

Evaluation policy and the federal workforce

Diana Epstein  | Erica Zielewski  | Erika Liliedahl 

U.S. Office of Management and Budget,
Washington, DC, USA

Correspondence

Diana Epstein, U.S. Office of Management
and Budget, Washington, DC, USA.
Email: Diana_F_Epstein@omb.eop.gov

Abstract

The federal evaluation workforce plays a central role in the development and execution of evaluation policy. This workforce performs critical functions that include identifying where evidence should be built in particular policy areas, determining key research questions to inform their agency's mission, and shaping the field more broadly through federal investments in evaluation. Other roles include designing evaluation studies, overseeing contracts to conduct evaluations, performing internal evaluations, and communicating results to decision-makers. For the most part, these are highly skilled and trained career staff responsible for overseeing and executing technical projects in a challenging bureaucratic and political environment. This chapter describes the role of the federal evaluation workforce in the executive branch and its importance in developing and executing evaluation policy. It also describes recent changes, including the passage of the Foundations for Evidence-Based Policymaking Act of 2018, that have affected the roles, responsibilities, and opportunities for this vital workforce.

INTRODUCTION

Across the executive branch of the federal government, there is a dedicated workforce devoted to using evaluation to design and improve programs, policies, regulations, and operations. From designing evaluations to conducting data collections to disseminating and using evaluation results, the federal evaluation staff, who we refer to as “federal evaluators” in this chapter, are at the center of the federal evaluation enterprise. From an external perspective, what is inside the “black box” of this workforce remains largely a mystery, as does how this cadre of technical experts and talented civil servants uses its duties to support and influence evaluation policy and the policymaking process. There is not one set of characteristics or roles that defines the federal evaluation workforce, but what is clear

is their invaluable role in executing the evaluation policies of their agencies and making consequential contributions to enable evidence-based policymaking.

Federal evaluators are unique within the larger evaluation profession and play many roles in the federal evaluation enterprise, including serving as technical experts, subject matter experts, and responsible civil servants required to follow federal policies and regulations for processes including acquisitions and contract management, information collection, privacy, and information security, among others. A skilled evaluator in the federal government is a strategic negotiator and coordinates with program managers, procurement (contract) offices, financial and administration offices, data offices, attorneys, funding award managers, and agency career and political leadership—all necessary touch-points to plan and execute a single evaluation project. One evaluation staff person can also be responsible for leading a portfolio of evaluation work for an entire policy area or a niche policy specialization, or may serve as a jack-of-all-trades across evaluation methods and policy areas.

At the same time, the federal evaluation workforce must balance principles such as scientific rigor, independence and objectivity, transparency, and ethics with policy relevance and utility for political officials, all while working within a highly bureaucratic structure where political leadership changes frequently. In recent years, this workforce has experienced increased attention, supported by the Foundations for Evidence-Based Policymaking Act of 2018 (“Evidence Act”). There is growing recognition that for the federal government to carry out credible and influential evaluations, it must hire, develop, and retain specialized staff with technical evaluation expertise across the spectrum of evaluation methods and techniques. Simply adding evaluation to a general analyst’s duties will no longer suffice; the federal evaluation workforce must be further elevated and professionalized.

Agency evaluation policies help codify how a skilled federal evaluation workforce enables evaluation activities to inform policymaking; given the unique environment these staff work in, this is far from guaranteed. Over the past 40 years, the links between evaluation and policy more generally have been well documented in the literature. As early as the 1970s, there was a robust discussion of the role that evaluation plays in influencing policy and programmatic decision-making, with general agreement that for evaluation to be useful, it has to take politics and policy processes into account (Chelimsky, 1987). Moreover, the discussions in the past four decades have also shone a light on how evaluators, particularly those in the federal government, must also operate within this context and consider these political and policy processes. Chelimsky (2009) discusses the three requirements of evaluators and evaluation units most affected by policy and political pressures: independence, credibility, and morale; these remain relevant today. In doing so, she highlights the environment in which the federal evaluation workforce operates. This workforce strives to promote the usefulness of evaluation findings while adhering to principles such as credibility and independence, often working to strike a balance between technical requirements, timeliness, and the information user’s needs (Chelimsky, 1987). Gamoran (2018) observed that the use of evaluation findings to inform policy and practice will not become widespread if decision-makers reject or fail to examine credible evidence, arguing that creating conditions to encourage more widespread use of evidence in policy decisions is a formidable challenge. If we are to achieve an evidence-based government, the federal evaluation workforce must play a central role in implementing their individual agency’s evaluation policies, but also by working to ensure that evaluations and evaluation findings are credible, relevant, and used in the policymaking context.

This chapter provides new insights into the federal evaluation workforce within the executive branch, highlighting the range of roles and responsibilities of federal evaluation staff and describing how these roles affect evaluation policy at the government- and

agency-wide levels. We first discuss the different organizational structures and funding mechanisms that support the federal evaluation workforce, focusing on their implications for policy. The chapter next provides a detailed look into the many roles and responsibilities of federal evaluation staff, highlighting the variation across agencies and how these roles have changed over time. This discussion draws particular attention to how these roles interact with and influence policy and the broader evaluation field. We then describe recent legislative attention and federal policy guidance on the expectations for federal evaluators concerning their competencies and tasks, as well as the new government-wide standards that will govern evaluation policy. Finally, we close by highlighting leading examples of centralized support and coordination across the federal evaluation workforce. This chapter opens up the “black box” about this critical workforce. It showcases how this cadre of technical experts and talented civil servants are using their duties to support and influence evaluation policy and the policymaking process.

THE FEDERAL EVALUATION WORKFORCE: KEY IMPLEMENTERS OF EVALUATION POLICY

As noted throughout this volume, “evaluation policy” is a broad term that encompasses an agency’s written, explicit evaluation policy and those unwritten or implicit norms or practices that an agency adheres to as it undertakes evaluation. Trochim (2009) defines an evaluation policy as “any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation” (p. 16). These policies, both implicit and explicit, then inform evaluation practice and activity within an organization or, more specifically for the purposes here, a federal agency.

Trochim (2009) proposes a taxonomy of evaluation policies, noting that they may focus on one or more of eight dimensions including: (1) evaluation goals, (2) evaluation participation, (3) evaluation capacity building, (4) evaluation management, (5) evaluation roles, (6) evaluation process and methods, (7) evaluation use, and (8) meta-evaluation (the evaluation of evaluation policy). While not every agency evaluation policy addresses each element of Trochim’s taxonomy, the federal evaluation workforce is a key influencer and implementer of evaluation policy in all its forms through the organizational structure and funding of evaluations and evaluation offices and the roles and responsibilities of federal evaluators.

Organizational structures and funding mechanisms

As the federal evaluation landscape has changed over time, how these functions are organized and funded has also changed. Consistent with what Trochim (2009) calls evaluation capacity-building and management policies, federal agencies execute the evaluation function through various organizational structures, reflecting both an agency’s history and specific legislation. One common model is a centralized office that oversees all, or most, of the agency’s evaluation activities. This idea of a centralized evaluation function has long been discussed as an approach that has important implications for executing evaluation activities and using evaluation findings (Chelimsky, 1977). One example is the central office that Congress established at the Department of Education (ED)—the Institute of Education Sciences (IES)—which has independence and other important aspects of an evaluation policy codified in statute. An alternative structure has “pockets” of evaluation expertise embedded in either programs or subagencies. Some agencies have a hybrid structure with

a strong, centralized office and additional evaluation expertise in some of the subagencies within the department. Practitioners and scholars have debated the merits of the differing models, placement, and structures for evaluation staff within executive branch agencies to maximize the value of these personnel (Basco, 2017). Regardless of location, having the director of the office report to a key decision-maker elevates the office's work and its relevance to decision-making. In some cases, agencies have a senior leader responsible for both the evaluation and performance functions, given the overlap or complementarity of these activities.

The evaluation workforce is led by and supported through agency senior officials. The Evidence Act created a new position called Evaluation Officer for the 24 Chief Financial Officers (CFO) Act agencies, similar to the role that a few agencies already had in place, that is, Chief Evaluation Officer, Director of Research and Evaluation, etc. ([Foundations for Evidence-Based Policymaking Act of 2018, Pub. L. 115–435, S. 313](#)).¹ Some higher-capacity agencies also have evaluation leaders at the operating division or bureau level, such as the Office of Planning, Research and Evaluation (OPRE) at the Administration for Children and Families (ACF) in the U.S. Department of Health and Human Services (HHS). Creating a senior position to provide evaluation leadership, with evaluation experience and expertise, elevates evaluation in the policymaking and decision-making processes. Although not guaranteed, it can help ensure that evaluation has a seat at the table with agency leadership.

Agencies vary in how they fund their federal evaluation workforce, often called full-time equivalent or FTE, across agencies or even within the same agency. This includes, but is not limited to, using departmental or component salaries and expenses funds, a portion of appropriations designated for evaluation activities that allow use for staff, authorities that establish shared services accounts (e.g., require contributions from components in an agency), or a combination of these. In some cases, funding for the evaluation workforce largely depends on transfers from other accounts or the authority to set aside a portion of program funds for evaluation, which may result in variable funding from year to year. The way that any authorities for set-asides (or transfers of a portion of funds) for evaluation are established in law will greatly influence the stability, or lack thereof, of the level of funding and whether an agency can use this authority to fund staff positions. Unless these authorities are structured to allow for a level of funding sufficient for a workforce and are consistently appropriated by Congress, the unpredictability of the funds can make it difficult to sustain FTE. When designated program funds can be used to fund evaluation staff, such as in the case of the Health Profession Opportunity Grants Program within ACF, it allows for staff capacity proportional to the evaluation portfolio without impinging on funding for other purposes. However, this funding structure comes with a risk, as there are often no guarantees that those staff salaries can be absorbed elsewhere if the larger program is no longer funded. As another example, Congress has given the Department of Labor (DOL) consistent authority for transferring up to a maximum percentage of DOL program appropriations across the agency for evaluation; however, the use of the authority fluctuates so greatly in practice that it has not been suitable for staffing.

The exact account(s) used to fund the evaluation workforce of a particular office can also change over time as an office evolves and matures. For example, when DOL established a centralized, departmental Chief Evaluation Office (CEO), the initial staff were funded from a central account that draws on contributions from each component of the Department. Within the next budget and appropriations cycle, the Department secured evaluation funding as part of a direct appropriation in the Departmental Management account for evaluation FTE (Department of Defense and Full-Year Continuing Appropriations Act, 2011, Pub. L. No. 112-10, S. 1808). After the passage of the Evidence Act, the Department of Homeland Security (DHS) repurposed an unfilled, senior-level FTE position to hire a PhD social

scientist with extensive evaluation expertise to support the new Evidence Act work. Over time, DHS plans to build in additional evaluation FTE and/or funds, potentially from other accounts, to strengthen enterprise-wide evidence and evaluation efforts.

In sum, there is no single preferred solution to fund the evaluation workforce to achieve the goal of evidence-based policymaking and implement an agency's evaluation policy effectively. Rather, the lesson from existing evaluation offices' funding arrangements is that agencies can and do use a variety of funding mechanisms to establish, maintain, or increase their evaluation capacity, including temporary or permanent measures. Decisions about how these positions are funded and their placement within an agency inherently reflect an agency's evaluation policy, what is published and both the formal and unwritten rules that guide budget policy. However funded, having a stable cadre of federal evaluation staff to build and maintain institutional memory and expertise over time is needed if an agency is to successfully carry out its evaluation policy.

One position, many roles

The elevation of program evaluation through legislation such as the Evidence Act and congressionally mandated studies has required the federal evaluation workforce in the executive branch to take on a range of responsibilities from evaluation design to evaluation implementation to translation and use of findings. As part of the agency's formal evaluation policy or through unwritten norms, evaluation role policies, per Trochim's (2009) taxonomy, vary across agencies. Federal evaluation staff play a variety of roles that overlap with, but extend beyond, those of an evaluator in a contract research firm or academic setting. Ideally, federal evaluators are technically trained, professional staff with advanced academic degrees in economics, public policy, sociology, or a related scientific discipline; as with other scientific roles, undertaking high-quality evaluation requires highly qualified staff. At IES, for example, the career professional staff and commissioners have expertise to function as professional coequals with the contractors and grantees that receive IES funding (Whitehurst, 2018). Across agencies, federal evaluators are responsible for conceptualizing and designing evaluations, including identifying methodological options, preparing solicitation requests for third-party independent evaluations, reviewing technical proposals, overseeing contracted evaluation projects, conducting technical reviews, and interacting with technical evaluators and expert technical work group members. Federal evaluators must also build and maintain collegial relationships with other stakeholders in their agency to accomplish their goals, including career staff who lead the programs being evaluated, procurement and contracting officials, attorneys, performance and statistical staff, and career and political leadership. Maintaining relationships with staff in the federal programs addressed by the evaluation is a particularly essential role that does not often receive the attention from the outside that it deserves.

As described earlier, federal evaluators must also balance the needs of policy actors within and outside of their agencies. Unlike a typical evaluator-funder relationship, federal evaluators must balance the information needs of external parties, such as Congress, with methodological rigor and carefully navigate the translation of policy questions to evaluative questions (Chelimsky, 1987). The scale of many federal evaluation projects and the associated complexities inherent in national policies implemented across varying contexts also present unique challenges—and opportunities—for federal evaluators to tackle. Moreover, federal evaluators must start every effort with a lens focused on the findings' ultimate use and then prepare to be active participants in how those findings are disseminated and used.

Across these many roles, there is variation in how each particular agency or evaluation office within an agency approaches these responsibilities, resulting in a continuum across duties. We identified five key facets of the federal evaluator role where evaluation offices must make critical policy decisions related to the conduct of the evaluation, including (1) evaluation management; (2) methodological and topical prioritization; (3) administrative requirements; (4) interactions with program staff and (5) dissemination and use. The following discussion highlights how these five facets reflect an agency's evaluation policy, both written and unwritten. It also builds on Trochim's (2009) work by presenting examples of how these policies manifest in federal agencies today. The examples demonstrate existing variation across the executive branch and its implication for policy.

Evaluation management

Agencies differ in how they manage the evaluation studies they undertake with respect to the mechanism used to undertake the evaluation (e.g., carried out by internal agency staff versus through a federal contract, grant, or cooperative agreement) and the level of federal staff involvement in the technical execution of the project. For example, all federal awards involve oversight and monitoring required by law and regulations; federal contracts require oversight by a Contracting Officer's Representative (COR).² Federal grants require monitoring and oversight by the Federal Project Officer. This oversight includes ensuring that the government's interests and goals are met, deciding how to adjust if circumstances change, and setting priorities if costs increase but available funding does not. However, depending on the specific award requirements, agencies may have flexibility in how much oversight and involvement is needed for each evaluation project.

An agency's choices about how to carry out and manage its evaluations reflect its evaluation policy both implicitly and explicitly, including the funding parameters set by legislation. In some cases, the federal staff person or team undertakes the evaluation from start to finish, eliminating the need for an external entity. The policies of other federal evaluation offices prioritize the use of external entities as evaluators to help ensure the independence of evaluation findings. In other cases, a federal agency may require entities with federal awards to procure and carry out their own evaluations. This is especially common when there is little lead time to plan a federally-managed evaluation or where service strategies vary widely across sites. For instance, from 2011 to 2014, DOL awarded almost \$2 billion in funding to over 700 community colleges in every state (Department of Labor, n.d.). These grants affected sixty percent of the nation's publicly funded community colleges. They involved the development or redesign of nearly 2600 programs of study to help unemployed, trade-affected, and other adults learn skills and training for in-demand industries with quality jobs. DOL selected a subset of grantees to participate in a mandatory national evaluation and required the awarded entities to carry out their own third-party independent evaluations to document outcomes across the full sample of participants and capture lessons learned from different sites. In this case, the federal evaluation workforce still played an important role in setting out uniform requirements and expectations for these state- or locally-controlled evaluations. Similarly, agencies including HHS, ED, AmeriCorps (formerly the Corporation for National and Community Service), and DOL have used federal evaluation staff time to provide direct evaluation technical assistance and/or oversee an evaluation technical assistance contract to provide support to entities with a federal award on their evaluations to increase the likelihood that these studies are both high-quality and credible.

Federal evaluators are typically experts in both evaluation methodology and the substance of the program being evaluated and how it is administered. Agencies use both high- and low-touch approaches to manage evaluations. At the high end, federal evaluation staff provide federal award oversight, such as the COR for federal contracts, and invest substantial intellectual capital across the evaluation lifecycle. From playing a key role in evaluation design (i.e., developing the research questions and selecting evaluation methods) to evaluation implementation (i.e., reviewing data collection tools and choices such as weighing response rate goals against costs) to reporting and dissemination, these staff go beyond contract or project management and play a significant role in the success of the evaluation. Importantly, however, they stand apart from both the program office and researchers contracted to evaluate it. They are a step removed from program operations, direct policy-making, and from the external evaluation team. This separation allows them to pair their evaluation expertise with a sophisticated understanding of an agency's mission and goals in ways that others cannot, and therefore identify key insights in designing evaluations and applying lessons from evaluation findings.

In contrast, some agencies employ a lower-touch strategy to manage federal awards for evaluation projects. In these cases, federal evaluation staff overseeing the contract fulfill all technical oversight and management functions required of a federal project officer while leaving discretion to the evaluation contractor on technical decisions. For example, some agencies use Statements of Objectives that focus only on an evaluation's goals and leave decisions about the particular design or execution of the evaluation to the bidders and ultimately the selected contractor. This strategy might be valuable when, for instance, an agency is seeking particularly innovative, cutting-edge, or creative evaluation approaches. In other cases, the federal evaluation staff take a lighter touch in reviewing final deliverables, ensuring the technical accuracy of the final product, but offering flexibility over content to the evaluation study's contracted authors. The reasons why an agency chooses a high- versus low-touch approach to contract management vary. The approach may differ even within an evaluation office based on office priorities, staffing levels, workload, and other factors. More concretely, these choices reflect an agency's or office's evaluation policy and, in this particular area, often implicit, unwritten norms.

Importantly, the type of approach used can have implications for the independence of the work that results. To the extent that federal evaluation staff are so engaged in all facets of the evaluation that they become *de facto* members of the evaluation team, there are risks to the independence and credibility of the study. Nonetheless, some level of engagement in technical decisions is typically necessary so that the evaluation meets the agency's quality standards and its need for relevant evidence that can inform policymaking, program improvement, and other needs.

Methodological and topical prioritization

Federal evaluation staff also play a crucial role in prioritizing topics for study and elevating particular methodological approaches. In this way, they help set the evidence-building agenda for an agency and its programs while also moving the evaluation field forward. Decisions about what to study happen in a few ways, and federal evaluation staff have substantial influence in most cases. For example, the large-scale experiments of welfare programs in the 1980s served to elevate the federal government's use of experiments in social policy and the field more broadly (Gueron & Rolston, 2013). In the early 2010s, through its Behavioral Interventions to Advance Self-Sufficiency study, ACF's OPRE started a series of experiments to test the use of behavioral nudges in human services, which other

agencies have since replicated. Where there is flexibility in design and large investments, federal evaluation staff can significantly influence the prioritization of evaluation topics and methods. This may mean that staff play a consequential role in the amount and quality of evidence built in a particular area. In these instances where flexibility is permitted, how agencies choose what to evaluate and how often reflect their evaluation policy.

In some cases, Congress requires an evaluation, but the statutory language does not mandate a particular study or method, leaving a large role for staff to shape the evaluation's goals and subsequent design. For instance, the Moving to Work Demonstration, authorized by the Omnibus Consolidated Rescissions and Appropriations Act of 1996, provided funding for technical assistance and evaluation of demonstrations to improve housing opportunities for low-income families but only specified that the U.S. Department of Housing and Urban Development (HUD) test the broad aims of the program and put no further limits on this evaluation (Omnibus Consolidated Rescissions and Appropriations Act of 1996, Pub. L. No. 104–134, S. 204). Over many years, this flexibility has allowed HUD staff to develop a rich and robust research portfolio around Moving to Work. For example, for a multisite, randomized control trial of a mobility demonstration intersecting with over 10,000 households, HUD's program evaluation division is very involved in all technical aspects of the evaluation, including determining the appropriate, feasible, and rigorous approach prior to procuring a contract for the study; drafting an evaluation study design; designing the survey instrument that will be administered; and finalizing the language of the informed consent form for data collection to permit future longitudinal data collection and analysis efforts.

At other times, Congress mandates an evaluation for a particular program or policy topic. This is often the case in demonstration projects authorized and funded by Congress. For instance, the Patient Protection and Affordable Care Act of 2010 authorized the Health Profession Opportunity Grants Program and issued a requirement that HHS evaluate the demonstration projects. This statutory language still left a large role for evaluation staff in shaping the evaluation within those broad specifications. In contrast, the Social Security Administration's Disability Insurance Demonstrations must adhere to restrictive statutory language that addresses consent and participation in the demonstrations (for example, participation in the demonstration must be voluntary), which imposes methodological constraints (Social Security Disability Amendments of 1980, Pub. L. No. 96–265, S. 505).

Administrative requirements

As Trochim (2009) notes, evaluation does not operate in a vacuum. Federal evaluation policy must integrate with and operate alongside a range of other federal policies that govern acquisitions, personnel, and data collection, among others. Given this, federal evaluation staff must adhere to and execute a number of administrative requirements, such as protocols for announcing and selecting a federal contract and grant opportunities and making awards, invoicing, funding approval and disbursement, and ensuring data security. This requires learning and implementing rules and procedures that may be outside the scope of their evaluation technical training. For example, federal evaluators must abide by the Federal Acquisition Regulations (FAR) to procure evaluation services and oversee contracts.

Similarly, federal agencies must navigate and consider award requirements for grants and cooperative agreements to determine how to effectively require participation in an evaluation project. The Paperwork Reduction Act (PRA), which governs how federal agencies collect information, including for evaluation, is another process that federal evaluators and their contractors must navigate as they undertake evaluation. Evaluation staff

must also work with others in their agency and contractors to guarantee adherence to federal data security requirements such as the Federal Information Technology Acquisition Reform Act (FITARA). New laws and revisions to prior law emerge over time, so federal staff must determine how to navigate and implement these changes.

Complying with these requirements, which are unique to the federal government, often makes up a substantial component of a federal evaluator's workload. They are also foundational to addressing transparency and ethics as part of an evaluation policy framework. Many members of the federal evaluation workforce come to the government with topical or technical expertise, which is more broadly consistent with evaluation policy. However, they often have little to no background in these administrative requirements and must learn and implement them. Furthermore, their contractors must also expend resources to meet the administrative requirements, such as in the case of the PRA or FITARA. Some evaluation offices, such as DOL's CEO and ACF's OPRE, have invested in building knowledge of these processes among technical staff who then develop tools, templates, and trainings to support others. As an extension of this, OPRE has specialized and centralized some of these processes, such as PRA preparation and review, to allow its staff to focus their time on technical matters while also meeting their requirements on procurement, data security, and other federal regulations and policies, such as the FAR.

Interactions with program staff

Building collaborative relationships with program staff is critical for ensuring that evaluations are aligned with program goals and that the results are used. Interactions between evaluation staff and program staff look different across and within agencies and depend on structural considerations, staffing levels, and history. Some agencies, such as DOL, have a formalized process in which staff from the CEO engage with program staff on a quarterly (if not more frequent) basis to develop a learning agenda and check in on progress for evaluations relevant to that program. Others, such as U.S. Agency for International Development's more decentralized model, have an evaluation expert/liaison in bureaus, in order to ensure compliance with USAID evaluation policy in country Missions and that the bureau's evaluation work is timely and relevant. Evaluation staff at AmeriCorps work closely with program staff at various phases of the evaluation lifecycle so that the evaluation asks the questions most needed by the program and that the results are understandable for program improvement decisions. Over time, these interactions with program staff become part of the informal—and at times formal—norms of evaluation policy at an agency.

Dissemination and use

Returning to Trochim's taxonomy of evaluation policies, dissemination and use are key components of an agency's evaluation policy. How agencies approach these issues has implications for the federal evaluation workforce. These staff play an important role in ensuring that results are translated into actionable language that programs can apply for their needs. Dissemination choices are informed by (and inform) an agency's evaluation policy, as agencies or bureaus within an agency typically make clear choices about dissemination in their evaluation policy and make conscious decisions in a dissemination strategy about how to frame final evaluation products. For instance, some agencies choose not to allow federal staff to coauthor reports from evaluations conducted by a contractor or have reports identified as agency products to draw clear boundaries of independence between

study findings and agency influence. Other agencies see value in branding evaluation studies and reports they fund to capitalize on findings and influence program improvements or policy changes; these agencies must still consider how to adhere to an evaluation policy that maintains transparency, rigor, and independence from undue influence.

A transparent, multifaceted approach to sharing findings with varied stakeholders is critical for the evaluation process. Federal evaluators' role in dissemination and use includes translating evaluation findings to different audiences, often in different ways for particular stakeholders' needs (Federal Research Division, 2018). They brief agency political leadership proactively and on request, sometimes on a regular cadence and sometimes in response to a particular evaluation or to help officials prepare for Congressional briefings. Staff must also work with agency communications and public relations officials to translate complex technical findings into plain language that can be shared with a broad audience. Their role in dissemination and use also goes beyond the agency, as staff often lead presentations and webinars to share results with stakeholders. Evaluation staff play a critical role in managing the credibility of how findings are communicated, with appropriate caveats, particularly when findings are unexpected, unfavorable, or mixed. Staff also discuss evaluation findings at professional conferences. In recent years, more agencies have started to take a proactive approach to dissemination by using office e-newsletters, regular briefings, blogs, and social media accounts to share results. Planning and executing these communications serves an important function in helping to share evaluation results widely, increasing the chance that they will be used to help improve policies and programs and subsequently expanding the influence of staff on evaluation policy.

OPPORTUNITIES FROM THE EVIDENCE ACT FOR STRENGTHENING EVALUATION POLICY

Federal evaluators' competencies and tasks

Congress recognized the importance of a strong federal evaluation workforce in the Foundations for Evidence-Based Policymaking Act of 2018. Notably, the Evidence Act creates a statutory system for program evaluation akin to those for performance and statistics by requiring agencies to designate an evaluation officer, produce an Annual Evaluation Plan, and create an agency Evaluation Policy. The law also tasks the Office of Personnel Management with working with the Office of Management and Budget (OMB) within the Executive Office of the President³ to “identify key skills and competencies needed for program evaluation in an agency; establish a new occupational series, or update and improve an existing occupational series, for program evaluation within an agency; and establish a new career path for program evaluation within an agency.”

The role of the Evaluation Officer represents the first time that agencies must designate a senior career official with substantive expertise in evaluation methods and practices to oversee evaluation activities and the development of the agency's Learning Agenda and Annual Evaluation Plan. A new position for many agencies, the Evaluation Officer must be appointed without regard to political affiliation and possess “demonstrated, senior-level technical expertise in evaluation methods and practices and ... appropriate expertise in the culture, disciplines, and policy areas of the agency” (Office of Management and Budget, 2019). The role should be filled by a “senior career employee with the skills and expertise to maintain principles of scientific integrity throughout the evaluation process, ensure adherence to the agency evaluation policy, and maintain the standards in OMB M-20-12... and have sufficient time and resources to lead and execute this work” (Office of Management

and Budget, 2021). Creating this specific role to lead and implement the agency's evaluation activities and policy is an important opportunity to elevate the evaluation function across the federal government.

As part of the Evidence Act, OMB was required to issue standards and leading practices for program evaluation. This guidance was issued in March 2020 as *OMB M-20-12: Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices*. This guidance established the first-ever government-wide standards for program evaluation and ten leading practices that agencies can use to implement those standards. The document also acknowledges the importance of building and maintaining evaluation capacity, as the first leading practice states that agencies should “staff the federal evaluation workforce with qualified personnel and support their continued professional development in order to effectively plan, manage, implement, and oversee high-quality evaluation activities” (Office of Management and Budget, 2020). These government-wide evaluation standards have important implications for Federal evaluators and those who receive Federal evaluation contracts, grants, and cooperative agreements. The standards and resulting agency policies and practices are expected to broadly influence the development, revision, nature, and implementation of evaluation policy.

Federal evaluation standards

An interagency group of federal evaluation experts worked together to develop the government-wide standards for program evaluation in OMB M-20-12. The working group, led by staff from OMB's Evidence Team, included evaluation experts from agencies with different missions, for example, human services agencies, international development agencies, and regulatory agencies. The development process included an extensive review of evaluation literature, evaluation policies, and evaluation frameworks both inside and outside the United States. Staff conducted listening sessions to solicit feedback, input, and expertise from a variety of experts internal and external to the government. Importantly, federal staff were included from agencies with mature evaluation capacity and incipient evaluation efforts; this helped ensure that the resulting standards and practices could be implemented government-wide. The workgroup landed on a set of standards that reflects the best thinking in the field as a whole: relevance and utility, rigor, independence and objectivity, transparency, and ethics.

As the guidance document notes, these standards should be used by federal agencies to continually improve, or in some cases maintain, evaluation activities and evaluation capacity broadly, as well as the specific requirements from the Evidence Act (i.e., agency evaluation policy, multiyear Learning Agenda, and Annual Evaluation Plan). The standards and practices should be used at the enterprise level and within subagencies, operational divisions, and bureaus where much evaluation activity occurs. This includes generating credible evidence from evaluations, applying that evidence to budget justifications, strategic plans, grant-making, and other processes, and identifying and addressing areas for improving programs, policies, operations, regulations, and other agency functions. Notably, the intention behind these standards is to create a higher baseline for evaluation across the federal government, not make processes more rigid, stifle innovation, or otherwise hamper high-quality evaluation work that is already happening in some agencies. While certainly not an all-inclusive list, the ten practices included in the guidance memo provide concrete mechanisms that agencies can use to implement the standards.

The process of developing these standards was led entirely by the federal evaluation workforce, thus providing a leading example of the direct influence that such staff can

have on agency evaluation policies and the field more broadly. Furthermore, the standards and practices will necessarily be critical as agencies consider how to hire and retain qualified evaluation staff. If federal evaluators fulfill their mandate to hold up these standards widely and implement leading practices such as those listed in the guidance, they will end up greatly shaping evaluation policy and practice. This influence will likely extend beyond the federal government; practice number three notes the benefit to an agency from “referencing and requiring adherence to the [agency] evaluation policy in funding opportunities and federal awards so that evaluation grants, contracts, and cooperative agreements are conducted according to the standards and principles contained in the policy” (Office of Management and Budget, 2020).

Central coordination and support

The federal evaluation workforce is supported by a dedicated team of senior staff at OMB, all with previous experience in agency evaluation offices, that sets evaluation policies government-wide. This Evidence Team supports the federal evaluation system more broadly by providing evaluation support and resources to agencies directly and to the rest of OMB. The Team organizes and chairs the Evaluation Officer Council, which meets monthly and serves as a forum for agency Evaluation Officers to share lessons learned, exchange information, and discuss challenges. Within OMB, the Team works with budget-side offices on funding notices and other ways to embed evaluation in programs, reviews of reports to OMB on mandated evaluations, and other activities. The Team also partners with both budget and management offices at OMB to review agencies’ submissions of the new deliverables required by Title I of the Evidence Act, including multiyear Learning Agendas and Annual Evaluation Plans. These reviews and the subsequent feedback provided to agencies are important opportunities for federal staff to influence evaluation as it is carried out across the government.

The importance of professional development and skills improvement for the federal evaluation workforce cannot be overlooked. The OMB Evidence Team convenes the Interagency Council on Evaluation Policy, which supports the federal evaluation community by providing technical assistance, developing papers and tools, and providing professional development opportunities for staff. The Evidence Team also organizes an online internal-to-government Community of Practice, which features curated resources—including reports, toolkits, presentations, and workshops—for the federal evaluation community on topics related to Evidence Act implementation, program evaluation, and evidence more generally. The team also developed the public-facing Evaluation.gov website, which is the centralized home for information about evaluation in the federal government, with an emphasis on work responsive to the requirements of the Evidence Act. The Evidence Team partners with the Office of Evaluation Sciences at the General Services Administration to run a regular workshop and training series for executive branch staff on a variety of evaluation, evidence-building, policy, and related topics. Given the elevation of program evaluation in the Evidence Act, professional development and training will assume even more importance in the coming years as the federal evaluation workforce is expected to grow and respond to these new mandates.

An important role the evaluation workforce can play in implementing an agency’s own evaluation policy is to lead and coordinate an evidence and evaluation community of practice (CoP) for their agency. Some CoPs have a core purpose of supporting professional development of the evaluation workforce, particularly when evaluation staff is spread throughout an agency such as the Department of State. Other agencies, such as the Small

Business Administration and the Office of Energy Efficiency and Renewable Energy in the Department of Energy, use CoPs to train nonevaluation staff on the value of evaluation, basic evaluation concepts, and understanding and applying evaluation results. The Centers for Disease Control and Prevention (CDC) at HHS has a long-standing evaluation CoP that, among many other topics, offers regular training on how to develop a logic model and use this program-planning tool; this supports all CDC staff in meeting their agency-wide requirement that all funding announcements include a logic model.

Growing influence on the field

As the roles of federal evaluators have expanded, there are increased opportunities for these staff to share their expertise and knowledge with the evaluation profession, thereby shaping policy for the field more broadly. One avenue is interagency workgroups where staff across agencies meet to share experiences and develop common processes across agencies. For example, a group of staff from ED and the National Science Foundation met from 2011 to 2013 to develop common guidelines for education research and development, including evaluation (Institute of Education Sciences and National Science Foundation, 2013). The Institute of Education Sciences (ED), Office of the Assistant Secretary for Planning and Evaluation (HHS), and Chief Evaluation Office (DOL) (2013) published a draft framework for consideration for “Exploring a Cross-Agency Platform for Judging Evidence: Resources for Federal Agencies.” The framework was the result of an interagency workgroup of evaluation leaders that recognized the need for a common way to approach the review, critique, and determination of credibility of evidence (Institute of Education Sciences, Office of the Assistant Secretary for Planning and Evaluation, and Chief Evaluation Office, 2013). In response to a 2016 Government Accountability Office report recommending more collaboration across agencies, OMB convened an interagency group on tiered evidence grants that meets regularly to share lessons learned around grant administration, levels of evidence, and evaluation criteria (Government Accountability Office, 2016). These interagency groups provide valuable venues for the federal government to work through common challenges and share promising practices that shape policy in the participating agencies and more broadly.

Federal agencies have also started to organize conferences for federal and external evaluators led by experts in evaluation offices. For example, OPRE runs an annual meeting on methods; recent topics have included core components research, promoting open science, and rapid learning methods. OPRE has also led the Research and Evaluation Conference on Self-Sufficiency for over 20 years and the National Research Conference on Early Childhood for 30 years, drawing hundreds of participants from across the country. For each of these conferences, OPRE evaluation staff not only organize and run panels, but also play prominent roles in sharing their own work in innovative methods and new results from the evaluations they have designed and overseen. DOL’s Employment and Training Administration previously held an annual research-focused conference, and in recent years has transitioned to participating in workshops and sessions focused on evidence and evaluation findings in policy- or program-focused conferences, such as national convenings hosted by the Department on the Workforce Innovation and Opportunity Act. DOL has also held roundtables with international partners such as the European Union to discuss differing evaluation strategies before, during, and after policy implementation. Even smaller agencies have started to organize conferences to share their work and collaborate with external researchers. For example, AmeriCorps has led conferences highlighting the work of both their grantees and evaluation projects designed and overseen by staff. Similarly, the Millennium Challenge Corporation hosts topical evaluation conferences where findings

can be presented and discussed. The ability of federal agencies to host large conferences that draw national audiences and researchers is constrained by limits on federal funding and the use of funds; however, federal staff can also play a role in shaping conferences led by external parties such as the National Association for Welfare Research and Statistics, which has for some time included a federal representative on the planning committee.

As the breadth of work and sophistication of federal evaluators has grown, staff are increasingly publishing articles in peer-reviewed journals and presenting their work at professional conferences such as the American Evaluation Association and the Association for Public Policy Analysis and Management. The number of federal evaluation staff participating in these conferences over the past decade reflects the increasingly prominent role that federal evaluators play in engaging with academic and industry audiences. Whereas it was once fairly rare for a federal evaluator to present work at the leading conferences for these professional associations, such staff are now featured in panels and roundtables, and their presentations can draw sizeable interest. This reflects an increasing professionalization of the workforce. Federal evaluation staff are viewed as key contributors and experts in their own right rather than simply paper-pushers signing off on invoices for contracted work.

Finally, federal evaluators can advance the field of evaluation and highlight the benefits of evidence-building by aligning their work to policy priorities that emerge from the Executive and Legislative branches, particularly at the start of a new administration. President Biden's focus on evidence-based policymaking, equity, and the American Rescue Plan offer new and renewed opportunities for federal staff to develop and execute evaluation policy, both government-wide and within individual agencies. The Presidential Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking stated that the administration is committed to making evidence-based decisions guided by the best available science and data and tasked OMB with issuing guidance to improve agencies' evidence-building plans and annual evaluation plans (Executive Office of the President, 2021b). In response, OMB issued memo M-21-27, which directs agencies to build and nurture a culture of evidence and the infrastructure needed to support it, including strengthening the federal workforce to ensure that staff with the right skills and capabilities are positioned across the federal government to build and use evidence to improve decision-making.

The Executive Order 13985 on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government names equity as a priority that should be embedded in all of the federal government's efforts (Executive Office of the President, 2021a). This extends to evidence-building and evaluation as well. OMB guidance to agencies laid out the expectation that equity be included in agencies' Strategic Plans and evidence-building activities as reflected in agencies' Learning Agendas and Annual Evaluation Plans. While it is still too early to assess the result of these efforts, it is clear that the focus on equity in federal evaluation has the potential to help evaluators meet what is often a key priority for their work—better understanding effective policy and practice for certain types of beneficiaries—while broadly shaping the field and its policies through the federal awards that are issued and the resulting sums of federal evaluation dollars that are invested. The American Rescue Plan injected billions of dollars into the American economy in key areas to mitigate economic damage from the COVID-19 pandemic. Federal evaluators are poised to play a key role in helping to understand the end result of these investments, as well as whether or not the desired focus on equitable distribution of funds was ultimately achieved. For example, the Department of Treasury's Coronavirus State and Local Fiscal Recovery Funds explicitly includes evaluation as an allowable use of funds, and the reporting guidance encourages recipients to pursue evidence-building through rigorous evaluations (U.S. Department of Treasury, 2021). This type of federal communication can have wide reach and influence on shaping evaluation policy at all levels of government.

CONCLUSION

While the public at large may be unaware of the federal evaluation workforce or (mis)perceive its role as largely bureaucratic, these civil servants play many important roles in shaping the evaluation profession and its policies. From shaping study designs to disseminating results, these staff influence what gets evaluated, with what methods, and how subsequent findings are shared. With the increased focus on federal evaluation in recent years, the importance of these staff members is expected to grow. The new government-wide requirements for evaluation leaders and standards for program evaluation may improve the quality of federal evaluations, leading to long-term effects on evaluation policy. Further professionalization of the workforce may follow, although the extent to which capacity is built and expanded may ultimately rest with the desire of Congress to invest resources in the evaluation function.

ORCID

Diana Epstein  <https://orcid.org/0000-0001-5415-8493>

Erica Zielewski  <https://orcid.org/0000-0002-4225-619X>

Erika Liliedahl  <https://orcid.org/0000-0003-4572-7441>

ENDNOTES

¹These agencies include: U.S. Departments of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services, Homeland Security, Housing and Urban Development, Interior, Justice, Labor, State, Transportation, Treasury, Veterans Affairs; Environmental Protection Agency; National Aeronautics and Space Administration; Agency for International Development; Social Security Administration; General Services Administration; National Science Foundation; Nuclear Regulatory Commission; Office of Personnel Management; Small Business Administration (*Chief Financial Officers Act of 1990, Pub. L. 101-576*).

²The Contracting Officer's Representative or COR is the staff member who supports the Contracting Officer by ensuring that the contractor meets the technical requirements of the contract. COR's must go through specialized training in order to be certified, and the amount and level of training depends on the contract size.

³This work was still in progress at the time of this chapter's publication.

REFERENCES

- Basco, D. (2017). *Summary for strengthening the capacity for federal research and development evaluation: A collaborative design workshop*. The American Association for the Advancement of Science and RAND Corporation. Retrieved from <https://www.aaas.org/sites/default/files/2018-10/Strengthening%20Evaluation%20Capacity%20Workshop%20Summary%201217.pdf>
- Chelmsky, E. (2009). Integrating evaluation units into the political environment of government: The role of evaluation policy. *New Directions for Evaluation*, 123, 199–213.
- Chelmsky, E. (1987). What have we learned about the politics of program evaluation. *Educational Evaluation and Policy Analysis*, 93, 199–213.
- Chelmsky, E. (Ed.) (1977). *The use of evaluation by federal agencies: Proceedings of a symposium*, 1976. MITRE Corporation.
- Chief Financial Officers Act of 1990, Pub. L. No. 101–576, § 901, 104 Stat. 2842.
- Department of Defense and Full-Year Continuing Appropriations Act, 2011, Pub. L. No. 112–10, § 1808, 125 Stat. 157.
- Executive Office of the President. (2021a). *Executive order on advancing racial equity and support for underserved communities through the federal government*. 86 Fed. Reg. 7009. Retrieved from <https://www.federalregister.gov/documents/2021/01/25/2021-01753/advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government>
- Executive Office of the President. (2021b). *Presidential memorandum on restoring trust in government through scientific integrity and evidence-based policymaking*. 86 Fed. Reg. 8845. Retrieved from <https://www.federalregister.gov/documents/2021/02/10/2021-02839/restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking>
- Federal Research Division. (2018). *A profile of federal research and evaluation offices*. Library of Congress.
- Foundations for Evidence-Based Policymaking Act of 2018, Pub. L. No. 115–435.

- Gamoran, A. (2018). Evidence-based policy in the real world: A cautionary view. *The ANNALS of the American Academy of Political and Social Science*, 6781, 180–191. <https://journals.sagepub.com/doi/10.1177/0002716218770138>
- Gueron, J. M., & Rolston, H. (2013). *Fighting for reliable evidence*. Russell Sage Foundation.
- Institute of Educational Sciences & National Science Foundation. (2013). *Common guidelines for education research and development*. U.S. Department of Education and National Science Foundation. Retrieved from https://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf?WT.mc_id=USNSF_1
- Institute of Education Sciences, Office of the Assistant Secretary for Planning and Evaluation, & Chief Evaluation Office. (2013, February). *Exploring a cross-agency platform for judging evidence: Resources for federal agencies*. U.S. Department of Education, U.S. Department of Health and Human Services, & U.S. Department of Labor. Retrieved from https://youth.gov/docs/Common/Evidence/Framework-Draft_508_3-11-13.pdf
- Office of Management and Budget. (2019). *Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance* (OMB Memorandum M-19-23). Executive Office of the President. Retrieved from <https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf>
- Office of Management and Budget. (2020). *Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices* (OMB Memorandum M-20-12). Executive Office of the President. Retrieved from <https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf>
- Office of Management and Budget. (2021). *Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans* (OMB Memorandum M-21-27). Executive Office of the President. Retrieved from <https://www.whitehouse.gov/wp-content/uploads/2021/06/M-21-27.pdf>
- Omnibus Consolidated Rescissions and Appropriations Act of 1996, Pub. L. No. 104–134, § 204, 110 Stat. 1321.
- Social Security Disability Amendments of 1980, Pub. L. No. 96–265, § 505, 94 Stat. 475.
- Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. In W. M. K. Trochim, M. M. Mark, & L. J. Cooks (Eds.), *Evaluation policy and evaluation practice. New Directions for Evaluation*, 123, 13–32.
- U.S. Department of Labor. (n.d.). *Trade Adjustment Assistance Community College and Career Training: The broad reach of TAACCCT Funding* [Fact sheet]. Retrieved from https://www.dol.gov/sites/dolgov/files/ETA/TAACCCT/pdfs/TAACCCT_Fact_Sheet_Broad_Reach_of_Funding_10.21.2016.pdf
- U.S. Department of Treasury. (2021). *Compliance and reporting guidance: State and local fiscal recovery funds*. Retrieved from <https://home.treasury.gov/system/files/136/SLFRF-Compliance-and-Reporting-Guidance.pdf>
- U.S. Government Accountability Office. (2016). *Tiered evidence grants: Opportunities exist to share lessons from early implementation and inform future federal efforts* (GAO-16-818). Retrieved from <https://www.gao.gov/assets/gao-16-818.pdf>
- Whitehurst, G. (2018). The Institute of Education Sciences: A model for federal research offices. *The ANNALS of the American Academy of Political and Social Science*, 6781, 124–133. <https://doi.org/10.1177/0002716218768243>

AUTHOR BIOGRAPHIES

Diana Epstein leads the Evidence Team at the Office of Management and Budget, having previously held positions at AmeriCorps, the American Institutes for Research, Abt Associates, and the RAND Corporation.

Erica H. Zielewski is a Senior Evidence Analyst at the Office of Management and Budget following prior roles at the U.S. Department of Health and Human Services and the Urban Institute.

Erika Liliedahl is a Senior Evidence Analyst at the Office of Management and Budget; her government service began at the U.S. Department of Labor where she worked at the Employment and Training Administration and the Chief Evaluation Office.

How to cite this article: Epstein, D., Zielewski, E., & Liliedahl, E. (2022). Evaluation policy and the federal workforce. *New Directions for Evaluation*, 2022, 85–100. <https://doi.org/10.1002/ev.20487>