluation of think tanks?

10- Are newer methodologies, like Bellwether surveys or social media analytics, being used in the field?
    ○ Can you give any example?

11- Does any foundation recommend to other grantees the use of policy research developed through a supported think tank? Does monitoring utilization have a role in promoting this integration?

12- Could you suggest members of those organizations with which you have worked closely and who may be willing to contribute to this study?
## Appendix E. List of respondents

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Affiliation</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jon Baron</td>
<td>Vice President of Evidence-Based Policy</td>
<td>Laura and John Arnold Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Meredith Blair Pearlman</td>
<td>Evaluation &amp; Learning Officer</td>
<td>Packard Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Richard Bush</td>
<td>Director of Center for Northeast Asian Policy Studies</td>
<td>Brookings Institution</td>
<td>Think Tank</td>
</tr>
<tr>
<td>Helena Choi</td>
<td>Program Officer, Global Development and Population Program</td>
<td>Hewlett Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Julia Coffman</td>
<td>Director</td>
<td>Center for Evaluation Innovation</td>
<td>Evaluator</td>
</tr>
<tr>
<td>Meredith Hanley</td>
<td>Director of Development</td>
<td>New America Foundation</td>
<td>Think Tank</td>
</tr>
<tr>
<td>Chantell Johnson</td>
<td>Managing Director, Evaluation</td>
<td>MacArthur Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Josh Joseph</td>
<td>Senior Officer, Planning and Evaluation</td>
<td>Pew Charitable Trusts</td>
<td>Foundation</td>
</tr>
<tr>
<td>Jackie W. Kaye</td>
<td>Director of Research and Evaluation</td>
<td>Wellspring Advisors</td>
<td>Evaluator</td>
</tr>
<tr>
<td>Tom Kelly</td>
<td>Vice President – Knowledge, Evaluation &amp; Learning</td>
<td>Hawai‘i Community Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Nick Lee</td>
<td>Program Officer</td>
<td>Gates Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Ruth Levine</td>
<td>Global Development and Population Program Director</td>
<td>Hewlett Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Sarah Lucas</td>
<td>Global Development and Population Program Officer</td>
<td>Hewlett Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Justin Milner</td>
<td>Senior Research Associate</td>
<td>Urban Institute</td>
<td>Think Tank</td>
</tr>
<tr>
<td>Johanna Morariu</td>
<td>Director</td>
<td>Innovation Network</td>
<td>Think Tank</td>
</tr>
<tr>
<td>Brian Quinn</td>
<td>Assistant Vice President, Research-Evaluation-Learning</td>
<td>Robert W. Johnson Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Chera Reid</td>
<td>Director of Strategic Learning, Research and Evaluation</td>
<td>Kresge Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Jane Reisman</td>
<td>Founder and Senior Advisor</td>
<td>ORS Impact</td>
<td>Evaluator</td>
</tr>
<tr>
<td>Name</td>
<td>Position</td>
<td>Organization</td>
<td>Affiliation</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Peter Taylor</td>
<td>Program Manager</td>
<td>IDRC</td>
<td>Think Tank</td>
</tr>
<tr>
<td>T'Pring Westbrook</td>
<td>Senior Research Associate</td>
<td>Casey Foundation</td>
<td>Foundation</td>
</tr>
<tr>
<td>Confidential</td>
<td>Evaluation Specialist</td>
<td>Confidential</td>
<td>Think Tank</td>
</tr>
</tbody>
</table>
Bibliography


Finnigan, Kara S., and Alan J. Daly. 2014. Using Research Evidence in Education. Springer: Cham


Mendizábal, Enrique. 2011b. “Evaluating a think tank from the AEA365 (but should we?)” *On Think Tanks*. May 20. [https://onthinktanks.org/articles/evaluating-a-think-tank-from-the-aea365-but-should-we/](https://onthinktanks.org/articles/evaluating-a-think-tank-from-the-aea365-but-should-we/)


Prewitt, Kenneth, Thomas A. Schwandt, and Miron L. Straf. 2012. Using Science as Evidence in Public Policy. Center for Education; Division of Behavioral and Social Sciences and Education; National Research Council.


