UNIVERSITY OF CALIFORNIA

Los Angeles

An Analysis of Evaluation Policy in the Philanthropic Sector

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor of Philosophy

in Education

by

Alana Kinarsky

2019

ABSTRACT OF THE DISSERTATION

An Analysis of Evaluation Policy in the Philanthropic Sector

by

Alana Kinarsky

Doctor of Philosophy in Education

University of California, Los Angeles, 2019

Professor Marvin C. Alkin, Co-Chair

Professor Christina A. Christie, Co-Chair

Foundations annually distribute over $60 billion to nonprofit organizations. They rely on evaluation to learn about the impact of their investment and make decisions about future investments. To govern evaluation practice, foundations are beginning to draft evaluation policies. To date, there is some research on the evaluation policies being written by international aid organizations and governments around the world. However, there is little research on the evaluation policies begin written in the philanthropic sector. It is unclear why these policies are being developed and what impact, if any, they have on evaluation practice. Furthermore, it is uncertain if the content of these policies is informed by the information and knowledge imbedded in evaluation theories. The present dissertation begins with an exploration of foundation spending on evaluation and tracks the prevalence of published evaluation policies among leading foundations. Next, extant evaluation policies are analyzed utilizing two proposed taxonomies that identify dimensions of evaluation practice that could be specified in an evaluation policy. Suggested modifications to the taxonomies are discussed and a modified taxonomy is presented. The third, and final, paper documents the intent and influence of evaluation policies in the philanthropic sector, according to evaluation directors. Then, how evaluation theory informs these policies is discussed. The dissertation concludes with suggestions for future research and implications for practice.

The dissertation of Alana Kinarsky is approved.

Leslie Fierro

Todd M. Franke

Mark P. Hansen

Marvin C. Alkin, Committee Co-Chair

Christina A. Christie, Committee Co-Chair

University of California, Los Angeles

2019

To my mom and dad: Thank you for demonstrating what can be accomplished through hard work and perseverance. I aspire to be as brave and resilient as you.

To Isaiah: Thank you for coming on this journey with me. You have given me the gift of seeing the world around me anew.

Table of Contents

List of Figures & Tables

ACKNOWLEDGEMENTS

I have had the privilege of learning with a remarkable community of faculty, students and staff at UCLA over the past five years. These interactions have influenced my thinking on social research methodology and evaluation in countless ways. I am particularly lucky, because many have become lifelong friends who nourish and inspire me.

I am immensely grateful for the guidance and mentorship of my advisors and co-chairs, Marv Alkin and Tina Christie. Marv guided me through decades of research on evaluation, always made time to meet and supported my growth. Tina steered me through my graduate work and helped me hone my ideas. Her expertise was invaluable in formulating this research agenda, and in particular the methodological approach. Together, Marv and Tina helped me balance the practical and theoretical aspects of evaluation and gave me many opportunities to grow as a scholar and practitioner. Also, I cannot underscore enough how grateful I am for their support as I sought to balance new parenthood with my academic and professional goals. Thank you for the space and support I needed to strike a work-life balance that I am proud of.

I am deeply appreciative of the members of my dissertation committee: Leslie Fierro, Todd Franke, and Mark Hansen for their feedback, guidance, and time. Your teaching and research agendas have greatly influenced my thinking and this research. There is no doubt that your questions and suggestions have strengthened this dissertation and for that I am very thankful. I look forward to continuing to work together.

I was fortunate to engage with an amazing community of PhD students that have made me a better scholar and person: Talia, Jenn, Sebastian, Lisa, Celina, Anne, Meredith, Jayashri, Priscilla, Joy, Julie, Adrienne and Kristen. Thank you for engaging, encouraging, and challenging me over the last five years. I look forward to being colleagues and friends for years to come. I would also like to thank Carmel for taking time to review every evaluation policy with me - ultimately making this research better.

Even when I questioned my ability to be successful in graduate school and to complete this dissertation, I had a fierce supporter at home who never doubted me. Thank you to my husband, Gabriel, for always being willing to talk through ideas, explore new analytical approaches, and remind me of the significance of my research, especially when I lost sight of it. Thank you for spending evenings helping me think through emergent ideas and moving post-its around to make sure this manuscript was well-organized. Your eagerness to engage with difficult questions and quest to learn motivate me to do the same.

To my in-laws, Pam and Dan, everything I have written is better because you have taken the time to review and refine it. Pam – thank you for carefully reading every word of this dissertation. You have made me a more confident and capable writer.

Last, but not least, I am indebted to my mother and father, Yulia and Leo, as well as my sisters, Natasha and Irina, for working harder than anyone else I know. I have always tried harder and done more because of the example you set. Thank you for having high expectations of me and for loving me unconditionally.

VITA

EDUCATION

2015	M.A. Social Research Methodology University of California, Los Angeles
2008	B.A. International Relations Loyola University, Chicago

PROFESSIONAL EXPERIENCE

2014 - 2019	Graduate Student Researcher University of California, Los Angeles
2016	Teaching Assistant, Educational Program Evaluation University of California, Los Angeles
2015	Evaluation Associate (Summer) Informing Change, San Francisco
2010-2014	Manager of Assessment and President's Interfaith and Community Service Campus Challenge Interfaith Youth Core, Chicago

SELECTED PUBLICATIONS

In press	Kinarsky, A.R. Evaluation Complexity Through A Theorist's Lens. *New Directions in Evaluation*.
2018	Kinarsky, A.R. The Evaluation Landscape: U.S. Foundation Spending on Evaluation. *New Directions in Evaluation*.
2018	Lemire, S., Fierro, L.A., Kinarsky, A.R., Fujita-Conrads, E., Christie, C.A. The U.S. Federal Evaluation Market. *New Directions in Evaluation*.
2018	Kinarsky, A.R., Alkin, M.C., and Christie, C.A. Volume Editors Introduction. In *D. Mertens Mixed Methods Design in Evaluation. Los Angeles, CA: SAGE*.

2017	Kinarsky, A.R. Fostering Success: Understanding the Experience of Foster Youth Undergraduates. *Children and Youth Services Review*.

SELECTED PRESENTATIONS

2018	"*Evaluation Policy: Mapping the Landscape and Charting its Influence on Practice*" at American Evaluation Association Conference. Cleveland.
2017	"Fostering Success: Evaluation of Campus Support Programs for Foster Youth" at American Evaluation Association Conference. Washington, D.C.

SELECTED AWARDS

2018	Alkin Evaluation Award
2017	UCLA Graduate Research Mentorship Fellowship
2016	UCLA Graduate Summer Research Mentorship Fellowship

Chapter 1.
Introduction

Organizations invest in evaluation to understand their impact and generate information that can aid decision making. However, as organizations and their work grow in complexity, so do evaluations. There is an increasing demand for quality evaluations that provide useful information for a wide range of stakeholders. In response to this need, organizations are turning to evaluation policies to organize and guide evaluation practice. Yet, there is currently "a paucity of good evaluation policy examples" (Trochim, 2009, p. 19). While there is reason to believe explicit policies are widespread, there has been no broad investigation to quantify their prevalence.

In recent years, governments have been writing evaluation policies to guide their evaluation practice. Government agencies in the Netherlands and across the European Union began to establish formal evaluation policies in the 1990s (Stern, 2009; Leeuw, 2009). As of 2016, every federal agency in Canada was required to have an evaluation policy that included outcome measures and performance indicators. Similarly, the United States government recently passed Foundations for Evidence-Based Policymaking Act of 2018 which stipulates that all government agencies must develop evaluation policies (H.R. 4174). In passing such legislation, Canada and the United States are relying on evaluation policies to enhance the government's capacity to generate quality evaluations that can be used to improve government operations.

The development of evaluation policies is not restricted to government agencies. Al Hudib (2018) identified 36 multilateral and bilateral aid agencies and government organizations, like the United Nations and World Health Organization, that have publicly available evaluation policies. There are also two publications tracing the evolution of evaluation policy in the philanthropic sector (Mattessich, Mueller & Holm-Hansen, 2009; Dillman & Christie, 2016).

Evaluation policies are being utilized by organizations around the world to guide their evaluation practice.

The focus of the present dissertation is on the evaluation policies in the philanthropic sector. Recent research on the philanthropic sector suggests that we are in "the second golden age of American philanthropy":

> "In 1930 approximately two hundred private foundation possessed aggregate assets of less than $1 billion. In 1959 there were more than two thousand, in 1985 just over thirty thousand private foundations…. the number ballooned to seventy-six thousand in 2004, and as of 2014 the number was nearly one hundred thousand with total capitalization close to $800 billion" (Reich, 2018, p. 141).

Of the $800 billion, foundations annually distribute over $60 billion to nonprofit organizations working in nearly every issue area around the world, from disease eradication to homelessness prevention (The Foundation Center, 2014). Foundations collect and share data to assess and demonstrate the extent to which the initiatives they fund are serving the public good. Additionally, in exchange for funding, foundations may mandate that their grantees engage in evaluation.

Evaluation is used in the philanthropic sector to generate evidence for accountability, learning, and knowledge sharing (Kramer & Bickel, 2004). It offers a systematic approach to studying an intervention to better understand its merits and challenges. The information generated by the evaluation is ideally used by foundation staff to understand its impact and inform funding and strategy decisions. These findings, positive or negative, may also be disseminated more broadly to encourage learning across the sector. While all three functions of evaluation in philanthropy are important, they are difficult to balance:

> “Accountability for use of funds is the most common form of evaluation but contributes little to the learning of the foundation or the broader field. Learning within the foundation is more valuable though less common and knowledge sharing with the field is critically important, though least common of all” (Kramer & Bickel, 2004, p. 53).

Foundations have begun to formalize their evaluation practices, demands and expectations by drafting evaluation policies that govern when, what, and how they engage in evaluation (Mattessich, Mueller & Holm-Hansen, 2009; Dillman & Christie, 2016). Evaluation policies authored by foundations likely guide evaluation practice across the civic sector, which impacts the interest in, demand for, and practices of evaluation. They ultimately influence the quality of information generated about the effectiveness of interventions being developed and executed by nonprofit organizations and foundations. While some evaluation policies may help streamline and improve evaluation practice, others may do more harm than good. For example, if a foundation’s policy prioritizes the use of randomized control trials (RCT) and de-incentivizes qualitative inquiry, then initiatives that are more targeted or experimental in nature may be forced to participate in an RCT before they are ready or not get evaluated at all. Despite the importance of the relationship between evaluation policy and practice, there is little empirical research.

What gets funded and the quality, or trustworthiness, of the findings are of great concern. Evaluation findings, at their best, impact funders’ perceived value of an initiative and help them prioritize what to fund (Datta, 2009). Evaluation policies can shape a foundation’s funding priorities, as well as the specific programs and services they fund. Given that foundations and non-profits work together to improve society, often providing vital services for the most vulnerable and disenfranchised communities, what gets funded is of great importance. Therefore, it is necessary to further investigate the evaluation policies being developed by foundations to guide their evaluation practice.

Evaluation Policy Defined

Evaluation policies are the "rules or principles that help set the content, characteristics, and context of evaluation itself" (Mark, Cooksy, & Trochim, 2009, p. 5). Stated slightly differently, an evaluation policy is "any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation" (Trochim, 2009, p. 16). These definitions of evaluation policy imply that the rules and principles a foundation follows in its evaluation practice constitutes an evaluation policy.

Trochim (2009) extends this definition by stating that enforcement is a pivotal feature of policy:

> "only if and when they are consciously adopted to guide decisions or actions about evaluation and when the organization institutes *consequences for encouraging or enforcing them*" (p. 17, italics added for emphasis).

However, Christie and Lemire (2019) challenge that enforcement is a defining characteristic of an evaluation policy. They set forth that:

> "an evaluation policy becomes an evaluation policy when it is formulated and made explicit as such—consequences or no consequences" (pg. 6).

Furthermore, in a review of explicit and implicit evaluation policies at the Robert Wood Johnson Foundation, Dillman (2014) found that even though the foundation did not enforce their policies, they still had a meaningful influence on evaluation practice. In the present dissertation, evaluation policy is understood as any written document that details an organization's rules and principles for evaluation practice, regardless of whether implementation is encouraged or enforced.

Evaluation policies can encompass and dictate nearly every aspect of evaluation practice. They can specify when and how an organization should conduct evaluations, delegate responsibilities to stakeholders, allocate resources, or outline how evaluations are to be used. Therefore, it is crucial that evaluation policies are well crafted:

> "developing well-informed evaluation policies that can guide evaluation practice may be the most important issue currently facing our field. It's more important than getting our methods right, or concerns about validity. It's more important than the ethics of evaluation. It's more important than making evaluation participatory or using it for empowerment. Why? Because evaluation policy encompasses all of those things and more. It touches virtually everything we think about or do in evaluation" (Trochim, 2009, p. 14).

Evaluation policies have a broad reach and significance. In addition to guiding a foundation's evaluation practice, they may impact the millions of nonprofit organizations that rely on foundation funding. To obtain or retain funding, nonprofits must abide by the rules and expectations set by funders. Therefore, depending on quality and scope, evaluation policies developed by foundations can also aid or constrain the crucial work of nonprofits.

Christie and Lemire (2019) expand on the transformative role of policies by advancing the evaluation theory–policy–practice nexus:

> "The position we hold is that evaluation policies serve as an important connector between evaluation theory and evaluation practice. More specifically, evaluation policies serve as a transformative mechanism for translating and perhaps more importantly situating

evaluation theory in the organizational, political, and cultural context of an organization" (p. 2).

The theory–policy–practice nexus makes clear that evaluation theory is not the only driver of evaluation practice. Organizations are adapting evaluation theory, the broad knowledge of the field, to their context through evaluation policies to help inform and guide evaluation practice (Figure 1). Despite its essential role, there is little research to understand, inform, or guide the development and use of evaluation policies.

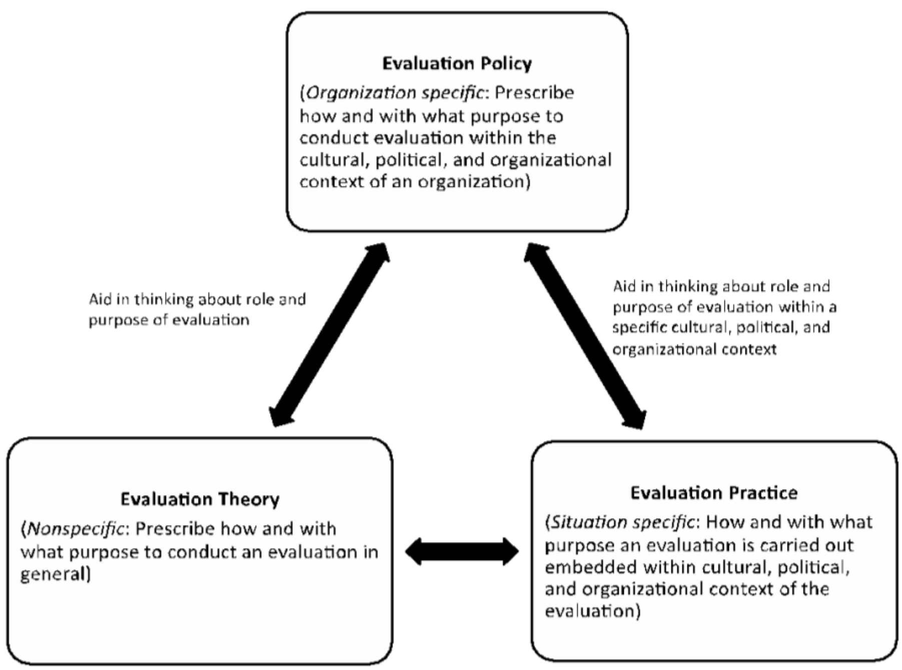


Figure 1. *Evaluation Theory-Policy-Practice Nexus Diagram from Christie & Lemire, 2019*

Research Gap

Evaluation policies are an important yet overlooked matter for research on evaluation. Trochim (2009) summarizes the broad reach and significance of evaluation policies by stating:

> "Evaluation policies profoundly affect the day-to-day work of all evaluators and ultimately the quality of the programs they evaluate. Many recent and current controversies or conflicts in the field of evaluation can be viewed, at least in part, as a struggle around evaluation policy. Because evaluation policies typically apply across multiple evaluations, influencing policies directly may have systemic and far-reaching effects for practice" (p. 14).

Policies guide evaluation practice, which, in turn, influences perceptions and decisions on program quality, and potentially on interventions that impact countless individuals. In spite of their importance, there remain significant theoretical and empirical questions about evaluation policy.

Mark, Cooksy, & Trochim (2009) opened their *New Directions for Evaluation* volume on evaluation policies by drawing attention to the lack research on the topic:

> "…there is a paucity of literature explicitly addressing the topic of evaluation policy, broadly construed—its origins, how it is developed, the evidence base for evaluation policies, examples of coherent evaluation policies in use, the consequences of having one evaluation policy rather than another, and so on" (p. 4).

Ten years later, very few publications have emerged on the topic. Meanwhile, evaluation policies continue to be used to guide evaluation practice by organizations and agencies around the world. Motivated by this gap, the present dissertation seeks to advance understanding of evaluation policy by addressing three key gaps in the literature.

First, Trochim (2009) argues that every entity that utilizes evaluation has implicit or explicit policies that influence how evaluation is conducted. Yet, extant research has not investigated how widespread evaluation policies are across sectors. This suggests that it is necessary and timely to survey the field more broadly to understand which organizations have policies and variations in policies across organizations. Furthermore, there is a need to catalog extant evaluation policies to facilitate research and knowledge-sharing across the philanthropic sector.

Second, two taxonomies of evaluation policies have been proposed. Both were conceptualized by experts in the evaluation field. In 2007, the American Evaluation Association convened an advisory group to discuss the intended scope for a task force on evaluation policy. Those discussions led to a list of dimensions that may be addressed by evaluation policies. Shortly after that, Trochim (2009) put forth a taxonomy that was intended to support the development and review of evaluation policies. These taxonomies are the foundation of our understanding of what constitutes an evaluation policy, yet both were developed conceptually. Trochim (2009) himself calls for empirical research on the composition of evaluation policies:

> "We need to develop taxonomies and corresponding audit methodologies that are better than the ones suggested here, and these need to be informed by systematic research" (p. 28).

There is some research addressing this need. Leeuw (2009) used Trochim's taxonomy to guide a review of evaluation policies written by government agencies in the Netherlands. Similarly, Al Hudib (2018) used it to analyze evaluation policies across aid agencies. Researchers at Claremont University are using Trochim (2009) and AEA (2007) taxonomies to review evaluation policies in agencies across the United States government. To contribute to this growing area of study, the current study uses the conceptual frameworks proposed by Trochim (2009) and AEA (2007) to analyze evaluation policies developed in the philanthropic sector.

Lastly, there is a need for more research on evaluation situated in the philanthropic sector. Prewitt (2006) observed that there is "little theory in foundation evaluation" (p. 372) and such research continues to be scarce more than 10 years later. Foundations are major commissioners of evaluation. They are also formalizing and codifying their evaluation practice. Yet, there remains little empirical research on evaluation policies and how they translate evaluation theory into practice in the philanthropic sector. Christie and Lemire (2019) propose that reviewing evaluation policies offers rich insight into how organizations translate theory into practice:

> "As with many significant developments in evaluation, evaluation policies emerge from practice, responding to a need for organizations and government agencies to formally guide evaluation practice. In this way, the evaluation policies importantly may be viewed as a reflection of an organization's values related to evaluation, its procedures, and its potential influences. Formal evaluation policies (especially those of major commissioners of evaluation) serve to frame a significant and still growing part of evaluation practice" (p. 2).

Examination of evaluation policies across the philanthropic sector has the potential to advance our understanding of evaluation practice, which will also benefit and propel evaluation theory since the two are intrinsically linked.

Dissertation Structure

Motivated by these gaps in the research, the present dissertation seeks to identify and investigate the evaluation policies guiding evaluation practice in the philanthropic sector. Extant research paints evaluation policy in broad strokes. Little is known about the prevalence of evaluation policies in the sector; what dimensions of evaluation practice are specified; what is motivating their development; how they are influencing practice; or to what extent evaluation theory is informing their development. To this end, the following research questions, which are arranged in the order of my inquiry, guide the present dissertation:

1. To what extent are US foundations investing in evaluation? (Paper 1)
2. To what extent do US foundations, who are major commissioners of evaluation, have publicly available evaluation policies? (Paper 1)
3. How do publicly available evaluation policies in the philanthropic sector compare to taxonomies proposed by Trochim (2009) and the American Evaluation Association (2007)? (Paper 2)
4. How do publicly available evaluation policies in the philanthropic sector compare to each other in terms of their comprehensiveness and completeness? (Paper 2)
5. According to evaluation directors, what are the primary goals of their foundation's policy? (Paper 3)
6. According to evaluation directors, in what ways do the foundation's formal and informal evaluation policies guide evaluation practice? (Paper 3)
7. According to evaluation directors, to what extent does evaluation theory inform the evaluation policies of US foundations? (Paper 3)

The purpose of the present dissertation is primarily empirical. There have been some notable conceptual studies conducted on evaluation policy. Those have motivated and shaped the present study, which is the first to look at evaluation policy adoption across the philanthropic sector.

Toward these aims, this dissertation is organized in a multi-paper structure that follows the exploration of related yet distinct research questions. Each paper addresses findings from one of the three phases of this research. The first phase was discovery. I set out to identify which philanthropic foundations are investing heavily in evaluation, and of those, how many have publicly available evaluation policies. Next, once evaluation policies were cataloged, I examined the policies with an eye to whether they were similar or different to the taxonomies proposed by Trochim (2009) and AEA (2007). Lastly, in phase three, I interviewed evaluation and research directors at foundations with evaluation policies. Interviews focused on the intent and influence of these policies, as well as what informed their development.

What's Ahead

The remainder of this dissertation is organized in four parts. The first paper is a discussion of the evaluation marketplace and the prevalence of evaluation policies amongst philanthropic foundations. Next, the findings from an analysis of 12 evaluation policies are discussed. In paper three, findings from interviews with evaluation and research directors are presented. The dissertation closes with conclusions drawn from this work and suggestions for future areas of research. Before turning to the first paper, abstracts of the three papers are presented below:

Paper 1. The Evaluation Landscape: U.S. Foundation Spending on Evaluation & Prevalence of Evaluation Policies

Foundations annually distribute over $60 billion to nonprofit organizations. Foundations rely on evaluation to learn about the impact of their investment and make decisions about future investments. In exchange for funding, nonprofits are often required to evaluate their work by collecting and sharing data that demonstrates their progress and impact. Accordingly, foundations and nonprofits represent a meaningful influence on the evaluation market. The present paper first describes the historical development of foundations, their important role in society, and their use of evaluation. Informed by this historical account, the paper then provides an overview of foundation spending on evaluation, detailing the number of evaluation-related grants and funds allocated by year and by select foundations. Next, this paper tracks the prevalence of published evaluation policies among leading foundations. In conclusion, implications for the profession and practice of evaluation are discussed and suggestions for future areas of research are shared.

Paper 2. A Closer Look: An Analysis of Evaluation Policies in the Philanthropic Sector

Evaluation policy encompasses most aspects of evaluation practice and theory, making it one of the most critical areas of research on evaluation (Trochim, 2009). Depending on quality and scope, evaluation policies can aid or constrain the crucial work of foundations and nonprofit organizations. Yet, extant research paints evaluation policy in only broad strokes. To date, two taxonomies have been proposed that identify dimensions of evaluation practice that could be specified in an evaluation policy. Little is known about how these taxonomies align with evaluation policies developed by philanthropic foundations. Through document analysis, this paper first compares evaluation policies from 12 foundations to the proposed taxonomies to extrapolate their similarities and differences. Then, suggested modifications to the taxonomies are discussed and a modified taxonomy is presented to guide future development and review of evaluation policies.

Paper 3. The Intent and Influence of Evaluation Policies in the Philanthropic Sector and What Informs Their Development

Evaluation policies are being written across the philanthropic sector. However, it is unclear why these policies are being developed and what impact, if any, they have on evaluation practice. Furthermore, it is uncertain if the content of these policies is informed by the information and knowledge imbedded in evaluation theories. Through interviews with 10 evaluation directors at foundations with evaluation policies, this paper first identifies the intent and influence of evaluation policies in the philanthropic sector. Then, how evaluation theory informs these policies is discussed. The paper concludes with suggestions for future research on evaluation policy.

Paper 1.
The Evaluation Landscape: U.S. Foundation Spending on Evaluation & Prevalence of Evaluation Policies

Abstract

Foundations annually distribute over $60 billion to nonprofit organizations. In exchange for funding, nonprofits are often required to evaluate their work by collecting and sharing data that demonstrates their progress and impact. Accordingly, foundations and nonprofits represent a meaningful influence on the evaluation market. The present paper first describes the historical development of foundations, their important role in society, and their use of evaluation. Informed by this historical account, the paper then provides an overview of foundation spending on evaluation, detailing the number of evaluation-related grants and funds allocated by year and by select foundations. In conclusion, the paper considers implications for the profession and practice of evaluation and offers suggestions for future areas of research.

Introduction—The History of Foundations and Their Use of Evaluation

The Important Role of Foundations in Society

The civic sector, comprised of nonprofit organizations and foundations, is focused on causes that improve society and correct injustices. There are over 1.4 million tax-exempt organizations that constitute America's civic sector (McKeever, 2017). Of those, 86,000 are foundations (The Foundation Center, 2014). In 2013, the civic sector reported $905.9 billion in revenue, which accounted for 5.4 percent of the United States' total gross domestic product (McKeever, 2017). The civic sector is vast and has significant influence on American society.

Foundations use their income from endowments and other sources to support nonprofit organizations. In addition to distributing direct fiscal grants, foundations may also offer nonprofit organizations funds to support capacity building, technology improvements, networking opportunities with other funders, and publicity. Foundations distribute fiscal grants or technical support to nonprofit organizations that align with their mission or strategy. Nonprofit organizations then implement programs or services intended to improve local, national or global communities. By working together, foundations and nonprofit organizations advance their mission of improving society.

Foundations are granted tax-exempt status to their assets and income because of a belief that philanthropy plays a special and vital role in American society. By diverting funds from government to the private sector, foundations are given the opportunity to “initiate thought and action, experiment with new and untried ventures, dissent from prevailing attitudes, and act quickly and flexibly” to improve well-being and “enrich the pluralism of the social order” (US Treasury Dept., 1965, p. 5). Having been given this preferential status, the Internal Revenue Service requires private foundations to pay out at least 5 percent of the previous year’s average net assets for charitable purposes (Internal Revenue Service, 2017). In total, it is estimated that nearly $20 billion in tax revenue is diverted from the federal government and infused into the civic sector (Fleishman, 2009). This leads to confusion about whether foundations are private entities or if they have some responsibility to demonstrate impact to the public. Fleishman (2009) summarizes this tension by stating:

> “Taxpayers shoulder part of the cost of the foregone revenue from the capital and income earned on that capital through the tax breaks that foundations enjoy. For this reason, the taxpayers, too, have an interest in how wisely and well foundation money is spent. Thus, the general public has a real, tangible, even proprietary interest in a foundation’s deployment of its assets and in all the ways it makes its decisions” (p. 232).

Foundations are uniquely positioned in American society to support charitable endeavors that enrich society. They have full control over who they fiscally support, for how long, and to what degree. Foundations annually distribute over $60 billion to nonprofit organizations working in nearly every issue area around the world, from disease eradication to homelessness prevention (The Foundation Center, 2014). In 2012, foundations in America had $715 billion in assets and distributed $52 billion to nonprofit organizations. The top 1,000 of 86,000 foundations awarded about 154,000 grants totaling $22.4 billion to nonprofit organizations (The Foundation Center, 2014; The Foundation Center, 2017). This reflects an increase in giving over the previous 10 years. Between 2003 and 2013, foundation grant making increased 44 percent (after adjusting for inflation) (McKeever, 2017). Foundations play a “priming role in starting new civic sector organizations, in nurturing them into self-sustainability, and in providing a continuing supply of social venture capital to the civic sector” (Fleishman, 2009, p. 83).

Unlike for-profit businesses and government agencies, foundations are able to pursue ambitious and long-term initiatives. Government agencies are expected to react to voter demands and regularly shift focus following elections. For-profit businesses are beholden to their shareholders and obligated to make decisions that maximize profits. Civic organizations, particularly foundations, however, have more flexibility and freedom to use their funds in accordance with their interests and priorities. They are also better positioned to innovate through small scale programming developed for targeted communities. Working in partnership, nonprofit organizations and foundations have made significant contributions to national and global well-being.

Many of the societal advancements we enjoy today were conceptualized or realized by foundations and the nonprofit organizations they fund. The Robert Wood Johnson Foundation funded the development of the 911 emergency call system (RWJF: Our History, 2017). The William T. Grant Foundation provided funding for the pioneering research by Dr. Spock on child rearing and Dr. Goodall on primates (WT Grant: History). More recently, The Bill & Melinda Gates Foundation is supporting vaccine research for AIDS and Malaria. The Gates Foundation has also partnered with the Rotary Foundation in the ambitious pursuit of eradicating polio by 2018 (Carl, 2017). This is just a synopsis of the significant ways in which society has, and continues, to benefit from innovations and improvements spearheaded and funded by foundations.

The Role of Evaluation in Foundations

The evaluation of programs and policies designed to ease social injustices has been an essential part of philanthropic foundations from the beginning. When founding his namesake foundation in the early 1900's, John D. Rockefeller outlined a vision for philanthropy that was rooted in knowledge creation and application: "The best philanthropy involves the search for cause, an attempt to cure evils at their source" (as quoted in Prewitt, 2006, p. 362). Rockefeller felt that an investment in evaluation was a vital endeavor because it allowed foundations to identify the root cause of injustices and develop strategies to right the wrong.

The initial motivation for evaluation was internal. When the Rockefeller Foundation was established in 1913, there were no tax benefits allocated for generous giving. Philanthropists of that time invested in evaluation and research stemming from intrinsic desires to utilize their funds well and achieve the greatest impact. Since then, the tax code has been modified to grant deductions for charitable contributions with the assumption that civic organizations can create benefit to the public. This change has led to increased external pressure on foundations to be more accountable to and transparent with the public about their impact.

In the 1960's, the federal government began to use evaluation as a tool for ensuring that federal funds were being used for programs that worked well. As government-sponsored social programs were established, requirements were instituted and standardized for regularly evaluating these programs. By the 1970's federal funding for evaluation grew tenfold (Carman, Fredericks, & Introcaso, 2008). By the early 2000's, 98 percent of all federal programs were evaluated by the Office of Management and Budget (Carman et al., 2008). The increase in legislation and government mandates for evaluation has also significantly impacted private foundations and the civic sector.

As the federal government made a push towards accountability and transparency in funding for social programs, foundations faced increased scrutiny that was rather new to the sector. Congress began to raise concerns that foundations were a mechanism for shielding the wealthy from taxes and that resources were being used to fund controversial activities, such as voter registration. The foundation sector, lacking both organization and evidence, had difficulty convincing government officials and the wider public of its positive impact (Wadsworth, 1975; Troyer, 2000). This debate led to the passing of the Tax Act of 1969, which, for the first time, imposed regulations on foundations. In particular, it established a minimum annual payout requirement and outlined various public disclosure rules (Frumkin, 1998). In response, foundations began to professionalize their staff, circulate annual reports, and disclose financial details.

Despite these pressures in the 1960's and 1970's, there is, at present, little formal oversight of foundations. The Internal Revenue Service, where government oversight of foundations is concentrated, has limited resources to monitor their spending. For example, in 2014, only .7 percent of all tax returns filed by nonprofit organizations were audited (US

Government Accountability Office, 2014). Although foundations are private entities, they have been afforded generous tax benefits in exchange for funding initiatives that improve social betterment. Therefore, foundations do bear a responsibility to those they serve and the broader public to demonstrate impact.

Apart from existing government oversight, there is minimal, if any, external regulation or accountability of foundation grant making. While foundations share annual reports, those tend to "read more like campaign documents than balanced accounts of modest success, false starts, and even a few failures" (Prewitt, 2006, p. 373). After closely examining the grant making of the Eli & Edythe Broad Foundation, the Ford Foundation, the Bill & Melinda Gates Foundation, and the W. K. Kellogg Foundation through staff interviews, Tompkins-Stange (2016) concluded that:

> "in the absence of formal bureaucratic mechanisms that are designed to ensure greater foundation transparency and accountability, foundations are unlikely to incorporate strategies that reflect these norms…the likelihood of compliance hardly seems promising, at least within the spirit of the law rather than the letter of the law" (p. 139).

While some posit that without external pressure foundations are unlikely to be more forthcoming, there is some evidence to suggest that foundations are shifting towards a more strategic approach, which may also increase transparency and accountability.

As more engineers and technologists (such as David Packard, William Hewlett, and Bill Gates) established foundations, they also brought their professional backgrounds, which emphasizes efficiency and return on investment, to their philanthropic work (Tompkins-Stange, 2016; Frumkin, 1998). These leaders were interested in justifying their spending by documenting and measuring social impact. To track outcomes, their foundations invest in the development of theories of change, formative and summative evaluation, as well as research. In fact, according to one study of the sector, 45 percent of all staffed foundations report conducting formal evaluations of their work, and 88 percent of those with more than $400 million in assets are requiring their grantees to do the same (Ostrower, 2004). There is growing interest in using evaluation to support internal learning and strategy, and further general knowledge on what approaches and programs work best.

The Hewlett Foundation, for example, embraced an "outcomes-oriented" approach to philanthropy in 2000 when Paul Brest became president. They sought "clearly defined goals", pursued "evidence-based strategies for achieving those goals", monitored "progress toward outcomes", and assessed "their success in achieving them" (Brest, 2012). They also expected their grantees to do the same. However, in 2012, as Paul Brest was ending his tenure as foundation president, he confided that, while "outcomes-oriented" was their priority since 2000, the foundation did not actually employ the "systematic monitoring and evaluation necessary to see if strategies are working and correct them when they're not" until very recently (Brest, 2012). The Hewlett Foundation is not alone in making changes to evaluation practice to increase its relevance and impact. A sector survey by Coffman, Beer, Patrizi, and Heid Thompson (2013), found that foundations:

> "have been challenged to reframe evaluation from an older model of "post hoc" assessment of grantee projects and programs for accountability purposes to one that is more focused on examining foundation strategy and informing learning about it from start to finish" (p. 44).

Perhaps to meet this challenge, foundations have begun to increase their investment in evaluation.

In a survey of senior foundation staff in the US and Canada, Buteau, Coffman, Loh, Glickman, and Beer (2016) found that a quarter of foundations spend $1 million or more on evaluation, while half of foundations spend at least $200,000 on evaluation annually (Buteau et al., 2016). Foundations have also established evaluation units and hired more staff dedicated to overseeing evaluation projects (Coffman et al., 2013; Buteau et al., 2016). While foundations have come a long way in their understanding and utilization of evaluation, emergent research indicates that they are still struggling to make use of evaluation findings. Foundations remain unsure how to plan and design evaluations that result in valuable insights for the field, their grantees, and the foundation (Buteau et al., 2016; Buteau & Glickman, 2018). Evaluation staff at foundations remain uncertain of their mandate, have limited authority and often feel unwelcome when working with program officers on evaluation projects (Kramer & Bickel, 2004).

At present, many foundations struggle to meaningfully use evaluation findings to improve practice and inform strategy. Senior foundation staff also report a need to increase transparency about internal evaluation practices and share what they are learning with external audiences. As foundations formalize and codify their evaluation practice, it is important to track progress and add to the scarce literature on evaluation in the philanthropic sector because to date there is very “little theory in foundation evaluation” (Prewitt, 2006, p. 372).

Research Gap

Research into evaluation practice, quality, or spending across foundations is limited. This is somewhat surprising given the growing role of evaluation in foundations. Promising research on evaluation practice in the philanthropic sector, like the work by Buteau & Glickman, 2018, Buteau et al. (2016) and Coffman et al. (2013) cited above, is being conducted by affinity organizations such as the Center for Effective Philanthropy, Grantmakers for Effective Organizations, and Evaluation Roundtable. These case studies, articles, white papers, and survey reports document the varied approaches to evaluation, and challenges encountered with evaluation across the foundation sector.

In extension, foundations are also beginning to document and share data on their evaluation spending and evaluation practice. For example, The Hewlett Foundation released a report in early 2018 reflecting on the utility and value of the external evaluations they commissioned between 2009 and 2016 (Arbreton, Trivedi, & Twersky, 2018).

Scholarly work in this area is also emergent, including research into the establishment of evaluation policies and their impact on evaluation practice. Mattessich, Mueller, and Holm-Hansen (2009) documented the development of evaluation policies at the Wilder foundation. Dillman and Christie (2016) documented how implicit and explicit evaluation policies of the Robert Wood Johnson Foundation impacted the evaluations of three initiatives. Christie and Lemire (2019) argued for more scholarly attention to the pivotal role of evaluation policy in translating evaluation policy into evaluation practice.

The present study prioritizes major commissioners of evaluation in the philanthropic sector. Prominent foundations, with their outsized financial resources, have considerable influence on the sector and its approaches to evaluation. What they evaluate and how they evaluate matters. Christie and Lemire (2019) emphasized their importance and called for more research on their evaluation policies: “Formal evaluation policies (especially those of major commissioners of evaluation) serve to frame a significant and still growing part of evaluation practice" (p. 2). To address these gaps in the literature, this paper sets out to answer the following research questions:

1. To what extent are philanthropic foundations investing in evaluation?
2. To what extent do US foundations, who are major commissioners of evaluation, have publicly available evaluation policies?

Since the early 1990’s, the Foundation Center has been maintaining a dataset on grant funding distributed by leading foundations in America. While not all evaluation spending is conducted through grants, the Foundation Center data offers an opportunity to further examine foundation investment in evaluation. The remainder of this paper begins with an overview of the procedures used for this study. Next, findings from the Foundation Center data on evaluation and research spending by foundations between 2010 and 2014 are discussed. Lastly, conclusions and areas of future research are briefly discussed.

Methodology

The analysis for this study was conducted in two phases. The first phase sets out to identify the major commissioners of evaluation in the philanthropic sector. At present, there is little research into evaluation spending by foundations. However, the Foundation Center has been maintaining a data set since 2008 that tracks grant making by leading foundations in America. Using this data set, trends in grant funding for evaluation by philanthropic foundations were tracked for the years 2010 through 2014.

Upon the identification of the top US foundations commissioning evaluations, then, in phase two, their websites are reviewed to identify publicly available evaluation policies. More specifically, foundation websites for top funders of evaluation in 2013 and 2014 were reviewed to locate publicly available evaluation policies. More detail on the analysis conducted in phase one and two follows.

Phase 1: Foundation Center Data on Grants Allocated for Evaluation

The Foundation Center is a nonprofit organization, funded by more than 500 different foundations and corporations, that gathers and analyzes data on the foundation sector. Established in 1956, the Foundation Center manages “the most comprehensive database on U.S. and, increasingly, global grantmakers and their grants” (R. Mukai, personal communication, April 2018). In 2008, the Foundation Center tracked 25 percent of all foundation grants apportioned that year. Today, the top 1,000 independent, family, corporate, and community foundations (measured by total annual giving) electronically report on the scope, duration, purpose, and beneficiaries of their grant making directly to the Foundation Center. The Foundation Center data is sourced from “grants reported directly by funders to the Foundation Center and the 990/PF tax returns filed by foundations” (R. Mukai, personal communication, March 2018).

The dataset used in this study was compiled by the Foundation Center and acquired by the author in March 2017. Grant level data for all grants over $10,000 tagged as “program evaluations” or “institutional evaluations” in the grant activity or recipient organization activity columns were included. The Foundation Center defines program evaluation as “evaluation of a specific project or program; includes awards both to agencies to pay for evaluation costs and to research institutes and other program evaluators” and institutional evaluation as “evaluation of an organization's specific mission, purpose, profile, and performance with an aim to improving stated objectives and outcomes” (R. Mukai, personal communication, March 2018). Grants that had the keyword “evaluation” in the text description were also included in this analysis. Some grants were allotted for evaluation alone (e.g., the Measurement, Learning and Evaluation (MLE) Project of the URHI, a multi-country program that aims to improve the health of the urban poor).

Other grants include evaluation as one component of the broader grant purpose (e.g., “To extend general operating support to recipients of NDP production grants and to fund an evaluation of the program as it approaches its twentieth anniversary”). Grants may include funding for more than one activity. For example, a grant may be issued for both program development and program evaluation. It is not possible, from this review, to reliably discern what share of the total grant was allocated for program development on its own, or program evaluation on its own.

The resultant dataset contains information on 3,360 “program evaluation” or “institutional evaluation” grants distributed by 345 foundations between 2010 and 2014 (see Table 1). When the dataset was acquired, no information on foundation spending after 2014 was available. The Philanthropy Classification System (PCS) developed by the Foundation Center which categorizes the work of foundations, recipient organizations and the philanthropic transactions between those entities was used to guide analysis and interpretation of the dataset.

Table 1. Overview of Grants Included in the Foundation Center Dataset by Year

	Year				
	2010	2011	2012	2013	2014
Foundations	169	171	176	178	154
Grants	713	711	727	662	547
Total Spending on Evaluation	$218,245,189	$317,072,933	$273,830,948	$280,101,685	$243,485,294

Phase 2: Systematic Review of Foundation Websites

The next phase of analysis focused on grant data from only two years: 2013 and 2014. These years were selected because the Foundation Center’s data collection methods, and therefore the integrity of the data, improved annually. First, to focus this study, the top 50 foundations who invested heavily in evaluation spending across in 2013 and 2014 were identified. This was done by aggregating all evaluation grants awarded in those two years by each foundation. This resulted in 67 unique foundations.

Next, a systematic review of all 67 foundation websites was conducted. Documentation related to evaluation practice and policy were identified and cataloged. Through examination of foundation websites, I sought to “uncover meaning, develop understanding, and discover insights” pertinent to each foundation’s evaluation approach (Merriam, 2009, p. 118). Keywords used in this search included: evaluation, evaluation policy, approach to evaluation, impact, learning, measurement, and assessment. Documentation such as web pages, reports, policy briefs, interviews, and framing documents were cataloged. The initial systematic review of foundation websites was conducted between July and August 2017. Eighteen documents were copied from the web and organized in foundation specific folders on Google Drive.

To determine if the cataloged document was an evaluation policy, a preliminary round of coding was conducted using Trochim’s evaluation policy taxonomy (2009). For this study, evaluation policy is being defined as “...any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation” (Trochim, 2009, p. 13). To qualify for inclusion, evaluation documentation met the following criteria: First, the written documentation was publicly available on the foundation’s website. Next, the document contained at least four of the eight dimensions identified in Trochim’s taxonomy of evaluation policies (2009). Regardless of the document’s title, it qualified as an evaluation policy in this study if four or more categories of an evaluation policy were present. The document title did not need to include the term “policy”. Documents with terms like “approach” “principles” or “strategy” were included if they met the other criteria. Of the 18 cataloged documents, 12 foundations were found to have an evaluation policy.

Limitations

At present, the Foundation Center is the only known source for tracking evaluation spending in the sector. While the dataset offers noteworthy insights, there are some important omissions and limitations. First, the data are either self-reported by foundations directly to the Foundation Center or are retrieved from the funder's tax return. It is unclear how consistently foundations upload their data or if all grants that include evaluation funds are reported. Therefore, detailed information necessary to index whether a grant is for activities related to evaluation may be missing.

Furthermore, while there are over 87,000 private and community foundations, only the top 1,000 report their charitable spending to Foundation Center. The dataset used in the present paper includes 345 foundations that awarded at least one charitable grant for evaluation between 2010 and 2014. Next, the Foundation Center currently only tracks charitable grant funding, which does not account for the evaluation services being directly commissioned by foundations through other funding strategies (such as administrative funds). Third, as mentioned in the prior section, evaluation spending is frequently bundled with other knowledge production activities, program development and program execution. Therefore, discerning funds dedicated solely to evaluation using the Foundation Center data is difficult. Despite these limitations, the dataset on evaluation funding used in this paper offers valuable insights into the evaluation marketplace.

In phase two, only publicly available evaluation policies from top commissioners of evaluation were cataloged. Top commissioners of evaluation are prioritized because, as discussed in the findings section, they are driving demand for evaluation in the philanthropic sector. Their policies and practices are most likely influencing the philanthropic sector's approach to evaluation. Furthermore, given that they are commissioning a lot of evaluations, they are also the most likely to be developing written guidance to streamline their practice. This is one of the first studies to investigate evaluation policy broadly across the philanthropic sector. Publicly available policies are prioritized because if foundations are sharing their evaluation policies then, ideally, these policies are an accurate representation of what informs and guides evaluation practice at the foundation. However, it is likely that foundations also have evaluation policies that are implicit or not disseminated broadly. While not the focus of this research, that is a promising area of future research.

Lastly, while the term “evaluation policy” is defined in the evaluation literature it is not commonly used by foundations. Only two of the evaluation policies identified in this study used the term “policy” to describe the document. Rather than determine qualification for this study just based on a document’s title, the contents were analyzed and compared to the evaluation policy taxonomy proposed by Trochim (2009). To increase reliability, the process of identifying and coding policies is described in detail (Merriam, 2009).

Findings

The finding section is organized in two parts. First, findings are shared from the analysis of the Foundation Center data on foundation’s charitable grant spending on evaluation between 2010 and 2014. Then, findings from the search for publicly available evaluation policies are discussed.

Part 1: Evaluation Spending by Foundations between 2010 and 2014

Spending on evaluation ebbed and flowed between 2010 and 2014, as shown in Table 2 below. While evaluation spending increased by nearly $100 million between 2010 and 2011, it began to level out by 2012. By 2014, evaluation spending was at approximately the same level as in 2010. However, looking at evaluation spending as a percentage of total annual giving, as depicted in Figure 2, foundations spent less on evaluation in 2014 (1.7%) than they did between 2010 and 2013. In 2011, foundations allocated the most (2.6%) of their annual giving on evaluation.

Despite some volatility in evaluation spending, the number of evaluation grants awarded remained steady between 2010 and 2012, before decreasing 9 percent between 2012 and 2013 and another 8 percent between 2013 and 2014 to a low of 547 grants. The median and mean size of grants that included funds for evaluation remained relatively unchanged. The median grant size peaked in 2011 and was lowest in 2014. Conversely, the mean grant size grew annually, with the exception of 2013. Between 2010 and 2014 the mean grant size increased 9 percent ($41,000). See Table 2 and Figure 2 for additional detail on evaluation grants and spending by foundations between 2010-2014.

Table 2. *Foundation Evaluation Grants and Spending by Year*

	Year				
	2010	2011	2012	2013	2014
Total Foundations	169	171	176	178	154
Total Grants	713	711	727	662	547
*Annual Giving (billions)	$11.2	$12.7	$12.2	$12.8	$14.4
*Annual Evaluation Spending (millions)	$237	$334	$282	$285	$243
*Median Size of Grant (thousands)	$126	$157	$137	$127	$114
*Mean Size of Grant (thousands)	$442	$452	$457	$410	$483

Source: Foundation Center.
**Dollars adjusted for inflation (2014 constant dollar).*
**The mean represents the average, while the median is the middle number.*

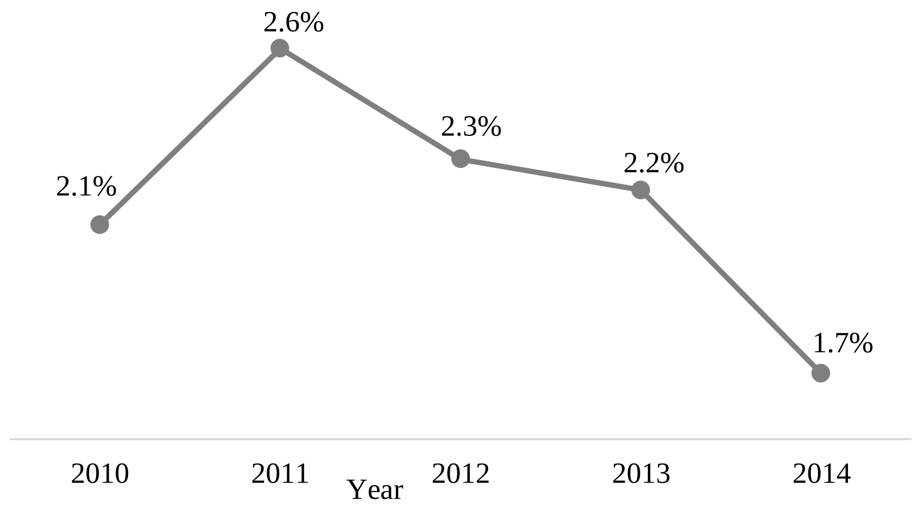


Figure 2. *Annual Evaluation Spending as a Percentage (%) of Annual Giving*

As the disparity between median and mean spending per evaluation grants indicates, evaluation spending was not equally distributed across foundations. Table 3 focuses on the top 10 foundations by evaluation-related spending across all five years. Cumulatively, the top 10 foundations were responsible for 69 percent, or $953.4 million, of all grant spending on evaluation between 2010 and 2014 ($1.38 billion). These foundations are: The Bill & Melinda Gates Foundation, The Robert Wood Johnson Foundation, The Edna McConnell Clark Foundation, The Gordon and Betty Moore Foundation, The Ford Foundation, The William and Flora Hewlett Foundation, The Rockefeller Foundation, The W. K. Kellogg Foundation, The Wallace Foundation, and The Annie E. Casey Foundation. In 2014, these foundations spent $166.7 million on evaluation-related grants, which corresponds to 68.5 percent of the $243.5 million in total spending on evaluation-related grants.

Kellogg (54 percent), Gates (75.3 percent) and RWJF (75 percent) spent the majority of their evaluation related spending on health-related grants. Hewlett (38.6 percent), Clark (40.3 percent), and Wallace (96.4 percent) spent the majority of their evaluation funds on education related grants. For Rockefeller and Ford, the largest percentage of their evaluation spending (32.3 percent and 31.5 percent) was dedicated to community and economic development grants. Annie E. Casey spent 57.4 percent of its evaluation spending on human services grants and Moore Foundation spent 35.7 percent of its evaluation spending on science related grants.

Even among these top 10 consumers of evaluation, large differences in evaluation spending are evident. Cumulative spending over five years ranged from as high as $537.2 million by the Gates Foundation to as low as $16 million by the Annie E. Casey Foundation. The mean evaluation spending across the 10 foundations was $95.3 million. This wide range is also reflected in the percentage of funds allocated for evaluation spending between 2010 and 2014. On average, this group allocated 7 percent towards evaluation. An outlier in the data, the Edna McConnell Clark Foundation reported that 35 percent of its grant funding included dollars for evaluation. Lastly, the average grant size between 2010 and 2014 was as low as $144.5 thousand and as high as $2.2 million, with an average grant size across the 10 foundations of $687.5 thousand.

The Gates Foundation was a notable outlier in this dataset. Its evaluation spending between 2010 and 2014 represented 39 percent of all evaluation-related spending. The Gates Foundation spent $121.1 million more on evaluation-related grants than the other nine leading foundations combined. Removing the Gates Foundation from the analysis reduced the average evaluation-related spending across five years by 33 percent from $317.5 million to $210 million.

Table 3. *Top 10 Foundations by Aggregated Evaluation Spending (2010-2014)*

	2010-2014 Aggregate				
	Evaluation spending (millions)	Total spending (billions)	% allocated for evaluation	Evaluation grants	Average grant size
Bill & Melinda Gates	$537.3	$17.0	3.2%	240	$2,238,745
Robert Wood Johnson	$116.1	$1.7	6.8%	229	$507,065
Edna McConnell Clark	$74.0	$207.3 millions	35.7%	90	$822,334
Gordon & Betty Moore	$58.3	$1.3	4.5%	89	$655,395
Ford	$56.2	$2.5	2.3%	119	$472,046
William & Flora Hewlett	$34.7	$1.3	2.7%	85	$408,673
Rockefeller	$24.6	$688.4 millions	3.6%	66	$372,907
W. K. Kellogg	$19.8	$1.6	1.2%	45	$440,561
Wallace	$16.3	$263.3 millions	6.2%	20	$813,365
Annie E. Casey	$16.0	$373.8 millions	4.3%	111	$144,592
Total	$953.3	$26.9		1,094	

Table 4 shows that foundations funded evaluation across a wide range of sectors between 2010 and 2014. The top three areas of evaluation-related giving were health, education, and human services. 44 percent, or $605.6 million, of all evaluation funds distributed between 2010 and 2014 was allocated for health-related causes such as HIV/AIDS, maternal and perinatal health, public health, and malaria. While health represented 44 percent of all spending, it represented only 24 percent of grants distributed. Education causes also received 24 percent of all grants, but only 22 percent, or $301.3 million, of total spending. Education-related causes included spending on elementary and secondary education, education services, university education, and reading promotion. The third largest recipient of foundation giving, both by grant count and total spending, were human services-related causes, such as child welfare, nursing care, and foster care. While human services represented 14 percent of all grants, they only attained 8 percent of total spending, or $105.3 million, between 2010 and 2014. Sports/recreation and religion each received meager investment (0.3 percent and 0.2 percent, correspondingly) of foundation funds between 2010 and 2014.

Table 4. *Total Evaluation Spending by Grant Purpose (2010-2014)*

	2010-2014 Aggregate			
	Evaluation Spending (millions)	Grants	% of Total Grants	% of Total Evaluation Spending
Health	$605.7	790	23.5%	43.9%
Education	$301.3	806	23.9%	21.8%
Human services	$105.4	456	13.6%	7.6%
Community & economic development	$80.4	318	9.5%	5.8%
Science	$45.5	73	2.2%	3.3%
International relations	$41.7	66	1.9%	3.0%
Environment	$40.5	151	4.5%	2.9%
Public affairs	$24.5	101	3.0%	1.8%
Human rights	$20.9	97	2.9%	1.5%
Agriculture, fishing and forestry	$19.8	44	1.3%	1.4%
Arts and culture	$19.8	119	3.5%	1.4%
Public safety	$19.4	91	2.7%	1.4%
Information & communications	$15.9	50	1.5%	1.2%
Social sciences	$15.1	49	1.5%	1.1%
Philanthropy	$14.7	104	3.1%	1.1%
Unknown	$4.4	10	0.3%	0.3%
Sports & recreation	$3.4	26	0.8%	0.3%
Religion	$2.7	9	0.3%	0.2%
Total	$1.38 (trillions)	3360		

At present, little is known about the evaluation procurement and contracting practices of foundations and nonprofits. An overview of the types of organizations that received evaluation-related grants disbursed by foundations between 2010 and 2014 can be found in Table 5. Grants that included evaluation funding were principally allocated to non-governmental organizations

(NGOs), such as public charities. NGOs received 65.7 percent, or $907.7 million. A little more than a quarter of all grants, or $365.5 million, went to schools or school districts. Considered collectively, NGOs and schools received over 92 percent of the evaluation spending. Government agencies and businesses, like those that specialize in evaluation and research, received less than 3 percent of total evaluation-related spending between 2010 and 2014.

While more research is needed to better understand who is awarded evaluation funding by foundations, these data suggest that the evaluation marketplace is decentralized and local. Only 3 percent or less of the evaluation giving documented in Foundation Center data went directly to a company that specializes in evaluation. Most of the funds were first awarded to a local NGO or school, which in turn solicited and procured evaluation services. Further research is needed into what happens with evaluation-related grants once they are distributed to the receiving organization. It is likely that the receiving organization, like an NGO or school, expend a portion of those funds to hire external evaluators to conduct and support their evaluation needs. It is unclear, however, what role and influence, if any, foundations have in deciding which evaluation providers are hired for these services. The current dataset limits our ability to discern how grantees utilized foundation funding for evaluations and which evaluation providers received these funds. More research into evaluation procurement and contracting is needed to better understand the interaction between the demand and supply side of this evaluation market.

Table 5. *Recipients of Evaluation-Related Grants (2010-2014)*

	2010-2014 Aggregate		
	Number of grants	Evaluation spending	% of total evaluation spending
Non-governmental organizations	2,190	$907,796,403	65.7%
Schools	736	$365,521,403	26.5%
Governments & agencies	130	$38,571,625	2.8%
Businesses	118	$34,471,414	2.5%
Religious institutions	111	$18,353,409	1.3%
Unknown	67	$11,969,198	0.9%
Charitable gift funds	1	$3,256,998	0.2%
Multipurpose centers	6	$1,160,658	0.1%
Intergovernmental organizations	1	$22,000	0.0%
Total	3,360	$1,381,123,108	100.0%

Part 2: Prevalence of Evaluation Policies

The broad analysis of foundation spending revealed that the majority of evaluation spending was concentrated amongst a few foundations. More specifically, of the $1.38 billion spent on evaluation between 2010 and 2014, the top 10 foundations were responsible for 69 percent ($953.4 million). Therefore, in phase two of the present study, I set out to identify foundations heavily investing in evaluation practice. To qualify as a major commissioner of evaluation, the foundation needed to be one of the top 50 foundations investing in evaluation in either 2013 or 2014. This analysis resulted in 67 unique foundations. Thirty-three foundations appeared in the top 50 list in both years. Seventeen foundations increased their standing in the top 50 list between 2013 and 2014 (For example the William and Flora Hewlett Foundation was #4 in 2013 but #3 in 2014), while 13 decreased their standing. Three foundations remained in the same position from 2013 to 2014. For more information on evaluation spending at the 67 foundations included in the sample, see Appendix A.

Of the 67 foundations, 55 did not have an evaluation policy posted on their company website (see Table 6). However, 12 were found to have written evaluation policies (see Appendix B for copies of policies). Table 7 lists the foundation, title of evaluation policy document, and the document's word count. It also includes the abbreviated name of the foundation that will be used throughout the remainder of the present dissertation. There are several important differences across the 12 policies worthy of elaboration. First, only Gates and Irvine actually use the term "policy" in the title of their document. More common terms are "approach", used by six foundations, and "principles", used by four foundations. The length of policies also differs greatly across foundations. The average length of the 12 policies is 1754 words. The longest policy, written by the Hewlett Foundation, is nearly six times longer than the average and three times longer than the second longest policy, written by Gates. The shortest policies found were written by Skillman (324 words) and Doris Duke Charitable Foundation (372 words).

The amount spent on evaluation-related programming varies widely across these 12 foundations. Table 8 details the evaluation-related and total spending between 2013 and 2014. Cumulative spending on evaluation-attached grants ranged from as high as $223.6 million by the Gates Foundation to as low as $881,700 by the California Wellness Foundation. The mean evaluation spending across the 12 foundations was $42.6 million. Between 0.7 percent and 6.2 percent of total spending was allocated for evaluation. On average, this group allocated 2.8 percent towards evaluation-attached grants.

The Gates Foundation was a notable outlier in this dataset. Their evaluation spending between 2013 and 2014 represented 81 percent of all evaluation-related spending. The Gates Foundation spent $170.1 million more on evaluation-related grants than the other 11 leading foundations combined. Removing the Gates Foundation from the analysis reduced the average evaluation-related spending across 2013 and 2014 from $42.6 million to $4.9 million.

Table 6. *Foundations with No Publicly Available Evaluation Policy Documentation*

Alfred P. Sloan Foundation	S.D. Bechtel, Jr. Foundation
Altman Foundation	Stuart Foundation
Barr Foundation	The Andrew W. Mellon Foundation
Ben and Catherine Ivy Foundation	The Annie E. Casey Foundation
Blue Shield of California Foundation	The Brown Foundation, Inc.
Buffett Early Childhood Fund	The Chicago Community Trust
Burroughs Wellcome Fund	The Cleveland Foundation
Bush Foundation	The Coca-Cola Foundation, Inc.
California Endowment	The Commonwealth Fund
Carnegie Corporation of New York	The East Bay Community Foundation
Charles Stewart Mott Foundation	The Edna McConnell Clark Foundation
Citi Foundation	The Harry and Jeanette Weinberg Foundation, Inc.
Claude Worthington Benedum Foundation	The Heinz Endowments
Conrad N. Hilton Foundation	The John and Lisa Pritzker Family Fund
Eli & Edythe Broad Foundation	The Joyce Foundation
Ford Foundation	The JPB Foundation
Foundation to Promote Open Society	The Kresge Foundation
GHR Foundation	The Leona M. and Harry B. Helmsley Charitable Trust
Gordon and Betty Moore Foundation	The Marcus Foundation, Inc.
Hawaii Community Foundation	The Michael and Susan Dell Foundation
Healthcare Georgia Foundation, Inc.	The New York State Health Foundation
Heising-Simons Foundation	The Noyce Foundation
Howard G. Buffett Foundation	The Rees-Jones Foundation
Laura and John Arnold Foundation	The Robert Wood Johnson Foundation
Lilly Endowment Inc.	The Skoll Foundation
MetLife Foundation	The William Penn Foundation
Omidyar Network Fund, Inc.	W. K. Kellogg Foundation
Raikes Foundation	

Table 7. *Foundations with Publicly Available Evaluation Policy Documentation*

Foundation	Abbreviation	Policy Document Title	Word Count
Bill & Melinda Gates Foundation	Gates	Evaluation Policy	3032
California Wellness Foundation	C.A. Wellness	An Approach to Evaluation: Principles for TCWF's Evaluation Activities	436
David & Lucile Packard Foundation	Packard	Guiding Principles and Practices for Monitoring, Evaluation, and Learning	1390
Doris Duke Charitable Foundation	Doris Duke	Grant-Making Process: Evaluation	372
Duke Endowment	Duke	Guiding Principles: Learning and Evaluation	577
James Irvine Foundation	Irvine	Evaluation Policies and Guidelines: Approach to Evaluation	966
John D. & Catherine T. MacArthur Foundation	MacArthur	Our Approach to Evaluation	540
Kansas Health Foundation	K.S. Health	Evaluation/Learning: Our Approach	621
Rockefeller Foundation	Rockefeller	Our Strategy: Monitoring and Evaluation	1720
Skillman Foundation	Skillman	Evaluation Philosophy and Approach	324
Wallace Foundation	Wallace	How We Work: The Wallace Approach	1198
William and Flora Hewlett Foundation	Hewlett	Evaluation Principles and Practices	9875

Table 8. *2013-2014 Aggregated Evaluation Spending for Foundations with Evaluation Policies*

Foundation	Evaluation spending (millions)	Total spending (millions)	% allocated for evaluation
Gates	$223.6	$7.5 billion	3.0%
C.A. Wellness	$881.7 (thousand)	$54.8	1.6%
Packard	$3.5	$484.2	0.7%
Doris Duke	$6.6	$126.2	5.2%
Duke	$2.4	$196.1	1.2%
Irvine	$4.6	$135.3	3.4%
MacArthur	$2.6	$317.0	0.8%
K.S. Health	$1.7	$27.1	6.2%
Rockefeller	$5.1	$267.3	1.9%
Skillman	$1.0	$32.2	3.2%
Wallace	$2.2	$115.6	2.0%
Hewlett	$23.0	$523.5	4.4%

Source: Foundation Center.

**Dollars adjusted for inflation (2014 constant dollar).*

The Foundation Evaluation Landscape

This paper began with a brief introduction to the foundation sector and then turned to a review of extant research into evaluation practice in the sector. The Findings section examined the Foundation Center data on charitable grant spending on evaluation between 2010 and 2014. While there are limitations to using this dataset, it also offers a rare opportunity to appraise spending on evaluation in the foundation sector. The findings section concludes with an examination of the prevalence of evaluation policies across leading foundations.

The empirical examination of foundation spending presented in this paper indicates that the demand for evaluation, and other knowledge production services, is a fixture of foundation funding. On average, the foundation sector apportioned approximately $276 million annually for grants that included evaluation services. Evaluation spending as a percentage of total annual giving peaked at 2.6 percent in 2011. As of 2014, it was at a low of 1.7 percent. However, no continual growth in evaluation spending was observed. That could be a reflection of the limited data available through the Foundation Center. Alternatively, the annual fluctuations may also be evidence of foundations struggling to identify a stable strategy for commissioning and utilizing evaluation.

The current composition of the data made it arduous to identify which evaluation providers received grant funding from foundations. Only $33 million (3 percent) of evaluation spending was apportioned directly to businesses, which includes evaluation and research providers. The majority of funds were awarded to nonprofit organizations (66 percent) and schools (26 percent), which then in turn solicit and procure evaluation services. This suggests that foundations are bundling evaluation and program funding and requesting that grantees procure and contract evaluation services directly.

Lastly, the analysis conducted for this paper revealed that of the 67 major commissioners of evaluation in 2013 and 2014, 12 foundations have publicly available evaluation policies. There are notable differences between the 12 foundations and their policies. First, evaluation spending varies greatly across foundations with evaluation policies. Between 0.7 percent and 6.2 percent of total spending was allocated for evaluation- attached grants. This suggests that foundations that spend about 1 percent or more of their total spending on evaluation dedicate resources to drafting an evaluation policy. However, more research is needed to better understand the connection between investment on evaluation and the development of evaluation policies.

Interestingly, only two foundations, Irvine and Gates, used the term “policy” to refer to their documents. Foundations instead preferred terms like “approach” or “principles”. These documents had similar dimensions as those outlined in Trochim’s evaluation policy taxonomy, however foundations did not choose to use the same terminology. Furthermore, the length of policies ranged significantly. The shortest policy was one-tenth (324 words) of the longest evaluation policy (3032 words).

Future Research: Looking Ahead to Paper 2

This paper is one of the first attempts at understanding how the demand for evaluation and research services differs across foundations. Extant research has begun to document the increase in evaluation training and staffing in the foundation sector. Some foundations, namely the Hewlett Foundation, have begun to conduct internal reviews of their evaluation practices and expenditures. However, there are few published studies that present empirical insights into differences and similarities across foundations. By charting existing trends in evaluation spending, this research offers foundation staff an opportunity to consider how their evaluation practice aligns with or deviates from that of the broader sector.

To better understand what is driving evaluation demand, more attention needs to be paid to the policies framing evaluation practice in the sector. At present, little is known about the policies that govern evaluation practice and spending across foundations. Yet, these policies likely impact a foundation's interest in and demand for evaluation. Given that evaluation policy encompasses most aspects of evaluation practice and theory, it is a promising area of research that can inform our understanding of evaluation practice in the sector (Trochim, 2009).

Now that 12 evaluation policies in the philanthropic sector have been cataloged, future studies could investigate how these policies compare to theoretical conceptions of evaluation policy. Evaluation research is in the initial stages of understanding evaluation policy. It has been defined and two taxonomies detailing components of evaluation policy have been put forward. However, it is unclear how these taxonomies compare to the policies guiding evaluation practice in the philanthropic sector. Furthermore, it is uncertain how evaluation policies in the philanthropic sector compare to each other. This is a promising area for future research.

Paper 2.
A Closer Look: An Analysis of Evaluation Policies in the Philanthropic Sector

Abstract

Evaluation policy encompasses most aspects of evaluation practice and theory, making it one of the most critical areas of research on evaluation (Trochim, 2009). Depending on quality and scope, evaluation policies can aid or constrain the crucial work of foundations and nonprofit organizations. Yet, extant research paints evaluation policy in only broad strokes. To date, two taxonomies have been proposed that identify components of evaluation practice that may be specified in an evaluation policy. Little is known about how these taxonomies align with evaluation policies developed by philanthropic foundations. Through document analysis, this paper first compares evaluation policies from 12 foundations to the proposed taxonomies to extrapolate their similarities and differences. Then, proposed modifications to the taxonomies are discussed and an updated taxonomy is shared to guide future development and review of evaluation policies.

Introduction – Toward an Understanding of Evaluation Policy

To date, few publications have focused on evaluation policy. A search of article keywords and titles published by prominent evaluation journals (New Directions for Evaluation, American Journal of Evaluation and Evaluation Review) resulted in a 2009 volume edited by Trochim, Mark, and Cooksy, *Special Issue: Evaluation Policy and Evaluation Practice* and 15 other publications between 2011 and 2019. Despite minimal attention given to evaluation policy in peer-reviewed publications, evaluation policy has continued to impact the practice of evaluation. The 2008 American Evaluation Association (AEA) conference theme was "Evaluation Policy and Evaluation Practice". This drew practitioners from around the world to present on topics such as “Moving Evaluation Policy and Practice Towards Cultural Responsiveness”; “Developing an Evaluation Policy For a Nonprofit, Youth-Serving Organization While Meeting the Needs of the Organization and its Funders”; and “Funder Evaluation Policy: Setting the Stage for Evaluation Practice and Use” (American Evaluation Association, 2008). The Board of Directors of AEA convened an Evaluation Policy Task Force (EPTF) around the same time to advance the use of evaluation policy. EPTF continues to meet and promote the use of evaluation policies, primarily across federal agencies.

According to Trochim (2009), evaluation policy is “any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation” (p. 19). Another definition proposed in the same volume by Mark, Cooksy, and Trochim (2009) states that evaluation policy “refers to rules or principles that help set the content, characteristics, and context of evaluation itself” (p. 5). More recently, Patton (2018) has been advancing the idea of principles-focused evaluation to “provide guidance and direction for those working for [societal] change” (p. 13). His definition of an evaluation principle is closely aligned to the above definitions of evaluation policy:

> “A good principle provides guidance for making choices and decisions, is useful in setting priorities, inspires, and supports ongoing development and adaptation” (p. 13).

When principles are made explicit, they create what Trochim (2009) calls an “evaluation policy”. These definitions suggest that through evaluation policies organizations and agencies tailor evaluation practice to their unique needs and contexts.

Evaluation policies are “consciously adopted” and “the organization institutes consequences for encouraging or enforcing them” (Trochim, 2009, p. 17). Therefore, anything that is encouraged, enforced, or used to guide decisions and actions should be treated as a policy, regardless of whether it is implicitly or explicitly stated. However, Christie and Lemire (2019) challenge the idea that implicitness and enforcement are defining characteristics of evaluation policy. They set forth that:

> "an evaluation policy becomes an evaluation policy when it is formulated and made explicit as such—consequences or no consequences" (pg. 6).

The Christie and Lemire (2019) understanding of evaluation policy is assumed in the present paper. Evaluation policy is understood as any written document that details an organization's rules and principles for evaluation practice, regardless of whether implementation is encouraged or enforced.

Evaluation policies can encompass and dictate nearly every aspect of evaluation practice. Two taxonomies have been proposed that explicate the components of an evaluation policy. The first taxonomy was developed in 2007 by AEA's board of directors when convening the Evaluation Policy Task Force (EPTF). The board of directors proposed seven general areas, or components, that are expected to be present across evaluation policies. The components and definitions they identified are as follows (AEA Evaluation Policy Task Force, 2007):

1. Evaluation definition. How, if at all, is evaluation defined in an agency or in legislation? In such contexts, how is evaluation formally distinguished from or related to other functions such as program planning, monitoring, performance measurement or implementation?
2. Requirements of evaluation. When are evaluations required? What programs or entities are required to have evaluations? How often are evaluations scheduled? What procedures are used to determine when or whether evaluation takes place?

3. Evaluation methods. What approaches or methods of evaluation are recommended or required by legislation or regulation, for what types of programs or initiatives?
4. Human resources regarding evaluation. What requirements exist for people who conduct evaluations? What types of training, experience or background are required?
5. Evaluation budgets. What are the standards for budgeting for evaluation work?
6. Evaluation implementation. What types of evaluation implementation issues are guided by policies? For instance, when are internal versus external evaluations required and how are these defined?
7. Evaluation ethics. What are the policies for addressing ethical issues in evaluation?

Two years later, Trochim (2009) suggested a different taxonomy that included eight components (p. 20). Trochim did not explicitly define each component. Therefore, the definitions detailed below were derived from my interpretation of the examples of components he provided. The taxonomy consists of the following components:

1. Evaluation goals: The goals of evaluation in the organization or context.
2. Participation policies: Who is involved in evaluations, when, and under what circumstances?
3. Capacity building policies: The intention of the organization to build evaluation capacity.
4. Management policies: General management of evaluation practice and the allocation of resources (such as: time, resource allocation, or scheduling).
5. Roles policies: The responsibilities of those involved in conducting evaluations.
6. Process and methods policies: Procedures for conducting evaluations including question identification, sampling, measurement, design, and analysis.

7. Use policies: How evaluation findings are communicated to encourage use.
8. Evaluation of evaluation (meta-evaluation) policies: How the quality of evaluations is assessed.

These taxonomies were created theoretically and reveal what evaluation theorists and practitioners expect evaluation policies to include. Trochim (2009) acknowledged the limitations of conceptual taxonomies in his call for more research:

> "It would be preferable to develop a classification system for evaluation policies empirically" (p. 20).

That is the aim of the present paper. There is a need to compare the proposed taxonomies to existing evaluation policies. Such research will help verify, alter, and improve the current thinking on what components shape evaluation policy.

To date, only a few studies have used these taxonomies to review extant evaluation policies. Leeuw (2009) used Trochim's taxonomy to explicate the evaluation policy of the Netherlands. Utilizing the taxonomy to guide analysis helped Leeuw surface shortcomings, namely that "roles" or "process and methods" are not adequately stipulated in the policy. This analysis also highlighted that the policy emphasized "goals", "capacity building", and "meta-evaluation". Trochim's taxonomy also guided analysis of 52 policies from 36 multilateral and bilateral aid agencies and government organizations, like the United Nations and World Health Organization (Al Hudib, 2018). Al Hudib (2018) modified the taxonomy with components that emerged from content analysis, namely "norms and standards" and "competencies and ethics". The findings from the analysis conducted by Al Hudib (2018) were not incorporated into the present paper because they were published after the analysis presented in this paper was conducted.

There is only one study that has used both frameworks to guide analysis of evaluation policy in the philanthropic sector. Dillman (2014) examined how the explicit and implicit evaluation policies of a single foundation, the Robert Wood Johnson Foundation, were implemented by stakeholders during three different evaluations. Dillman (2014) integrated the AEA (2007) and Trochim (2009) taxonomies into a single framework, titled the Evaluation Policy Coding Framework. It was used to identify and examine the implicit and explicit evaluation policies at the foundation.

Research conducted by Leeuw (2009), Dillman (2014), and Al Hudib (2018), demonstrated that these two taxonomies are a promising analytic tool for examining evaluation policies. The present paper utilizes a similar approach to analyze the content of evaluation policies guiding evaluation practice in the philanthropic sector. Little is known about the composition of evaluation policies being written and used by foundations to guide their evaluation practice. To address this gap, foundation websites were examined to identify explicit and public evaluation policies (Kinarsky, paper 1). Of the 67 foundation websites reviewed, 12 evaluation policies were cataloged. The present research analyzes those 12 policies and addresses the following research questions:

1. How do publicly available evaluation policies in the philanthropic sector compare to taxonomies proposed by Trochim (2009) and the American Evaluation Association (2007)?
2. How do publicly available evaluation policies in the philanthropic sector compare to each other in terms of their comprehensiveness and completeness?

Methodology

The present paper builds on the work discussed in paper one, which cataloged publicly available evaluation policies at 12 foundations (see Appendix B for links to these policies). Analyses began with a deductive approach driven by the taxonomies proposed by Trochim (2009) and American Evaluation Association (2007). Then, through inductive inquiry, emergent themes were incorporated into the analysis. The following describes the three rounds of coding that were employed.

To guide the first round of coding, a codebook consisting of a priori codes was developed. The content of the coding manual was heavily informed by the evaluation policy taxonomies proposed by Trochim (2009) and American Evaluation Association (2007). The components of these taxonomies are identified in the introduction to this paper.

Coding was first conducted using the comment feature in Google Docs. When possible, full sentences were coded, rather than just key words or phrases. This consistency helped ensure that the word count was a reliable measure across policies and that the full sentiment of the policy component was captured. Following the initial round of coding, extensive changes were made to the codebook. The refinement of the codebook included consolidation of codes, development of "child codes" and the addition of emergent codes. This is discussed in greater detail in the findings section.

After the coding manual was revised, a pilot test was conducted. During a research seminar, a group of graduate students were asked to apply the coding manual to an evaluation policy. Their code application and questions were then discussed as a group. Following the seminar, minor modifications were made to the codebook. Definitions were refined based on feedback and example text was added to encourage consistent interpretation and application of codes.

Then, a second round of content analysis began. Evaluation policies and the codebook were imported into MaxQDA, a qualitative data analysis software. All 12 evaluation policies were coded again. When uncertainties arose, a “review” code was applied to the text. At the end of the second round of coding, all excerpts tagged “review” were re-analyzed and assigned a primary code. Also, code application in MaxQDA was compared to that which was conducted in phase one of coding in Google Docs. This analysis identified inconsistencies with code application and aided further refinement of the codebook.

The third, and final, round of content analysis was conducted by the author and another graduate student with expertise in program evaluation. Researchers worked independently but convened four times over two months to discuss progress. After four policies were coded, researchers gathered in-person to discuss the policies and application of codes. In particular, the definition and boundaries of each code were deliberated. Codes were refined before commencing with the remaining policies.

Team coding was implemented to ensure that codes were being consistently interpreted and applied (Saldaña, 2013; Bazeley, 2013). This analysis was not intended to calculate inter-rater reliability. Influenced by the work of Fierro (personal communication, December 2018), this approach to content analysis facilitated a rich discussion about the content of each policy and resulted in a codebook that best captures the content of extant evaluation policies. The final codebook includes information on all the codes used, their origin, definition and example text. It is discussed in greater detail in the findings section.

The 12 policies included in this study are detailed in Table 9. This table also contains information on the policy document title, total word count, the number of selected excerpts and codes applied to each policy. In total, 998 codes were applied across 475 excerpts. The number of excerpts and codes differ because in some cases several codes were applied to the same excerpt. Nearly half of all excerpts (44 percent and codes (42 percent were applied to the two longest policies: Gates Foundation (3032 words) and Hewlett Foundation (9875 words). The Hewlett Foundation had the greatest number of excerpts and codes (25 percent of excerpts and 22 percent of codes) and the Gates Foundation was a close second (19 percent of excerpts and 20 percent of codes).

Table 9. *Descriptive Information on Coding of Evaluation Policies*

Foundation	Abbreviated Name	Policy Document Title	Word Count	Excerpts	Codes
Bill & Melinda Gates Foundation	Gates	Evaluation Policy	3032	88	199
California Wellness Foundation	C.A. Wellness	An Approach to Evaluation: Principles for TCWF's Evaluation Activities	436	16	32
David & Lucile Packard Foundation	Packard	Guiding Principles and Practices for Monitoring, Evaluation, and Learning	1390	27	45
Doris Duke Charitable Foundation	Doris Duke	Grant-Making Process: Evaluation	372	22	63
Duke Endowment	Duke	Guiding Principles: Learning and Evaluation	577	25	58
James Irvine Foundation	Irvine	Approach to Evaluation	966	31	65
John D. & Catherine T. MacArthur Foundation	MacArthur	Our Approach to Evaluation	540	23	57
Kansas Health Foundation	K.S. Health	Evaluation/Learning: Our Approach	621	23	44
Rockefeller Foundation	Rockefeller	Our Strategy: Monitoring and Evaluation	1720	47	91
Skillman Foundation	Skillman	Evaluation Philosophy and Approach	324	31	78
Wallace Foundation	Wallace	How we work: The Wallace Approach	1198	25	45
William and Flora Hewlett Foundation	Hewlett	Evaluation Principles and Practices	9875	117	221
		Total:	21,051	475	998

Limitations

Before advancing to a discussion of findings, there are a few limitations worth detailing. First, the philanthropic sector is vast, some estimate that it consists of over one hundred thousand foundations. This study discusses trends based on 12 of the top commissioners of evaluation in the philanthropic sector. This selection was strategic but represents a very limited sample of foundations. While not representative of the sector, examination of 12 evaluation policies does surface valuable findings that can inform future studies of the sector.

Next, the content analysis and interpretation in this study is dependent on the researcher. To address researcher bias, the study design and, more specifically, the codebook were informed by extant research on evaluation policies. During content analysis, the codebook was reviewed by others knowledgeable about evaluation theory and practice to increase reliability and validity of findings (Saldaña, 2013; Bazeley, 2013). The codebook was piloted in a research seminar and then team coding was used during the third and final round of coding to ensure that codes were being interpreted and applied consistently. Throughout the process analytic memos were generated to track code definitions, emergent patterns, and challenges.

Lastly, policies are living documents that are updated and revised in response to new processes or learnings. Since the start of this study, some foundations have updated or removed their evaluation policies from the foundation's website. The motivation for and substance of these changes were not investigated for the present study.

Findings

The following section addresses findings from content analysis of 12 evaluation policies. First, adaptations to current evaluation policy taxonomies are discussed. Then a modified taxonomy is proposed for guiding future research. Finally, the comprehensiveness of components and completeness of evaluation policies are examined to identify important similarities and differences across policies.

Evaluation Policy Components: A Modified Taxonomy

The first aim of this paper is to investigate how evaluation policies written by major commissioners of evaluation in the philanthropic sector compare to those proposed by Trochim (2009) and AEA's board of directors (2007). After three rounds of content analysis, a number of modifications surfaced. I begin by describing refinements made to the taxonomies before concluding this section with a modified taxonomy that can be used to guide future research. A summary of components proposed by AEA (2007) and Trochim (2009), their definitions, and a brief summary of how they were modified following three rounds of content analysis is summarized in Table 10. The modified taxonomy, including component definitions and example text from evaluation policies, is presented in Table 11.

First, similar components across the two taxonomies were merged. Two components proposed by Trochim (2009), "Participation" and "Roles", as well as the "Human Resources" component proposed by AEA, capture information on who should be involved in an evaluation and in what capacity. The nuance among them was not meaningful during content analysis so they were ultimately merged into one component: "Participation and Roles". This component was then further refined to capture information shared about different stakeholders: foundation staff, grantees, and external evaluators.

Seven components across the taxonomies describe different aspects of evaluation implementation and management. These included: "Evaluation Implementation"; "Requirements of Evaluation"; Evaluation Budgets; "Evaluation Implementation; "Process and Methods"; and

“Management”. During content analysis, the same text was simultaneously labeled with several of these components. While some overlap is expected and appropriate, it became apparent that there was a need to distinguish between managing an evaluation, or project management, and the design or methodology used for an evaluation. Therefore, these codes were simplified into two components: “Management” and “Design and Methods”.

The “Management” component was re-defined to include three sub-components: the resources necessary to support a vibrant evaluation practice (ex: budgets, staff time); when an evaluation is required (ex: at the start or end of a program); and the management of an evaluation, or project management. For example, Hewlett’s policy specifies that staff should “request periodic memos to document progress and any obstacle the evaluator is facing in data collection, data quality, or other areas” (p. 19). Text like this was coded as “Management” because it relates to evaluation implementation and process.

The “Process and Methods” component in Trochim’s taxonomy was renamed to “Design and Methods” to more narrowly denote guidance on evaluation design and methodology. For example, The Doris Duke policy encouraged particular methods: “Through site visits, grantee consultations and review of annual progress reports, staff determines compliance with grant agreements and assesses initial outputs, outcomes and indicators that are likely to correlate with the future impact of the grant” (p. 1). Other policies outlined different types of evaluations and specified when they should be used. This information was also captured under “Design and Methods”. For example, the Gate’s policy encouraged staff to utilize the following three types of evaluations (p. 3): “Evaluations to understand and strengthen program effectiveness”; “Evaluations to test the causal effects of pilot projects, innovations, or delivery models”; and “Evaluations to improve the performance of institutions or operating models”.

In addition to the components proposed by AEA (2007) and Trochim (2009), three new components emerged during content analysis: “Background”; “Avoid”; and “Example”. “Background” captures text that conveys the history of evaluation practice at the foundation and motivation for drafting an evaluation policy. Neither of these sub-components were present in prior taxonomies. AEA’s “evaluation definition” component was situated under “Background”.

Evaluation policies also specified practices that could be detrimental to an evaluation and thus should be avoided. This information was captured by the “Avoid” component. Some foundations, like Skillman, included broad “Avoid” statements such as: “[Evaluation should be] integrated in the work (not a stand-alone activity)” (p. 1). Other foundations, like Gates, detailed when an evaluation was not necessary: “Evaluation is a low priority when the results of our efforts are easily observable or when our partners are conducting basic scientific research, developing but not distributing products or tools, or creating new data sets or analyses” (p. 3). “Avoid” was not originally proposed by AEA (2007) or Trochim (2009), however Bullock, Anderson, and Brady (1983) specified that policies can encourage and discourage actions. Adding this component emphasizes the ways in which evaluation policies discourage particular evaluation practices.

Lastly, several policies integrated examples or anecdotes that further explained their evaluation principles and practices. For example, the C.A. Wellness policy guides readers to an external source for more information: “One of the most recent examples of such an analysis is Jennie Schacht’s report on grants made to regional clinic associations and consortia, which is posted in this section of the website” (p. 3). The Hewlett Foundation policy integrates callout boxes throughout the document to illustrate how components were or could be applied in practice. For example, one callout box on page 15 describes a hypothetical evaluation and discusses different guiding questions that could be used depending on the aims of the evaluation. The "Example" code captures how this approach was implemented across evaluation policies.

Table 10. *Modifications Made to the Evaluation Policy Components Proposed by AEA (2007) and Trochim (2009)*

Taxonomy Components	Component Definition	Modifications
AEA (2007)		
Evaluation Definition	How, if at all, is evaluation defined in an agency or in legislation? In such contexts, how is evaluation formally distinguished from or related to other functions such as program planning, monitoring, performance measurement or implementation?	None
Requirements of Evaluation	When are evaluations required? What programs or entities are required to have evaluations? How often are evaluations scheduled? What procedures are used to determine when or whether evaluation takes place?	Merged with “Management” component
Methods	What approaches or methods of evaluation are recommended or required by legislation or regulation, for what types of programs or initiatives?	Merged with “Design and Methods” component
Human Resources	What requirements exist for people who conduct evaluations? What types of training, experience or background are required?	Merged with “Participation and Roles” component
Evaluation Budgets	What are the standards for budgeting for evaluation work?	Merged with “Management” component
Evaluation Implementation	What types of evaluation implementation issues are guided by policies? For instance, when are internal versus external evaluations required and how are these defined?	Merged with “Management” and “Design and Methods” components
Evaluation Ethics	What are the policies for addressing ethical issues in evaluation?	None
Trochim (2009)		
Goals	The goals of evaluation in the organization or context	None
Participation	Who is involved in evaluations, when, and under what circumstances	Merged with “Roles” component & changed title to “Participation and Roles”
Capacity Building	The intention of the organization to build capacity to conduct evaluations	None
Management	Inform the allocation of resources (financial and otherwise) and oversight	None
Roles	the varying responsibilities of actors involved in conducting evaluations	Merged with “Participation” component & changed title to “Participation and Roles”
Process & Methods	procedures for conducting evaluations including question identification, sampling, measurement, design, and analysis	Changed title to “Design and Methods”
Use	How evaluation results will be communicated and acted upon	None
Meta-evaluation	How the quality of evaluations is assessed	None

Table 11. *Modified Taxonomy for Evaluation Policies*

Components	Sub-Component	Definition	Example from Evaluation Policy
Background		Context that helps situate the evaluation policy.	
	History	History of evaluation or of the evaluation policy.	"The Evolution of Monitoring and Evaluation" (Rockefeller, p. 4)
	Policy Motivation	Motivation and intent behind writing an evaluation policy.	"Our evaluation policy is a starting point for strengthening how we use evaluation within the foundation and with our partners." (Gates, p. 6)
	Evaluation Definition	How evaluation is defined.	"Evaluation is the systematic, objective assessment of an ongoing or completed intervention, project, policy, program, or partnership." (Gates, p. 1)
Goals and Purpose		Desired outcomes of evaluations or why the foundation chooses to evaluate.	"Our focus is on evaluation as a tool for learning. We also recognize the importance of accountability and improving performance." (Duke, p. 1)
Management		How different aspects of evaluation practice are managed.	
	Resources	Resources allotted for evaluation (e.g. staff time, budgets).	"In general, program officers will spend 5 to 20 percent of their time designing and managing evaluations…" (Hewlett, p. 7)
	Schedule	Any mention of timelines for evaluation practice.	"We also conduct an annual foundation-wide assessment to measure our progress against our organizational goals." (Irvine, p. 1)
	Necessity	Under what circumstances evaluations are required and preferred, and how this is determined.	"Evaluation is a high priority when program outcomes are difficult to observe, and knowledge is lacking about how best to achieve results…" (Gates, p. 2)
Participation and Roles		Who is involved, when, and how?	
	Grantee	Roles and responsibilities of grantees for evaluation.	"Core grant partners systematically provide output and outcome data…" (Skillman, p. 2)
	External Evaluator	Roles, responsibilities, and selection criteria for external evaluators.	"The ideal evaluator is strong technically, has subject matter expertise, is pragmatic, and communicates well…" (Hewlett, p. 18)
	Foundation Staff	Roles and responsibilities of foundation staff for evaluation. This includes all internal stakeholders, such as: evaluation staff, program officers, leadership, board.	"an initiative team partners with the Foundation's Monitoring and Evaluation (M&E) team to scan the existing evidence in a given field…" (Rockefeller, p. 1)

Table 11. *Continued*

Components	Sub-Component	Definition	Example from Evaluation Policy
Design and Methods		How evaluations are designed, and methodological decisions are made.	"We develop customized evaluation designs for each of MacArthur's programs based on the problem, opportunity, and approach to the work, as well as Foundation priorities around outcomes, impact, and learning." (MacArthur, p. 1)
	Types of Evaluations	Specifies different types of evaluations (e.g. outcome, impact).	"The DDCF employs three main evaluation strategies…" (Doris Duke, p. 1)
Use of Evaluations		How evaluation results are communicated and acted upon.	"The final report…should be no more than 50-60 pages and written in "plain English" in a user-friendly graphic format." (C.A. Wellness, p. 3)
	Design Phase	What ought to happen before an evaluation begins to ensure use of evaluation findings.	"Consider the intended use of the evaluation…" (Irvine, p. 1)
	Internal	How evaluation findings are shared or used internally.	"At all points in the strategy lifecycle, we create the space for reflective practice and learning from our monitoring and evaluation data…" (Packard, p. 4)
	External	How evaluation findings are shared with an external audience.	"When planning how to share results broadly, program staff should consult with the Foundation's Communications staff about the best approach." (Hewlett, p. 8)
Capacity Building		Any discussion of building evaluation capacity to conduct evaluations.	
	Internal Capacity	Capacity building guidance or resources for internal stakeholders (foundation employees and board).	"The foundation's central Strategy, Measurement and Evaluation team is responsible for… providing foundation staff and partners with training and skill-building support..." (Gates, p. 7)
	External Capacity	Capacity building guidance or resources for external stakeholders (grantees, affiliates, etc.).	"Building the data capacity of core partners is an essential component of the work…" (Skillman, p. 2)
Meta-evaluation		How quality of evaluations is assessed.	"Debrief every evaluation with the appropriate program staff…" (Hewlett, p. 9)
Avoid		Guidance on what should be avoided when conducting evaluations.	"External evaluations on their own don't drive decisions for the Packard Foundation." (Packard, p. 2)
Example		Example, anecdote, or quote.	"To get an idea of this approach, see this article..." (Wallace, p. 3)
Ethics		Discussion of ethical issues in evaluation.	"To avoid undue burdens on grantees and inappropriate use of resources, we will measure only what we will use, and use everything that we measure." (Duke, p. 1)

Frequency of Components

Once content analysis was completed, the frequency with which components appeared across all 12 evaluation policies was examined. This analysis offers insight into which components were most readily included, and those that were lacking, in the evaluation policies written by foundations. Table 12 lists the modified evaluation policy component and in which policies it was present. Five components appeared in all 12 evaluation policies: "Goals and Purpose"; "Management"; "Participation and Roles"; "Design and Methods"; and "Use". These components were included in the evaluation policy taxonomies proposed by AEA (2007) and Trochim (2009) but were slightly modified for this study. Their inclusion in all evaluation policies suggests that there is consensus that these components are essential to evaluation practice and therefore important aspects of evaluation policies.

The components least present in evaluation policies included: "Background" (6 policies); "Avoid" (6); "Ethics" (3); and "Meta-evaluation" (2). "Background" captures content related to the history of evaluation at the foundation, the motivation for the document, or the definition of evaluation. While some foundations chose to include it in the policy document, other foundations may have such background information documented in other forms (supplementary materials, videos, articles, or webpages). Additionally, given that this information is not likely to have a direct impact on evaluation practice, it may be thought of as an optional component by policy writers.

"Avoid"; "Ethics" and "Meta-evaluation" address evaluation practice more directly. Discouraging particular practices is an aspect of policy identified by Bullock et al. (1983), but not found in Trochim's or AEA's taxonomies. While it was added to the modified taxonomy presented in this study, given its infrequent use in evaluation policies it may be better as a sub-component. The “Ethics” component was proposed by AEA. During coding it was often difficult to discern which aspects of a policy addressed issues of evaluation ethics. There are two possible reasons for this. First, ethical considerations can be subtly communicated and embedded in the other components. For example, how stakeholders are, or are not, engaged may address issues of ethics. However, it is difficult to discern aspects of ethics in brief policy statements that don’t make them explicit. The second challenge with the “Ethics” component is that the definition proposed by AEA is too broad to guide content analysis. AEA defined ethics as any “discussion of ethical issues in evaluation”. The “Meta-evaluation” component’s omission from policies may be more of an oversight or a practice that foundations have not adopted. While “Meta-evaluation” is seen as an important aspect of evaluation practice by evaluation scholars, it doesn’t appear to be a common practice employed by foundations.

Table 12. *Frequency of Components Across 12 Evaluation Policies*

Component	Frequency
Background	6
Goals and Purpose	12
Management	12
Participation and Roles	12
Design and Methods	12
Use of Evaluations	12
Capacity Building	9
Meta-evaluation	2
Avoid	6
Example	7
Ethics	3

A Review of Evaluation Policies Across the Philanthropic Sector

With the discussion of policy components as a backdrop, the next section reviews findings from the analysis of extant evaluation policies in the philanthropic sector. Findings are presented in two ways: component comprehensiveness and policy completeness. Component comprehensiveness gauges the proportion of an evaluation policy dedicated to a particular component. The completeness measure looks at how many components are included in a particular policy. Table 13 presents the distribution of components across policies. The total word count for each policy and each component is shared, as well as what percentage of each policy is dedicated to a particular component. The column at the end of the table, titled "Completeness" shows how many of the 11 components were present in a foundation's policy. This measure is referred to as completeness. A discussion of component comprehensiveness and policy completeness follows. To help visualize how policies differed across foundations, Appendix C includes spider charts, also called radar charts, of the data represented in Table 13.

Component Comprehensiveness

To gauge the extent to which components are comprehensive, or fully describe the ways in which this component might be implemented, word count is used as a proxy. As a policy expands on general policy statements and offers guidance on operationalization, the word count is likely to also increase. While not a perfect measure, reviewing word count across policies and components revealed important insights about what aspects of evaluation practice are prioritized by foundations.

Component comprehensiveness differed across evaluation policies. Five foundations prioritized issues related to "Design and Methods": Gates (67 percent of their policy), Doris Duke (78 percent), Irvine (43 percent), MacArthur (94 percent), and Skillman (74 percent). Duke Endowment and Kansas Health prioritized detailing the goals and purposes for their evaluation practice (35 percent and 49 percent). Two foundations dedicated the majority of their policy to sharing examples or anecdotes that elucidate different components of their evaluation policy:

Rockefeller (27 percent) and Hewlett (30 percent). The remaining three foundations had different areas of focus. Wallace focused on “Participation and Roles” policies (28 percent), Packard prioritized “Use” (27 percent) and California Wellness emphasized “Ethics” (38 percent).

Of the 11 components reviewed, more text was dedicated to “Design and Methods” than any other component. Approximately one-third (29 percent) of all the text analyzed was about guidance on evaluation design and methodology. Less than five percent of all policy text related to the following components: “Avoid” (4 percent); “Capacity Building” (3 percent); “Ethics” (1 percent); and “Meta-evaluation” (< 1 percent).

Policy Completeness

Completeness is a measure of how many components are included in each evaluation policy. It offers insight into how comprehensive an evaluation policy is, and therefore also identifies components that may be lacking from a policy. Completeness, however, is not intended to communicate quality. It may be that some components are simply more relevant to one foundation than another.

Hewlett and Duke have the most complete policies, both having policies missing only one of the 11 components. Hewlett’s policy lacked explicit mention of ethics, although as mentioned earlier, “Ethics” was a difficult component to define and code. Any mention of ethics was found in just three policies and captured only one percent of total word count across all policies. Duke’s policy did not include the “Meta-evaluation” component. Only two policies included the “Meta-evaluation” component, which accounted for less than one percent (68 words) of all policies.

While the Hewlett policy is the longest at 9875 words, Duke’s policy is one of the shortest at 577 words. This suggests that there isn’t a linear relationship between number of components in an evaluation policy and its length. Two evaluation policies stipulated six or fewer components: MacArthur (6) and Doris Duke (5).

Table 13. *Overview of Components in Foundations' Evaluation Policies*

Foundation	Policy Word Count	Background		Goals and Purpose		Management		Participation and Roles		Design and Methods	
		#	%	#	%	#	%	#	%	#	%
Gates	3032	329	11%	387	13%	488	16%	798	26%	2026	67%
C.A. Wellness	436	0	0%	110	25%	97	22%	68	16%	44	10%
Packard	1390	44	3%	329	24%	362	26%	110	8%	159	11%
Doris Duke	372	0	0%	92	25%	102	27%	156	42%	292	78%
Duke	577	64	11%	202	35%	155	27%	87	15%	110	19%
Irvine	966	0	0%	190	20%	53	5%	121	13%	415	43%
MacArthur	540	0	0%	151	28%	112	21%	59	11%	508	94%
K.S. Health	621	159	26%	302	49%	39	6%	102	16%	36	6%
Rockefeller	1720	393	23%	178	10%	193	11%	281	16%	279	16%
Skillman	324	0	0%	83	26%	20	6%	111	34%	240	74%
Wallace	1198	0	0%	194	16%	34	3%	339	28%	149	12%
Hewlett	9875	1188	12%	413	4%	1181	12%	1623	16%	1780	18%
Total:	21051	2177	10%	2631	12%	2836	13%	3855	18%	6038	29%

Table 13. *Continued*

Foundation	Policy Word Count	Use		Capacity Building		Meta-evaluation		Avoid		Example		Ethics		Completeness
		#	%	#	%	#	%	#	%	#	%	#	%	#
Gates	3032	170	6%	120	4%	0	0%	280	9%	0	0%	0	0%	8
C.A. Wellness	436	61	14%	20	5%	0	0%	0	0%	0	0%	165	38%	7
Packard	1390	373	27%	84	6%	0	0%	12	1%	0	0%	0	0%	8
Doris Duke	372	78	21%	0	0%	0	0%	0	0%	0	0%	0	0%	5
Duke	577	100	17%	32	6%	0	0%	25	4%	69	12%	25	4%	10
Irvine	966	223	23%	0	0%	28	3%	0	0%	0	0%	43	4%	7
MacArthur	540	39	7%	0	0%	0	0%	0	0%	111	21%	0	0%	6
K.S. Health	621	32	5%	19	3%	0	0%	72	12%	22	4%	0	0%	9
Rockefeller	1720	330	19%	104	6%	0	0%	0	0%	463	27%	0	0%	8
Skillman	324	38	12%	38	12%	0	0%	17	5%	5	2%	0	0%	8
Wallace	1198	105	9%	86	7%	0	0%	0	0%	99	8%	0	0%	7
Hewlett	9875	992	10%	159	2%	40	0%	437	4%	2951	30%	0	0%	10
Total:	21051	2541	12%	662	3%	68	<1%	843	4%	3720	18%	233	1%	

The Content of Evaluation Policies in the Philanthropic Sector

This paper opened with a review of present research into the components of evaluation policy. After a discussion of methodology, a revised evaluation policy taxonomy was presented in the Findings section. Then key findings from the analysis of 12 evaluation policies developed by foundations were discussed. Particularly, how publicly available evaluation policies of US foundations compared in terms of component comprehensiveness and policy completeness were detailed.

The empirical examination of evaluation policies in the present chapter resulted in an updated evaluation policy taxonomy. Approximately 10 years ago, AEA (2007) and Trochim (2009) independently theorized which aspects of evaluation practice may be stipulated in evaluation policies. Following three rounds of content analysis of 12 policies, modifications were made to their taxonomies which resulted in a single taxonomy consisting of 11 primary components and 15 sub-components. Between the AEA (2007) and Trochim (2009) taxonomies there were three components addressing how stakeholders engage in evaluation practice and six components dedicated to aspects of managing and designing an evaluation. In the revised taxonomy, these were streamlined. There is now one component, “Participation and Roles”, that captures information related to roles and responsibilities of stakeholders, primarily foundation staff, grantees, and external evaluators. One component, “Management”, is dedicated to all aspects of managing evaluations and another, “Design and Methods”, for issues related to evaluation design, including methodology. In addition to merging existing components, three new components, “Background”, “Avoid”, and “Example” were added to the taxonomy.

Several key findings surfaced from analysis of 12 evaluation policies using the revised taxonomy. First, no policy contained all 11 primary components. Hewlett and Duke had the most complete policies, each policy only missing one component. Of the 11 components in the revised taxonomy, five appeared in every evaluation policy: “Goals and Purpose”; “Management”; “Participation and Roles”; “Design and Methods”; and “Use”. More specifically, 29 percent of all text (21,051 words) was devoted to “Design and Methods”, more text than any other component. Together, these five components accounted for 84 percent of text across the evaluation policies. Their prominence across policies suggests that these are essential aspects of evaluation practice for foundations.

Conversely, the following four components cumulatively accounted for less than 9% of all policy text: “Avoid” (4 percent); “Capacity Building” (3 percent); “Ethics” (1 percent); and “Meta-evaluation” (< 1 percent). Given that policies are primarily positive documents, detailing what should be done, it is not surprising to see so little text devoted to what is to be avoided. In future research, the “avoid” component may be best conceptualized as a sub-component rather than a primary component. It is concerning that policies rarely address issues related to evaluation ethics, capacity building and meta-evaluation. Ethics is a far reaching and important consideration in evaluation, yet it was only addressed by three policies and accounted for only 1 percent of total word count. Its infrequency may be attributed to how it is defined or conceptualized in the policy. The definition for evaluation ethics in AEA’s taxonomy is very broad which made it difficult to use as a primary code during content analysis. Second, issues of ethics may be imbedded, and thus difficult to differentiate, in various aspects of evaluation practice. Therefore, “Ethics” may be better conceptualized as a sub-component situated under various components, rather than a primary component of the taxonomy.

In the Netherlands, federal evaluation policies prioritized issues related to “capacity building”, and “meta-evaluation” (Leeuw, 2009). However, they were only briefly, if at all, discussed in evaluation policies guiding practice in the philanthropic sector. To date, it is unclear why foundations are not including these components. It may be that these are not aspects of their evaluation practice. Alternatively, they may be part of their evaluation practice, but were just not stipulated in the policy document reviewed in this study.

Future Research: Looking Ahead to Paper 3

The findings discussed in the present chapter have a number of potential implications for future research on evaluation policies. In particular, three areas of future research are addressed in the third and final paper of the present dissertation. Examination of how policies differ across foundations revealed that policies range in length, comprehensiveness and completeness. However, it is unclear if these differences are intentional and how they reflect an organization's context and priorities. There is also a need to understand what impact these differences have on evaluation practice. Lastly, it remains unknown if, or to what extent, evaluation theory and the taxonomies discussed in this paper are informing the substantive content of policies.

Discussions with evaluation directors at these 12 foundations may provide valuable insights into how or why evaluation policies are developed and revised, and if the content of policies are informed by the information and knowledge imbedded in evaluation theories. Furthermore, evaluation directors may offer insight into how these policies influence the foundation's evaluation practice.

Paper 3.
The Intent and Influence of Evaluation Policies in the Philanthropic Sector and What Informs Their Development

Abstract

Evaluation policies are being written across the philanthropic sector. However, it is unclear why these policies are being developed and what impact, if any, they have on evaluation practice. Furthermore, it is uncertain if the content of these policies is informed by the information and knowledge imbedded in evaluation theories. Through interviews with 10 evaluation directors at foundations with evaluation policies, this paper first identifies the intent and influence of evaluation policies in the philanthropic sector. Then, how evaluation theory informs these policies is discussed. The paper concludes with suggestions for future research on evaluation policy.

Preface

Evaluation provides foundation leaders with the information they need to learn from and improve their grantmaking. In 2014, 154 of the top independent foundations in the United States spent a cumulative $243.5 million on evaluation-related charitable grants (Kinarsky, paper 1). Of the top commissioners of evaluation in 2013 and 2014, 12 foundations had publicly available evaluation policies. These policies are thought to "help set the content, characteristics, and context of evaluation itself" yet little is known about the content or impact of these policies (Mark, Cooksy, & Trochim, 2009, p. 5). Furthermore, it is unclear why policies are being written and what role, if any, evaluation theory has in their development.

To advance research on evaluation policies, the second paper of the present dissertation examined similarities and differences across policies in the philanthropic sector. This analysis revealed that policies range in length, comprehensiveness and completeness. While some policies were short (only 324 words) others were much longer (9875 words), yet none contained all of the components delineated in evaluation policy taxonomies. All 12 policies were found to include guidance on five of the 11 components: “Goals and Purpose”; “Management”; “Participation and Roles”; “Design and Methods”; and “Use”. Conversely, very few policies had any content related to the following components: “Avoid”; “Capacity Building”; “Ethics”; and “Meta-evaluation” (Kinarsky, paper 2).

These findings suggest that foundations are developing evaluation policies that are unique to how evaluation is practiced at the foundation. However, it remains unclear why these policies are being developed and what impact, if any, they have on evaluation practice. Furthermore, it is uncertain if evaluation theory is informing the content of these policies. This paper is the first to address these gaps through interviews with evaluation directors at foundations with evaluation policies. The following research questions guide the present inquiry:

1. According to evaluation directors, what are the primary goals of their foundation’s evaluation policy?
2. According to evaluation directors, in what ways do the foundation’s formal and informal evaluation policies guide evaluation practice?
3. According to evaluation directors, to what extent does evaluation theory inform the evaluation policies of US foundations?

Introduction – Research on the intent and influence of policy and the role of theory

Understanding the Intent of Policy

Policies typically arise when problems can benefit from standardized actions, rather than allowing staff to individually choose next steps (Cochran & Malone, 2010; Munger, 2000;

Bardach & Patashnik, 2016). According to Munger (2000), policy formation tends to follow a three-stage process. First, a problem is identified. Once the problem is defined, a discussion ensues as to whether it should be resolved collectively or individually. If it is determined that the resolution is collective, then the policy gets written either by the stakeholders or with the guidance of experts.

Policy research stresses that prior to exploring the impact of a policy, it is important to understand what it intends to accomplish (Bullock, Anderson, & Brady 1983; Bardach & Patashnik, 2016). Once motivations for writing a policy are surfaced, that policy's impact can be appropriately gauged. In this way, evaluation of policies is similar to the evaluation of programs. Before attempting to evaluate the impact of a program, an evaluator seeks to understand why this program was established and what it intends to accomplish. Once the program's goals are identified, the evaluator can begin to explore connections between goals, practices, and outcomes.

There are notable challenges that may motivate foundations to write evaluation policies. Evaluation is used in the philanthropic sector to generate evidence for accountability, learning, and knowledge sharing (Kramer & Bickel, 2004). While all three functions of evaluation in philanthropy are important, they are often difficult to balance:

> "Accountability for use of funds is the most common form of evaluation but contributes little to the learning of the foundation or the broader field. Learning within the foundation is more valuable though less common and knowledge sharing with the field is critically important, though least common of all" (p. 53).

The challenge of using evaluation findings for accountability, learning and knowledge sharing may be because evaluation practice in the philanthropic sector tends to be "ad hoc, not aligned to strategy, underfunded, and underused" (Preskill & Mack, 2013, p. 4). There is growing consensus amongst researchers and practitioners that foundations need to adopt a more strategic approach to evaluation (Preskill & Mack, 2013; Kramer & Bickel, 2004; Buteau & Glickman, 2018).

Evaluation policies can provide foundations with conceptual clarity in their evaluation practice. In the opening chapter, evaluation policy was defined as "any written document that details an organization's rules and principles for evaluation practice, regardless of whether implementation is encouraged or enforced." In other words, they articulate an organization's "mental model" for evaluation, the ideas, theories, and assumptions that inform evaluation practice within an organization:

> "None of us can carry an organization in our minds - or a family, or a community. What we carry in our heads are images, assumptions and stories… Our "mental models" determine not only how we make sense of the world, but how we take action" (Senge, 2006, p. 164).

Policies make mental models explicit. Doing so can help stakeholders "define a method of grappling with the situation and ruling out a vast array of possible actions" (Rumelt, 2017, p. 84). Furthermore, they offer coherence to diverse activities, concentrate an organization's efforts, reduce complexity, and lessen ambiguity (Senge, 2006; Rumelt, 2017). The present paper begins by exploring the goals and objectives motivating the development of evaluation policies in the philanthropic sector.

What Informs Policy

While examining the intent of evaluation policies, the present paper also seeks to illuminate the extent to which the content of these policies are informed by evaluation theory.

Ideally, policy formulation is informed by and consistent with relevant theory. However, it is not uncommon for policy to deviate from theory and yield to more immediate concerns or needs (Cochran & Malone, 2010). Another concern is that there may not be adequate research available to inform policy at the time it is being crafted. Quality research takes a long time to conduct and corroborate, which may prompt policymakers to move forward without consulting theory.

Evaluation theory, or the evaluation approaches and models that prescribe why and how evaluation should be conducted, is often informed by evaluation practice (Alkin, 2013). Shadish, Cook, and Levitan (1991) state that evaluation theory has “evolved in response to experience in doing evaluation” (p. 33) which involves “the tactics and strategies evaluators follow in their professional work, especially given the constraints they face” (p. 32). While we know that theory benefits from the study of practice, we are still in the early stages of understanding how theory does, or can, improve practice.

Despite the challenges of connecting theory to practice, many argue that as the field grows so does the need to establish a coherent link between them (Shadish et al., 1991; Christie, 2003; Dillman, 2013; Schwandt, 2013). Chelimsky (2012) suggests that, “best practices’ cannot be developed separately because evaluation theory and practice are interdependent: each one learns from the other and, in that learning process, both are inspired to stretch, to bend a little, and to grow” (p. 91). In fact, theory is essential to evaluation practice. Shadish and Epstein (1987) argued that without theory “evaluators are deprived of knowledge of the options they have available to use in their practice” and that for evaluation to be a professional practice it must “be concerned with its own level of knowledge of its core literature” (p. 586).

Understanding how and if evaluation theory is being intentionally integrated into evaluation policies is critical. According to Christie and Lemire (2019) policies, at their best, translate evaluation theory into practice:

> “The position we hold is that evaluation policies serve as an important connector between evaluation theory and evaluation practice. More specifically, evaluation policies serve as a transformative mechanism for translating and perhaps more importantly situating evaluation theory in the organizational, political, and cultural context of an organization. In order to do this, in order to bridge the theory–practice divide, we argue that the theory integration has to be purposeful and explicit, emphasizing adaptation over adoption” (p. 2).

Evaluation theory tends to be broad. It is generally intended to be applicable across a range of contexts. Evaluation policies, ideally, apply these theories within the context of the foundation, guiding its stakeholders. However, little is known about if and how evaluation theory actually informs the formulation of evaluation policies.

Understanding the relationship between evaluation theory and evaluation policies is important for several reasons. First, it may explain some of the differences across evaluation policies in the sector discussed in paper 2. Second, it can identify challenges foundation leaders experience adapting evaluation theory to their organizational context. Lastly, it can surface gaps in current evaluation theory that need to be addressed to better support evaluation practice in the philanthropic sector.

The Influence of Evaluation Policies on Evaluation Practice

Additionally, the present paper begins to identify the influence evaluation policies have on evaluation practice at foundations. Emergent theory suggests that there is an important relationship between evaluation policy and practice. Trochim (2009) posits that written evaluation policies are important because they encourage consistency of practice, make public internal opinions on evaluation, encourage learning, and, lastly, allow for adjustments to and innovation in evaluation practice. Policy not only impacts how evaluation is conducted but also influences the quality of information generated. Mark et al. (2009) discuss the importance of policy quality by asserting that:

> "Evaluation policy is a critical concern for evaluation practice, in part because it shapes evaluation practice thereby both *enabling and constraining the potential contribution evaluation can make*… At the extreme, evaluation policy may not only enable some forms of contribution and constrain others; bad evaluation policy can have serious negative consequences" (p. 3, italics added for emphasis).

The scope and quality of an evaluation policy can have a constructive or adverse impact on evaluation practice.

As noted in the opening chapter, evaluation policies authored by foundations likely guide evaluation practice across the civic sector, which impacts the interest in, demand for, and practices of evaluation. They ultimately influence the quality of information generated about the effectiveness of interventions being developed and executed by nonprofit organizations and foundations. While some evaluation policies may help streamline and improve evaluation practice, others may do more harm than good. For example, if a foundation's policy prioritizes the use of randomized control trials (RCT) and de-incentivizes qualitative inquiry, then initiatives that are more targeted or experimental in nature may be forced to participate in an RCT before they are ready or not get evaluated at all. As the demand for evaluation to inform program improvements and demonstrate impact increases in the civic sector, so does the need to establish policies that support evaluation practice and advance the field, rather than hold it back (Liket, Rey-Garcia, & Maas, 2014; Behrens & Kelly, 2008).

Several empirical studies have examined the impact of evaluation policy on practice. This research offers critical insights into the importance of evaluation policy, its impact on evaluation practice and suggestions for future areas of inquiry.

First, two researchers traced the impact federal evaluation policies have had on national evaluation practice. Stern (2009) reviewed the evolution of evaluation policy in the executive branch of the European Union and concluded that as evaluation policy shifted so did the purpose or goals of evaluation practice. Evaluation was initially used to inform local project management and program delivery. However, as evaluation policy shifted towards a more national focus, so did practice. Evaluation is no longer being used "to make transparent to the citizens what they are getting for their money" (Bosch quoted in Stern, 2009, p. 70). It is now a tool for appraising policy proposals from an "economic, social and environmental perspective" (p. 74). In a review of how Dutch evaluation policy has evolved since the seventies, Leeuw (2009) concluded that evaluation policies contributed to an increase in evaluation demand and an expansion of local evaluation units. National evaluation policy was broad enough to allow local evaluators to make technical and methodological decisions independently. Both studies demonstrate the significant impact evaluation policy can have on evaluation practice.

A study conducted by Christie and Fierro (2012) inquired into how evaluators interpret and enact evaluation policy. Through their case study on the implementation of a California state evaluation policy, they identified an "inherent tension" between written evaluation policy and how evaluators interpret that policy (p. 71). An evaluator's outlook on what "constitutes credible evidence in evaluation" in addition to challenging logistics faced during an evaluation resulted in evaluations that deviated from what was intended by policymakers.

Only one study was identified that has explored the connection between policy and practice in the philanthropic sector. Dillman and Christie (2016) appraised how evaluation policies of the Robert Wood Johnson Foundation were applied across three initiatives. They determined that while evaluation policy standardized some aspects of evaluation practice, there were also informal policies and guidelines that informed staff decisions and actions. Because they were informal, their application was inconsistent and unpredictable.

Together, these studies demonstrate that evaluation policies influence evaluation practice in important ways. Policies can determine the demand for evaluations as well as their strategic focus. At foundations, policies can strengthen evaluation practice by improving alignment. However, policies tend to be interpreted and applied differently by stakeholders. More inquiry is needed to better understand the relationship between evaluation policy and evaluation practice.

Methodology

The present paper extends the work conducted in paper one and two. Once policies were cataloged and analyzed, evaluation directors at each of the 12 foundations were contacted to schedule a one-hour phone interview. These interviews helped further understanding of the intent and influence of evaluation policies, along with what informed their development, across the philanthropic sector. The remainder of this section begins with an introduction to the participants, followed by a discussion of the instruments and data collection procedures. Lastly, the analytic procedures are outlined.

Participants

Evaluation directors from the 12 foundations discussed in paper 1 were invited to participate in the present research. To identify contact information of these evaluation directors, a review of foundation websites was conducted. Individuals were eligible for inclusion if they were listed on the foundation's website with a title indicating that they were in charge of evaluation at the foundation. Actual titles varied across foundations, but, for simplicity, in this study all participants are referred to as an "evaluation director".

A request to participate in an interview was emailed to evaluation directors between December 2017 and February 2018 (see appendix D). If there was no response, up to two follow up emails or phone calls were made. Ultimately, two directors declined to participate in the study. One director chose to not be interviewed because the foundation no longer uses its evaluation policy and they were not at the foundation when the policy was developed. The second evaluation director who declined did not disclose their reasoning.

Interviews with evaluation directors at 10 foundations were conducted between February and May 2018. Interview participants bring years of academic and professional experience with evaluation to their current roles. Four directors have graduate degrees in fields related to evaluation. Prior to their current role, four worked at large research organization or university; three led evaluation efforts at a non-profit; and three worked at another foundation. Six directors were and four directors were not employed at the foundation when the policy document was written. This background information is summarized in Table 14. Each foundation has been assigned an abbreviated name which is also listed in Table 14 and used for the remainder of the paper.

Table 14. *Background Information on the Foundation, their Evaluation Policy and Interview Participants*

Foundation	Abbreviation	Policy Document Title	Year Policy Written	Year Participant Joined
Bill & Melinda Gates Foundation	Gates	Evaluation Policy	2008	2016
David & Lucile Packard Foundation	Packard	Guiding Principles and Practices for Monitoring, Evaluation, and Learning	2004	2013
Doris Duke Charitable Foundation	Doris Duke	Grant-Making Process: Evaluation	early 2000's	2000
Duke Endowment	Duke	Guiding Principles: Learning and Evaluation	2010	2009
John D. & Catherine T. MacArthur Foundation	MacArthur	Our Approach to Evaluation	2010	2012
Kansas Health Foundation	K.S. Health	Evaluation/Learning: Our Approach	in process	2015
Rockefeller Foundation	Rockefeller	Our Strategy: Monitoring and Evaluation	2017	2015
Skillman Foundation	Skillman	Evaluation Philosophy and Approach	2007	2005
Wallace Foundation	Wallace	How We Work: The Wallace Approach	2012	1996
William & Flora Hewlett Foundation	Hewlett	Evaluation Principles and Practices	2012	2013

Interview Procedures

Semi-structured phone interviews were conducted with evaluation directors from 10 foundations. A master interview protocol was designed based on the study's research questions. Interview questions were organized under four domain areas informed by the study's guiding research questions:

- Motivation for Establishing an Evaluation Policy
- Role of Evaluation Theory in Developing Evaluation Policy
- Goals and Purpose of Evaluation Policy
- Role of Evaluation Policy in Guiding Evaluation Practice

Protocols were modified for each foundation based on findings from the analysis of evaluation policies discussed in paper 2. For example, if a policy made explicit mention of a particular theorist or evaluation theory in their policy, this was noted in the protocol and discussed during the interview. An example of an interview protocol can be found in Appendix E.

Once interview protocols were drafted, two pilot interviews were conducted to understand how questions work under realistic conditions. The first pilot was with a graduate student with expertise in program evaluation. The second pilot was conducted over the phone with an evaluation staff member at one of the foundations in this study. The pilot interviews helped ensure that the protocol used terms that were familiar and consistently understood by the target population. Following the second pilot, the introduction to the interview was modified to make explicit that this interview is about the explicit policy document. Also, questions were revised to produce an interview protocol that was brief yet detailed enough to answer the research questions.

A semi-structured interviewing approach was employed to allow deviation from the protocol to "respond to the situation at hand" while still keeping consistency across interviews (Merriam, 2009, p. 90). In advance of every interview, interview protocols were re-read, and occasionally adjusted, based on another review of the foundation's evaluation policy, as well as other relevant information from its website (such as interviews, publications, or blog posts). At least two days before the interview, an external version of the interview protocol was shared with participants (similar to that in Appendix E). This was done to help participants understand the focus of the interview and give them an opportunity to review the evaluation policy.

Interviews were conducted over the phone and generally lasted sixty minutes. Participation was voluntary, and participants understood that they could refuse to answer any questions and terminate the interview at any time. With participant permission, interviews were recorded. They were then transcribed verbatim within two weeks of completion.

At the close of each interview, participants were invited to submit additional documentation related to the foundation's evaluation practice and policy. Documents from eight foundations were cataloged in the study database and used to better understand how evaluation is practiced at the foundation. Analytic memos were written after each interview to reflect on emerging findings, future directions and unanswered questions (Saldaña, 2013).

Analysis of Interview Transcripts

Data analysis occurred "simultaneously with data collection" and focused on answering the research questions (Merriam, 2009, p. 171). Between interviews, preliminary findings helped narrow the focus of inquiry. Once interviews were completed and transcribed, data was imported into and analyzed using MaxQDA. Before formal coding began, transcripts were reviewed for emerging patterns and key phrases.

For the first cycle of coding, a priori codes were generated based on the guiding research questions (motivation for establishing an evaluation policy; role of evaluation theory; and role of evaluation policy in guiding evaluation practice) as well as findings from the analysis of evaluation policies conducted in paper 2 (Saldaña, 2013). Then, the data was coded using initial coding to closely examine and compare data across interviews (Saldaña, 2013). Codes included "component: capacity building"; "evaluation theorist"; "affinity group". During second level coding, pattern coding was utilized to "attribute meaning" to the emerging findings (Saldaña, 2013, p. 209). During this coding cycle, major themes surfaced, such as internal and external motivations for policy development. Analytic memos of emergent findings and initial reflections were documented in MaxQDA during coding and reviewed after the second cycle of coding was completed.

Once formal coding was completed, summary grids were generated in MaxQDA. Every coded segment was reviewed, and short summary statements were written. These segments were grouped into categories informed by the research questions and studied for thematic connections (Seidman, 2013). Visual maps of emerging themes were created to facilitate the transition from analysis to writing. Analysis continued during the writing process to help further distill themes.

To increase internal validity, member checks with participants were conducted to ensure there were no misunderstandings or inaccuracies (Merriam, 2009). Once a draft of the paper was completed, a document for each foundation was generated. The document included background information on the participant and policy, as well as interview quotes and my interpretations. Participants were invited to review the materials and suggest modifications or clarifications. Member checks were completed in May 2019.

Limitations

Careful attention during interviewing, coding and interpretation was given to ensure "trustworthiness" of findings (Lincoln & Guba quoted in Seidman, 2013, p. 27). First, to ensure that the experiences of participants were not misrepresented or misinterpreted, member checks were conducted. Quotes and interpretations used in this paper were shared with participants from whom they were derived. Participants were given two weeks to review. Suggested revisions were reviewed and addressed in the final manuscript. Next, analytic memos were written regularly during interviews and analysis to ensure that emergent patterns, challenges, questions and areas for future research were tracked. Lastly, the findings are presented with quotes and rich descriptions to help readers "determine the extent to which their situations match the research context, and, hence whether findings can be transferred" (Merriam, 2009, p. 229).

Findings

At present, there is limited research into why philanthropic foundations are developing and sharing evaluation policies. Munger (2000) posits that the development of a policy is only warranted when a problem emerges that needs to be resolved collectively. Therefore, before trying to understand in what ways, if any, these policies are influencing evaluation practice at foundations, the motivation for developing evaluation policies is discussed. Table 15 summarizes the findings related to the intent and influence of policies. I first expand on internal intents (1-3) motivating the development of evaluation policies before discussing the influence (4-7) these policies have had within the foundation. I then turn my attention to external intent (8-9) and influence (10-11). The findings section concludes with a discussion of what informed the development of evaluation policies.

Table 15. *Intent for & Influence of Developing an Evaluation Policy*

	Intent	Influence
Internal	1. Communicate evaluation philosophy and practices to an internal audience. 2. Identify when evaluation practice should remain consistent. 3. Identify when evaluations should vary and how.	4. Evaluation philosophy is consistently understood, and practice is more consistent. 5. Staff know when and how to vary evaluation practice. 6. Evaluation staff have more clarity on how to guide organization-wide evaluation practice. 7. Process of developing a policy is beneficial.
External	8. Explain evaluation requirements & expectations to grantees. 9. Communicate evaluation philosophy and practices to external stakeholders.	10. Current & prospective grantees better understand foundation's evaluation expectations and practice. 11. Foundations are benefiting from reviewing and reflecting on each other's policies.

Internal Intent (1-3)

The most widely stated intent for the development and adoption of evaluation policies is the desire to articulate a shared vision for how evaluation should be practiced at the foundation.

All foundations stated, during the interview or in their policy document, that the policy's intent was to achieve better alignment around the purpose and practice of evaluation. The need for a policy that creates a shared mental model for evaluation practice emerged for two reasons. First, evaluation design, execution and use are dispersed across stakeholders such as the board, president, and program officers. Second, of the 10 foundations only one had a central evaluation department that was responsible for executing evaluation and research across the foundation. Two foundations have evaluation staff embedded with program teams and seven foundations have central evaluation departments with minimal staff or financial resources. Evaluation tends to be decentralized so decision making around evaluation is dispersed. These challenges prompted the development of an organizational policy that had guidance on how to perform evaluation at the foundation. These three intents are discussed in greater detail below:

1. Communicate evaluation philosophy and practices to an internal audience

A foundation's mission, strategy and culture are heavily influenced by their founder(s), history, current leadership, and internal expertise. This context shapes how a foundation executes its grant making, as well as its approach to evaluation. Therefore, foundations are writing evaluation policies that aim to connect the foundation's distinct culture and values to how they execute and use evaluation. As one evaluation director stated:

> "Drafting guiding evaluation principles is important because it sets the tone. The principles have to match your culture at the foundation and since every foundation has their own culture, I think being explicit about your assumptions and values around evaluation is important." – Packard Foundation

Evaluation policies aim to make explicit the ways in which a foundation's culture and values impact the practice of evaluation.

In fact, this connection between organizational context and evaluation practice is often the first thing to be probed when conceptualizing an evaluation policy. Internal stakeholders examine how the values and culture of the foundation inform their evaluation practice:

> "In the beginning, we sat down to discuss what kind of evaluation practice represented the values we had as a foundation." – MacArthur Foundation

Once connections between organizational values and evaluation are established other considerations are discussed. For example, MacArthur sought to identify and discuss issues that staff across program areas were experiencing in their evaluation practice. In developing their policy, MacArthur outlined the correct way, for the foundation, of navigating these challenges or decisions. At the Duke Endowment, leaders wanted to deter staff from thinking of evaluation narrowly as an audit function or method for measuring impact. They sought to reframe evaluation as a practice that can help grant makers increase their effectiveness, potentially leading to greater impact in the areas they fund.

> "… I wanted to broaden the conversation and make sure people knew that evaluation was there to serve their effectiveness. It should be considered as a service to better grant making, not as an audit function and it can be useful in ways other than just measuring impact. When people are not familiar with the field, they tend to think of it in very narrow terms." – Duke Endowment

In addition to aligning evaluation practice with the foundation's unique culture and values, evaluation policies also aim to position evaluation as a vital function of effective grant making. Beyond articulating how evaluation is practiced, evaluation policies also seek to shape how evaluation is thought about at the foundation. These policies are reflective of an organization's beliefs regarding the importance of evaluation, when and how it should be conducted, and who should be involved.

At foundations with centralized evaluation staff, program officers are typically the ones funding and leading evaluation efforts. Evaluation staff tend to be available for program staff as technical advisors or internal consultants. At these foundations, evaluation policies seek to educate internal stakeholders and set common parameters for evaluation practice. For example, at Kansas Health Foundation, whose evaluation policy is still under development, the evaluation policy is a tool to communicate what evaluation is and how it should be used to support the work of the foundation:

> “The [evaluation policy] is a chance for me to write down this is our work; this is why we do it; and this is how we do it. Otherwise there is a lack of understanding and assumptions are made. That is not conducive to a learning environment. It is important that all of us start on the same page about what evaluation is and what it is not. The word “evaluation” alone, people have different interpretations. It helps to have a shared understanding. Then it can help folks understand how this work can and cannot meet their needs.” - Kansas Health Foundation

In other words, if evaluation is to support a learning environment at the foundation, then first “internally, we need to have the same mental model” (K.S. Health). Policies seek to make explicit the mental model guiding evaluation practice.

2. Identify when evaluation practice should remain consistent

As noted above, a foundation's responsibility for evaluation is frequently dispersed across the organization. In most of the foundations in this study, program officers situated in different program areas of the foundation are responsible for commissioning, designing, and overseeing evaluations. To support the varied and ongoing evaluation work occurring across the foundation, evaluation policies aim to identify cross-cutting functions of evaluation. They outline norms, expectations, and best practices to encourage some consistency in evaluation practice. By doing so, policies seek to provide stakeholders with enough direction to make decisions that align with organizational expectations and goals.

At Hewlett, program officers are charged with commissioning and using evaluations. Evaluation staff serve principally as internal consultants and technical advisors. Therefore, their policy aims to communicate why and how evaluation is conducted at Hewlett in hopes that internal stakeholders can commission, conduct, and use evaluations in accordance with internal expectations. The policy communicates Hewlett's evaluation philosophy and suggests actions in an effort to help simplify decision-making:

> "The purpose of this document is to advance the Foundation's existing work so that our evaluation practices become more consistent across the organization. We hope to create more common understanding of our philosophy, purpose, and expectations regarding evaluation as well as clarify staff roles and available support. With more consistency and shared understanding, we expect less wheel re-creation across program areas, greater learning from each other's efforts, and faster progress in designing meaningful evaluations and applying the results" (Hewlett policy, p. 3).

The intent to detail internal processes and how they vary may explain why Hewlett has the longest evaluation policy of all 12 reviewed in this study. About two-thirds of the 30-page document is dedicated to specifying evaluation practice (Kinarsky, paper 2). The policy begins by outlining roles and responsibilities of foundation staff, and then guides them through how to plan, implement and, ultimately, use evaluations. Examples and anecdotes are integrated throughout the document. They illustrate how a suggested practice has been performed by other programs or propose questions that could be asked when hiring an external evaluator.

By surfacing their approach to evaluation and specifying cross-cutting functions, foundations provide a shared language for discussing evaluation. For example, at the MacArthur Foundation, where evaluation staff are imbedded with program teams, the evaluation policy intends to identify and detail aspects of evaluation practice that should remain consistent across the foundation at the enterprise level:

> "We have standards that sit outside of each individual program area and are developed at the enterprise level. They outline expectations for evaluation that allow for consistency at an enterprise level and flexibility at a program level." – MacArthur Foundation

The evaluation policy at MacArthur connects a decentralized evaluation team around a shared set of values and standards for evaluation. The policy sets a broad vision for how evaluation should be practiced, helping stakeholders understand the foundation's evaluation approach.

> "I felt having an articulated approach and guidance were important so that our evaluation practice didn't change in response to every shift in the internal or external environment. It was critical to have evaluation as a deeply held value. We're an established, institutionalized, and operationalized entity that exists within the foundation that has all the pieces in place that a business would have to service the programs in the way that we think we need to be servicing them." – MacArthur Foundation

In addition to standardizing evaluation across program areas, MacArthur also views the evaluation policy as a way of standardizing practice over time. Ideally, by developing an evaluation policy, there won't be a need to make major alterations to evaluation practice as inevitable leadership or strategy changes occur.

3. Identify when evaluations should vary and how

In addition to specifying how evaluation practice should remain consistent across the foundation, evaluation policies also intend to outline when and how evaluation needs to adjust to the local context. The Gates policy explicitly states that the intent of their policy is to help stakeholders adapt their evaluation practice to their context:

> "Because the foundation supports a diverse range of partners and projects, it is necessary to have a clear organizational understanding of how evaluation should vary to best inform decision-making across each of these areas" (p. 1).

Given that Gates funds global initiatives ranging from malaria prevention to improving student outcomes, their policy seeks to give stakeholders a clear understanding of when and how to vary evaluations. In fact, most of the foundations interviewed have diverse funding priorities. To have an evaluation practice that generates meaningful findings, evaluation policies need to specify when and how practice should vary:

> "We don't have a standard approach to evaluation that is foundation-wide. The reason for that is our program areas are disparate. They're each trying to achieve very different goals so one program's measure of success is meaningless in another context." – Doris Duke Charitable Foundation

To accommodate varied programmatic goals and approaches, it is important to delineate between aspects of evaluation that need to be held consistent across an organization and those that should be varied. Directors found that once cross-cutting components of evaluation practice were identified, it was then necessary to determine how context can and should amend evaluation practice:

> "We had to think through when we vary. Are we varying to customize something that meets the needs, context and environment that we're working in? Or are we varying something that's a principle? We needed to be really clear, so we had something to fall back on." – MacArthur Foundation

The desire to encourage adaptation to the local context was not only an intent of the evaluation policy written at Rockefeller; it also motivated policy development and informed the policy design. Their new evaluation policy adapted in November 2017, is a set of principles seeking to communicate a vision for evaluation practice that encourages adaptation:

> "We don't want to be tied to one process because we understand that every new programmatic work will probably look a little bit different in terms of how we measure, monitor and evaluate…So rather than having a policy that is very prescriptive, we've developed a set of principles that we're holding ourselves accountable to, but that might drive the way this work ends up differently for different types of programs." – Rockefeller Foundation

Foundations tend to have disparate funding priorities that change over time. This requires the foundation to develop broad expertise across sectors. It also impacts how foundations evaluate. To support adaptation, evaluation policies aim to identify aspects of evaluation practice that need to adapt to the context and purpose of each evaluation.

Internal Influence (4-7)

Core to an evaluation policy is the relationship between an entity's evaluation policies and their evaluation practice. Extant research has begun to detail how policies support or constrain evaluation practice across a range of sectors. This section summarizes how policies are influencing evaluation practice, according to evaluation directors interviewed for this study.

Four areas of internal influence were identified. First, internal stakeholders better understand why and how evaluation is practiced. Additionally, staff are now better at adapting evaluations to their unique context. Policy has helped evaluation directors and their staff better understand and execute their duties. Finally, policy formulation was beneficial for surfacing implicit understandings of evaluation and delineating how evaluation is practiced at the foundation.

4. Evaluation philosophy is consistently understood, and practice is more consistent

As noted in the section on internal intent, the most commonly stated motivation for developing an evaluation policy was to identify cross-cutting functions of evaluation. Foundations want to make clear how evaluation is conducted at the foundation, in particular specifying aspects that need to remain consistent and those that should vary depending on context. Establishing an evaluation policy has created more clarity and cohesion for foundations:

> "It's led to complete consistency. There is something we are all driving towards that's a certain threshold of rigor, a certain threshold of the types of things or the range of things that we're doing in each space." – MacArthur Foundation

Policies have unified internal stakeholders in why and how evaluation is to be practiced.

Policies have clarified evaluation's importance and purpose within philanthropic organizations. At the Duke Endowment, according to the evaluation director, their policy has helped bring more clarity to their evaluation practice by making explicit why it is an important practice:

> “…there are all kinds of mismatched goals at different levels of the organization and those all connect to or impact what you can do with evaluation. Having them down on paper certainly helps. It helps to have something of a shared understanding of what the purposes are for evaluation and by extension the purpose of a philanthropic organization. We're here to have a social benefit and evaluation can help drive that very basic point home when it's not always so obvious." – Duke Endowment

Their policy makes clear that evaluation is a core function of foundations that helps advance their mission to improve the social good.

In addition to being more visionary, foundations also compel very specific evaluation practices that are “non-negotiable” through their policies. At Hewlett, stakeholders are found less likely to question decisions if they are prescribed in the policy. For example, the policy encourages stakeholders to widely share their evaluation findings:

> “Our intention is to share evaluation results—both the successes and failures— so that others may learn from them” (Hewlett policy, p. 21).

Therefore, when discussions transpire about how an evaluation report should be managed, stakeholders reference the evaluation policy. According to the evaluation director, it has helped make sharing a fixture of their evaluation practice:

> "It's amazing. Being able to point to the policy and say we have a principle of sharing that says we share what we're learning. People don't question it." – Hewlett Foundation

Having a policy minimizes decision-making and increases efficiency. The same conversation does not need to occur across multiple program areas or evaluations. Key features of evaluation practice have been determined and detailed allowing stakeholders to focus their attention on other, often more context-dependent, concerns.

5. Staff know when and how to vary evaluation practice

Another important aspect of generating a more consistent evaluation practice is specifying whether it can and should vary. Evaluation policies support internal stakeholders with tailoring their evaluations to the local context without compromising rigor. As stated before, foundations typically have disparate funding priorities. The same foundation may have dozens of evaluations occurring at the same time. These evaluations may be in very different program areas – such as education or climate change – and have different purposes – to determine the success of a six-year strategy or collect baseline data for a new initiative.

The design and methodology aspect of evaluation practice is prioritized in evaluation policies written by foundations. The "Design and Methods" component was the only component present in every evaluation policy. Furthermore, across all policies, more text was devoted to this component than any other (Kinarsky, paper 2). This emphasis on design and methodology may be because foundations are recommending that the method of inquiry follow the purpose of the evaluation. At Gates, for example, this approach to evaluation is referred to as "fit to purpose" which "allows for a range of methods, including qualitative and quantitative data collection and analysis" (Gates policy, p. 3). The policy reframes evaluation quality and rigor by affirming that evaluation should prioritize alignment between purpose and method:

> "Recently, there has been a belief that if it is not an RCT, then it's not useful. However, the policy also prevents against just getting a couple quotes from a couple of satisfied people. I think it's useful that there is a reminder to think about the appropriate methodology. It is a simple but important reminder that the answer to a question on methodology is and will always be, 'it depends'." – Gates Foundation

Policies encourage evaluations that respond to the local context and embrace broad methodological choices. They have decentralized decision-making and provided guidance on how to vary evaluation practice.

During my conversation with an evaluation director at Gates, he acknowledged that while the policy helped him understand the vision for evaluation at the foundation when he first started, it is not something he often revisits. For him, the policy is less relevant on a day-to-day basis because it disperses autonomy to staff working in different strategy areas:

> "I was reminded why I don't think about the evaluation policy when I looked at it again today in preparation for our call. The policy says, 'We do a lot of different things so it's up to the individual strategies to figure out their evaluation plans.' That's how I interpret the policy… As I read it, it doesn't really offer much guidance except to say we think evaluation is important, but how, why, when and with who you do evaluation depends on the kind of work you are doing. The policy puts trust in the folks doing in the individual strategies to figure it out…and that's my job for the K-12 strategy." – Gates Foundation

The policy creates a structure for evaluation practice that prioritizes the local context. It gives evaluation staff situated in different strategy areas the responsibility of determining the correct evaluative approach.

The Hewlett policy also does not enforce how evaluation is practiced across the foundation. In addition to setting a broad philosophy for evaluation practice, their policy has prompts for internal stakeholders to consider when doing evaluations. These prompts may guide how stakeholders conduct or use evaluation, but it does not enforce it:

> "It is a set of guidelines. It doesn't force people to do A, B or C. Instead, it says these are the questions you should be asking yourself as you go through strategy development,

> implementation, refresh, and sometimes exit…Since they are principles and not demands, we're not entirely living by our principles." – Hewlett Foundation

Rather than strive for strict adherence to the guidelines or principles set forth in their policies, foundations favor an evaluation practice that is tailored to the unique context of each program area and each evaluation.

6. Evaluation staff have more clarity on how to guide organization-wide evaluation practice

By outlining how evaluation is practiced at the foundation, policies have given evaluation directors and their teams greater clarity and more authority. Designing, conducting, analyzing, and using evaluations can be an arduous process. It involves a lot of decision-making often by various stakeholders over a long time. Having key components of evaluation practice detailed in a policy can make it easier for directors to offer consistent guidance.

For example, one foundation regularly holds strategy meetings for their grant making where evaluation is frequently discussed. As of late, internal stakeholders have been eager to conduct randomized control trials of their grant making. Rather than repeatedly explain the need to first identify the purpose for an evaluation and then select an appropriate methodology, evaluation directors can point stakeholders to the policy document:

> "We often have conversations related to strategy or evaluation. In those discussions, I reference the policy. For example, when we have a debate about if there is a hierarchy of evidence…that is something that keeps coming up. I say, "Let's look at our guiding principles". Then they say "Oh! Look at that. In our own guiding principles, we have the idea that methodology has to follow purpose. So, I guess there isn't just one methodology that is at the top of the heap, right?" So, we do use the policy. It's nice to point to our

principles and say, “Remember, this is the framework we work within”. It sometimes helps solve disagreements." – Duke Endowment

Policies enable evaluation directors to better support the evaluation efforts of the organization. Directors say that policies help them better understand their roles, influences the guidance they give, and informs the materials they generate:

> “Our evaluation principles really permeate everything. They inform any advice I offer to staff, as well as all the tools that I provide. The policy helps me be really clear in my advice and my guidance." – Hewlett Foundation

Evaluation policies help evaluation staff be clear and consistent when guiding the various evaluation efforts happening across the organization.

7. Process of developing an evaluation policy is beneficial

Aside from having the policy, foundation directors indicated that the process of developing an evaluation policy was advantageous for the organization. Developing the policy presented an opportunity to reflect on the principles of evaluation practice that the foundation deems important and then discuss the practices that are associated with that principle.

Rockefeller recently revised its evaluation policy to reflect organizational changes prompted by a new president. Once the evaluation team drafted the policy, they began to share it with stakeholders from across the organization. Those conversations surfaced aspects of evaluation practice that are valued at the foundation, which impacted the content of the policy. For example, one of their principles is that evaluations need to be objective and independent so that findings are credible to a broad range of stakeholders. However, another principle is that evaluations need to fit the purpose of the evaluation. Conversations about the policy surfaced a tension between these two principles:

> “Internal conversations have been really helpful in terms of getting us on the same page. We had a nice back-and-forth around the independence principle. At first, it fell off because we were like, ‘Well, if evaluation is user centric and impact oriented and the user wants an independent evaluation, then it'll be independent. But if they want a developmental evaluation, then it won’t be independent." However, during our internal conversations it turned out that right now we actually value independence. So, it circled its way back to being a principle." – Rockefeller Foundation

Discussing a draft of the evaluation policy with internal stakeholders helped clarify their understanding of how evaluation should be practiced across the organization.

At Kansas Health Foundation, the process of writing an evaluation policy has been instrumental to defining the role and responsibilities of an evaluation director. The current director is only the second to have this role, and the first to exclusively focus on evaluation:

> "Writing it down is beneficial in a lot of ways. It has forced me to have more clarity, that I wouldn't have otherwise. It has also provided time for reflection which has been so important in figuring out what is working and what is not. It has helped me hone in on what is a value add." – Kansas Health Foundation

Relatedly, the development of Hewlett’s evaluation policy surfaced the need for a full-time staff person dedicated to upholding their evaluation principles and helping staff apply them in their practice. The job description was integrated into the policy document and also used to hire the current evaluation officer:

"The policy development process ended with a new job description for an evaluation director. It looks very similar to the job description that I responded to. I'm the first dedicated evaluation officer at the foundation. The foundation has been doing evaluations over the years, but there was no dedicated support person until my role. My role is really to hold the institutional knowledge of those principles and to look for ways to help our staff apply the principles in practice." – Hewlett Foundation

External Intent (8-9)

In addition to improving internal alignment, foundations develop and share their evaluation policies to explain their evaluation approach to external stakeholders. This is done for two reasons. First, foundations want their current and prospective grantees to better understand how evaluation is practiced and used at the foundation. By being forthright, they seek to set realistic expectations about how the foundation evaluates and what impact it may have on grantees. Second, foundations believe that how they perform evaluation is important to share with the sector. They have invested resources in their evaluation practice and want to disseminate their internal beliefs and lessons learned with the broader field.

8. Explain evaluation requirements & expectations to grantees

Evaluation directors identified two broad reasons for sharing evaluation requirements and expectations with current and prospective grantees. First, disclosing this information can help prospective grantees write better grant applications. Grantees are often applying for funding from a number of different sources. Their applications are generally tailored to address the funder's stated needs and expectations. Understanding how and why a foundation does evaluation can help a grantee draw parallels between their own practices and that of the foundation:

> "I think it is helpful for external people to understand our mental model. So, if they want to apply for grants, maybe this is helpful for them as they write it." – Kansas Health Foundation

Through the evaluation policy, foundations aim to provide external stakeholders, including grantees, insight into grant requirements and expectations.

Second, sharing evaluation requirements from the outset may help foster more transparent and positive relationships between funders and their grantees. For example, if a foundation is committed to tracking student attendance data, but the grantee does not have the capacity to conduct or manage such data collection, the resulting tension could lead to an adverse relationship, with negative funding consequences for the grantee. Therefore, a foundation's evaluation requirements and expectations are important considerations for grantees to consider before applying for or accepting funds.

The Wallace Foundation published its evaluation policy online after realizing that its evaluation expectations and practices were sometimes unexpected or overwhelming to their grantees. Now, prospective grantees are invited to review the evaluation policy before applying for or accepting a multi-year grant:

> "Having grantees find out about all the research requirements after they've gotten the grant led to some really difficult, awkward, & problematic conversations and relationships. It didn't take many iterations of that before we started letting people know in the earliest possible minute about our approach. We found that being explicit causes some organizations to not go ahead with the application process because it's not a fit with them. That is totally appropriate, and we always tell them they are eligible to be considered for future initiatives. The folks who do go ahead are not surprised and are o.k.

with participating in the research. They are O.K. with having their organization publicly identified in the research reports.” – Wallace Foundation

After some trial and error, the foundation recognized that not having an explicit document led to confusion and frustration. Rather than sharing information piecemeal throughout the multi-year grant, the evaluation policy is intended to give grantees and their foundation contacts a shared mental model and language to discuss evaluation expectations.

9. Communicate evaluation philosophy and practices to external stakeholders

The second external audience for evaluation policies is rather wide-ranging. As foundations continue to invest in their evaluation and research practice, they are generating new knowledge that may be of value to their colleagues. In addition to sharing their evaluation and research findings, foundations are also sharing broadly their evaluation philosophy and practice. One reason for doing so is to facilitate dialogue and learning with their peers.

“We believe there are things that we're doing that would be good to share with others so that they can learn from our experience." – Packard Foundation

“We felt that the way that we did work - the way we made grants and the process for monitoring and evaluating them - was the thing that was unique to the foundation.” – Rockefeller Foundation

This aspiration to support the sector by circulating internal opinions, processes and learnings is not a new instinct for foundations. As one evaluation director shared, the desire to write and share an evaluation policy was motivated by the foundation’s long history of investing in the field of evaluation:

> "We have a great commitment to evaluation. We helped establish the local association for evaluation as a state partner with the American Evaluation Association, provide funds in our grants for evaluation, and created a handbook for grantees to guide logic model framing. Long before I came to the foundation, it was an important figure in the evaluation space in this city and in the state… I wanted to build on that history. There are many ways people think and practice evaluation. We wanted to be very clear on how and why we have the practice of evaluation." – Skillman Foundation

The development of evaluation policies is a way for foundations to demonstrate their commitment to evaluation. Evaluation policies enable foundations to share internal practices and processes with the field and contribute to an evolving understanding of how evaluation should be utilized in the philanthropic sector.

External Influence (10-11)

In the fourth and final segment of this section, ways in which evaluation policies are influencing external stakeholders are discussed. First, publicly disseminating evaluation policies has helped external stakeholders better understand the evaluation expectations and practices of their partner foundations. Second, evaluation directors are reading the evaluation policies of other foundations and reflecting on their internal practice.

10. Current & prospective grantees better understand foundation's evaluation expectations and practices.

Evaluation directors reported that having an explicit evaluation policy has led to more satisfied grantees. For example, Wallace has a very particular and intensive approach to evaluation and research that imposes a number of requirements and expectations on their grantees. Their evaluation policy is shared and discussed a lot with grantees because it helps set realistic expectations. In particular, their grantees are more satisfied:

> "It is burdensome and a pain in the neck to be part of a Wallace initiative. But if you know what you're in for, then it is pretty manageable and not that surprising. It winds up being O.K. with people." – Wallace Foundation

Evaluation and research are important functions for the Wallace Foundation, but places some burden on their grantees. Being upfront and explicit with grantees about what they should expect helps facilitate a more positive working relationship.

Policies are not read only by prospective grantees. They are also reviewed by external evaluators hired by foundations to support their evaluation practice. For example, program officers at Packard rely on external evaluators to develop and execute evaluation plans. The evaluation policy communicates the foundation’s expectations and processes so external evaluators' work is in alignment with internal expectations:

> "I think our contractors look at and use the guiding principles. Just this week, I received an email from one of them about the guiding principles. External consultants help teams draft and implement their measurement, evaluation and learning plans. So, at that time they are referencing the evaluation principles.” – Packard Foundation

Evaluation policies offer external stakeholders’ insight into how evaluation is done at the foundation. Since each foundation articulates a unique vision and approach to evaluation, the policies orient external stakeholders to the foundation’s context.

11. Foundations are benefiting from reviewing and reflecting on each other's policies.

Foundations look to what their peer institutions are doing when developing or revising their evaluation policies, in part to identify what they do or do not want to emulate in their own policies.

> "As I prepare to update our policy, I first did a scan of what other folks are doing." – Packard Foundation

> "I knew I don't want anything as long as the Gates or Hewlett policies... those are influential policies, but we don't need anything that detailed." – Duke Endowment

Evaluation directors are looking across the sector for examples of evaluation policies. They are critically analyzing policies and reflecting on how their foundation's policy may look different.

What Informs Policy

Above, insights into the intent for developing evaluation policies and their influence on practice are discussed. The findings section concludes with a discussion of what informs the development of evaluation policies.

Evaluation research on evaluation policy is not motivating or informing the development of evaluation policies in the philanthropic sector. During interviews, evaluation directors did no mention any publications on evaluation policy or the taxonomies proposed by AEA (2007) and Trochim (2009). However, when conceptualizing the first draft of an evaluation policy, evaluation theory helps delineate the ideas and practices that should be integrated into a policy. In addition to clarifying key concepts, theory provides a common language that can be consistently understood across stakeholders. The evaluation policy at MacArthur was influenced by theorists like Rossi and Freeman, Patton, and Scriven:

> "Our approach and guidance (policy) are based on theory…We adopted existing theory and operationalized it for our working environment and culture at the foundation." – MacArthur Foundation

Michael Quinn Patton also influenced the evaluation policies written by Hewlett and Rockefeller. At Hewlett, the evaluation policy was influenced by Patton's Utilization- Focused Evaluation, a framework for designing and implementing an evaluation to fit a particular situation. Rockefeller's updated evaluation policy was influenced by Patton's (2018) more recent work on Principles-Focused Evaluation:

> "We're very much aligned with Michael Quinn Patton's thinking of how philanthropic evaluation should be set up. We're following his Principles-Focused Evaluation approach. Rather than having a policy or standards that are very prescriptive, we've developed a set of principles." – Rockefeller Foundation

At the beginning, when a policy is being conceptualized, evaluation theory is often consulted to identify key ideas, practices and concepts. In addition to providing shared language, evaluation theory can inform the format and structure of an evaluation policy.

However, evaluation theory is consulted less frequently when making modifications to an extant evaluation policy. Instead, personal experiences, internal developments and organizational culture inform alterations. At MacArthur, now that its policy is used regularly, the evaluation staff review evaluation theory less regularly and frequently when making adjustments to the policy. Updates are largely influenced by internal developments:

> “We looked externally more at first. We continue to refine in response to our evolving cultural change. In the beginning, we were starting from nothing. We had no history or experience doing this as an institution, so we looked elsewhere to help us develop our parameters. Now, we have lived with these parameters. After implementing them we moved into a new phase. We know which parameters are working for us and which are not. So now changes to the approach and guidance (policy) are more around shifts in culture. Some things, like whether we should be evaluating a grantee or not, it didn't matter to us if every other foundation was doing it or not doing it. We have made some decisions and defined how we want to operate and developed a comfort and culture around that.” – MacArthur Foundation

As a foundation lives with its evaluation policy, it identifies areas of growth and revision unique to its organization. Rather than consult evaluation theory again, revisions are developed internally, reflecting the organizational culture and context.

Interviews revealed that it can be difficult for evaluation directors to use theory as a primary source because theory is not always easily understood or applicable to their context:

> "I don't use the term evaluation theory very much. I don't really know what it is. I've always thought of evaluation in pretty practical terms, so I've never really gotten into the academic aspect." – Duke Endowment

> "Preskill and Patton are two of the few evaluation thought leaders that have been able to crack the philanthropy nut. We are a really interesting bunch. We are not the largest commissioners of evaluation, that is definitely still government. So, most of the evaluation trainings and education programs are geared towards training you to be an

> evaluator for government. There are very few that look at the nuance of philanthropy." – Rockefeller Foundation

Foundation leaders are not relying on evaluation theory, in part because it is often not relevant to their context. While evaluation theory that emerges from and addresses the unique challenges faced by the philanthropic sector may be lacking, there are some theorists, namely Patton and Preskill, that have been able to generate a literature base that is valuable to evaluation directors.

Industry groups that understand the practice of evaluation in the philanthropic sector are translating evaluation theory for the foundation audience. They are sharing developments and trends in the evaluation community that are then influencing policies and practice at foundations. Evaluation directors attend evaluation conferences, such as those hosted by AEA, and meet with other evaluation directors from across the philanthropic sector. They also participate in industry groups that discuss emerging issues in philanthropy, evaluation, and evaluation in philanthropy. During interviews, evaluation directors noted the importance of organizations like the Center for Evaluation Innovation, Evaluation Roundtable, and Left Coast Evaluators, GEO, Center for Effective Philanthropy.

> "I am a very active participant in AEA and GEO. I find that GEO can be helpful. But most importantly, three or four years into my work at Skillman, I became part of an ad hoc group of philanthropic evaluation directors. We piggyback on AEA and GEO and have a meeting of about 20 evaluation directors in philanthropy. That group has been really, really important in shaping my thinking." – Skillman Foundation

These conferences, conversations, and industry groups inform their thinking, and in turn their policies.

An example of an issue that is gaining prominence in evaluation research and among evaluation industry groups is the need to integrate diversity, equity and inclusion (DEI) into evaluation practice. DEI discussions are motivating amendments to evaluation policy and practice at several foundations. For example, in addition to adapting how evaluation is used for internal learning, Skillman is considering the diversity of their evaluation contractors:

> "When the policy is updated there will probably be a stronger emphasis on the use of evaluation in changing systems. Also, a focus on equity and social justice. Our internal practices and practices with our grantees have to get developed and embedded around learning. I think the learning questions with any evaluation projects will change. Perhaps the way we screen contractors may change in relation to our focus on diversity, equity and inclusion." – Skillman Foundation

Hewlett is also responding to the calls for evaluation to adopt a more intentional orientation towards DEI. They are modifying their policy to help educate stakeholders on the importance of DEI and offer examples of how to integrate it into their evaluation practice:

> "When I look at the questions we give as example in our guidebook, none of them bring out anything related to diversity, equity or inclusion. I feel that's a real clear example where I think we need to update in terms of an issue that's really present in the evaluation field right now." – Hewlett Foundation

Given that evaluation theory is often furthered by trends happening in evaluation practice, it will be worthwhile to track how these efforts influence evaluation practice and theory in years to come.

Conclusion

The present paper is one of the first to investigate why foundations are developing evaluation policies and how these policies influence their evaluation practice. Through interviews with evaluation directors at foundations with evaluation policies the present research begins to document why evaluation policies are being developed and what impact, if any, they have on evaluation practice. Furthermore, interviews begin to explicate how the information and knowledge embedded in evaluation theories are informing the content of evaluation policies.

The intent and influence of publicly available evaluation policies can be divided into two broad categories: internal and external. The responsibility for commissioning, conducting, and using evaluations is often dispersed across a broad range of internal stakeholders, such as board members, the president, evaluation staff, and program staff. Evaluation policies seek to improve internal alignment around evaluation practice by communicating a broad philosophy and purpose for evaluation. Policies also specify when and how evaluation practices should remain consistent and vary across the foundation.

The process of writing an evaluation policy offers foundations opportunities to clarify their approach to evaluation and articulate a shared mental model. Upon implementation of the policy, evaluation directors believe that evaluation practice is improved in a few key ways. First, there is more consistency in how different stakeholders conduct and use evaluation. Next, when and how evaluation practice varies has improved. Lastly, evaluation directors are better able to support evaluation across the organization. The policy organizes and simplifies how evaluation is practiced at the foundation resulting in clearer and more consistent guidance offered to stakeholders.

The second intent of evaluation policies is to explain a foundation's evaluation approach to external stakeholders. External stakeholders include current or prospective grantees, external evaluators, the general public, or other foundations. Policies seek to explain internal processes to grantees, offer direction to external evaluators, share internal opinions on evaluations with other foundations, and inform the public on the foundation's internal practices. As a result, foundations have improved relationships with their grantees and external consultants. They are also using the policies written by others to reflect on how evaluation is practiced at their foundation.

Interviews with evaluation directors revealed that many drafted a policy to help foster a strong evaluation practice within the foundation. There was a need to articulate their foundation's purpose for and approach to evaluation. Policies tend to outline a shared approach to the cross-cutting functions of evaluations while still leaving room for stakeholders to make context specific decisions. By specifying how to use and manage evaluations, policies foster a more intentional evaluation practice across the foundation that aims to generate more meaningful evaluations. Notably, not just the evaluation policy but also the discussions during policy development enabled foundations to be more intentional in evaluation practice.

The present study also found that evaluation policy development across the sector is not directly motivated or informed by emerging research on evaluation policy. However, evaluation theory and research, notably the writings of Patton and Preskill, are referenced by some directors when evaluation policies are being conceptualized and written. Evaluation directors noted that conferences, like those hosted by AEA, and industry groups, such as Center for Evaluation Innovation and Center for Effective Philanthropy, are more influential than evaluation literature. These meetings and groups help evaluation directors prioritize and situate developments in evaluation to their context.

Overall, findings on the intent and influence of evaluation policies are largely consistent with those outlined by Trochim (2009). Policies specify and encourage consistency of practice across the foundation, which is necessary because responsibility for evaluation is typically dispersed across various stakeholders. Policies are also an educational tool - helping explain the foundation's approach to evaluation, as well as best practices. Foundations view others' explicit policies to reflect and adapt their own evaluation practice. Grantees and external contractors also review policies to understand and align to internal expectations. Lastly, as foundations prioritize diversity, equity, and inclusion, they are also revising evaluation policies to reflect this focus in their evaluation practice.

Chapter 5. Conclusion

Evaluation policies have a growing influence on evaluation practice, yet currently little is known about them. Evaluation policies have been informing evaluation practice in European countries since the 1990's. In 2016, the Canadian government passed legislation requiring every federal agency to establish an evaluation policy. More recently, the United States passed the Foundations for Evidence-Based Policymaking Act of 2018 requiring agencies to develop evaluation policies. Multilateral and bilateral aid agencies and government organizations, like the United Nations and World Health Organization, have also written evaluation policies. These policies impact evaluation practice, as well as the quality of information generated:

> "Evaluation policy is a critical concern for evaluation practice, in part because it shapes evaluation practice thereby both enabling and constraining the potential contribution evaluation can make… At the extreme, evaluation policy may not only enable some forms of contribution and constrain others; bad evaluation policy can have serious negative consequences" (Mark et al., 2009, p. 3).

The quality of evaluation is particularly important in the philanthropic sector where findings are frequently used to inform funding decisions. Annually, US foundations grant over $60 billion to nonprofit organizations working in nearly every issue area around the world, from disease eradication to homelessness prevention (The Foundation Center, 2014). Given that foundations and nonprofits work together to improve society, often providing vital services for the most vulnerable and disenfranchised communities, what gets funded is of great importance. However, researchers and practitioners in the philanthropic sector have observed that evaluation practice tends to be "ad hoc, not aligned to strategy, underfunded, and underused" resulting in evaluations that don't generate useful and credible evidence (Preskill and Mack, 2013, p. 4).

Extant research on evaluation policy reveals that foundations are writing evaluation policies (Mattessich et al., 2009; Dillman, 2014). However, little is known about the prevalence, contents, intent or influence of these policies. The present dissertation sought to catalog and analyze the evaluation policies of the philanthropic sector, guided by the following research questions:

1. To what extent are US foundations investing in evaluation? (Paper 1)
2. To what extent do US foundations, who are major commissioners of evaluation, have publicly available evaluation policies? (Paper 1)
3. How do publicly available evaluation policies in the philanthropic sector compare to taxonomies proposed by Trochim (2009) and the American Evaluation Association (2007)? (Paper 2)
4. How do publicly available evaluation policies in the philanthropic sector compare to each other in terms of their comprehensiveness and completeness? (Paper 2)
5. According to evaluation directors, what are the primary goals of their foundation's policy? (Paper 3)
6. According to evaluation directors, in what ways do the foundation's formal and informal evaluation policies guide evaluation practice? (Paper 3)
7. According to evaluation directors, to what extent does evaluation theory inform the evaluation policies of US foundations? (Paper 3)

The present chapter begins with an overview of findings. Then, the implications of this study and directions for future research are discussed.

Overview of Findings

Evaluation is a fixture of foundation spending, but evaluation policies are not prevalent among top commissioners of evaluation

The findings in this study suggest that evaluation is a fixture of foundation spending. While evaluation spending ebbed and flowed between 2010 and 2014, on average, the sector apportioned approximately $276 million annually for grants that included evaluation services. While this represents only between 1.7 and 2.6 percent of total annual giving, it is still a significant annual investment into a practice that is intended to generate evidence for accountability, learning, and knowledge sharing (Kramer & Bickel, 2004).

Despite being a fixture of foundation giving, few foundations have publicly available evaluation policies informing their evaluation practice. Of the 67 top commissioners of evaluation, only 12 foundations had publicly available policies. Evaluation policies are intended to bring conceptual clarity to evaluation practice. They align internal and external stakeholders around a purpose for evaluation, as well as how it should be practiced across the foundation. Furthermore, public evaluation policies foster an open and ongoing conversation about evaluation practice and use:

> "Evaluation policies help make evaluation a more transparent and democratic endeavor. They constitute a public stance that an organization takes regarding evaluation. Because they are public, written policies, they can be known by everyone in the organization and thus criticized and challenged. Participation and dialogue can occur about which policies make the most sense under which circumstances" (Trochim, 2009, p. 17).

Millions of dollars are allocated by foundations towards evaluation-related grants every year. Writing and sharing evaluation policies can help foundations use their limited resources wisely, leading to a better evaluation practice that results in more rigorous and credible evaluation findings.

A modified evaluation policy taxonomy

To date, two evaluation policy taxonomies have been proposed to elucidate the components of an evaluation policy. The AEA evaluation policy task force proposed seven components: a definition of evaluation; evaluation requirements; methodology recommendations; evaluator requirements; budget considerations; implementation standards; and ethical considerations (AEA, 2007). In 2009, Trochim offered a slightly different, more expansive taxonomy of evaluation policy with the following eight components: evaluation goals; participation; capacity building; management; roles; process and methods; use; and evaluation of evaluation (meta-evaluation) (p. 20). However, these taxonomies were created theoretically: “We need to develop taxonomies and corresponding audit methodologies that are better than the ones suggested here, and these need to be informed by systematic research” (Trochim, 2009, p. 28).

In response to this call for more research, the examination of 12 evaluation policies in the present study led to alterations to these taxonomies. The two taxonomies were integrated into one and components were streamlined. Three new components that emerged from content analysis were added: “Background”; “Avoid”; and “Example”. The modified taxonomy more accurately represents the components of evaluation practice detailed in evaluation policies written in the philanthropic sector.

Key motivations for the development of evaluation policies

This study identified several challenges to evaluation practice in the philanthropic sector that are motivating the development of evaluation policies. First, evaluation practice is complex in the philanthropic sector. Analysis of the Foundation Center data found that evaluation-related grants are distributed across 17 different sectors, such as health, education, and human services. Relatedly, the responsibility for evaluation is dispersed across a wide range of stakeholders. Given that most foundations tend to have several strategic priorities with initiatives at different stages of maturity, the types of evaluations that are appropriate vary greatly.

The second motivator for evaluation policy relates to managing external stakeholders. 92 percent of all evaluation-related grants are awarded to foundation grantees, such as nonprofit organizations (66 percent) and schools (26 percent). This suggests that foundation staff and grantees are regularly discussing and using evaluation. Foundation staff have found that not being explicit about evaluation expectations results in tension and even adverse relationships with grantees. Therefore, evaluation policies are intended to explain their evaluation approach and requirements from the outset to potential and current grantees. Relatedly, foundations contract with external evaluators to conduct and support their evaluation needs. Evaluation policies are crafted to orient contractors to internal expectations and processes.

To address these internal and external challenges, foundations are developing evaluation policies that capture the foundation's thinking about why and when evaluation is important, how evaluation is organized and practiced at the foundation, and how findings are used. Policies are also intended to specify the practices that support useful and credible evaluations.

Evaluation theory is informing policy, but more can be done to strengthen this connection

Christie and Lemire (2019) propose that foundations may be adapting evaluation theory and literature to their context through evaluation policies to help inform and guide evaluation practice. However, interviews with evaluation directors revealed that the connection between evaluation theory and evaluation policy is typically not direct.

First, research on evaluation policy is not motivating or informing the development of evaluation policies in the philanthropic sector. During interviews, evaluation directors did not mention publications on evaluation policy, or the taxonomies proposed by AEA (2007) and Trochim (2009). This may explain why evaluation policies differed in several ways from the conceptual ideas of evaluation policy outlined in evaluation literature. For example, while favored in the literature, the term “evaluation policy” is not commonly used by foundations. Foundations favored terms like “principles” and “approach” when naming their evaluation documentation. Also, the components prioritized within a policy differed from those proposed by AEA (2007) and Trochim (2009). In particular, less than five percent of all policy text addressed the following three components: “Ethics”, “Capacity Building”, and “Meta-evaluation”.

Second, while policy formation should be informed by relevant theory, evaluation theory is not directly informing the policies being written in the philanthropic sector (Cochran & Malone, 2010). Evaluation directors said it can be challenging to identify evaluation approaches and models that are relevant to their context. Rather than reference evaluation literature directly, they rely on conferences and industry groups to help surface trends and information that can improve their evaluation practice. Furthermore, most evaluation directors interviewed in the present study have years of academic and professional experience with evaluation. They are experts in the field and have developed opinions on how evaluation should be conducted in the philanthropic sector which inform their policies.

More collaboration between researchers, industry groups, and practitioners could help standardize and improve evaluation policies. For example, AEA (2018) has developed standards that specify professional ethics for evaluators in order to identify and promote ethical practices in evaluations. These could be integrated into policies to ensure that internal and external stakeholders are aware of and acting in accordance with these standards. Additionally, as foundations begin to adapt their policies and practice to be more responsive to issues of diversity, equity and inclusion, the literature on evaluation related to social justice and empowerment could help generate ideas and recommendations.

Evaluation design & methodology is prioritized in evaluation policies

Trochim (2009) suggested that "evaluation policy is important because many of the controversies in evaluation today are essentially about such policy" (p. 18). He went on to list a few of these controversies, all which relate to policies that prioritize the widespread use of randomized control trials to determine evaluation effectiveness. Today, there is a growing consensus that design and methodological choices should be informed by the evaluation context and purpose. This is particularly important for foundations who fund initiatives in the social sector that may not be well suited for evaluations with an experimental design.

To help instill an interest in and capacity for using a broader methodological approach, evaluation policies are focusing on issues related to design and methodology. More text was dedicated to "Design and Methods" than any other component. Approximately one-third (29 percent) of all the text analyzed was guidance on evaluation design and methodology. In interviews, evaluation directors noted that stakeholders frequently begin an evaluation by picking a method, such as a randomized control trial or survey, rather than first identifying the evaluation purpose. To limit this practice and educate stakeholders, evaluation policies outline when and what aspects of evaluation practice should vary and those that should remain consistent. They also explain practices that lead to rigorous and credible evaluations.

Implications for Research and Practice

The present dissertation has several implications for research and practice. First, implications for future research on evaluation are discussed. Next, implications for practitioners in the philanthropic sector are considered.

Implications for Research

Evaluation policy addresses many aspects of evaluation practice, presenting numerous opportunities for future research. Below are a few ideas that emerge from the present study.

This study is one of the first attempts at tracking the demand for evaluation in the philanthropic sector by examining charitable grant spending on evaluation, and other knowledge production services, in the philanthropic sector. There were two limitations of the Foundation Center data that could be addressed in future research. First, evaluation spending is frequently bundled with other knowledge production activities, program development and program execution. Second, in addition to using charitable grant funding, foundations also utilize administrative funds to commission evaluations. The Foundation Center data did not account for these aspects of evaluation spending, which made it challenging to discern and track funds dedicated solely to evaluation.

Foundations are annually spending an average of $276 million on evaluation-related charitable grants, yet very little grant funding is being apportioned by foundations directly to evaluation providers. In other words, according to Foundation Center data, while foundations are the ultimate source of revenue, they are less often the direct buyers of evaluation services. A charitable grant that requires evaluation is most often first awarded to a grantee and then that grantee retains evaluation services. This contributes to a decentralized and localized evaluation marketplace that is not driven, at least directly, by a fixed number of funders, but rather by a broader network of NGOs. Given that 66 percent of all evaluation funding is actually expended by nonprofits, more research is needed into how these funds are disbursed, and any influence foundations have on how funds are utilized. At present it is also unclear how evaluations in the foundation sector are procured, the number of providers receiving these funds, and the type of evaluation providers selected (e.g. large research firms, independent consultants).

To date, little empirical research exists on evaluation policies, and even less on those being developed in the philanthropic sector. The present study is one of the first to examine the prevalence of evaluation policies in the philanthropic sector. However, only top commissioners of evaluation services with publicly available policies were included in the study. Broader surveys of the sector may help inform our understanding of the prevalence of policies at various foundations and surface explicit evaluation policies that are not publicly posted. Expanding the catalog of evaluation policies will offer a more comprehensive understanding of the scope and scale of explicit evaluation policy and reveal variations in evaluation policies across foundations.

Content analysis of evaluation policies concluded that the following five components were present in every policy: “Goals and Purpose”; “Management”; “Participation and Roles”; “Design and Methods”; and “Use”. A more detailed examination of what practices these components stipulate is a promising area of future research. Examination of how these components differed across foundations may reveal important differences in how evaluation is practiced. The Evaluation Theory Tree developed by Alkin and Christie (2004; Christie & Alkin, 2008, 2012) which categorized evaluation theories into areas of emphasis, could be used to identify how components of evaluation policies align with different theorists.

Relatedly, more investigation is needed into why issues related to “ethics”, “capacity building” and “meta-evaluation” were so rarely integrated into evaluation policies. For example, a broader examination of evaluation documentation at a foundation may uncover that these components are detailed elsewhere. Future research could also pursue questions such as: How many components are necessary and which ones? Should this differ based on the organizational context? Such research has the potential to align a wide range of stakeholders from the evaluation and philanthropy communities, ensuring a meaningful relationship between theory, policy and practice.

Findings from interviews with evaluation directors point to a number of areas for future research. Prior research showed that evaluation policies can be interpreted inconsistently (Christie & Fierro, 2012). Evaluation directors identified numerous internal and external stakeholders that may benefit from evaluation policies. Yet it is unclear how these stakeholders interpret and use foundations' evaluation policies. Interviews with various stakeholders, such as external evaluators, program officers and grantees, may produce valuable insights. Alternatively, simulation studies could be used to address this gap. For example, external evaluators can be given an evaluation scenario and an evaluation policy, and then asked to design an appropriate evaluation. Simulation studies have revealed fundamental differences across evaluation theorists and are a promising next step for research on evaluation policy (Eisenberg, Winters & Alkin, 2005; Mark et al., 2009).

Lastly, while there are many new areas of inquiry, there remains a need to continue learning about how policies get developed and applied in organizations. More case studies on evaluation policy in the philanthropic sector will improve our understanding of how policy impacts practice in particular contexts. Qualitative researchers could also follow an organization through the process of developing a study to deepen our understanding of how policy is set.

Implications for Practice

The present study has a number of implications for practitioners in the philanthropic sector. First, research on evaluation spending offers foundation staff an opportunity to consider how their evaluation spending aligns with or deviates from that of the broader sector. However, at present, there are some shortcomings with the data available on the philanthropic sector. As noted above, reported evaluation spending is currently bundled with other charitable funding, and administrative funds used to commission evaluations are not tracked. During interviews with evaluation directors, several indicated a willingness to refine internal accounting practices to better track industry trends. Adjusting accounting and reporting practice to address these concerns would greatly improve data on the sector.

Next, the conceptual frameworks proposed by Trochim (2009) and AEA (2007) were integrated and modified to capture findings from content analysis of evaluation policies in the philanthropic sector. Practitioners may find that the modified taxonomy can serve as a useful framework to guide conceptualizing and writing evaluation policies. It may also serve as a reflection tool for organizations that already have explicit evaluation policies.

Furthermore, foundations that already have an evaluation policy may want to conduct a similar analysis to that described in paper two to examine their policy’s completeness and comprehensiveness. This analysis may elicit conversations about how the evaluation policy reflects the organization's goals and practices.

Interviews with evaluation directors revealed that industry groups, such as Center for Evaluation Innovation and Center for Effective Philanthropy, are influencing their evaluation practice, and therefore their evaluation policies. These groups help evaluation directors prioritize and situate developments in evaluation to their context. Therefore, the ideas and materials advanced by these groups have a great impact on evaluation practice in the civic sector. Industry groups have an opportunity to expose evaluation directors to a broad range of evaluation theory helping to bridge the theory, policy, and practice nexus. The ideas and people industry groups endorse have a lasting and direct impact on how evaluation is practiced.

The present study makes a case for the development of publicly available evaluation policies across the philanthropic sector. Evaluation practice in foundations is complex which makes obtaining meaningful and useful evaluation findings challenging. There are many internal and external stakeholders. Responsibility for evaluation is often decentralized, with foundation funding often dispersed across various sectors to projects at differing stages of maturity. This has contributed to foundations struggling to plan and design evaluations that produce meaningful insights for the field, their grantees, and the foundation (Buteau et al., 2016; Kramer & Bickel, 2004). Foundation grantees are also overwhelmed by the growing demand for evaluation (Behrens & Kelly, 2008).

Well-crafted evaluation policies that specify procedures and goals for evaluation practice may help address these concerns. The modified taxonomy developed in the present dissertation, as well as the cataloged 12 evaluation policies, can assist evaluation directors in developing policies that align internal and external stakeholders around a shared purpose, and offer procedural advice on how evaluation should be practiced. Policy development that is guided by a shared taxonomy and standard language may aid practitioners to more effectively communicate their evaluation practice with external stakeholders. Grantees will benefit from better understanding evaluation expectations from the various foundations who fund their initiatives. Foundations can review and discuss evaluation policies from a range of foundations to identify new ideas and reflect on shared practices. Standardizing the structure and key terms of evaluation policies may help improve evaluation practice and lead to more rigorous and credible evaluation findings across the sector.

Appendix A.
Foundation spending data for 2013 and 2014

	2013				2014			
Foundation	**Evaluation Spending**	**Total Spending**	**% Allocated to Eval**	**Rank**	**Evaluation Spending**	**Total Spending**	**% Allocated to Eval**	**Rank**
Bill & Melinda Gates Foundation	$119,771,933	$3,525,106,663	3.40%	1	$101,877,134	$3,901,340,835	2.61%	1
The Robert Wood Johnson Foundation	$34,368,173	$373,324,447	9.21%	2	$11,905,458	$385,103,419	3.09%	4
The Edna McConnell Clark Foundation	$20,781,708	$35,715,208	58.19%	3	$13,483,489	$42,151,489	31.99%	2
The William and Flora Hewlett Foundation	$9,408,892	$203,720,006	4.62%	4	$13,448,600	$316,444,525	4.25%	3
Ford Foundation	$7,598,649	$521,796,430	1.46%	5	$10,337,731	$432,495,922	2.39%	5
Conrad N. Hilton Foundation	$5,045,000	$103,781,000	4.86%	6	$4,700,000	$123,280,433	3.81%	9
The Annie E. Casey Foundation	$4,068,432	$53,800,016	7.56%	7	$1,836,878	$34,725,656	5.29%	19
Laura and John Arnold Foundation	$3,192,104	$123,305,334	2.59%	8	$549,026	$102,785,643	0.53%	42
The California Endowment	$2,996,243	$107,457,986	2.79%	9	$2,798,499	$229,729,065	1.22%	14
The Rockefeller Foundation	$2,835,000	$154,535,813	1.83%	10	$2,195,000	$110,276,874	1.99%	17
The Skoll Foundation	$2,800,000	$22,690,256	12.34%	11	$234,000	$62,839,750	0.37%	N/A
The John and Lisa Pritzker Family Fund	$2,500,000	$13,059,280	19.14%	12	N/A	N/A	N/A	N/A
The David and Lucile Packard Foundation	$2,341,291	$229,799,070	1.02%	13	$1,082,226	$250,722,604	0.43%	31
The Joyce Foundation	$2,237,881	$32,023,535	6.99%	14	$1,675,000	$37,109,780	4.51%	24
W. K. Kellogg Foundation	$2,223,055	$167,758,330	1.33%	15	$7,625,837	$344,191,951	2.22%	6
MetLife Foundation	$2,000,000	$41,007,849	4.88%	16	N/A	N/A	N/A	N/A
The Cleveland Foundation	$1,960,000	$63,271,326	3.10%	17	$760,000	$93,536,626	0.81%	38
Stuart Foundation	$1,940,000	$14,570,200	13.31%	18	N/A	N/A	N/A	N/A
Lilly Endowment Inc.	$1,880,000	$258,246,331	0.73%	19	$1,076,657	$342,469,978	0.31%	32
The William Penn Foundation	$1,821,000	$86,882,073	2.10%	20	$4,713,068	$122,253,919	3.86%	8
Blue Shield of California Foundation	$1,676,861	$34,607,843	4.85%	21	$3,609,750	$31,612,455	11.42%	11
Buffett Early Childhood Fund	$1,547,153	$18,348,685	8.43%	22	$1,751,081	$24,793,733	7.06%	21

The Michael and Susan Dell Foundation	$1,528,600	$63,467,132	2.41%	23	$2,930,306	$61,014,407	4.80%	13
Charles Stewart Mott Foundation	$1,498,200	$107,996,981	1.39%	24	$2,270,000	$72,034,453	3.15%	16
Heising-Simons Foundation	$1,431,182	$26,040,717	5.50%	25	N/A	$29,061,245	N/A	N/A
The JPB Foundation	$1,399,636	$77,242,833	1.81%	26	$995,358	$113,344,861	0.88%	34
Alfred P. Sloan Foundation	$1,310,674	$75,961,390	1.73%	27	$20,000	$70,963,361	0.03%	N/A
Citi Foundation	$1,260,000	$77,827,150	1.62%	28	$87,000	$78,325,000	0.11%	N/A
Howard G. Buffett Foundation	$1,225,000	$80,782,105	1.52%	29	$577,896	$158,388,585	0.36%	41
The Coca-Cola Foundation, Inc.	$1,132,000	$90,953,138	1.24%	30	$60,000	$89,807,677	0.07%	N/A
Gordon and Betty Moore Foundation	$1,102,372	$278,574,142	0.40%	31	$2,696,719	$291,216,057	0.93%	15
The Rees-Jones Foundation	$1,100,000	$21,756,136	5.06%	32	$1,000,000	$34,532,800	2.90%	33
Foundation to Promote Open Society	$1,094,000	$378,796,129	0.29%	33	$1,762,936	$310,059,577	0.57%	20
Bush Foundation	$1,078,102	$21,284,798	5.07%	34	$1,646,174	$23,210,843	7.09%	25
The Duke Endowment	$1,009,372	$98,305,412	1.03%	35	$1,393,444	$96,158,911	1.45%	26
Omidyar Network Fund, Inc.	$1,000,000	$45,565,298	2.19%	36	N/A	N/A	N/A	N/A
The Wallace Foundation	$912,486	$54,894,886	1.66%	37	$1,365,336	$59,843,368	2.28%	28
The Commonwealth Fund	$879,165	$14,894,884	5.90%	38	$1,324,407	$27,865,757	4.75%	29
The John D. and Catherine T. MacArthur Foundation	$850,000	$127,796,181	0.67%	39	$1,690,000	$187,109,952	0.90%	23
The New York State Health Foundation	$816,902	$9,565,427	8.54%	40	N/A	N/A	N/A	N/A
Barr Foundation	$784,000	$49,845,605	1.57%	41	$300,000	$48,134,950	0.62%	N/A
The California Wellness Foundation	$720,000	$35,725,000	2.02%	42	$150,000	$18,511,000	0.81%	N/A
Doris Duke Charitable Foundation	$711,000	$72,523,303	0.98%	43	$5,866,000	$52,489,548	11.18%	7
Healthcare Georgia Foundation, Inc.	$690,000	$4,252,000	16.23%	44	$525,000	$3,893,000	13.49%	44
The Harry and Jeanette Weinberg Foundation, Inc.	$678,000	$75,756,111	0.89%	45	$320,000	$127,915,000	0.25%	N/A
Eli & Edythe Broad Foundation	$650,000	$101,009,530	0.64%	46	$200,000	$162,871,874	0.12%	N/A

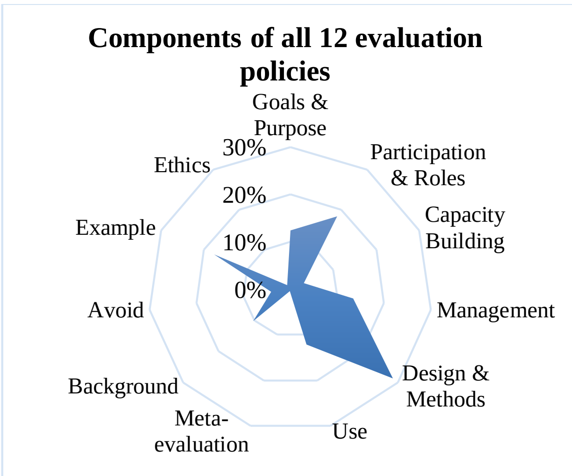
Hawaii Community Foundation	$638,185	$25,237,001	2.53%	47	N/A	N/A	N/A	N/A
The Brown Foundation, Inc.	$620,431	$51,575,366	1.20%	48	N/A	N/A	N/A	N/A
The East Bay Community Foundation	$575,000	$57,134,186	1.01%	49	N/A	N/A	N/A	N/A
The James Irvine Foundation	$567,000	$64,558,679	0.88%	50	$3,975,000	$69,730,600	5.70%	10
Carnegie Corporation of New York	$349,900	$123,874,500	0.28%	N/A	$3,500,000	$117,490,900	2.98%	12
S.D. Bechtel, Jr. Foundation	N/A	N/A	N/A	N/A	$2,061,647	$97,041,189	2.12%	18
The Leona M. and Harry B. Helmsley Charitable Trust	$470,000	$210,238,949	0.22%	N/A	$1,720,303	$224,391,999	0.77%	22
Kansas Health Foundation	$300,000	$7,671,949	3.91%	N/A	$1,365,945	$19,286,614	7.08%	27
The Kresge Foundation	$60,000	$123,662,375	0.05%	N/A	$1,139,410	$244,049,019	0.47%	30
Raikes Foundation	N/A	N/A	N/A	N/A	$962,475	$7,659,499	12.57%	35
GHR Foundation	N/A	N/A	N/A	N/A	$888,254	$20,594,455	4.31%	36
The Chicago Community Trust	$360,000	$141,659,253	0.25%	N/A	$790,000	$142,898,419	0.55%	37
The Heinz Endowments	$93,124	$73,171,942	0.13%	N/A	$630,000	$63,349,098	0.99%	39
The Skillman Foundation	$393,000	$14,819,750	2.65%	N/A	$625,000	$17,177,500	3.64%	40
Burroughs Wellcome Fund	$45,000	$34,204,982	0.13%	N/A	$535,000	$32,232,653	1.66%	43
Ben and Catherine Ivy Foundation	$500,000	$9,018,974	5.54%	N/A	$500,000	$9,156,097	5.46%	45
Claude Worthington Benedum Foundation	N/A	N/A	N/A	N/A	$497,000	$16,875,606	2.95%	46
Altman Foundation	$565,000	$10,603,000	5.33%	N/A	$470,000	$10,960,350	4.29%	47
The Andrew W. Mellon Foundation	$247,000	$233,383,289	0.11%	N/A	$448,500	$264,020,287	0.17%	48
The Marcus Foundation	N/A	N/A	N/A	N/A	$445,778	$65,397,395	0.68%	49
The Noyce Foundation	$62,961	$14,741,659	0.43%	N/A	$413,014	$48,398,332	0.85%	50

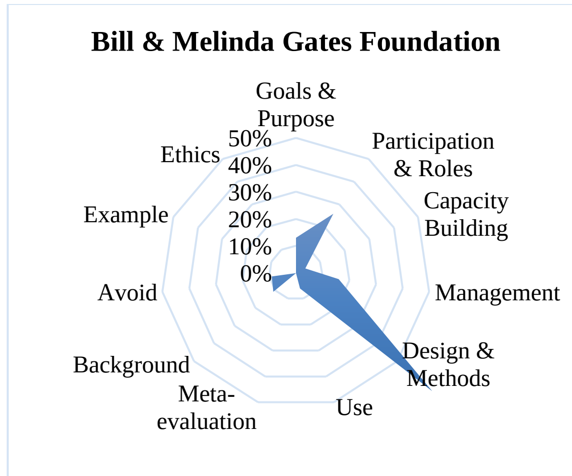
Appendix B.
Links to evaluation policies included in the present study

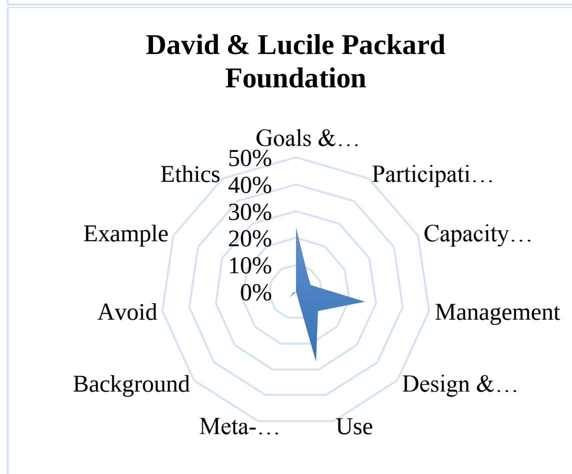
Foundation	Policy Document Title	Link
Bill & Melinda Gates Foundation	Evaluation Policy	https://tinyurl.com/y5zner8a
California Wellness Foundation	An Approach to Evaluation: Principles for TCWF's Evaluation Activities	https://tinyurl.com/yxgt97ft
David & Lucile Packard Foundation	Guiding Principles and Practices for Monitoring, Evaluation, and Learning	https://tinyurl.com/y55ont6d
Doris Duke Charitable Foundation	Grant-Making Process: Evaluation	https://tinyurl.com/yxd68l2n
Duke Endowment	Guiding Principles: Learning and Evaluation	https://tinyurl.com/y3ujcu83
James Irvine Foundation	Evaluation Policies and Guidelines: Approach to Evaluation	https://tinyurl.com/y5u83xvs
John D. & Catherine T. MacArthur Foundation	Our Approach to Evaluation	https://tinyurl.com/yyjxeqpk
Kansas Health Foundation	Evaluation/Learning: Our Approach	https://tinyurl.com/yyt47zsz
Rockefeller Foundation	Our Strategy: Monitoring and Evaluation	https://tinyurl.com/yytuw6w7
Skillman Foundation	Evaluation Philosophy and Approach	https://tinyurl.com/y4ml754a
Wallace Foundation	How We Work: The Wallace Approach	https://tinyurl.com/y57aeqlm
William and Flora Hewlett Foundation	Evaluation Principles and Practices	https://tinyurl.com/y5mpbgbf

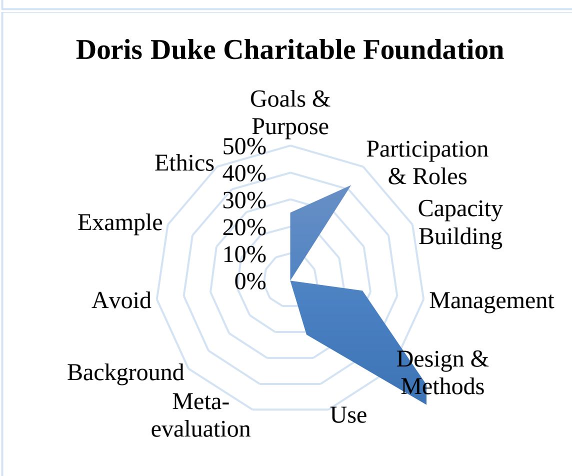
Appendix C.
Spider charts representing distribution of evaluation policy components

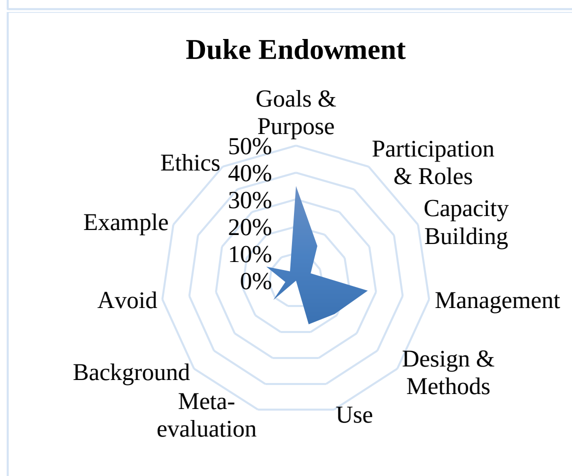
(Data summarized in Table 13)

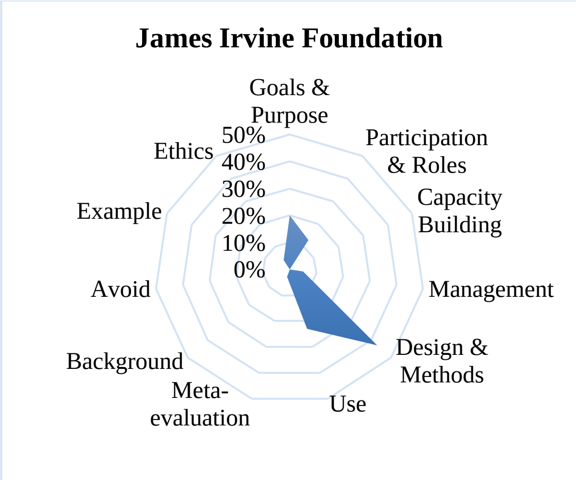


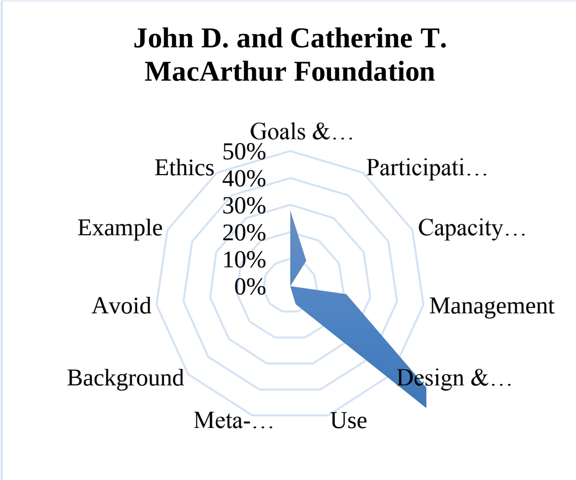


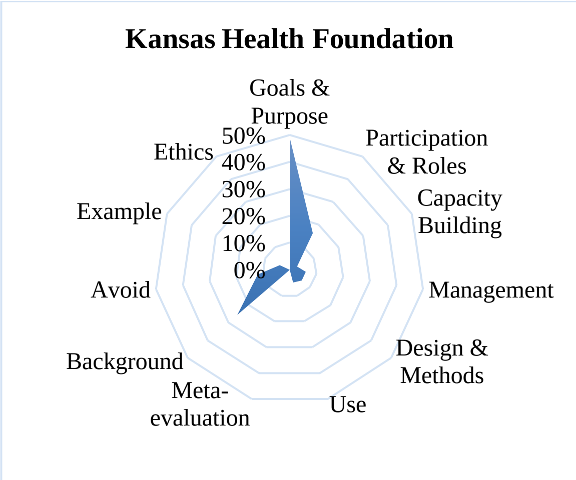


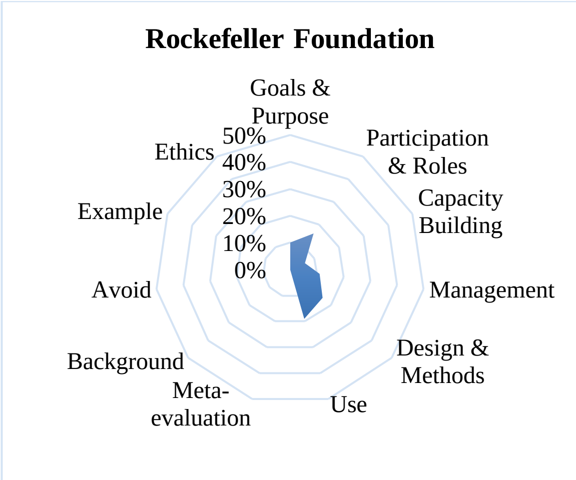


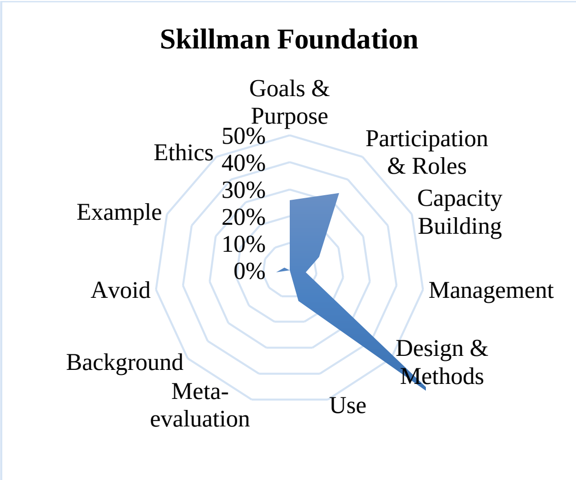


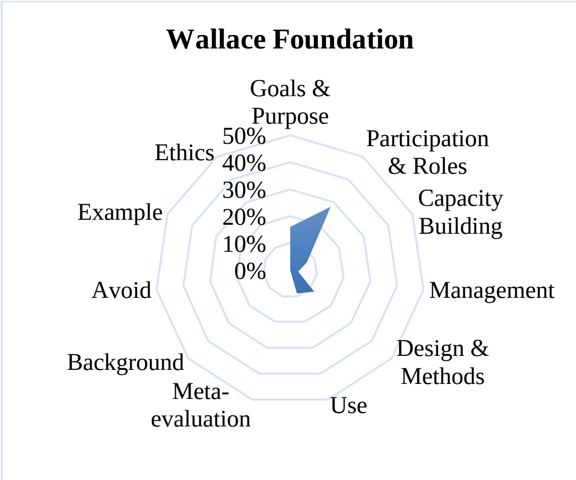

James Irvine Foundation
Goals & Purpose
Participation & Roles
Capacity Building
Management
Design & Methods
Use
Meta-evaluation
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%

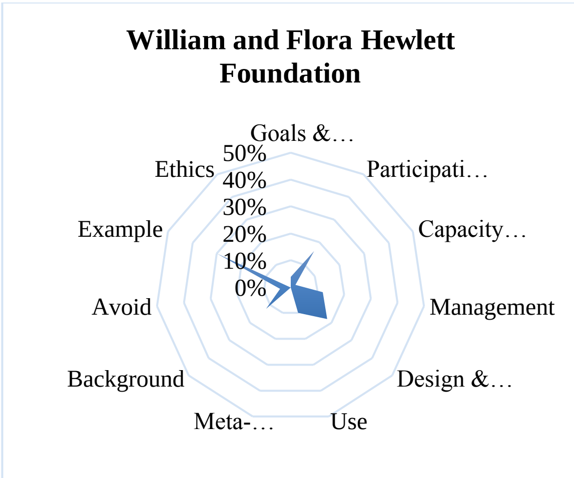

John D. and Catherine T. MacArthur Foundation
Goals &…
Participati…
Capacity…
Management
Design &…
Use
Meta-…
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%


Kansas Health Foundation
Goals & Purpose
Participation & Roles
Capacity Building
Management
Design & Methods
Use
Meta-evaluation
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%


Rockefeller Foundation
Goals & Purpose
Participation & Roles
Capacity Building
Management
Design & Methods
Use
Meta-evaluation
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%


Skillman Foundation
Goals & Purpose
Participation & Roles
Capacity Building
Management
Design & Methods
Use
Meta-evaluation
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%


Wallace Foundation
Goals & Purpose
Participation & Roles
Capacity Building
Management
Design & Methods
Use
Meta-evaluation
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%


William and Flora Hewlett Foundation
Goals &…
Participati…
Capacity…
Management
Design &…
Use
Meta-…
Background
Avoid
Example
Ethics
50%
40%
30%
20%
10%
0%

Appendix D.
Recruitment letter

Hi (insert name) -

I am a fourth year PhD student at UCLA in Social Research Methodology with a focus on program evaluation. My dissertation is a study of how 12 leading foundations, including (insert: name of foundation), are organizing and implementing their evaluation practice.

I recently defended my dissertation proposal and have started conducting interviews. To date, I have spoken with your colleagues at the Gates Foundation, Hewlett Foundation, Duke Endowment and a few others.

Are you available in the next few weeks to speak for about an hour to talk about evaluation practice at your foundation?

I am happy to share the interview questions in advance. Generally, I am interested in the rules, policies, guidelines, or principles that guide decisions and actions when conducting evaluation at foundations.

So far, I have had a chance to review your foundation's website (in particular the pages on your approach to evaluation and learning), as well as (insert names of additional documents). If there are other documents that may give me a better sense of how evaluation is organized at (insert: name of foundation), I would be happy to review those in advance of our call.

Please let me know if you have any questions about the dissertation or interview. I hope you will consider speaking with me for an hour in the next few weeks.

I look forward to hearing from you,
Alana Kinarsky

Appendix E.
Interview protocol shared with participants

(Name of Foundation): (Name of Evaluation Policy Document)

Background: Evaluation theorists argue that every entity that utilizes evaluation has implicit or explicit policies that influence how evaluation is conducted. Extant research has not revealed how widespread evaluation policies are nor how they differ across foundations. This dissertation seeks to understand variations in policies (rules, principles, or guidelines) across foundations to reveal important differences in factors such as requirements for data collection, what counts as proof for efficacy, or how evaluation data is used to inform decision making. This research will afford foundations the opportunity to review, streamline, and coordinate evaluation policies. The guiding research questions, which are arranged in the order of my inquiry, are:

1. To what extent do leading US foundations have publicly available evaluation policies?
2. How are publicly published evaluation policies of leading US foundations similar and different in terms of their goals, purpose, focus and specificity?
3. To what extent are aspects of evaluation practice that are identified in the evaluation theory literature evident in evaluation policies?
4. What do evaluation and research directors of foundations view as the primary goals of their foundation's policy?
5. According to evaluation and research directors, in what ways do the foundation's formal and informal evaluation policies guide evaluation practice?

Intro

1. Tell me a bit about your role at the foundation.
   How long you have been there?
   What led you to this role?

Establishing an Evaluation Policy

1. Why did the foundation decide to draft (name of policy)?
2. In your experience, how has standardizing evaluation practice across the foundation helped in addressing this problem?
3. Is this a problem the foundation still faces?

Impact of Evaluation Policy

1. Tell me a little about how, if at all, the (evaluation policy document) impacts and informs your work at the foundation.
   Program staff? Board? External contractors/evaluators?
2. How frequently, if at all, do you reference the document?
3. In particular, what aspects of evaluation practice have been impacted? (RFP process, spending, schedule)
4. How are the principles outlined in the document enforced across the foundation?
5. What have been the benefits of the policy? Shortcomings?

Role of Evaluation Theory in Developing Evaluation Policy

6. How often is evaluation theory discussed within the foundation? In your view, does it inform evaluation practice?
7. How do you know when an evaluation commissioned by the foundation is "good" or "bad"?

Broader Questions

8. Beyond the written document, what guides evaluation practice at the foundation?
9. How has evaluation practice changed in the foundation since you have been there/ since the policy was developed?

Closing:

10. Is there anyone else in the org I should speak with to learn more?
11. Any additional documents that you think would help me better understand the foundation's approach to evaluation?

Appendix F.
Sample member check document

Any amendments to the following table that details your job title, the year you joined the foundation and a few key facts about your educational background and professional experience?

Job Title	Year joined foundation	Relevant Background
		Education: Professional:

Any amendments to the following information about the policy document or evaluation at the foundation? Has the evaluation policy been updated since we last talked?

Policy document:	
Year developed:	
Written by:	
Document updated since 1/2018?	
Info on evaluation unit at foundation	

The next table has quotes from our conversation and my brief interpretation or explanation of the quote. The quotes have been slightly edited to make them easy to read, but I tried to not alter the meaning at all.

If the quote or the interpretation seem inaccurate to you, please let me know. You can add a comment to the far-right column, or just email back your response.

Edited quote from interview	AK's Interpretation	Comments?
(Quote from interview)	Evaluation policies aim to make explicit the ways in which a foundation's culture and values impact the practice of evaluation.	

(optional) During our conversation you mentioned a few things that were in progress. Please let me know if you have any updates you can share on the following. You can add a comment to the far-right column, or just email back your response.

Follow up on…	AK's Notes	Comments?
(quote from interview)	Can you share an update on this document? Has this been completed?	

References

Al Hudib, H. (2018). *The role of evaluation policy in organizational capacity to do and use evaluation* (dissertation).

Alkin, M. C., & Christie, C. A. (2004). *An evaluation theory tree*. In M. C. Alkin (Ed.), Evaluation roots (pp. 12–65). Thousand Oaks, CA: SAGE Publications.

Alkin, M. C. (2013). *Evaluation roots: A wider perspective of theorists' views and influences*. Los Angeles: SAGE.

American Evaluation Association Evaluation Policy Task Force. (2007). Retrieved from: http://www.eval.org/p/cm/ld/fid=151http://www.eval.org/p/cm/ld/fid=151.

American Evaluation Association. (2008). Retrieved from: http://archive.eval.org/eval2008/

American Evaluation Association Guiding Principles for Evaluators. (2018). Retrieved from: https://www.eval.org/p/cm/ld/fid=51

Arbreton, A., Trivedi, P., and Twersky, F. (2018). *The value of our evaluations: assessing spending and quality*. Retrieved from: https://www.hewlett.org/wp-content/uploads/2018/02/Evaluation-Quality-and-Spending-Analysis-2018.pdf.

Bardach, E., & Patashnik, E. M. (2016). *A practical guide for policy analysis: The eightfold path to more effective problem solving*. Washington, D.C.: CQ Press.

Bazeley, P. (2013). *Qualitative data analysis: Practical strategies*. SAGE.

Behrens, T. R., & Kelly, T. (2008). Paying the piper: foundation evaluation capacity calls the tune. *New Directions for Evaluation*, 2008(119), 37-50. doi:10.1002/ev.267.

Brest, P. (2012). *A decade of outcome-oriented philanthropy (SSIR)*. Retrieved from SSIR: https://ssir.org/articles/entry/a_decade_of_outcome_oriented_philanthropy.

Bullock, C. S., Anderson, J. E., & Brady, D. W. (1983). *Public policy in the eighties*. Monterey, CA: Brooks/Cole Pub. Co.

Buteau, E., Coffman, J., Loh, C., Glickman, J., & Beer, T. (2016). *Benchmarking foundation evaluation practices*. Retrieved from: http://research.cep.org/benchmarking-foundation-evaluation-practices.

Buteau, E., & Glickman, J. (2018). *Understanding & sharing what works: the state of foundation practice*. The Center for Effective Philanthropy.

Carl, A. (2017). *Gates and Rotary announce $450 million commitment to end polio*. Retrieved from: https://www.rotary.org/en/press-release-gates-and-rotary-announce-450-million-commitment-end-polio

Carman, J. G., Fredericks, K. A., & Introcaso, D. (2008). Government and accountability: paving the way for nonprofits and evaluation. *New Directions for Evaluation*, 119, 5-12. San Francisco, CA: Jossey-Bass.

Chelimsky, E. (2012). Balancing Evaluation Theory and Practice in the Real World. *American Journal of Evaluation*, 91-98.

Christie, C.A. (2003). What Guides Evaluation? A Study of How Evaluation Practice Maps onto Evaluation Theory. *New Directions for Evaluation,* 7-36.

Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, 34, 131–135.

Christie, C. A., & Alkin, M. C. (2012). *An evaluation theory tree*. In M. C. Alkin (Ed.), Evaluation roots (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Christie, C. A., & Fierro, L. A. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38(2), 65–72. https://doi.org/10.1016/j.stueduc.2012.05.003.

Christie, C. A., & Lemire, S. T. (2019). Why Evaluation Theory Should Be Used to Inform Evaluation Policy. *American Journal of Evaluation.*

Cochran, C. L., & Malone, E. F. (2010). *Public policy: perspectives and choices*. Boulder, CO: Lynne Rienner.

Coffman, J., Beer, T., Patrizi, P., & Heid Thompson, E. (2013). Benchmarking evaluation in foundations: do we know what we are doing? *The Foundation Review, 5*(2), Article 5, 36-51.

Datta, L. (2009). Golden is the sand: Memory and hope in evaluation policy and evaluation practice. *New Directions for Evaluation*, 2009(123), 33–50.

Dillman, L. (2013). Evaluator Skill Acquisition: Linking Educational Experiences to Competencies. *American Journal of Evaluation,* 270-285.

Dillman, L. M. (2014). *Alignment between intention and implementation: A case study of the Robert Wood Johnson Foundations evaluation policies* (dissertation).

Dillman, L. M., and Christie, C. A. (2016). Evaluation policy in a nonprofit foundation. *American Journal of Evaluation, 38*(1), 60-79.

Eisenberg, N., Winters, L. and Alkin, M. C. (2005), The case: Bunche–Da Vinci Learning Partnership Academy. *New Directions for Evaluation*, 2005: 5–13. doi:10.1002/ev.148

Fleishman, J. L. (2009). *The foundation: a great American secret; how private wealth is changing the world*. New York, NY: Public Affairs.

Frumkin, P. (1998). The long recoil from regulation. *The American Review of Public Administration, 28*(3), 266-286.

H.R. 4174 - Foundations for Evidence-Based Policymaking Act of 2018, 115th Cong. (2019) (enacted). Retrieved from: https://www.congress.gov/bill/115th-congress/house-bill/4174/text

Internal Revenue Service. (2017). Tax Information for Private Foundations. Retrieved from https://www.irs.gov/charities-non-profits/private-foundations.

Kramer, M.R., & Bickel, W.E. (2004). Foundations and evaluation as uneasy partners in learning. In Braverman, M. T., Constantine, N., & Slater, J. K. (Eds.). *Foundations and evaluation: Contexts and practices for effective philanthropy (51-75)*. San Francisco: Jossey-Bass.

Liket, K. C., Rey-Garcia, M., & Maas, K. E. (2014). Why Aren't Evaluations Working and What to Do About It. *American Journal of Evaluation*, 35(2), 171-188. doi:10.1177/1098214013517736.

Leeuw, F. L. (2009). Evaluation policy in the Netherlands. *New Directions for Evaluation*, 2009(123), 87–102.

Mark, M. M., Cooksy, L. J., and Trochim, W. M. K. (2009). Evaluation policy: An introduction and overview. *New Directions for Evaluation*, 2009(123), 3–11.

Mattessich, P. W., Mueller, D. P., and Holm-Hansen, C. A. (2009). Managing evaluation for program improvement at the Wilder Foundation. *New Directions for Evaluation, 121*, 27–42. San Francisco, CA: Jossey-Bass.

McKeever, B. (2017). *The nonprofit sector in brief 2015: public charities, giving, and volunteering*. Retrieved from the Urban Institute: https://www.urban.org/research/publication/nonprofit-sector-brief-2015-public-charities-giving-and-volunteering.

Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Munger, M. C. (2000). *Analyzing policy: choices, conflicts and practice*. New York: W W Norton.

Ostrower, F. (2004). *Attitudes and practices concerning effective philanthropy: survey report*. Retrieved from the Urban Institute: https://www.urban.org/sites/default/files/publication/58076/411067-Attitudes-and-Practices-Concerning-Effective-Philanthropy.PDF.

Patton, M. Q. (2018). *Principles-focused evaluation: The guide*. New York, NY: The Guilford Press.

Preskill, H., & Mack, K. (2013). *Building a Strategic Learning and Evaluation System for Your Organization*. Retrieved from FSG: :https://www.fsg.org/publications/building-strategic-learning-and-evaluation-system-your-organization

Prewitt, K. (2006). Foundations. In W. W. Powell and R. Steinberg (editors), *The nonprofit sector: a research handbook* (pp. 355-376). New Haven, CT: Yale University Press.

Reich, R. (2018). *Just Giving: Why Philanthropy Is Failing Democracy and How It Can Do Better*. Princeton, NJ: Princeton University Press.

Robert Wood Johnson Foundation (2017). *Our history*. Retrieved from http://www.rwjf.org/en/library/collections/our-history.html.

Rumelt, R. P. (2017). *Good strategy, bad strategy: The difference and why it matters*. London: Profile Books.

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Los Angeles, CA: Sage.

Schwandt, T. (2013). On the Mutually Informing Relationship Between Practice and Theory in Evaluation. *American Journal of Evaluation, 35*(2), 231-236.

Seidman, I. (2013). *Interviewing as qualitative research*. New York City: Teachers College Press.

Senge, P. M. (2006). *The fifth discipline: The art and practice of the learning organization*. New York, Ny: Doubleday/Currency.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of Program Evaluation: Theories of practice*. Newbury Park: Sage.

Shadish, W., & Epstein, R. (1987). Patterns of Program Evaluation Practice Among Members of the Evaluation Research Society and Evaluation Network. *Evaluation Review*, 555-590.

Stern, E. (2009). Evaluation policy in the European Union and its institutions. *New Directions for Evaluation*, 2009(123), 67–85.

The Foundation Center. (2014). *Key facts on U.S. foundations: 2014 edition*. Retrieved from http://foundationcenter.org/gainknowledge/research/keyfacts2014/index.html.

The Foundation Center. (2017). *2012 foundation stats.* Retrieved from http://data.foundationcenter.org/about.html#quick-start).

Tompkins-Stange, M. E. (2016). *Policy patrons: philanthropy, education reform, and the politics of influence*. Cambridge, MA: Harvard Education Press.

Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation, 123*, 13–32.

Troyer, T. A. (2000). The 1969 private foundation law: historical perspective on its origins and underpinnings. *The Exempt Organization Tax Review, 27*(1), 52-65.

U.S. Government Accountability Office. (2014). *Tax-exempt organizations: better compliance indicators and data, and more collaboration with state regulators would strengthen oversight of charitable organizations*. Retrieved from: https://www.gao.gov/assets/670/667595.pdf.

US Treasury Dept. (1965). Report on Private Foundations: Committee on Finance United States Senate (Rep.). Washington: U.S. Government Printing Office. Retrieved from: https://www.finance.senate.gov/imo/media/doc/SPrt101.pdf.

Wadsworth, H. C. (1975). Private foundations and the Tax Reform Act of 1969. *Law and Contemporary Problems, 39*(4), 255-262. Retrieved from https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=3467&context=lcp.

William T. Grant Foundation (n.d.). *History*. Retrieved from: http://wtgrantfoundation.org/history.