

The importance of implementation: Putting evaluation policy to work

Leslie A. Fierro¹  | Alana R. Kinarsky²  |
Carlos Echeverria-Estrada³  | Nadia Sabat Bass²  |
Christina A. Christie² 

¹Max Bell School of Public Policy, McGill University, Montreal, Canada

²University of California, Los Angeles, California, USA

³NORC at the University of Chicago, Illinois, USA

Correspondence

Leslie A. Fierro, Max Bell School of Public Policy, McGill University, Montreal, Canada.
Email: leslie.fierro@mcgill.ca

Abstract

Federal agencies are increasingly expected to write and implement guidance for program evaluation, also known as evaluation policies. The Foundations for Evidence-Based Policymaking Act required such policies for some federal agencies, and guidance from the White House Office of Management and Budget outlined an expectation that all agencies develop evaluation policies. Before these expectations, many federal agencies were already developing such policies to suit organizational needs and contexts. This chapter details findings from interviews with stakeholders at ten federal agencies and offices that developed and implemented evaluation policies before enacting the Foundations for Evidence-Based Policymaking Act. These organizations represent early adopters of evaluation policies that can support future guidance and implementation of evaluation frameworks and capacity building in government. The study provides insight into the breadth and depth of the various strategies they used as well as their experiences with implementation.

Much has been posited in *New Directions for Evaluation*—in both the current issue and previous issues (Mark, Cooks, & Trochim, 2009)—and in the recent and growing body of empirical studies (Al Hudib, 2018; Al Hudib & Cousins, [this issue](#); Christie & Lemire, 2019; Dillman & Christie, 2017; Kinarsky, 2019; Kinarsky & Christie, 2021) about the promise of evaluation policy for promoting effective evaluation practice within organizations. For instance, Trochim (2009) suggested that “Because evaluation policies typically apply across multiple evaluations, influencing policies directly may have systemic and far-reaching effects for practice” (p. 14). A decade later, Christie and Lemire (2019) called our attention to the potential promise of translating the prescriptions of evaluation theory into evaluation

The authors thank the Data Foundation for funding for open access.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](#) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *New Directions for Evaluation* published by American Evaluation Association and Wiley Periodicals LLC.

practice through policy, noting that evaluation policies could “serve as a guiding mechanism for evaluation practice” (p. 494).

The importance of evaluation policy extends into other domains of evaluation scholarship, featuring prominently in the literature on evaluation capacity building. In 2008, Preskill and Boyle published their multidisciplinary model of evaluation capacity building. Here, evaluation policies and procedures played a key role in fostering sustainable evaluation practice within organizations by enabling the effective transfer of learning about evaluation through evaluation capacity-building approaches in day-to-day organizational activities. Preskill and Boyle posited that the “evaluation policies and procedures an organization develops are intended to provide the structures and means by which evaluation becomes *institutionalized* in the organization” (Preskill and Boyle, 2008, p. 454, emphasis added).

Likewise, Bourgeois and Cousins (2013) provided empirical support for the importance of evaluation policies for an organization’s capacity to do evaluation. They found that federal agencies in Canada that demonstrated exemplary capacity to do evaluation were typified in part by supportive organizational infrastructure that included evaluation policies. Agencies with this exemplary capacity had not only developed but also implemented “organizational policies on evaluation and performance measurement” (p. 308). This contrasts with federal agencies that had intermediate, developing, or low capacity to do evaluation in which no policies on evaluation were established.

STUDYING IMPLEMENTATION IN CONTEXT

For organizational evaluation policies to realize their potential, they must be effectively socialized and *implemented* within an organization. Despite implementation being the crux of evaluation policy success, little has been written on this component of the evaluation policy lifecycle. Dillman (2014) and Dillman and Christie (2017) examined the implementation of evaluation policies within the Robert Wood Johnson Foundation. They found that videos and documents were used to explain evaluation policy to internal and external stakeholders. As a result, stakeholders had a shared understanding of the evaluation goals, but there was less alignment on how to implement evaluation policies across programs. Furthermore, even though the foundation did not enforce its policies, it still had a meaningful influence on evaluation practice.

Studies examining the implementation of government legislation that focused on evaluation—such as the Government Performance and Results Act of 1993 and the No Child Left Behind Act of 2001—do exist (Christie & Fierro, 2012); to our knowledge, however, few studies have been published that examine the implementation of *organizational-level* evaluation policies in the public sector (Al Hudib & Cousins, [this issue](#)). In light of the recent passage of the Foundations for Evidence-Based Policymaking Act (the Evidence Act), which mandates that the 24 departments/agencies subject to the Chief Financial Officers Act develop evaluation policies, examining the implementation of such policies is both relevant and timely (Evidence Act, 2019). Thus, we share findings from a research study that explored the initial stages of implementing evaluation policies within US federal agencies and offices, focusing on agencies that developed such policies before implementing the Evidence Act.

METHODS

Sample

The first step in this study was conducted between January and December 2018, and it consisted of locating existing evaluation policies in US government agencies. We

downloaded for content analysis evaluation policies that generally conformed to the definition provided by Trochim (2009)—“any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation” (p. 13). Since the term *policy* frequently has a very specific meaning within the context of government, documents did not necessarily include the term in their names; documents with titles that included *procedure*, *recommendations*, *guidance*, *principles*, or *framework* also qualified for inclusion. Educational documents that were developed with the intention of building evaluation capacity among grantees were excluded.

Though the team took a multipronged approach to identify the evaluation policies included in this study, all except one (which we were made aware of during the 2019 American Evaluation Association Conference following a presentation of preliminary findings from this study) stemmed from a systematic review of US federal agency websites. All US federal agencies and offices included in the Survey of Federal Offices—conducted by the US Commission on Evidence-Based Policymaking (2017)—were eligible for the review. This list comprises 31 agencies (11 of which are “single bureau agencies and commissions”) and 198 offices nested within 20 of these agencies. Team members used multiple methods for the website search and documented findings in a brief summary for each agency and office; we also archived documents, screenshots, videos, and any audiovisual information that constituted direct evidence of the agency or office’s evaluation policy to a team repository shared on DropBox.

Our team identified a total of 17 evaluation policies that were written in 2018 or earlier. These policies are sometimes applicable across an entire federal agency ($n = 7$) or office ($n = 6$); in other cases, the policy is applicable only to a relatively small unit within a federal agency or office (e.g., a specific evaluation division or program) ($n = 4$). We attempted to interview at least one individual within each of these agencies, offices, and units. We identified initial points of contact during the web search, which were documented in the summaries that team members developed. When points of contact were not evident as part of the search for evaluation policies, multiple team members conducted broader online searches for additional insights about possible points of contact. We could not identify a potential point of contact for only one of the agencies in which we located an evaluation policy.

Data collection

Between June and December 2019, the principal investigator disseminated email invitations to each identified point of contact to request participation in a 60-min telephone interview concerning the history of and general processes involved in creating the policy, how the content of the policy was developed, and what the policy adoption and implementation process had been to date. The invitation included a general description of the purpose of the study and mentioned that our team had identified them as “someone who may be knowledgeable about the evaluation policy.” The email included the title of the evaluation policy and provided the policy as an attachment. The email clarified that we did not intend to “rate” or “grade” existing policies, and the interview questions were included with the invitation for their information. If the contact felt they were not the most appropriate person to participate in the interview, we asked them to recommend current or prior employees and/or contractors to contact.

Ultimately, representatives from 11 of the 17 agencies and offices where we found evaluation policies participated in interviews. For the analysis presented in the current chapter, we removed one of these interviews because the discussion primarily focused on the

process involved in creating a learning agenda rather than on an evaluation policy as defined above. Therefore, the results presented in this chapter stem from a content analysis of 12 interviews representing ten agencies and offices. For seven of the agencies and offices, we spoke with a single individual. For the other three organizations, we conducted interviews with two individuals each, for a total of 13 individuals (note: one interview had two interviewees). Eleven of the individuals we interviewed were current employees, and two were past employees.

We developed a semi-structured interview protocol based on the study's research questions, primarily focusing on the motivation for developing an evaluation policy, policy content and development, and policy implementation. The semi-structured interview allowed for deviation from the protocol to "respond to the situation at hand" while keeping consistency across interviews (Merriam, 2009, p. 90). The principal investigator, who has a background of service with the US federal government, conducted all interviews. For the majority of interviews, another team member joined the conversation to take detailed notes. Interviews were digitally recorded with the permission of the interviewee and generally lasted 1 hour each. A professional transcription service provided verbatim transcripts for content analysis.

Data analysis

Once the interviews were transcribed, the data were imported into and analyzed using MaxQDA 2020, a qualitative and quantitative data analysis software (VERBI Software, 2021). First, we generated a priori codes based on the guiding research questions and findings from the analysis of evaluation policies (Saldaña, 2013). We then coded transcripts using initial coding to closely examine and compare data across interviews. During second-level coding, we utilized pattern coding to "attribute meaning" to the emerging findings (Saldaña, 2013, p. 209). During coding, we documented analytic memos of emergent findings and initial reflections in MaxQDA, and we reviewed these after the second cycle of coding was completed.

The final round of coding was conducted by two of the authors who worked independently but convened regularly over 2 months to discuss progress. After two policies were coded, the researchers gathered to discuss the policies and application of codes. In particular, we deliberated on the definition and boundaries of each code. We refined the codes before commencing with the remaining policies. Team coding ensured that the codes were consistently interpreted and applied (Bazeley, 2013; Saldaña, 2013). Once formal coding was completed, we generated summary grids in MaxQDA. We reviewed every coded segment and wrote short summary statements. We categorized emergent themes by the types of activities the agencies and offices conducted for policy implementation and then reviewed the coded segments to determine whether they fit within each of those implementation activities. Finally, we tabulated the presence or absence of each activity type in the interview narrative of a particular agency or office.

FINDINGS

The findings section outlines the different activities undertaken to implement evaluation policies across the federal agencies and offices included in our sample. We begin with an overview of implementation activities and a discussion of the most common implementation efforts mentioned by interviewees, providing quotes that highlight these activities. The

numbers in parentheses capture how many of the ten agencies and/or offices expressed the activity. The findings section concludes with three overarching categories that summarize agencies and offices by type of implementation.

Implementation activities

We broadly classified implementation activities according to whether they were conducted with a lens towards socializing individuals internal or external to the agency or office who may have responsibilities associated with implementing the evaluation policy. Our analyses surfaced seven distinct implementation activities: six were intended for stakeholders working internally to the organization, while three were intended for external audiences. The distribution of activities across agencies and offices is detailed in Table 1.

Internal audiences

The activities most commonly performed for internal audiences included conducting meetings with staff or leadership to discuss the evaluation policy once it was completed ($n = 7$) and featuring the evaluation policy in an interagency or interoffice blog or newsletter ($n = 4$). Representatives of between one and three agencies or offices described four other internally oriented activities: incorporating policy into new staff orientation, hosting trainings, creating office-related materials that noted the evaluation principles/standards, and offering individual consultations on the evaluation policy. The last two activities were unique to a single agency or office and are grouped in the “other” category (Table 1) to keep those organizations de-identified.

Meetings to discuss evaluation policy ($n = 7$)

Representatives from seven of the ten agencies and offices described sharing the completed evaluation policy with internal stakeholders, mainly through staff meetings. They all described efforts to distribute the policy to staff or leadership through meetings and webinars. In these instances, they prioritized making direct contact with staff to deliver and discuss the completed evaluation policy. For example:

There was a bunch of publicity about the evaluation policy. We went to every single [organizational unit] and asked to participate in their regularly scheduled senior leadership meetings. We gave a 30-min presentation and answered questions.... Almost everybody came away from those meetings more inspired and excited. Our evaluation staff was happy because somebody, who was not them, was meeting with leadership preaching what they've been saying for years, but not getting a lot of traction for [it].

Often, presentations on the policy were integrated into long-standing meetings and they created an opportunity for staff to discuss the purpose and process of conducting evaluation.

Most of the representatives from these seven agencies and offices talked about the need to develop an ongoing communications plan that would help institutionalize the evaluation policy. For example, one representative said, “We talked a lot about the evaluation policy in every learning agenda meeting and all the meetings in between. This is not just a once-a-year thing.” To ensure that the evaluation policy was understood and applied

TABLE 1 Policy implementation activities across 10 federal agencies and offices

Implementation activity	Federal agency/office										TOTAL (N = 10)
	Passive					Active					
	A	B	C	D	E	F*	G	H	I	J	
Internal audience											
Staff meetings include discussion of evaluation policy	1			1	1		1	1	1	1	7
Internal blogs/newsletters feature the policy							1	1	1	1	4
Staff are informed of the policy during new staff orientation						1			1	1	3
Training workshops/classes about evaluation policy are held								1		1	2
Other: Office-materials with policy language; one-on-one consultations are held about policy								1		1	2
External Audience											
Policy language is embedded into external documentation				1	1	1	1		1	1	6
Policy is shared with external stakeholders and other federal agencies/offices						1		1	1	1	4
Policy is shared via multimedia										1	1
TOTAL	0	0	1	2	2	3	3	5	5	8	8

*Two interviews were conducted with Agency F—one with an individual who was engaged in developing guidance that preceded and was used for the evaluation policy our team identified and one with an individual who came on board after the creation of the evaluation policy. Thus, we do not have definitive information about how the policy we surfaced was specifically developed. We may be missing some key activities as a result.

across the organization, leaders tried not to restrict discussion to one meeting. While some regularly discussed it across meetings, others found different ways to reinforce the information being shared. For example, representatives from one organization noted that they divided the policy into components and shared aspects of it with staff during meetings and over email. In addition to presentations at staff meetings, a weekly email—distributed across the organization—focused on a particular principle in the policy and reminded staff of what they had learned in the previous week:

We have disseminated it pretty assertively. Leadership really wanted to do coordinated communications within the [organization]. So, we sent a series of messages on each principle to the whole [organization]. Like, “Dear staff, here’s a reminder. We have an evaluation policy. This is the purpose of it. Next week we’ll tell you about rigor.” And then the next week we’d say, “Last week we talked about rigor. This week we’re going to talk about relevance.” We coordinated those messages with presentations at senior staff meetings.

Breaking the evaluation policy into sections reduced the burden staff would have borne if they had to review and make sense of the policy in its entirety. Instead, staff got short, educational insights weekly in their inbox, which were also discussed during senior staff meetings.

Agency/office blog or newsletter featuring the evaluation policy (n = 4)

Representatives from four agencies and offices reported that the policy was included on an internal blog or newsletter. Especially for large, dispersed organizations, the blog or newsletter was a way to reach staff who may not have been exposed to the policy in any other way: “At the time the evaluation policy came out, the agency had a blog where anyone in the agency, not just our office, could contribute an essay. So we did a blog on the evaluation policy.” For all these agencies and offices, the blog post was just one of several other activities intended to circulate the evaluation policy.

Information during new staff orientation (n = 3)

At three organizations, an introduction to the evaluation policy was included in new staff orientation, albeit differently. At one, a supervisor used the evaluation policy as an introduction to the organization’s evaluation approach and procedures: “[Our evaluation policy] gets a lot of exercise. I go over it with new staff in my office. I always brief new appointees about it.” At another, where new staff attend an orientation program, the policy was introduced during one of the required sessions. Finally, a representative from a third organization noted that the evaluation policy was one of the documents distributed to new staff when they joined. It was expected that new staff would become familiar with it, but there was not always an individual or group conversation about its significance.

Training workshops and classes about evaluation policy (n = 2)

Two organizations created hands-on workshops to help their staff understand the various components of the evaluation policy and how to apply it to their work. These organizations had existing learning infrastructure that was adapted to incorporate their evaluation policies. They developed new course materials to help internal stakeholders understand how the policy applied to their work. As one representative noted:

We have classroom courses and workshops on evaluation and performance monitoring....If we just had an evaluation policy that says “You need sound

program design, monitoring, and evaluation,” it would never happen. You have to tell people how to do it. Our job is to advise and help them do these things. That involves tools, one-on-one conversations, and consultations.

Providing in-depth learning opportunities to staff educated them on the evaluation policy and gave them an opportunity to reflect on how the policy shaped their evaluation practice. Doing such reflection while learning the tools to strengthen their evaluation practice offered staff an opportunity to align their practice with agency/office standards and expectations. Given the significant investment in training on evaluation broadly and the evaluation policy more specifically, it was perhaps not surprising to hear that staff at the organization described in the previous quote had “a high level of awareness of the evaluation policy.”

Other internal activities in brief

Two activities were unique to a single organization each. One had allocated resources and time to offering one-on-one consultations with staff about its evaluation policy. Another socialized staff to the evaluation policy through office-related materials that noted the evaluation principles/standards.

External audiences

Another potential means for implementing evaluation policy includes ensuring that external audiences, particularly evaluation contractors, are aware of applicable, relevant content from the policy. Compared to activities focused on internal audiences, fewer organizations—seven of the ten—intentionally carried out activities that could inform external audiences who may play a role in implementing the policy about its content. Across these federal agencies and offices, three activities were intended to reach these external stakeholders. The most common was embedding evaluation policy language into documentation intended for an external audience ($n = 6$). Agencies and offices also shared their evaluation policies with external audiences, including other federal agencies and offices ($n = 4$); one organization created multimedia presentations outlining key principles/standards from its evaluation policy. We discuss the two most popular external-facing activities in greater detail below.

Policy language embedded in external documentation (n = 6)

Representatives of six agencies and offices reported embedding aspects of their evaluation policies in key organizational documents intended for external audiences. In these instances, the ideas and principles prioritized in each evaluation policy were threaded into important documentation, such as monitoring and evaluation requests for proposals (RFPs) and grant announcements. Agency and office representatives wanted to create a shared understanding of their approach to evaluation to ensure more successful partnerships and evaluations. For example:

Our mission and principles are laid out in all of our external documents, including *Federal Register* notices and [our] annual notice to Congress....Doing so gave us a way to pull the community that we work with, not only within the agency but outside of the agency, in a common direction and say, “These are the things that are really important to us. If you’re going to work with us in partnership to do research, these things need to be part of the process.”

By being forthright about why the agency or office conducted evaluations and what they believed to be best practices, these organizations aimed to set realistic expectations with external stakeholders who would be consultants and partners in this work. Agencies and offices tended to have template contracts that got revised to accommodate each project being commissioned, but the sections describing the agency or office's approach to evaluation tended to remain consistent across contracts. This language was often modified to align with the text finalized in the evaluation policy.

Policy shared with external stakeholders and other federal agencies and offices (n = 4)

Beyond finding ways to embed policy language into documentation intended for external stakeholders, four agencies and offices intentionally shared their evaluation policies with external stakeholders. This looked different across organizations and served different purposes. In some instances, these efforts focused on offering counsel and insight to colleagues in other agencies and offices as they began crafting their evaluation policies. Agencies and offices looked at what their peers were doing when developing or revising their evaluation policies. We discuss this activity in more detail in a separate manuscript (Fierro et al., forthcoming) where we describe the creation and content of the evaluation policies in our study. However, we highlight this external-facing activity here as it can help harmonize the implementation of evaluation policies across federal agencies. This intent was raised by some interviewees who noted that they circulated their evaluation policies to the federal agencies and offices they worked closely with to help create a shared understanding of their approach to evaluation. As one representative noted, "We thought that having a coordinated set of plans and guidelines between the two agencies would be highly beneficial."

Evaluation policies were also shared beyond the federal government. Agency and office leaders found themselves sharing aspects of these policies during conference presentations and smaller gatherings of stakeholders. For example, one noted, "Sometimes I'm asked to speak at events about our evaluation policy. Other times, when I'm asked to speak about our work, the policy turns out to be a really easy framework to use."

Evaluation policies enabled federal agencies and offices to share internal practices and processes with external stakeholders and contribute to an evolving understanding of how evaluation should be implemented and used. As noted by one study participant, sharing externally can go beyond federal colleagues to those providing evaluation services, such as consultants or contractors:

All these resources are public facing, so they can be shared with implementers. We definitely see that happening. Just this morning, I was part of an implementers conference that one of our [agency units] was sponsoring. They had posters about monitoring and evaluation. One of the posters quoted the policy...An evaluation specialist was talking through it for contractors implementing the [agency unit's] programs. "This is our policy. This is what we need from you in order to implement this and to do good evaluations." I was really happy to see that.

Prevalence of implementation activities

To better understand the suite of activities agencies and offices may be performing to implement their evaluation policies, we collated the activities reported by representatives for each organization (Table 1). The average number of implementation activities reported per agency or office was three. The highest number of implementation activities for an

agency or office was eight, spanning internal and external audiences. In contrast, for two agencies/offices, no implementation activities were reported. To foster a deeper understanding of the constellation of implementation activities in which agencies and offices engaged, we describe them as undertaking *passive*, *active*, or *extensive* implementation efforts. We describe each in more detail below.

Passive implementation

One or no activities were taken to implement the evaluation policy in three organizations (Agencies A, B, & C). One organization made internal staff aware of the evaluation policy through typical distribution channels:

I made reference to it in various management team meetings to make sure people are aware it's up on our website. I referenced it in different contexts, like in a webinar series. Beyond that, it was made known to people in the same way that the other millions of updates get made: an email blast. It gets pushed out in various updates to [leadership], which is our usual internal mechanisms for raising awareness. Whether or not those are successful, I don't know.

In each of these instances, the evaluation policy was shared passively, meaning that they all had posted the evaluation policy to their websites or casually shared information about the policy through typical channels within the agency or office. This created a situation where the policy was accessible but the initiative to view it would have to come from internal and external stakeholders.

Active implementation

In the four organizations (Agencies D, E, F & G) where active policy implementation occurred, there was a greater expenditure of time and resources towards implementation than in the passive implementation scenarios. Most mentioned engaging staff and leadership throughout the organization to discuss the policy, and they all went one step further to familiarize internal or external stakeholders with it. Internally, they familiarized colleagues with the evaluation policy in various ways, including announcements of the policy or components through socialization mechanisms (such as a blog or newsletter) or mechanisms to ensure new hires were familiar with the policy. In all agencies and offices with active implementation efforts, we saw clear attempts to ensure the principles reflected in the evaluation policies were shared with external audiences through integration into funding documentation, such as RFPs. These efforts made it more likely that the principles articulated in the evaluation policies may impact practice.

Despite more progressive movement towards policy implementation than described in the organizations where passive implementation efforts occurred, some of the agencies and offices we classified as "active" felt their implementation efforts fell short, suggesting there was room to grow. One indicated the need to translate the principles articulated in the evaluation policy into standard operating procedures:

We share the institutional knowledge that these are our values. We have a shared understanding of what rigor means. We have a shared understanding of what transparency means and how that looks. But we have not done the best job of actually writing that down. Many people in the organization today are

founders. They've been with us since [we] began. But over the coming decades, they'll be retiring. They'll be moving on. And we don't have standard operating procedures where these things are talked about in detail.

Another suggested that more could have been accomplished with the policy if they had effectively leveraged the early support provided by their leadership:

If you've got leadership backing and support for doing this, really push the heck out of it. Try to move the ball across the goal line while you've got people's attention. Also, early on, create an instant test mechanism for reviewing progress to see what progress you're making. The initial goal had been to do a baseline survey really early in the game so that a couple years later we can see if the [evaluation policy itself] had any impact.

Extensive implementation

Three organizations (Agency H, I & J) intentionally engaged in five or more efforts to integrate the content of the evaluation policy into their evaluation practice. These organizations went well beyond the activities seen in the "active implementation" category – especially with respect to internal audiences who were engaged in a suite of activities to ensure familiarity with the evaluation policy. In one organization, staff and leadership had opportunities to become familiar with the policy through general staff meetings following its adoption. They were also socialized to it as it was integrated into workshops, courses, staff handbooks, and newsletters and other organizational announcements. Moreover, they fostered communities of practice. Their efforts intentionally reached beyond the headquarters location to staff located elsewhere geographically.

Agencies and offices with extensive implementation efforts threaded the ideas in their evaluation policies into important internal and external documentation, such as evaluation RFPs, staff handbooks, and new hire orientations. They took extensive efforts to create "wraparound" activities to educate their colleagues inside the organization about evaluation (in alignment with the evaluation policy principles) through guidance, training, and other tools. In at least one instance, they even shared the principles articulated with potential future evaluation consultants. Such efforts have the potential to weave the ideas articulated in the evaluation policy seamlessly into evaluation procedures and practice. For example, as one representative explained, "We have disseminated [our evaluation policy] pretty assertively.... We've done a lot to get it out there, internal to our office, our agency, and then externally to various stakeholders. I have been surprised at how much attention it's gotten."

STRENGTHS AND LIMITATIONS OF THE STUDY

One strength of this study is that, at least to our knowledge, it is the first empirical examination of the implementation of evaluation policies within federal agencies and offices to date. Additionally, despite being unable to secure participation from a representative within each agency or office where we identified an evaluation policy, our overall acceptance rate for interviews was quite high. At least one individual from ten of the 15 agencies and offices where we were able to find an initial point of contact agreed to an interview—a 67% response rate. Most individuals with whom we spoke were pleased to participate, perhaps in the hopes that others would benefit from their experiences, given that so many of them had learned extensively from their peers in developing their evaluation policies.

One limitation of the study is that the insights provided were self-reported, and for some interviewees years had elapsed since policy creation and implementation. We did not corroborate their insights through, for instance, a review of supporting documents or interviews with others engaged in policy drafting or implementation efforts. Additionally, the activities in Table 1 emerged from the interviews rather than from a set of predefined categories that we specifically asked about during the interviews. It is highly likely that some agencies and offices engaged in the activities represented in Table 1 but did not mention them during their interviews. Nevertheless, the list in Table 1 is a comprehensive depiction of implementation activities we heard about during interviews with agency representatives.

CONCLUSIONS

Decades of research (e.g., Pressman & Wildavsky, 1984; Waterman & Meier, 1998) indicate the fundamental importance of policy implementation in the broader realm of public policy as well as the intricacies and difficulties with moving policy language into action. Despite these long-standing concerns, it can be difficult for practitioners to know what options are available when attempting to move the ideas articulated in a written policy into practice. By examining the implementation of evaluation policies across the federal government (agencies, offices, and organizational units within), we offer current federal officials ideas for future implementation strategies.

Understanding various implementation options is particularly relevant at this moment. In 2019, the Evidence Act mandated that 24 federal departments/agencies write and implement evaluation policies. In March 2020, the Office of Management and Budget (OMB) issued a memorandum outlining program evaluation standards to guide the development of these policies (Vought, 2020). OMB expanded on these standards in a June 2020 memorandum that detailed why and how federal agencies should build a culture and infrastructure dedicated to evidence-based decision making (M-21-27). As agencies put the final touches on their evaluation policies and aim to “ensure that evaluation activities adhere to an evaluation policy and that stakeholders are aware of its content and use” (Vought, 2020, p. 5), it is critical that public servants carefully consider what these implementation strategies will contain. The evaluation community, more broadly, can assist by providing forums for discussions across agencies and disseminating findings from research and evaluation studies of implementation efforts.

As seen in this study, there are several implementation options, not only in terms of the breadth and depth of activities that can be used to socialize internal and external stakeholders to the content of an evaluation policy but also in terms of how to integrate the principles articulated in the policy into everyday evaluative functions. For an evaluation to reflect the principles included in an evaluation policy (e.g., relevance and utility, rigor, independence and objectivity, transparency, ethics; Vought, 2020), program and contracting officers commissioning the evaluation—as well as evaluators who design and perform it—need to be aware of the key features of that policy. Socializing agency staff and leadership to these key features by engaging them in existing meetings and other forms of regular communications (e.g., blogs, newsletters); familiarizing new hires with the evaluation policy as part of onboarding; including information about the evaluation policy in training materials, technical assistance, and workshops; and integrating language from the evaluation policy into externally facing documents, such as RFPs for evaluation services, would go a long way to moving principles into practice.

Although our study findings provide a good starting point for discussion, it is important to consider that the evaluation policies examined in this study were developed

before implementing the Evidence Act. As a result, the activities described in this chapter demonstrate the implementation steps agencies and offices took *before* evaluation policies became a requirement. There will inevitably be much more to learn about implementation as agencies move forward in fulfilling the requirements of the Evidence Act.

Future scholars might conduct a similar study to examine implementation efforts after the Evidence Act. For instance, do the Evidence Act and associated OMB guidance help agencies see more clearly how to translate evaluation policy into evaluation practice through (formal and informal) organizational processes and documents that prescribe practice, such as RFPs, learning agendas, and annual evaluation plans? Given that prior research has demonstrated evaluation policies can be inconsistently understood (Dillman, 2014; Dillman & Christie, 2017), future scholarship might also examine how stakeholders interpret and use agency evaluation policies and the role of specific implementation activities in clarifying these interpretations (Christie & Fierro, 2012). Such lines of research have the potential to inform stronger implementation of evaluation policies — and improved implementation of evaluation policies should increase the likelihood that their potential benefits are achieved.

ACKNOWLEDGEMENTS

We would like to acknowledge the efforts of others on our team who made this research study possible, including Emi Fujita-Conrads, Ann Marie Castleman, Minji Cho, Piper Grandjean Targos, and Kandice Ocheltree. In addition, we thank Mel Mark and Nick Hart for their advice, insights, and consultation as we considered how to share the findings from this work.

ORCID

Leslie A. Fierro  <https://orcid.org/0000-0001-7236-1022>

Alana R. Kinarsky  <https://orcid.org/0000-0002-2731-7724>

Carlos Echeverria-Estrada  <https://orcid.org/0000-0002-8185-749X>

Nadia Sabat Bass  <https://orcid.org/0000-0002-0569-0755>

Christina A. Christie  <https://orcid.org/0000-0002-2842-5006>

REFERENCES

- Al Hudib, H. (2018). *The role of evaluation policy in organizational capacity to do and use evaluation*. Ottawa, Ontario: [Doctoral dissertation, University of Ottawa]. uO Research. Retrieved from: https://ruor.uottawa.ca/bitstream/10393/38117/1/Al_Hudib_Hind_2018_thesis.pdf
- Al Hudib, H., & Cousins, J. B. (2022). Evaluation policy and organizational evaluation capacity building: A study of international aid agency evaluation policies. *New Directions for Evaluation*, 2022.
- Bazeley, P. (2013). *Qualitative data analysis: Practical strategies*. Thousand Oaks, California: SAGE Publications.
- Bourgeois, I., & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation*, 34(3), 299–319. <https://doi.org/10.1177/1098214013477235>
- 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*, 40(4), 490–508. <https://doi.org/10.1177/1098214018824045>
- Commission on Evidence-Based Policymaking. (2017). The promise of evidence-based policymaking. Retrieved from <https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2019/03/Full-Report-The-Promise-of-Evidence-Based-Policymaking-Report-of-the-Commission-on-Evidence-based-Policymaking.pdf>
- Dillman, L. M. (2014). *Alignment between intention and implementation: A case study of the Robert Wood Johnson Foundation's evaluation policies*. Los Angeles, California: [Doctoral dissertation, University of California, Los Angeles]. eScholarship. Retrieved from <https://escholarship.org/uc/item/6hw4r7dg>
- Dillman, L. M., & Christie, C. A. (2017). Evaluation policy in a nonprofit foundation. *American Journal of Evaluation*, 38(1), 60–79. <https://doi.org/10.1177/1098214016642864>
- Fierro, L. A., Echeverria-Estrada, C., Kinarsky, A., Fujita-Conrads, E., Ocheltree, K., Castleman, A. M., ... Christie, C. A. (Forthcoming). Drawing upon our past to inform a successful future: U.S. federal evaluation policies.

- Foundations for evidence-based policymaking Act of 2018, Pub. L. No. 115–435, 132 Stat. 5529. (2019). Retrieved from: <https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf>
- Kinarsky, A. R. (2019). An analysis of evaluation policy in the philanthropic sector. UCLA.
- Kinarsky, A. R., & Christie, C. A. (2021). Analysis of evaluation policies in the philanthropic sector. *American Journal of Evaluation*, *https://doi.org/10.1177/1098214021989435*
- Mark, M. M., Cooksy, L. J., & Trochim, W. M. K. (2009). Evaluation policy: An introduction and overview. *New Directions for Evaluation*, *2009(123)*, 3–11.
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29(4)*, 443–459. <https://doi.org/10.1177/1098214008324182>
- Pressman, J. L., & Wildavsky, A. (1984). *Implementation: How Great Expectations in Washington Are Dashed in Oakland; Or, Why It's Amazing that Federal Programs Work at All, This Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation*. United Kingdom: University of California Press.
- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). SAGE Publications.
- Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation*, *123(2009)*, 13–32. <https://doi.org/10.1002/ev.303>
- VERBI Software. (2021). *MAXQDA 2020 online manual*. Retrieved from: <https://www.maxqda.com/download/manuals/MAX2020-Online-Manual-Complete-EN.pdf>
- Vought, R. T. (2020). Phase 4 implementation of the foundations for Evidence-Based Policymaking Act of 2018: Program evaluation standards and practices [Memorandum M-20-12]. Office of Management and Budget. Retrieved from: <https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf>
- Waterman, R. W., & Meier, K. J. (1998). Principal-agent models: An expansion? *Journal of Public Administration Research and Theory*, *8(2)*, 173–202.

AUTHOR BIOGRAPHIES

Leslie A. Fierro, PhD, is the Sydney Duder Professor of Program Evaluation in the Max Bell School of Public Policy at McGill University. She is also the owner of Fierro Consulting, Inc.

Alana R. Kinarsky, PhD, is a postdoctoral scholar at University of California, Los Angeles, and an independent evaluator in Portland, Oregon.

Carlos Echeverria-Estrada is a Research Scientist at NORC at the University of Chicago where he designs and strengthens ME&L platforms in developing countries.

Nadia Sabat Bass is a PhD student in social research methodology at the University of California, Los Angeles, and holds an MA in global policy studies.

Christina A. Christie, PhD, is professor of social research methodology and Wasserman Dean at the School of Education and Information Studies at the University of California, Los Angeles.

How to cite this article: Fierro, L. A., Kinarsky, A. R., Echeverria-Estrada, C., Sabat Bass, N., & Christie, C. A. (2022). The Importance of Implementation: Putting Evaluation Policy to Work. *New Directions for Evaluation*, *2022*, 49–62. <https://doi.org/10.1002/ev.20490>