INTRODUCTION

How does any organization educate its members about the potential value of evaluation, design the infrastructure and systems to support evaluative activities, use evaluation in a sound fashion, and sustain support for and use of evaluation over time? These outcomes are rarely achieved by accident, and evaluation policy is an important part of the answer. Thoughtful, intentional approaches to designing and implementing entire evaluative systems have occurred in many organizations, including government agencies. Especially over the last decade, growing attention to evaluation policy and an emphasis on evaluation capacity building have taken place in conjunction with a new emphasis on evaluative systems at the federal, state, local, and tribal levels of government as well as among many nonprofit and philanthropic organizations.

Evaluation policy has been defined as “any rule or principle that a group or organization uses to guide its decisions and actions” with respect to evaluation (Trochim, 2009, p. 16). Importantly, the focus of evaluation policy is on policies about evaluation, not the substance or methods of a particular evaluative product or inquiry. Nor does evaluation policy refer to the policies made regarding a particular program area (such as early childhood education policy). One can think about an evaluation policy in terms of the formal policies and other directives guiding evaluation at the organizational level, such as the...
Department of Labor’s evaluation policy or a legal framework that outlines expectations across multiple agencies. Additionally, evaluation policy comprises a subfield of evaluation that includes both empirical and conceptual scholarship on the topic, as well as activities aimed at encouraging policies that facilitate the institutionalization of evaluation activities and evaluation use.

In practice, evaluation policy is important because it can “enable and constrain the potential contributions evaluation can make” (Mark et al., 2009, p. 3). Put another way, an evaluation policy is intended to define the rules, procedures, approaches, resources, infrastructure, personnel, and direction for evaluation capacity as well as for individual evaluations. Typically, an evaluation policy is somewhat abstract, setting up general expectations while deferring to practitioners and others involved in evaluative activities to determine evaluation questions and particular methods and design choices. Nevertheless, the guidance of an evaluation policy, however general, will likely have substantial implications for when and how evaluation is done, who it is for, and whether and how it is used.

The absence of a clear, coherent, and transparent evaluation policy may also be limiting in organizations and governments. Indeed, when a sound evaluation policy is absent, the expectation to produce and use evaluation may not be established by senior leaders or funders. Evaluation policy is central to the sustained success of evaluation activities: Evaluation policy outlines the expectation that organizations, programs, and policies should prioritize evaluative activities and use the results appropriately.

BRIEF HISTORY OF EVALUATION POLICY IN THE UNITED STATES

Evaluation policy has been a topic of discussion for more than 50 years, though it has often been unaddressed in the literature or addressed implicitly in the context of evaluation capacity building. A major period of growth of evaluation in the 1960s was stimulated by congressional mandates for, and funding of, evaluation in sectors such as education, job training, crime control, and substance abuse (Shadish et al., 1991, pp. 22–28). The enabling legislation established early evaluation policies, though ones that can be characterized as limited in scope and selectively focused only on particular programs or funding streams. For example, legislatively mandated evaluation activities in the areas of housing and welfare programs largely focused on the evaluation of particular demonstration projects.

More recent efforts to address evaluation policy in practice build on decades of contributions to the design of evaluation systems in some federal agencies that focus largely on health, employment, housing, and antipoverty programs (Newcomer & Hart, 2022). In the mid-1990s, enactment of the Government Performance and Results Act (GPRA) in the US federal government, for example, established a new focus on monitoring and assessing outputs and some outcomes of a broad swath of federal policies and programs. Federal laws like the Paperwork Reduction Act and the Information Quality Act also had major implications for how evaluation policy evolved in practice (Chenok & Hart, 2020). For example, the laws affect how and what types of information can be collected, including the number of individuals that can be surveyed, which leads some evaluation officials in government to cite the Paperwork Reduction Act as one of the most critical barriers to their work (Fletcher & Howell, 2017).

Major shifts in approaches to evaluation policy materialized at the turn of the 21st century. The administration of George W. Bush designed a system intended to better integrate performance and evaluation information into budgetary decision-making at the federal level of government, creating the Performance Assessment Rating Tool (PART). While PART had limitations, it promoted evaluative thinking and productively led to some new evaluative activities in federal agencies (Hart & Newcomer, 2018). The Barack Obama
Administration outlined approaches for evaluation capacity building and prioritized funding for new evaluations in some parts of the federal government. The GPRA Modernization Act of 2010 placed a new emphasis on prioritizing major goals and performance indicators, as well as transparently reporting information to the public and decision-makers.

Federal evaluation policy within particular policy domains also evolved over the past 50 years, often spurred by legislative activities and with major implications for state and local government. For instance, efforts at the Department of Education to prioritize particular methods for evaluation were a cause for conflict and concern across the evaluation community (Donaldson, 2009). Reauthorizations of major health, disability, employment, workforce, and antipoverty programs periodically included new directions for specific evaluation activities, sharing of results, or processes to build new evidence over time (Newcomer & Hart, 2022).

Evaluation policy efforts at the federal level in the United States rapidly accelerated with the publication of the unanimous recommendations from the US Commission on Evidence-Based Policymaking in 2017. In its final report, the Evidence Commission recommended that federal agencies establish chief evaluation officer positions, develop learning agendas, and publish evaluation policies and plans. In short, the Evidence Commission suggested a comprehensive federal evaluation system, based on the premise that evaluation is central to effective implementation of government policies in programs (CEP, 2017).

Within several years, the government-wide evaluation function recommended by the Evidence Commission would largely become a requirement for many agencies. Shortly after the Evidence Commission’s recommendations were published, champions in Congress advanced the Foundations for Evidence-Based Policymaking Act (Evidence Act) which incorporated the recommendations. Even as the legislation advanced in Congress, the administration of Donald Trump also touted the benefits of evaluation, offering a proposal as part of a government reorganization plan to direct agencies to have evaluation officers and learning agendas (Hart, 2019; OMB, 2018).

The Evidence Act was ultimately signed into law in early 2019, requiring large federal agencies to appoint evaluation officers, publish annual evaluation plans, and develop multiyear learning agendas. The law also set expectations for the establishment of an evaluation workforce and required that the White House Office of Management and Budget (OMB) publish standards and principles for use in federal evaluation. Today the Evidence Act provides the most comprehensive framework yet for federal evaluation policy, while also connecting that evaluation policy to other disciplinary activities in government such as statistics, economics, and policy research.

In the years following enactment of the Evidence Act, federal agencies began appointing senior leaders into the evaluation officer role in accordance with the legal requirement and guidance from OMB (OMB, 2019). OMB’s initial guidance also specified qualifications for federal evaluation officers, including the ability to manage organizational culture change and build capacity for government evaluation. In 2020, following consultation with evaluation experts, OMB issued guidance outlining five standards for government evaluation: Relevance and utility, rigor, independence and objectivity, transparency, and ethics (OMB, 2020). OMB also prescribed ten evaluation practices for government evaluation, including capacity building, stakeholder engagement, and use of agency-specific evaluation policies, among others.

In 2021, OMB issued renewed guidance on evaluation in government, outlining more direct expectations for federal agencies. These were aligned with President Joe Biden’s directive that decision-making in government would be informed by evidence (Biden, 2021; OMB, 2021). OMB’s 2021 guidance expanded the requirements from the Evidence Act to encourage the entirety of the federal government to build evaluation capacity and implement an evaluation policy.
Because the landscape for federal evaluation policy has evolved rapidly in recent years, it is the basis of much of the discussion in this volume. However, there have also been notable advances in evaluation policy among state, local, and tribal governments as well as nonprofits and philanthropy, and we anticipate this will only accelerate in upcoming years. Chapters in this issue of NDE do address some other sectors, including a global sample of international development organizations (Chapter 3). Still, the US federal government dominates the presentations.

We believe this emphasis is appropriate, in part because federal government activities influence other evaluation policy endeavors in the United States and perhaps beyond. Given that the federal government issues grants and contracts to other governmental entities and often sets expectations that over time diffuse to nonprofits and philanthropy, the Evidence Commission and Evidence Act set a tone and framework that will likely affect evaluation policy beyond the federal government for years to come. Although foundations, international organizations, and state, local, and tribal governments will not have the same timeline specified by the Evidence Act, they are likely to experience increasing pressures to have written evaluation policies. Even independent evaluation practitioners will increasingly need to be informed about evaluation policy, for they are increasingly likely to work with organizations that have such policies.

**ROLE OF THE AMERICAN EVALUATION ASSOCIATION IN EVALUATION POLICY**

As evaluation policy became more formalized in the United States over the past decade, the American Evaluation Association (AEA) also continued to increase its ability to influence and advise on how evaluation policy was implemented in government. In 2007, AEA launched a pilot project called the Evaluation Policy Task Force (EPTF) to support initial organization and planning for how AEA could devise more timely and influential engagements on evaluation policy. Following an evaluation of the EPTF’s activities, and taking into consideration the increasing activities in the policy environment that effect evaluation, in 2019, the AEA Board of Directors designated EPTF as a permanent standing task force.

Shortly after EPTF was created, the theme of the AEA annual conference, chosen by then President Bill Trochim, was evaluation policy. This led to a 2009 issue of *New Directions for Evaluation* on evaluation policy and evaluation practice (Trochim et al., 2009). Soon after, AEA published the first version of “An Evaluation Roadmap for a More Effective Government.” This followed an extensive and iterative process, including review and approval of the document by the AEA leadership and membership (AEA 2009). The document served as a foundation for the EPTF’s activities and AEA’s engagement with policymakers in positions that affect evaluation policy over the years. More recently, the Roadmap was provided as an input to the Evidence Commission. In 2019, the EPTF updated the Roadmap, in large part to reflect enactment of the Evidence Act and to align the document with more recent theory and practice that affected evaluation policy in federal agencies (AEA, 2019). In light of the strong alignment of both the Evidence Commission and the Evidence Act with concepts and practices espoused by AEA over the years, AEA praised both the recommendations from the Evidence Commission and the resulting law, the Evidence Act (AEA, 2017, 2018).

The activities of AEA and the EPTF are intentionally strategic, seeking to capitalize on the skills and expertise of the all-volunteer committee of experts. In recent years, the EPTF has convened meetings and briefings on evaluation with government officials, recommended AEA endorsements of legislative proposals and recommendations, identified
opportunities to offer technical assistance on policy matters that affect AEA members and the field of evaluation, and generally promoted improvements to evaluation policy in government. The EPTF originated in large part because of a sense that AEA too often did not have a seat at the table when decisions about evaluation were being made (Mark et al., 2009); this has changed dramatically in the ensuing years.

The membership of the EPTF has changed over the last decade, though some original members remain (Katherine Dawes and Mel Mark) and serve with the other current members (Lisa Aponte-Soto, Diana Epstein, Nick Hart, Kathy Newcomer, Demetra Nightingale, and Maurice Samuels). Through 2021, they served alongside longtime luminaries such as George Grob, the late George Julnes, Stephanie Shipman, and Tom Chapel. Collectively, the EPTF members provide a wealth of expertise about evaluation policy and practice for AEA, including many with tours through government agencies. The EPTF members work to share knowledge about policy activities across AEA, including through annual sessions at the AEA conference, an annual AEA town hall, periodic “Policy Watch” columns in the AEA newsletter, and, importantly, through public statements from AEA about important evaluation policy matters. In 2021 alone, AEA issued public statements on the prioritization of evidence-based policymaking by the Biden Administration (AEA 2021a), the new OMB guidance for evaluation policy (AEA 2021b), the role of equity and engagement in evaluation policy (AEA 2021c), and the importance of independence in evaluation (AEA 2021d). AEA also cohosted a public event to discuss the implications of the OMB guidance for evaluators (Data Foundation and AEA, 2021). Finally, the EPTF also supported a survey of federal evaluation officials in 2021 – the first since enactment of the Evidence Act – to understand emerging challenges and early successes from new capacity-building efforts (Mumford & Hart, 2021).

One challenge facing EPTF and AEA’s efforts to engage on evaluation policy is simply the breadth and volume of topics that could be addressed. At its inception, a strategic choice was made to focus the responsibilities of the EPTF primarily on federal evaluation policy (Mark et al., 2009). That focus remains today as a predominant, though nonexclusive, driver of EPTF activities. Each year, EPTF works with the AEA Board to outline an annual plan, though emerging opportunities and challenges often result in modifications to the plan because the policy environment can rapidly evolve.

Another challenge is representing the broad range of interests and the varied views of AEA members. The EPTF works closely with the AEA Board and Executive Committee to determine how best to reflect those broad interests and varied views, while also balancing the desire of AEA to influence evaluation policy in positive ways – even if some changes are improvements but not perfect. Better evaluation policy benefits the entire field of evaluation. Accordingly, the AEA and EPTF have a role in trying to ensure that evaluation policies align with the organization’s values and principles while also strengthening evaluation practice and evaluation use and influence.

In sum, AEA and the EPTF have worked strategically to formalize the Association’s processes and capabilities to influence evaluation policy, especially at the US federal level. In the process, AEA has taken steps to embrace the EPTF capabilities and institutionalize ways to consider its advice. We believe this has had positive returns in the recent advances in evaluation policy.

ORGANIZATION OF THIS ISSUE

This issue of NDE builds upon and updates the 2009 issue, Evaluation Policy and Evaluation Practice, edited by Bill Trochim, Mel Mark, and Leslie Cooksy. By 2022, much of the environment for establishing evaluation policy in the US government has dramatically
changed, largely due to the federal Evidence Act. The Evidence Act has also created a flurry of activity related to evaluation policy in federal agencies. Increased attention to evaluation policy is also evident outside the US federal government, such as in the philanthropic sector (Kinarsky & Christie, 2021). Among other changes since the earlier 2009 NDE on evaluation policy is a small but growing body of empirical research on the topic (Al Hudib, 2018; Al Hudib & Cousins, 2021; Christie & Fierro, 2012; Christie & Lemire, 2019; Dillman & Christie, 2017; Fierro, 2019; Kinarsky & Christie, 2021; Newcomer & Hart, 2022). The scholars who have conducted empirical research on evaluation policy are well represented among the contributors to this issue.

The chapters in the current volume are intended to provide an overview about the current landscape, largely focused on the federal government in the United States. The authors also delve into issues that lead to predictions about where evaluation policy may need further refinement, development, and attention in coming years. The roster of chapter authors includes individuals who have played central roles in developing and implementing evaluation policies in federal agencies, in laying the groundwork for and drafting of the Evidence Act, and in providing key guidance as to its implementation.

Although most chapters in this NDE issue focus primarily on the US federal government, implications exist for the entire evaluation community. Many organizations in other sectors already have evaluation policies (Al Hudib & Cousins, 2021; Kinarsky & Christie, 2021). Work at the federal level is likely to influence activity elsewhere, stimulating the formulation and revision of evaluation policies in other sectors. In Chapter 2, we reprint AEA’s Roadmap, updated by AEA in 2019 to reflect enactment of the Evidence Act. The Roadmap has broad applicability to evaluation policy in government at all levels.

In Chapter 3, Hind Al Hudib and J. Bradley Cousins report results of their systematic review of the written policies that have been developed to guide evaluation practice in organizations in a global sample of international development organizations, including those in the US federal government. Al Hudib and Cousins advance a new definition of evaluation policy, describe and illustrate the different content areas that may be included in an evaluation, and discuss possible connections between an evaluation policy and evaluation capacity building. In Chapter 4, drawing on their own research on evaluation policy, Leslie A. Fierro, Alana Kinarsky, Carlos Echeverria-Estrada, Nadia Bass, and Christina A. Christie explore the initial implementation of evaluation policies that were in use in the United States prior to the new Evidence Act requirements. Both Al Hudib and Cousins and Fierro et al. look backward at how evaluation policy evolved prior to recent sweeping mandates and guidance.

The fifth chapter of this issue, by Kathryn Newcomer, Karol Olejniczak, and Nicholas Hart, focuses on a requirement of the Evidence Act for agencies to establish learning agendas, or evidence-building plans. While some federal agencies and other organizations created these strategic plans for evaluation and evidence prior to the Evidence Act, the new requirement and expectation that agencies are using this tool offers an opportunity to identify strategies to apply evaluative thinking, effective engagement, and learning to evidence-building and evaluative activities. Similarly, a sixth chapter from Diana Epstein, Erica Zielewski, and Erika Liliedahl outlines the human role in evaluative activities and evaluation policy. Epstein et al. identify the role of the federal workforce in enabling evaluation, as well as the partnerships across the evaluation community and program partners to support effective evaluation. Both the Newcomer et al. and Epstein et al. chapters focus on the US federal government; again, evaluators working in other sectors should see implications for evaluation policies in the organizations they work with.

Because the concepts of evaluation policies, learning agendas, and the role of the workforce could seem somewhat abstract to some evaluators, a seventh chapter is included to
tie such elements together and present them in concrete terms. Former Chief Evaluation Officers at the US Department of Labor, Molly Irwin and Demetra Nightingale, coauthored this case study. It brings together the various concepts to illustrate how an agency could effectively implement evaluation policy in practice. This case study, examining the Department of Labor’s activities prior to the requirements of the Evidence Act, provides an example for agencies working to implement new requirements, by identifying lessons learned and approaches that may result in rapid scaling and effective adoption of evaluation policy in any agency.

This issue concludes with a short synopsis by the issue editors. Drawing on the chapters in this issue as well as other work and experiences, they identify several key issues and themes, project where evaluation policy will move next, and offer recommendations for the next wave of evaluation policy in the United States.

ORCID
Nicholas R. Hart https://orcid.org/0000-0003-3541-4337
Melvin M. Mark https://orcid.org/0000-0003-4545-1553

REFERENCES


U.S. Commission on Evidence-Based Policymaking. (2017). The promise of evidence-based policymaking: Final report of the commission on evidence-based policymaking. GPO.

**AUTHOR BIOGRAPHIES**

Nicholas R. Hart is President of the Data Foundation and Chair of the American Evaluation Association Evaluation Policy Task Force. He served as the Policy and Research Director of the U.S. Commission on Evidence-Based Policymaking and previously worked at the White House Office of Management and Budget during the Bush, Obama, and Trump Administrations.

Melvin M. (Mel) Mark, Professor of Psychology at the Pennsylvania State University, was co-editor, with Bill Trochim and Leslie Cooksy, of the 2009 NDE issue on evaluation policy, collaborated with Trochim on the establishment of the American Evaluation Association’s (AEA’s) Evaluation Policy Task Force, and has been a member since its inception, including time as co-chair.