This report is presented in its original form to preserve transparency. As such, the Department has not made any changes or edits to the language, which may vary from current policy. Thus, the content reflects the terminology used in the original research and the notion that political values, such as equity, are on par with objective scientific principles. The content contained therein may not reflect current DOL priorities, positions, or policies.

SUPPORTING AN EVIDENCE-DRIVEN CULTURE:
U.S. DEPARTMENT OF LABOR'S
EVIDENCE CAPACITY ASSESSMENT

January 2022

Carolyn Corea
Samia Amin
Amanda Ahlstrand
Jessica Cadima
Jessica Stockham
Gauri (Gee) Rege

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

TABLE OF CONTENTS

TABLE OF EXHIBITS

ACKNOWLEDGEMENTS

Many individuals contributed to the preparation of this report. We would like to thank Jody Dougherty and John Wendt for their contributions to the survey development and administration; Robert McMahon for his support in analyzing the survey data; Colleen McLelland for the preparation of all report exhibits and formatting; Mason Miller for his contributions to the evidence review; Raquel Sanchez for her support drafting the report; and Scott Davis for serving as the Quality Assurance Reviewer.

We also appreciate the input and participation of staff and leadership at the U.S. Department of Labor in support of this initial Evidence Capacity Assessment. We would especially like to thank Janet Javar, Chayun Yi, Meg Ferris, Christina Yancey, and Lauren Damme for their input, guidance, and assistance in the design and implementation of the assessment.

DOL'S EVIDENCE CAPACITY ASSESSMENT HIGHLIGHTS

The U.S. Department of Labor's (DOL) Chief Evaluation Office (CEO) oversees periodic assessments of DOL's capacity to produce and use evidence, with the aim of helping the Department and its agencies identify concrete steps for continuous improvement. Under CEO, the American Institutes for Research (AIR)/IMPAQ International, LLC (IMPAQ) (research team) designed and conducted this independent assessment.[1]

This evidence capacity assessment included the 16 DOL agencies in the Department's Strategic Plan.[2] It reflects data collected through a survey of targeted DOL staff, focus groups with selected DOL staff, and a review of selected evidence documents. The research team used a strengths-based approach that recognizes and leverages DOL's and agencies' progress and embodies a continuous improvement mindset.

The CEO and the research team collaborated to develop this summary of key survey and focus group findings that seem particularly actionable. In reviewing the findings, it is important to note that they are not representative of the Department as a whole. The survey and focus group findings have several important limitations: (1) the survey used a convenience sample of GS-13s and above even though staff in other grade levels also use and produce evidence; however, future surveys may include a broader sample of GS levels; (2) the survey targeted staff in occupations for which evidence use and production were assumed to play larger roles compared to other occupations; however, future surveys may include a broader list of relevant occupations; (3) the total number of staff in the survey was 3,446, but only a quarter responded to the survey[3]; and (4) the focus groups included 125 purposively selected staff from 15 agencies; however, future focus groups may be expanded to provide additional perspectives on evidence use and production.

Nonetheless, the findings offer important insights on baseline patterns and opportunities to improve capacity for evidence use and production going forward. It is important to note that this initial capacity assessment is intended, in part, to educate, facilitate, and engage staff for a more in-depth dialogue about evidence use and production within the Department over time. Future capacity assessment efforts should consider broadening the scope and sample of the assessment to build from this baseline effort, capitalize on progress made across the Department, and address more specific and nuanced research questions. Below, we present selected highlights relevant to the broadest array of stakeholders.

Do Staff Use Evidence Frequently? Around half of the survey respondents indicated frequent use of statistics, research, and analyses, and one third reported frequent use of evaluations (Exhibit

---

[1] This independent assessment was led by AIR Project Director, Samia Amin, under contract #1605C2-20-F-00033. The contents of this publication do not necessarily reflect the views or policies of the Department of Labor, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.

[2] This assessment uses "Department" or "DOL" when referring to the overarching entity of the Department of Labor. It uses "agency" when referring to individual components of the Department. "Agencies" are also referred to as "subagencies."

[3] The survey sample also excluded additional Bureau of Labor Statistics (BLS) staff after recognizing that BLS would be overrepresented in the sample. Additional information on the BLS sampling is discussed in Chapter 1.

1). Given that this survey is restricted to occupations for which evidence use is relevant, these findings suggest opportunities to improve in the future.

Exhibit 1: Frequency of Evidence Use Reported by Survey Respondents

53%
of respondents frequently use statistics, research, and analyses

48%
of respondents report that their ***teams*** frequently use statistics, research, and analyses

34%
of respondents frequently use evaluations (individuals and teams)

Are Staff Using Evidence in Key Areas of Decision-Making? Survey respondents selected a wide array of areas in which they use evidence (Exhibit 2). Over half of the respondents indicated using evidence for critical areas such as program strategy and goals, operating plans, process improvements, policy coordination, and communication with stakeholders. Focus group respondents indicated that capacity for evidence use had increased relative to two to three years ago. However, important areas remain where evidence use can and should be scaled. For example, a third or less of the survey respondents report using evidence for service improvements to constituents, research agendas or research questions, and resource allocation and agency budget recommendations.

Exhibit 2: Scope of Evidence Use Reported by Survey Respondents (Areas in Which Teams Typically Apply Evidence)

AREAS SELECTED BY MORE THAN HALF OF RESPONDENTS

- Program strategy and goals (**59%**)
- Operating plans (**53%**)
- Process improvements (**53%**)
- Policy coordination/ communication with stakeholders (**53%**)
- Program development or updates to programs (**52%**)
- Corrective to solve problems (**50%**)

AREAS SELECTED BY LESS THAN HALF OF RESPONDENTS

- Policy development or updates to policy (**44%**)
- Response to oversight inquiries (**44%**)
- Management coordination/ communication with stakeholders (**43%**)
- Resource allocation (**39%**)
- Service improvements for constituents (**34%**)
- Research agendas or research questions (**33%**)
- Agency budget recommendations (**29%**)

What Are Strengths of DOL's Evidence Capacity?

- **Awareness of agency's mission and strategic goals.** Almost all survey respondents report that they would be able to describe their agency's mission (95 percent) and strategic goals (76 percent), providing a strong foundation for focusing evidence activities.
- **Strong support for evidence use.** Almost three quarters of survey respondents feel supported by their supervisor (74 percent) and their agency (72 percent) in using evidence. Less than a fifth selected stronger political support as a key factor necessary for improving their team's ability to use evidence (18 percent) or produce evidence (16 percent). More than 80 percent of survey respondents did not express the need for political support to improve the use and production of evidence, suggesting that many respondents already feel supported by political leadership in this area. During focus groups, however, staff emphasized the importance of continued and consistent leadership support. They viewed it as critical for creating an environment that encourages using and producing evidence, and a necessary ingredient for investments in evidence infrastructure, processes, and staffing.
- **A precedent of using evidence in all important areas of decision making.** While there are clear opportunities for scaling evidence use in distinct areas, DOL agencies are not starting from scratch in any area and have notable proportions of staff reporting evidence use in every domain (Exhibit 2).

What Are Opportunities for Improvement?

- **Meeting high levels of need for training and other supports.** Over two thirds of survey respondents indicated *professional development/training* was needed for increasing capacity to use and produce evidence (Exhibit 3). Over half indicated needing *more time* and *improved access to software and/or hardware* to facilitate use and production of evidence. These three themes also surfaced prominently in staff focus group discussions.
    - **Training:** Most participants agreed that training on program evaluation or data analytics would help them be more successful in their current positions; they would make it a priority to attend trainings on topics relevant to their position, if given the time.
    - **Time:** Most focus group participants shared that heavy workloads and lack of time made it difficult to use and produce evidence. While some participants noted recent increases in staffing resources, participants across agencies most often noted that additional staff resources (e.g., new hires, improving access to existing staff's specialized skills, and supporting shifts in the allocation of staff time across tasks) are needed.
    - **Technology and data tools:** Participants noted that improvements to data infrastructure—especially software and hardware tools—are needed to help the staff better organize and sort through existing data. Specific needs included computer hardware and analytical software appropriate for data analysis; improved user interface with new and existing software; improved data infrastructure and systems to allow more

analysis; and data visualization tools that are more accessible for individuals with disabilities.

Exhibit 3: Survey Responses on Priority Supports Needed to Improve Evidence Capacity

- **Improving access to evidence.** While 70 percent of survey respondents agreed or strongly agreed that their team has access to statistics, research, analyses, or evaluations it needs to improve programs, policy, or operations, this means that 30 percent may lack such access. When probed on specific aspects of the quality of evidence available, satisfaction rates were lower: half or less of the respondents indicated being satisfied or very satisfied with their access to evidence that was comprehensive (50 percent), independent (49 percent), and appropriate in use of methodological approaches (49 percent). During staff focus groups, participants highlighted the need for timelier and time-efficient access to existing data and evidence, enhanced ability to access raw data and link relevant datasets, and more granular data collection (particularly on subpopulations of interest). Staff also expressed a desire for evidence products that are more user-friendly to agencies' wide range of stakeholders so they can obtain meaningful insights more quickly and easily.
- **Understanding disparities in the experience of leadership and staff related to relevance and use of evidence; and improving staff experience.** The survey revealed differences in responses from GS-13, -14, and -15–level respondents compared with those from Senior Executive Service (SES) respondents. Respondents at the GS-13, -14, and -15 levels were less likely to agree than SES respondents that there is an appropriate balance in the time spent using and producing evidence. Non-SES survey respondents were also less likely to be familiar with agency learning agendas and how they relate to their work. Finally, GS-13, -14, and -15 respondents were less likely to perceive support from their supervisor and agency to use evidence. They were also more likely to note the need for political support. During focus groups, staff repeatedly highlighted the importance of leadership support for the production and use of evidence and noted challenges that arise when leadership does not fully recognize the time needed and tasks involved in using or producing evidence.

How Will Findings be Used? DOL's overall capacity assessment approach, which comprises activities beyond the requirements of the Evidence Act, includes disseminating findings to, and obtaining additional feedback from, agencies through leadership discussions/briefings and internal agency capacity summaries. This provides a powerful opportunity for a collective and concerted effort to improve DOL capacity to use and produce evidence, and for building on the foundation the research team has documented in this assessment.

1. EVIDENCE CAPACITY ASSESSMENT DESIGN

This chapter begins with a description of the purpose of the Department's evidence capacity assessment, then provides an overview of the assessment approach, and concludes with a description of the data collection and analysis efforts.

1.1 Purpose and Guiding Principles of the Evidence Capacity Assessment Design

This assessment develops a baseline understanding of DOL's capacity to produce evidence that the Department and its agencies can use to guide decisions and next steps for improving capacity.[4] Under the direction of the Chief Evaluation Office, the research team used the following guiding principles to design its approach:

- **Demonstrate usefulness of evidence activities to DOL agencies.** The research team intentionally focused on agency strengths, capacity, and current and future needs. The team also integrated agency-level reporting so that the assessment findings would be meaningful and actionable for agencies, given that they each have unique missions, goals, and resources.
- **Plan for learning and iteration over time.** The research team approached this assessment as a first step–a foundation–that the Department could build and expand upon over time. As such, it explores a foundational set of strengths, capacity needs, and constraints that DOL can feasibly address, rather than an exhaustive (and potentially overwhelming) capacity assessment. Opportunities for improvement are highlighted to support evidence building over time.
- **Tailor Evidence Act requirements to DOL context.** The team organized and addressed Evidence Act and Office of Management and Budget (OMB) requirements and domains to (1) allow for a systematic and independent assessment that generates useful and actionable findings; (2) gather information on current DOL staff's familiarity and knowledge of evidence; and (3) minimize the burden on participating staff to sustain support for this effort long-term. This tailoring effort involved adapting the assessment domains and definitions such that they met both Evidence Act and OMB requirements, as well as Department stakeholder needs related to ease of understanding, feasibility, and efficient use of participant time for the broader set of stakeholders included in the capacity assessment. We were also mindful of rightsizing the capacity assessment effort to generate interest and secure buy-in for future iterations of the capacity assessment. We see this as a first step of an assessment that can expand in depth and breadth over time. Appendix A lists Evidence Act and OMB requirements for the capacity assessment, demonstrates how they have been organized and operationalized for DOL capacity assessment, and notes where to find relevant findings for each requirement in this report.

[4] Certain agencies such as the Bureau of Labor Statistics (BLS), CEO, and the Chief Data Office (CDO) have a specialized capacity to use and produce evidence because those tasks are central to their mission. This initial capacity assessment was designed to efficiently encompass a broader set of agencies in order to set a common baseline of capacity to use and produce evidence across the Department. If feasible, DOL may consider supporting future capacity assessments tailored to specific advanced capabilities, such as BLS and its partner statistical agencies conducting a targeted assessment that focuses on the innovative skills needed to meet their mission. These could include additional staff surveys and/or targeted focus groups.

DOL's capacity assessment focuses on the following types of evidence:

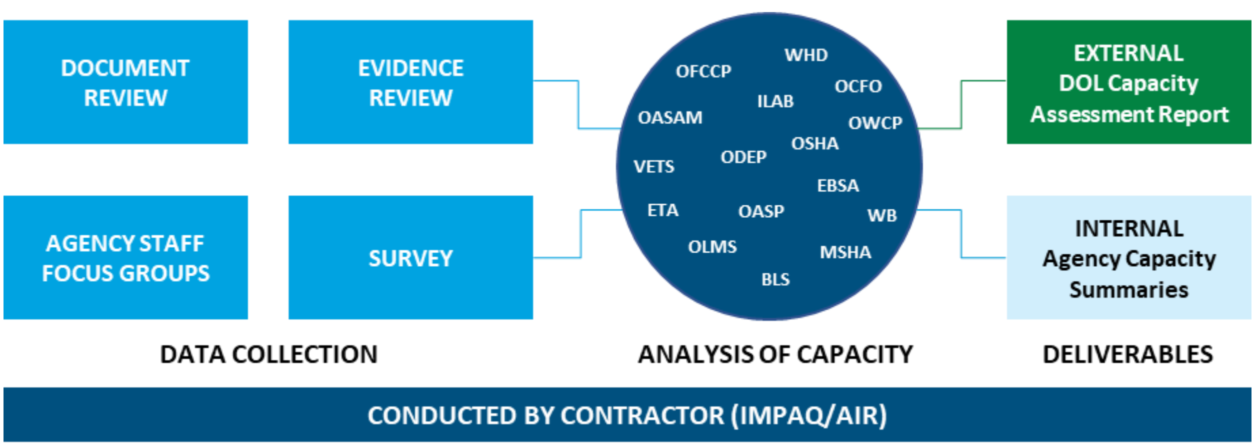
- **Statistics, research, or analyses** such as data analytics, trend analyses, or literature and/or document reviews on specific programs or policy issues.
- **Evaluations** that provide a formal assessment of implementation, impacts, effectiveness, or efficiency of a program, policy, or organization.

The research team's knowledge development and survey cognitive testing activities found it was not viable to assess capacity separately for "statistics," "research," "analysis," and "evaluations" because of the interrelated conceptual understanding among staff. In an effort to develop data collection instruments (survey and focus group protocols) that minimized the time burden and reflected existing knowledge of types of evidence among staff, the research team clustered all types of evidence together for most questions. A small subset of survey questions separately explores (1) statistics, research, or analyses; and (2) evaluations.

1.2 Overview of Evidence Capacity Assessment Approach

The research team undertook knowledge development activities to understand the scope of evidence activities that may occur, agencies' prior patterns of evidence production and use, and levels of maturity in understanding evidence concepts to fine tune data collection instruments for this assessment.[5] As shown in Exhibit 4, the approach to the capacity assessment draws on four sources of data (document review, review of evidence-related activities, web-based survey, and staff focus groups). These data sources inform the analysis of capacity for 16 agencies, providing insights on 10 domains of interest to DOL and producing several DOL-wide and agency-level deliverables.

Exhibit 4: Overview of Evidence Capacity Assessment Approach



[5] Knowledge development activities included interviews with leadership from the CDO, Performance Management Center, and BLS to understand their roles in supporting agencies in using and producing evidence and whether their data collection activities are relevant to this effort. Knowledge development activities also included interviews with CEO liaisons who work with various agencies across the DOL to inform data collection protocols.

DOL included the 16 agencies that contribute to DOL's Strategic Plan in this evidence capacity assessment. Exhibit 5 lists these agencies as three general types based on their functions: (1) employment and training programs; (2) compliance/enforcement programs; and (3) statistical, policy, and management support for DOL agencies (support agencies). We categorized agencies by type to analyze similarities and differences in evidence use and production, recognizing that each agency serves a unique and specific mission.

Exhibit 5: DOL Agencies Involved in Evidence Capacity Assessment, by Type

Employment and Training	Compliance/Enforcement	Statistical, Policy, and Management Support[6]
Bureau of International Labor Affairs (ILAB)[a]	Employee Benefits Security Administration (EBSA)	Bureau of Labor Statistics (BLS)[b]
Employment and Training Administration (ETA)	Mine Safety and Health Administration (MSHA)	Office of the Assistant Secretary for Administration and Management (OASAM)
Office of Disability Employment Policy (ODEP)	Occupational Safety and Health Administration (OSHA)	Office of the Assistant Secretary for Policy (OASP)
Veterans' Employment and Training Service (VETS)	Office of Federal Contract Compliance Programs (OFCCP)	Office of the Chief Financial Officer (OCFO)
Women's Bureau (WB)	Office of Labor-Management Standards (OLMS)	
	Office of Workers' Compensation Programs (OWCP)	
	Wage and Hour Division (WHD)	

[a] ILAB has been categorized as an employment and training agency, but also it plays an important role in protecting the rights of workers internationally.
[b] See Appendix B: Assessing the Statistical Capacity at the U.S. Department of Labor prepared by BLS.

Data Collection Activities. Below is a description of the four data collection activities that form the basis for the overall assessment:

- **Document review.** The research team conducted a review of current, relevant documents to glean important context for the assessment and inform the overall approach to the capacity assessment. In addition, this review provided agency-specific context for the research team that was relevant for focus group discussions and helped the team coordinate the assessment with related and concurrent DOL efforts/initiatives. We reviewed public and nonpublic documents, including:
  - DOL Strategic Plans (Fiscal year [FY]2018–2022 and draft FY2022–2026),
  - DOL Annual Performance Report (FY2019),

[6] This evidence capacity assessment limited its focus on the following offices within the Statistical, Policy, and Management Support agencies: Performance Management Center (PMC) within OASAM, Enterprise Risk Management (ERM) within OCFO, and the Office of Regulatory and Programmatic Policy within OASP. Future DOL evidence capacity assessments may include additional offices from these and other DOL agencies that were not included in this assessment.

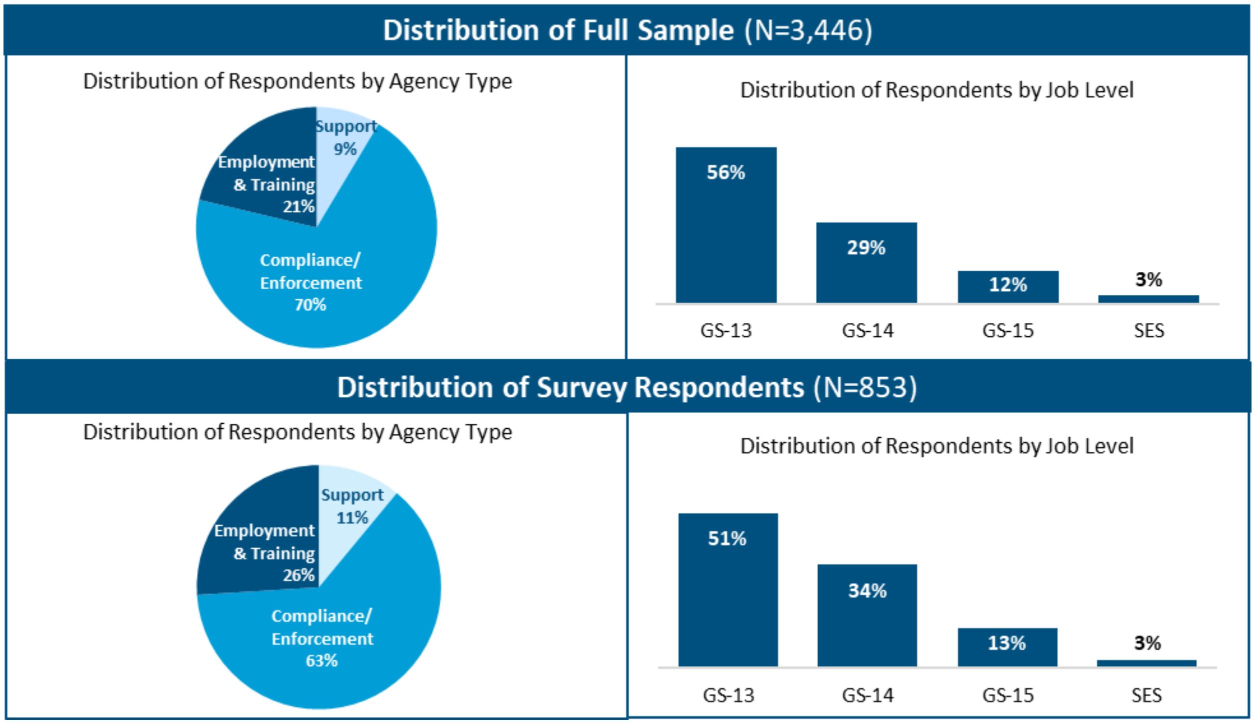
  - DOL Agency Financial Report (FY2020),
  - DOL Agencies' Annual Operating Plans (FY2021),
  - DOL Interim Capacity Assessment (September 15, 2020),
  - DOL Evaluation Plan (FY2019), and
  - resources from other federal agencies, including related capacity assessment surveys and a cross-agency paper on Evidence Act implementation (FY2020–2021).

- **Survey.** We administered an online survey to selected departmental career staff members at the 16 agencies. Eligible respondents included staff members who were at the GS-13, -14, -15, and SES levels and in occupations that the research team and DOL identified as most likely to regularly use or produce statistics, research, analysis, or evaluations. Appendix C presents additional information on the sample selection criteria. We designed the survey to capture a baseline understanding of:
  - the scope and composition of evidence activities,
  - each respondent's personal capacity for evidence use,
  - each respondent's perception of their office's capacity for evidence use,
  - perceived barriers to agency use and production of evidence that can help inform opportunities for improvement, and
  - each respondent's priority training needs related to using and producing evidence.

The survey was fielded between March 2, 2021, and April 26, 2021, to the survey sample described above. The survey was left open through May 7 to allow members of the survey sample who were also focus group participants a final opportunity to complete the survey (See Appendix C for survey sample selection criteria and Appendix D for the survey instrument.) The survey completion rate was 24.8 percent (including six percent of the sample that were partial completes).[7] As such, findings from the survey must be interpreted keeping these limited response rates and the potential for bias in mind. Exhibit 6 presents the distribution of the survey sample and survey respondents across agency and staff levels.

[7] The response rate (number of respondents divided by the number of individuals in the sample frame) by job level is as follows: 22 percent of GS-13s; 29 percent of GS-14s, 27 percent of GS-15s, and 22 percent of SESs. Similarly, the response rate by agency type is as follows: 30 percent for employment and training agencies; 22 percent for compliance/enforcement agencies; and 32 percent for support agencies.

Exhibit 6: Distribution of Survey Sample and Survey Respondents by Agency and Staff Level



Note: Distribution of respondents by job level does not total 100% due to rounding.

- **Focus groups with agency staff.** The capacity assessment also draws on focus group discussions with agencies. From May through November 2021, the research team convened 16 focus groups with agency employees from 15 agencies who were likely to use data and evidence in their work. In total, 125 national and regional employees participated in the staff focus groups, representing ODEP, WHD, ILAB, OWCP, VETS, ETA, OFCCP, EBSA, OSHA, WB, MSHA, OLMS, OASAM, OCFO, and OASP. During these discussions, the research team explored the following:
  - Examples of capacity for evidence use and production
  - Areas for improvement
  - Perceived barriers to and facilitators of evidence use and production
  - Training and other supports to improve evidence use and production
  - Vision and goals for evidence use and production and the supports needed to achieve them
- **Evidence review.** The research team conducted a review of a selection of publicly available evidence products to understand coverage, effectiveness, quality, methods, and independence. The review leveraged a data inventory that the CEO compiled for the interim capacity assessment. It includes a thorough review of selected evaluations and/or foundational fact-finding reports published in 2018 or 2019 for each agency. Appendix E includes findings from the evidence review along with the methodology for the review.

Exhibit 7 maps the assessment domains to their relevant data collection activities.

Exhibit 7: Assessment Domains Mapped to Data Collection Activities

	Data Collection Activities		
Assessment Domains	**Survey**	**Focus Groups**	**Evidence Review**
Coverage			
Types of evidence use reported	■		■
Frequency of evidence use and production	■		■
Satisfaction with access to evidence	■		
Uses			
Frequency of different types of evidence use activities	■	■	
Strengths and examples of evidence use and production	■	■	
Examples of investments in capacity to use and produce evidence		■	
Context of evidence use		■	
Effectiveness			
Satisfaction with access to evidence that is relevant, useful, and effectively disseminated	■		
Degree to which evidence meets staff needs		■	
Evidence activities that identify relevance of findings, present results that are easy to understand, and can be readily found on a public website			■
Balance			
Assessment of appropriate amount of time spent using or producing evidence	■	■	
Quality, Methods, and Independence			
Satisfaction with access to high-quality evidence	■		
Satisfaction with access to evidence that is appropriate in its use of rigorous methodological approaches	■		
Satisfaction with access to evidence that is independent (i.e., free from bias)	■		
Factors that determine whether staff have access to high-quality evidence		■	
Evidence activities that document their data sources and analysis methods well, use appropriate methods, note strengths and weaknesses of their approach, and are conducted by a third party			■
Equity			
Satisfaction with access to evidence that is helpful in improving underserved populations' access to the benefits and services that an agency offers	■		
Individual and Office Capacity			
Staff experience creating logic models and using DOL's Clearinghouse for Labor and Employment Research (CLEAR)	■		
Staff capacity to keep up with research relevant to their work	■		

	Data Collection Activities		
Assessment Domains	**Survey**	**Focus Groups**	**Evidence Review**
Staff capacity to assess whether information they read represents strong or weak evidence	■		
Staff capacity to describe their agency's mission, strategic plan goals, operating plan, and learning agenda	■		
Staff perceptions of their team's capacity to use and produce evidence	■	■	
Office and Agency Context			
Staff perceptions of support from their direct supervisor and agency management in using evidence	■		
Staff reporting that increased political support would increase/improve their ability to use/produce evidence	■		
Insights from staff about enabling conditions for using and producing evidence		■	
Aspirations for Using and Producing Evidence			
Insights from staff about desired goals for using and producing evidence		■	
Supports and Training Needed			
Resources and supports that staff identify as helpful in improving their team's ability to use and produce evidence	■	■	
Training topics and/or modes staff identify as helpful in improving their team's ability to use and produce evidence	■	■	

Data analysis and synthesis of findings. After systematically analyzing the data, the research team synthesized the findings to document DOL's and agencies' baseline capacities to use and produce evidence within each of the assessment domains. To do this, the team integrated analyses of the different data collection activity findings and grouped key takeaways by assessment domain in the deliverables described below. These syntheses provide the information DOL needs to assess the Department as a whole and to initiate discussions with agency leadership about the status of each agency's progress (i.e., the desired state appropriate for that agency given its mission, operations, resources, and needs; and the appropriate milestones given the baseline). To facilitate these iterative discussions with agencies, the research team is helping DOL develop a framework for an evidence maturity model suited to DOL's mission and context (See Appendix F).

Capacity assessment findings. Findings from the capacity assessment activities culminate in this **Department-wide Capacity Assessment**, which presents results from the department-wide survey and the selected evidence review as well as insights from analysis of qualitative data from agency-level focus groups. Findings will also be used to disseminate findings to, and obtain additional feedback from, agencies through leadership discussions/briefings and **internal agency capacity summaries.** The goals of the agency discussions are to understand each agency's vision for evidence use in the future; and identify actionable steps, existing strengths, desired supports, and lessons learned that can help move them toward that vision. DOL's overall approach to the capacity assessment is intended to improve agencies' capacity to use and produce evidence, and to build on the foundation the research team has documented in this assessment.

2. COVERAGE, USES, EFFECTIVENESS, AND BALANCE

This chapter focuses on four assessment domains: coverage, uses, effectiveness, and balance.[8] The chapter provides findings from the survey and describes how focus group discussions supplement these findings.

2.1 Coverage

COVERAGE:

- Frequency of staff evidence use and production
- Staff satisfaction with access to comprehensive evidence

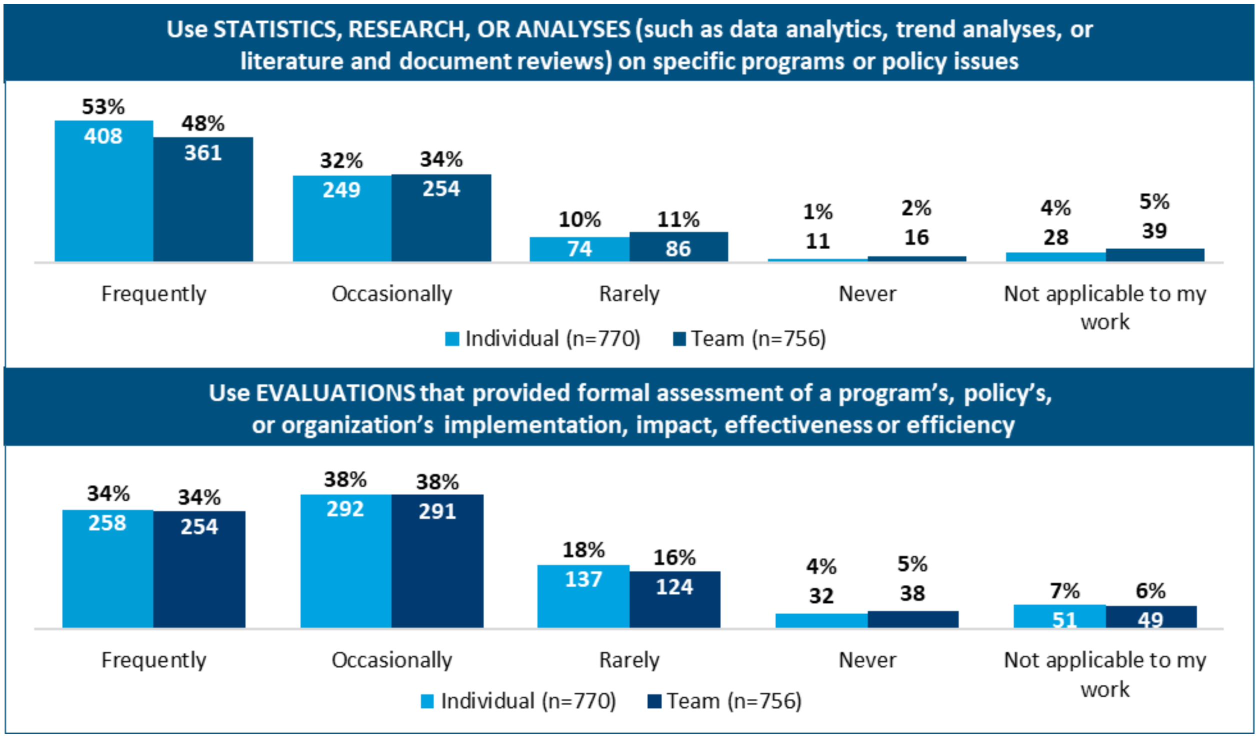
To complement separate DOL activities that tabulate evidence activities by agency and probe learning needs for areas of DOL investment, we focused on understanding staff contributions to and experiences with evidence coverage. Coverage of evidence activities is defined as the frequency at which DOL staff self-report using and producing evidence and staff satisfaction with having access to comprehensive evidence. This section presents findings from the survey on both topics, calling out notable variations by agency type and staff level when relevant.[9] Appendix E presents insights on the frequency and types of evidence activities conducted based on our review of DOL's evidence inventory (which contains published evaluations and research reports). Exhibit 8 presents the frequency of respondents using different types of evidence for themselves and their team over the past year.[10]

[8] The assessment domains and definitions were adapted to meet both the Evidence Act and OMB requirements, as well as Department stakeholder needs related to ease of understanding, feasibility, and efficient use of participant time for the broader set of stakeholders included in the capacity assessment. See Section 1.1 for more information.

[9] Focus groups offered limited opportunity to probe issues specific to coverage. (See Sections 2.2 [Uses] and 2.3 [Effectiveness] for focus group findings on how staff use evidence and their access to effective evidence. Findings on the availability of resources to support evidence use and production are presented across the relevant focus group sections.)

[10] The survey included some questions that referenced respondents' experiences over the "past year." Given that the survey was fielded in spring 2021, the responses roughly cover the period from spring 2020 to spring 2021.

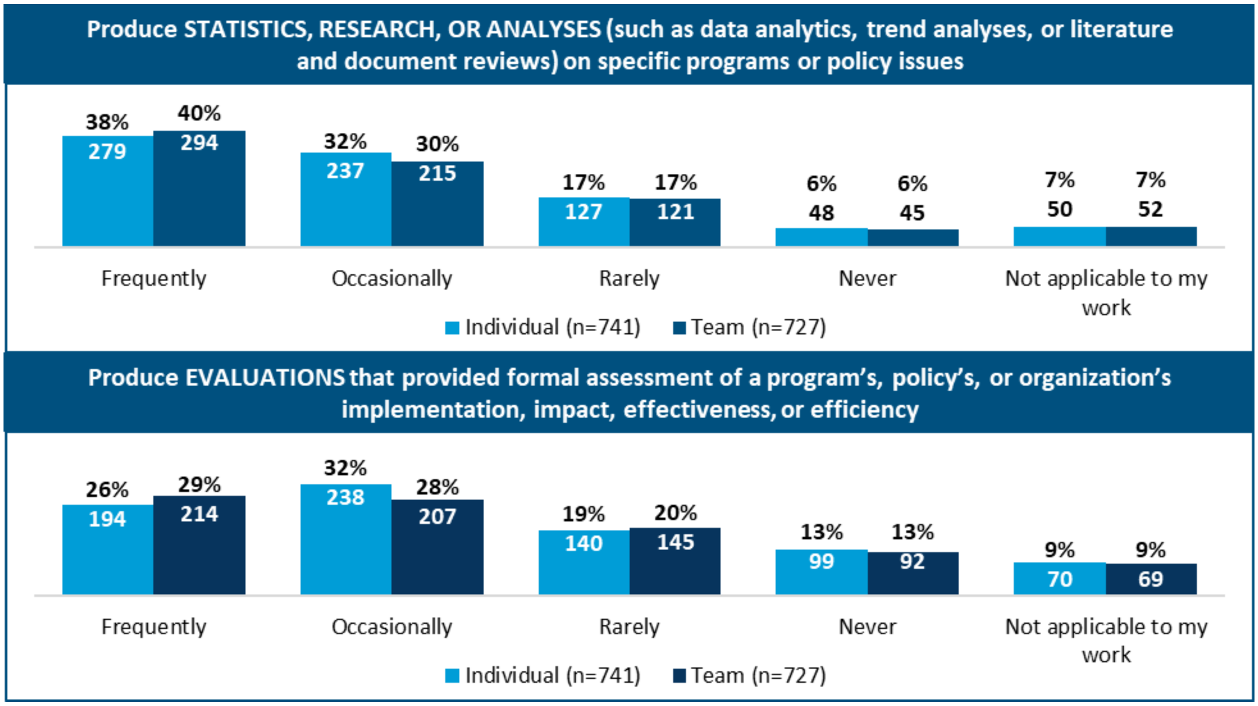
Exhibit 8: Frequency of Evidence Use Over the Past Year



- **Respondents used statistics, research, or analyses more frequently than evaluations.** As shown, around half of the respondents reported frequent use of statistics, research, and analyses for themselves and their teams compared with one third reporting frequent use of evaluations. There was less variation in the proportion of respondents selecting occasional use of both types of evidence (ranging between 32 and 38 percent). This trend remained consistent across different agency types (employment and training, compliance/ enforcement, and support agencies) and job levels. The more frequent use of statistics, research, and analyses could be due, at least in part, to that category of evidence being broader than evaluations, encompassing a wider range of activities. The more occasional use of evaluations may reflect the specificity of evaluation findings relative to the broad set of policy or program issues an agency may face. Another possible explanation for this finding may be the lack of relevant evaluations.
- Notably, **more than a quarter of respondents report rarely or never using evaluations or consider them not applicable to their work**, suggesting there may be room for improvement in incorporating the use of evidence, given that the survey focused on occupations for which evidence use could reasonably be considered a part of their roles and responsibilities.
- Additionally, **the reported use of statistics, research, or analyses increased by job level among the respondents.** About 63 percent of GS-15 and SES respondents, 54 percent of GS-14s, and 49 percent of GS-13s said they frequently used statistics, research, or analyses for themselves. These variations by job level do not surface in the reported frequency of using evaluations.

Exhibit 9 presents the frequency of respondents producing different types of evidence for themselves and their team over the past year.

Exhibit 9: Frequency of Producing Evidence Over the Past Year



- **Respondents were more likely to use evidence than produce evidence.** The majority of respondents reported frequent or occasional evidence production at the individual and team levels, although at relatively lower levels than those reported for evidence use. Less than half of the staff reported frequently producing statistics, research, or analyses compared with more than three quarters that reported using this type of evidence. Frequency of production of evaluations was similarly lower than use, although the differences were less notable. More than 40 percent of staff reported never or rarely producing evaluations or considered these not applicable to their work; more than 30 percent reported the same when asked about producing statistics, research, or analyses. These results could be a function of agency partnerships with CEO, Chief Data Office (CDO), and BLS to produce the evidence they need to inform their work. They also suggest an opportunity to expand agency capacity to produce both types of evidence, for example through training, professional development, and greater engagement of agencies in the design of evidence produced by CEO, CDO, BLS, and other agencies. Further investigation into the reasons for lower levels of evidence production as well as an exploration of how agencies partner with CEO, CDO, and BLS may help shed light on how to interpret these findings and increase capacity for evidence production across the Department.

- **Support agencies reported producing statistics, research, and analyses at slightly higher rates** than other agency types (i.e., employment and training, compliance/enforcement); 47 percent of staff from support agencies reported frequently working to produce statistics, research, and analyses, and 49 percent reported that their teams frequently worked to produce statistics, research, and analyses over the past year.[11]

- **Some variation surfaced in production of statistics, research, or analyses by job level**, with senior-level staff reporting more frequent production of this evidence type than lower GS-levels. For example, about 50 percent of GS-15 and SES staff reported that their teams frequently produce statistics, research, or analyses compared with 42 percent of GS-14 and 36 percent of GS-13 respondents. Overall, the data suggests that senior positions were more likely to report using and producing this type of evidence than lower GS-levels. The responses on producing evaluations were similar across agency types and job levels.
- **Only half of the respondents are satisfied with their access to comprehensive evidence.** Staff satisfaction with access to comprehensive evidence forms another important aspect of the coverage domain.

Exhibit 10: Staff Satisfaction With Comprehensive Evidence

Exhibit 10 shows that half of respondents said they were satisfied or very satisfied with their access to comprehensive evidence, with another 29 percent reporting they were somewhat satisfied. It also shows 14 percent of respondents as not satisfied with their access to comprehensive evidence, with another 7 percent reporting they did not know or it was not applicable. These findings were consistent across agency types and job levels.

2.2 Uses

To understand the capacity to use and produce evidence from a strengths-based perspective, and the contexts in which evidence activities occur, the research team examined the types of decisions informed by evidence, the contexts of evidence use, and examples of investments in evidence. This section first presents findings from the survey on how staff report using evidence to inform their work and concludes with a summary of related findings from the focus groups on how staff use evidence in what contexts.

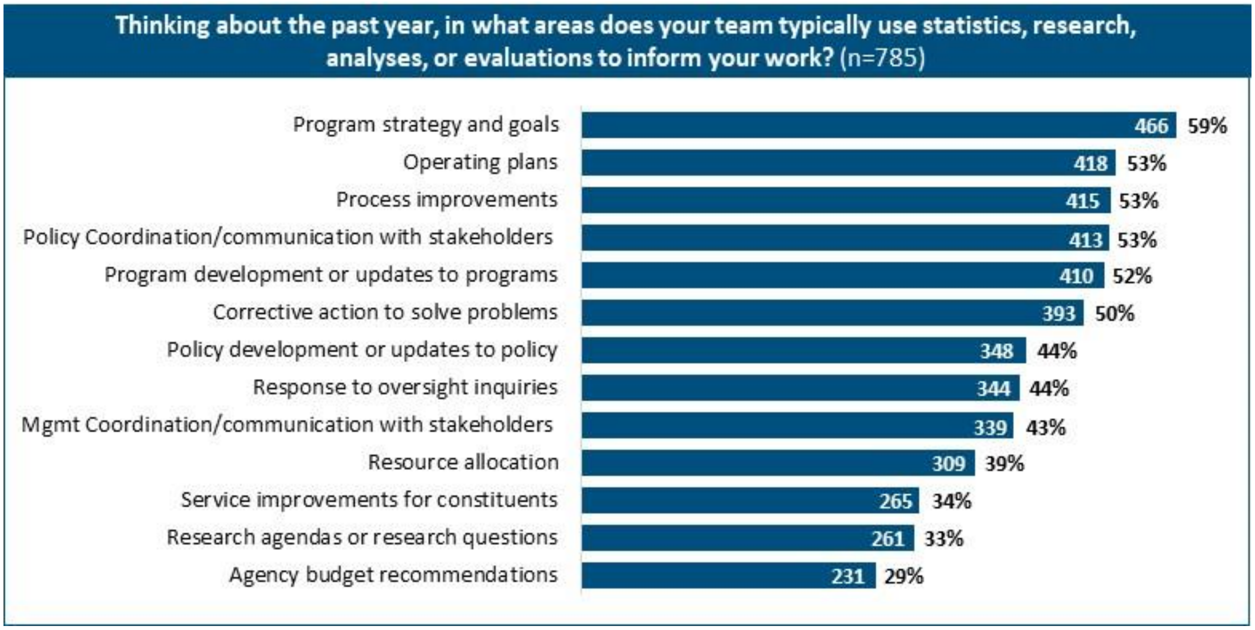
Uses:

- How staff use evidence
- Examples of evidence use
- Context of evidence use
- Strengthening factors for evidence use

[11] The percentages of respondents reporting that they frequently worked to produce statistics, research, and analysis by agency type were 35 percent (compliance/enforcement) and 39 percent (employment and training). The percentages of respondents reporting that their teams frequently worked to produce statistics, research, and analysis by agency type were 37 percent (compliance/enforcement) and 45 percent (employment and training).

Exhibit 11 presents the areas in which survey respondents across all agencies reported that their teams use evidence to inform their work.

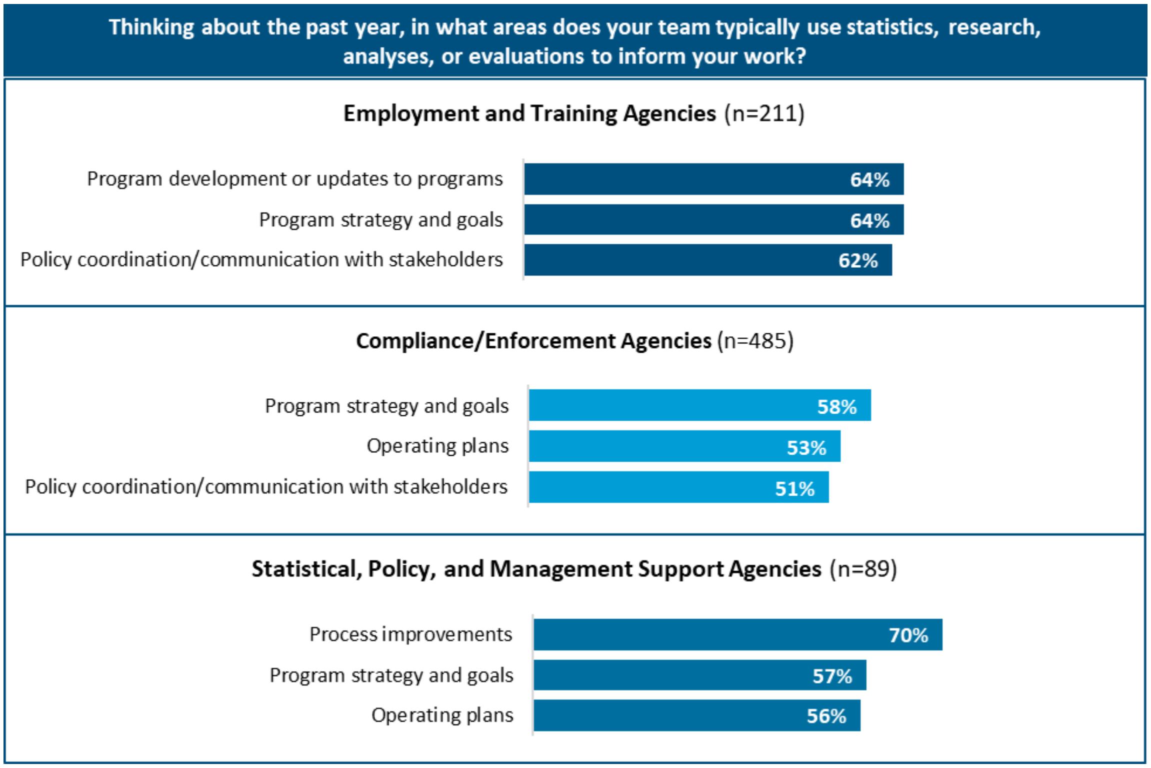
Exhibit 11: Areas That Typically Involve Evidence Use Across Agencies



Note: Staff members were encouraged to select as many areas as they thought applicable.

- **Respondents use evidence across a wide range of areas, but areas remain for improved use of evidence.** Looking across agencies (Exhibit 12), respondents reported using evidence in a wide range of areas, with half or more reporting that they use evidence in 6 of the 13 areas listed as response options, and no fewer than a quarter of respondents reported using evidence in any single area. The most frequently selected areas for evidence use were to inform program strategies and goals, operating plans, process improvements, coordination, or communication efforts with stakeholders on policy, and also program development or updates to programs. There are several important areas where the use of evidence could be enhanced, including service improvements for constituents, research agendas or research questions, and agency budget recommendations; all areas selected by a third or less of respondents.

Exhibit 12: Top Three Areas Where Staff Typically Use Evidence, by Agency Type



- **Respondents use evidence in similar areas, by agency types.** Program strategy and goals are among the top three areas selected by respondents from each agency type (Exhibit 12). Respondents from support and compliance/enforcement agencies also included operating plans in their top three, while respondents from employment and training and compliance/ enforcement agencies included coordination/communication with stakeholders on policy among their top three areas for using evidence. Respondents from employment and training agencies uniquely included program development or updates to programs in their top three areas for evidence use, and respondents from support agencies uniquely included making process improvements. Respondents from employment and training agencies were more likely to report evidence use in most areas compared with compliance/enforcement and support agencies (i.e., greater proportions of respondents from employment and training agencies reported that they use evidence in 9 of the 13 areas).
- **Respondents at higher GS or at SES levels are more likely to use evidence in areas related to strategic planning and budget.** Responses to this survey question by staff level reveal some interesting patterns. Those who were SES and, in some cases GS-15-level, were more likely to report using evidence to inform their work in areas mostly related to strategic planning and budget (program strategy and goals, operating plans, and agency budget recommendations), as well as some areas of program management and operations (program development or updates, resource allocation, and process improvements). In areas related to policy, similar proportions of respondents at each staff level reported using evidence. These differences may be driven by differences in roles at different GS-levels or other factors. Future capacity assessments may seek to gather more granular data on staff roles and patterns of evidence use to provide more meaningful interpretations of the data.

Focus Group Insights on Uses of Evidence

Focus group participants discussed the context of their evidence-use activities and the areas they or their teams are most likely to use evidence to inform decisions, and they provided illustrative examples of instances in which they use or produce evidence to inform decision making. Participants also shared strengthening factors that support their use and production of evidence. We have summarized insights from those discussions to better understand the ways evidence use and production inform the Department's work.

Focus Groups in 15 DOL agencies (n= 125).
Not a random sample.

EVIDENCE USE EXAMPLES

Participants reported using evidence in many ways to inform their work, citing examples across three categories: *Strategic Planning and Budget*, *Policy*, and *Program Management and Operations*.

Strategic Planning and Budget. Examples of using evidence in the area of *Strategic Planning and Budget* were concentrated on problem solving to better craft program strategies and goals, create operating plans, and develop agency budget recommendations. A few examples included how personnel actively use evidence to inform research agendas or research questions.

- **Program strategy and goals.** Participants noted they use findings to help make more informed evidence-based decisions for program strategies and goals. Using existing evidence, staff members identify trends and the specific areas of an industry or population where they need to target their initiatives. Participants mentioned using evidence to understand how to do things differently in the future.

"Performance evaluations…how we might be able to do things differently, next time we fund a similar project… to make them productive for the grantee."

- **Agency management plans, or "AMP" (previously called operating plans).** Participants have used the evidence from research studies and literature reviews as supporting references in operating plans. Agencies have created ways to collect and track program service activities to track their progress in meeting planned targets and redirect efforts as needed.
- **Agency budget recommendations.** Some participants said they rely heavily on using their statistical data (from past performance, current production numbers, and historical trends) to better estimate office needs and allocate Department resources, including staffing for upcoming fiscal years. Doing so helps the participants ensure they can meet agency goals with their proposed and allocated budget resources.

"We can extract data in real time, and we report our quarterly 'production' to the Office of the Assistant Secretary."

- **Research agenda or research questions.** Staff examples of evidence use in the area of informing research agendas included prioritizing research investments for projects that are working with the populations or issues of interest that address the broader learning priorities, as well as using existing evidence to understand how the evidence base has changed. Some respondents mentioned coordinating across offices within their agency to understand how best to address gaps in evidence and make efficient use of resources.

Policy. Respondents reported using evidence to inform policy in diverse ways:

- **Policy development and updates to policy.** Staff members review quantitative and qualitative evidence to inform policy-related questions that have been brought to their attention, which also helps the staff identify where gaps exist in the data. The evidence available is used to help clarify policy-related questions from stakeholders, update FAQs, and make needed policy statements and clarifications. Additionally, staff members have been able to formalize certain data collection processes on stakeholder interactions with the agency that aides the staff members in shaping their policy work.
- **Coordination and communication efforts with stakeholders.** Participants said they have used evidence/ data to review trends, monitor performance, and inform policy-related communications with stakeholders including grantees, employers, practitioners, and policy makers.

"We would look for trends, which helps us focus [stakeholder] education. ... We try to intervene before more serious issues arise."

- **Response to oversight inquires (Congressional, OMB, Government Accountability Office [GAO]/Office of Inspector General [OIG]).** Participants frequently described the use of evidence to support congressional testimony or speeches and respond to GAO and other oversight organization inquiries/recommendations.
- **Program Management and Operations.** Participants cite examples of using evidence for *Program Management and Operations* primarily for deciding staff resource allocations, process improvements, and corrective action to solve problems. A few participants shared examples of using evidence for service improvements to constituents.
- **Resource allocation.** Participants said they have used evidence to determine how to allocate resources including staff, tools, and materials.

"We do four things with the same data –1) target violations and assign cases; 2) [monitor] case performance progress; 3) measure results of the cases; and 4) measure impact – how many people got it right the next time – see if behavior has changed. It all comes from the same database. We have tools we put together to enable data analyses of the performance data and the roles we enforce."

- **Process improvements.** Participants said they have used evidence to track and monitor trends to identify areas for improvement and inform changes that are needed.
- **Service improvements to constituents.** Participants shared that they have used evidence to look for ways to better serve their program constituents/stakeholders and adjust program services.

"Every year we keep a Top 10 list of what standards are most cited—and things that are persistently on the top are focused on. We then use BLS data to see if the injury trends are also going up, to give us a holistic sense of maybe this is where we want to focus."

- **Corrective action to solve problems.** Participants discussed using evidence to monitor and measure compliance and performance and make informed decisions about targeted enforcement strategies.

Strengthening Factors

Across different agencies, participants shared many examples of strengths in evidence use and the factors that have helped support and build up their (and/or their agency's) evidence-use capacity.

INVESTMENTS IN STAFFING AND OTHER RESOURCES

- Creating and hiring for new positions requiring strong technical skills.
- Improving access to data analysis software.

> *"Access to software has improved in recent years. In general, there is an openness to providing the software we need."*

> *"Without the data we would have been shooting in the dark. Collection of data made it possible to communicate."*

- Establishing data collection systems that create fresh opportunities for improved data collection planning and use.
- Increasing management's support for data initiatives and purchasing software to better support data infrastructure.

LEVERAGING INSTITUTIONAL KNOWLEDGE/EXPERTISE

- Using staff members who possess strong institutional knowledge for using and producing evidence.

COLLABORATING ACROSS AGENCIES TO SHARE DATA AND EXPERTISE

- Accessing data collected by other agencies (internal and external to DOL) to help agencies meet needs for evidence.
- Using evidence to support efforts to build interagency collaboration.
- Improving agency data teams' efforts to increase their capacity by building networking relationships across the Department and other agencies to share reporting tools and data visualization techniques.

> *"That team has had a lot of opportunity for relationship building and networking across Department of Labor and even with other agencies – they also did a great job of building capacity before the team was as big as it is."*

RELATIONSHIP-BUILDING BETWEEN PROGRAM AND DATA TEAMS

- Increasing collaboration between agency program and data teams so staff truly understand the data, the needs around it, and what is feasible.

> *"If we know some things are not possible, we try to go in early or give our input and knowledge which saves a lot of time."*

SUPPORT FROM THE CHIEF DATA OFFICE (CDO) AND CHIEF EVALUATION OFFICE (CEO)

> *"The partnership between planning and data is really kind of critical to our general efforts around evidence, for sure."*

- Establishing the CDO has enabled agencies to build up the ways in which they can present, visualize, work with, and analyze their own data, which was not possible previously.
- CDO efforts to improve data sharing and collaboration.
- Partnerships with the CEO office to learn what data sources are available and how they can be distributed.

VERIFICATION OF DATA

- Making extra effort to verify data and use data to support decisions.

> *"In order to fulfill our responsibility, we must make our case using data – whether the stakeholders agree or not."*

2.3 Effectiveness

Effectiveness:

Staff satisfaction with access to evidence that is:

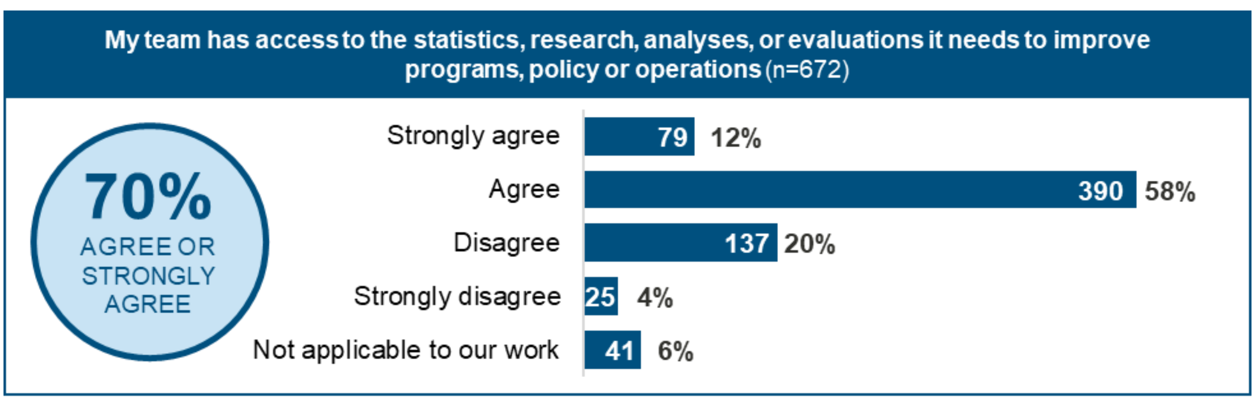
- Necessary to meet their needs
- Relevant
- Useful
- Effectively disseminated

This evidence capacity assessment defines effectiveness as the level of satisfaction that staff members have with their access to evidence that is necessary to address their needs; relevant to outcomes and issues that matter to their office; useful for making program, policy, or operational decisions; and effectively disseminated to internal and external stakeholders.

This section first presents a discussion of survey findings across the Department, calling out any notable variations by agency type and staff level. It concludes with a summary of findings from the focus groups on effectiveness. Appendix E presents further insights from the review of selected evidence on the degree to which those activities are accessible, relevant, and easy to use.

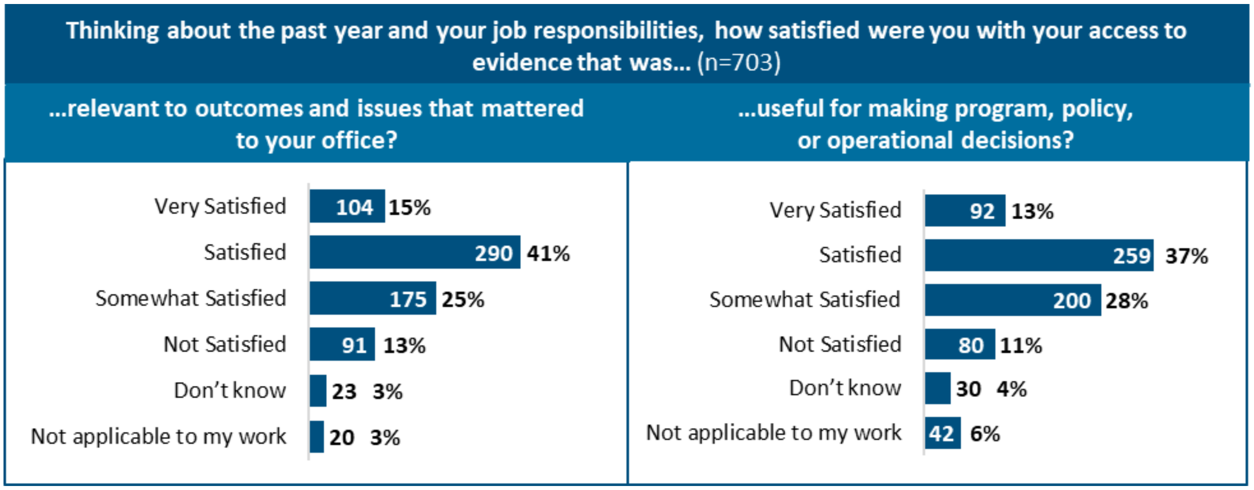
- **Although a majority of respondents have access to evidence that meets their needs, a notable proportion lacked such access.** As shown in Exhibit 13, almost three fourths of respondents agree or strongly agree with the statement, “My team has access to the statistics, research, analyses, or evaluations it needs to improve programs, policy, or operations.” At the other end, nearly a quarter (24 percent) of respondents disagree or strongly disagree with this statement. A small number (6 percent) reported that the statement was not applicable to their work. We observed very similar patterns when looking at responses to this question by agency type and staff level.

Exhibit 13: Access to Evidence to Improve Programs, Policy, or Operations



- **When we explore the adequacy of available evidence in more granular ways, reported satisfaction levels were lower.** We asked staff members about satisfaction with their access to evidence over the past year that was *relevant* to outcomes and issues that mattered to their office. We also asked about satisfaction with access to *useful* evidence for making program, policy, or operational decisions (Exhibit 14). In both instances, we offered the option of “somewhat satisfied” in addition to “satisfied” and “very satisfied.” In these instances, we see that around half of the staff members were satisfied or very satisfied with their access to evidence.

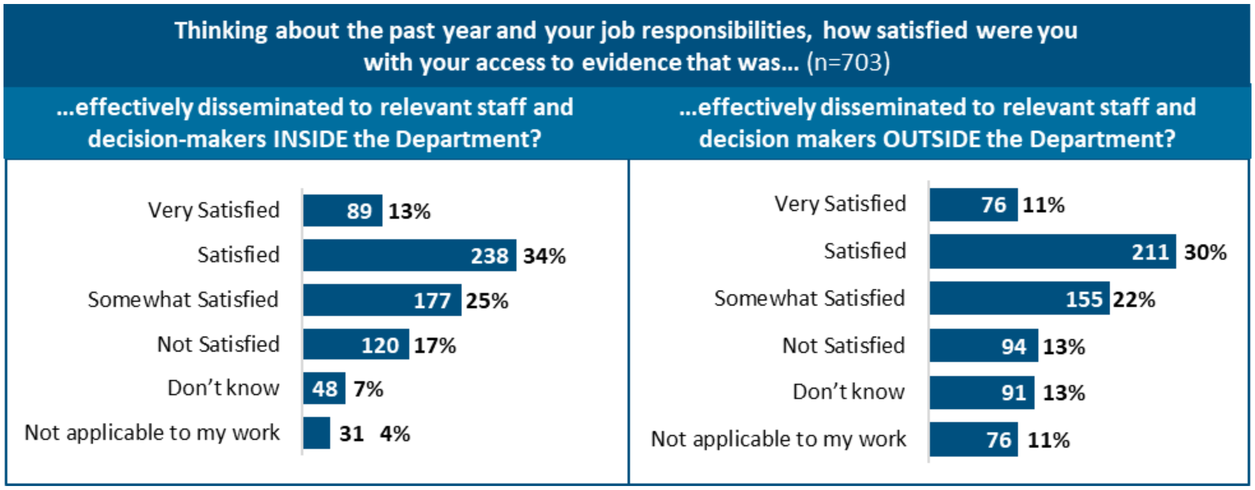
Exhibit 14: Satisfaction with Access to Evidence that is Relevant and Useful



- **SES respondents reported higher levels of satisfaction with access to relevant evidence and useful evidence relative to respondents at lower GS-levels**, which may affect SES staff responsiveness to evidence deficits at their agency. When reviewing responses by staff level, we see responses at the GS-13, -14, and -15 levels aligned closely with the department-wide findings, with SES respondents standing out as more satisfied with their access to relevant evidence on outcomes and issues that matter to their office (70 percent of SES respondents were satisfied or very satisfied compared with 54 to 60 percent of lower staff levels) and evidence useful for making policy decisions (60 percent of SES respondents were satisfied or very satisfied compared with 48 to 53 percent of lower staff levels).[12]
- **Respondents reported similar levels of satisfaction with disseminating evidence within and outside the Department.** The survey asked respondents to reflect on their satisfaction with evidence that was effectively disseminated to relevant staff members and decision makers inside and outside the Department. As shown in Exhibit 15, respondents reported similar levels of satisfaction with the dissemination of evidence within (47 percent satisfied or very satisfied) and outside (41 percent satisfied or very satisfied) the Department, but more respondents reported that they are not aware of dissemination *outside* of the Department or it is not applicable to their work. As with the other survey findings for this domain, we observed patterns similar to the department-wide patterns observed in responses to these questions by agency type and staff level. Respondents at the SES level reported somewhat higher levels of satisfaction with access to evidence that is effectively disseminated within the Department, with 60 percent of SES respondents reporting that they were satisfied or very satisfied compared with 45 to 52 percent of GS-13, -14, and -15–level staff.

[12] Thirty percent of SES-level respondents reported that they were very satisfied with their access to relevant evidence, with very similar proportions reporting that they were satisfied (40 percent) or somewhat satisfied (25 percent). No SES respondents reported that they were not satisfied. Similarly, for useful evidence, SES respondents reported higher satisfaction levels (30 percent of SES respondents reported that they were very satisfied with their access and another 30 percent reported that they were satisfied).

Exhibit 15: Satisfaction With Access to Evidence Effectively Disseminated Within/Outside the Dept.



Focus Group Insights on Effectiveness

Focus Groups in 15 DOL agencies (n= 125). Not a random sample.

While focus group participants noted many successful examples of how they use and rely on evidence to inform their work (see Section 2.2 Uses), they also offered the following insights on the challenges related to accessing, using, and producing evidence that is relevant and useful to their specific programs, policies, and operations.

CHALLENGES

- **Accessing evidence and data can be challenging and time-consuming.** Participants described bureaucratic barriers that make it challenging to request and receive access to evidence, including time-consuming processes for requesting journal articles and establishing data sharing agreements across agencies.
- **Staff members have challenges linking datasets.** Participants noted they often need data from across disparate datasets that are time-consuming and burdensome to link for analysis. Staff shared it is challenging to link data that address different outcomes of interest and to link datasets produced or hosted by different federal agencies or, at times, by different districts, regions, or field offices from within an agency.
- **Staff members have limited evidence relevant to agency practices and strategies.** Participants noted challenges with access to evidence that is relevant and specific to their programs, practices, or strategies. They identified the following types of hurdles:
  - Limited ability to conduct studies that estimate or demonstrate the impact or effectiveness of their work due to challenges isolating their effects and the time required to conduct rigorous studies.
  - Underinvestment in evaluation due to budget constraints and competing priorities.
  - Reporting requirements for programs and other standard data collection/tracking efforts that often focus on performance metrics rather than data specific to the intended outcomes of interventions.
  - The reactive nature of how their agencies use evidence limits their ability to plan for producing relevant and usable evidence specific to their agency's needs.

- **Data/evidence on subpopulations is lacking.** Participants described difficulties accessing data/evidence specific to populations or subpopulations that they serve due to a lack of data collected and limited studies of these groups. This lack of specificity makes it difficult to inform decisions and direction in policies and programs.

> *"Looking at subpopulations is not always a core value, so that data collection is either insufficient, or it is not publicly available. We measure what we value, and we value what we measure."*

- **Staff members are limited in their ability to use existing evidence/data effectively.** Participants face several challenges using existing evidence effectively, including:
  - limited access to raw data generated outside the Department,
  - challenges combining and analyzing data due to errors in key variables,
  - limited access to and training in using analytical software such as Tableau,
  - inadequate databases that do not support analyses and/or do not allow for new data/variables, and
  - difficulty mining large amounts of qualitative data across different sources.

> *"We have a lot of evidence at our disposal but lack the tools and technology to make anything meaningful from it."*

- **Evidence/findings are not always presented in user-friendly formats appropriate for a diverse audience.** Participants explained that there are varying levels of technical skills for interpreting evidence among staff members within the Department, including people in regional and field offices implementing programs and policies. The evidence available is often not easy to interpret and apply in their day-to-day work.

SUGGESTED SOLUTIONS

In discussing challenges related to effectiveness, focus group participants described the following steps to improve the effectiveness of the evidence they access, use, and produce.

Evidence and Data Access

- Develop and improve processes for requesting evidence from other agencies, especially within DOL.
- Facilitate easy access to journal articles.
- Establish/improve accessible data management tools/platforms that facilitate data consolidation.

Evidence and Data Production

- Require data collection on specific subpopulations and provide guidance and support for doing so.
- Increase proactive planning for evidence production, including developing more targeted research questions; and improve upfront design of reporting/ data collection requirements that support evaluation.

> *"There is a desire to want evidence right now, and sometimes it can take a year to generate good information. People lose interest. We need more commitment to putting the time in to get good evidence, and we need longer time horizons on projects."*

- Develop reporting requirements and data collection/ tracking efforts that tie back to the intended outcomes of the program/activity.
- Improve data tracking tools to enable addition of new variables and minimize data entry errors.

Tools and Resources to Improve Evidence Use

- Improve access to and training in using analytical software tools.
- Develop evidence-related products that communicate findings for a range of stakeholders.

As noted earlier (Section 2.2 Uses), staff members reported recent progress in data management and access department-wide. They also described efforts to establish data quality standards within the Department and across federal agencies. Across the focus groups, however, staff members noted that sustained leadership support is needed to make further strides toward these effectiveness goals.

2.4 Balance

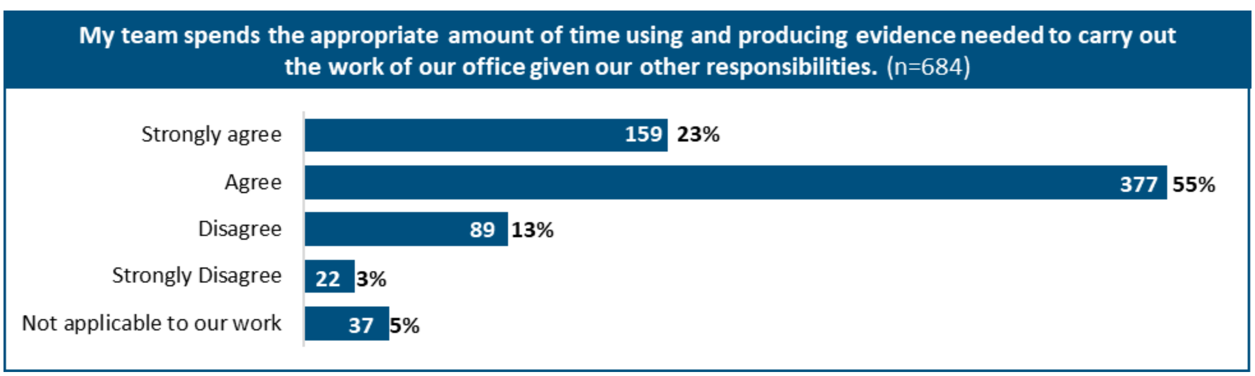
Balance:

- Staff assessment of time spent using and producing evidence

This evidence capacity assessment defines balance as staff spending an appropriate amount of time using and producing evidence needed to carry out the work of their office, given their other responsibilities. This section first presents survey findings on respondents' assessments of whether their teams spend an appropriate amount of time using and producing evidence, and it then examines patterns by agency size, function, and staff level. The section concludes with a summary of focus group findings on Balance.

Exhibit 16 presents the degree to which respondents agreed or disagreed with the statement, "My team spends the appropriate amount of time using and producing evidence needed to carry out the work of our office given our other responsibilities."

Exhibit 16: Agreement With Balance of Time Using and Producing Evidence



- **The majority of respondents believe that they spend an appropriate amount of time using and producing evidence.** Department-wide findings show that over three quarters (78 percent) of respondents either agreed or strongly agreed that their team spends the appropriate amount of time using and producing evidence within the context of their other responsibilities. Sixteen percent of department-wide respondents disagreed or strongly disagreed with the statement, demonstrating that opportunities exist for improving the balance of time spent using and producing evidence for some staff members. Five percent of respondents reported that this question was not applicable to their work, which suggests that some staff members do not see the use and/or production of evidence as part of their job responsibilities.

- **Respondents from support agencies (33 percent) were more strongly in agreement with this statement than respondents from employment and training (22 percent) or compliance/ enforcement agencies (also 22 percent).** Additionally, survey results show that a slightly larger percentage (21 percent) of respondents from employment and training agencies either disagreed or strongly disagreed with this statement, as opposed to the percentage of respondents from the support agencies (15 percent) or compliance/ enforcement agencies (14 percent) who responded they disagreed or strongly disagreed with this statement. It is unclear if respondents who disagreed or strongly disagreed with this statement think the amount of time their team spends is too little or too much.
- **Leadership and staff members have different perceptions about balance.** Examining survey responses across staff levels for this question, we see that 78 percent of respondents from each of the levels GS-13, -14, and -15 either agreed or strongly agreed with this statement, which aligned exactly with department-wide findings. Meanwhile, respondents at the SES level had a much higher rate of agreement, with 94 percent either agreeing or strongly agreeing with this statement and only 6 percent disagreeing (responding disagree or strongly disagree) as opposed to the 15 to 21 percent range found with the GS-13, -14 and -15 staff levels. The difference in responses could be due to the different roles and responsibilities of staff members at these levels and/or to a lack of leadership understanding of time spent by staff members on these activities. This divergence could lead to a less-than-ideal distribution of agency resources.

Focus Group Insights on Balance

While focus group participants reported a number of strengthening factors/investments that have helped improve how their agencies use and produce evidence (see Section 2.2 Uses, particularly the discussion on strengths and recent investments in staff capacity at some agencies), the vast majority of focus group participants reported they are still limited in their ability to spend the time needed to use and produce quality evidence. Participants shared insights on challenges related to balance and identified potential solutions to improve the balance of time spent using and producing evidence.

Focus Groups in 15 DOL agencies (n= 125).
Not a random sample.

CHALLENGES

"I think one limitation I personally have is not having enough time to gather all the evidence that could be out there that could be used to inform our work. (...) unless evidence is already available and sitting there waiting for me to use it."

- **Time constraints limit agencies' capacity to use and produce evidence.** Focus group participants explained that the task of using and producing evidence is labor-intensive and requires a high level of concerted effort. Most focus group participants, when reflecting on their ability to use and produce evidence, agreed that "time is tight" and they are often "too overworked" to prioritize the use and production of evidence. In recognizing the importance of using and producing evidence, some participants shared that it would be helpful if, in addition to the analyses done at national and regional offices, local/district office staff members also had more time to conduct analyses as well as review and reflect on data to inform their work.

- **Staffing-related challenges limit the time spent using and producing evidence.**
  - Participants cited staff turnover (and shrinking staff sizes) as placing significant limitations on their agencies' capacity to use and produce evidence. Some agencies described their staff capacities as "bare bones" and having a "skeleton crew" even as requests for evidence were increasing.
  - Participants also shared more staff is needed, such as analysts and interns, to support using and producing evidence, especially the labor-intensive aspects of preparing data for analysis.

"Well, some of it's extremely labor intensive, right? I mean, you're talking about pulling individual case files, and I don't know if the benefits would outweigh the costs – it would depend on a specific project – if you could get interns – when you talk about staffing you don't have people to do the early labor-intensive stuff."

- **Data quality and access issues pose barriers to achieving balance.** Participants cited time-consuming processes for requesting and receiving data and transforming it into a usable form. They indicated that time spent navigating these hurdles comes at the expense of time available to analyze, interpret, and use these data (see Effectiveness Section).
- **Responding to quick turnaround evidence requests makes it harder to achieve balance and quality.** Participants indicated that they often face concurrent short-turnaround evidence requests/projects and there are no standard timeframes for responding to such requests. Participants felt that leadership tends to have a limited understanding of the level of effort associated with data analysis and often has unrealistic expectations of how long a specific request would take to fulfill. Participants also shared that misalignment between the timing of receiving data and reporting cycles results in the agencies not having adequate time to produce solid evidence.

SUGGESTED SOLUTIONS

Participants expressed a need for additional support as well as process and system efficiencies to help alleviate the time pressures they face in using and producing evidence. Participants proposed the following solutions for improving balance:

- Increase staff capacity by onboarding new staff members dedicated to using and producing evidence.
- Improve data quality and access to reserve time for analyses and interpretation.
- Establish buy-in for creating and using evidence use and production schedules to allow for better planning and allocation of limited staff time.
  - Create and share standard timeframes for common types of data requests to align expectations.
  - Coordinate timelines for strategic planning with OMB reporting cycles.

3. QUALITY, INDEPENDENCE, METHODS, AND EQUITY

This chapter describes survey findings for the assessment domains of quality, methods, independence, and equity.

3.1 Quality, Methods, and Independence

The Evidence Act requires an assessment of the quality, independence, and methods of evidence activities. These concepts are interrelated. This evidence capacity assessment defines quality as evidence that is credible and objective. Credible evidence is appropriate in its use of rigorous methodological approaches, and it documents data sources and analysis methods well. Objective evidence is independent, conducted by a third party, and notes strengths and weaknesses of approach.

Quality, Methods, and Independence

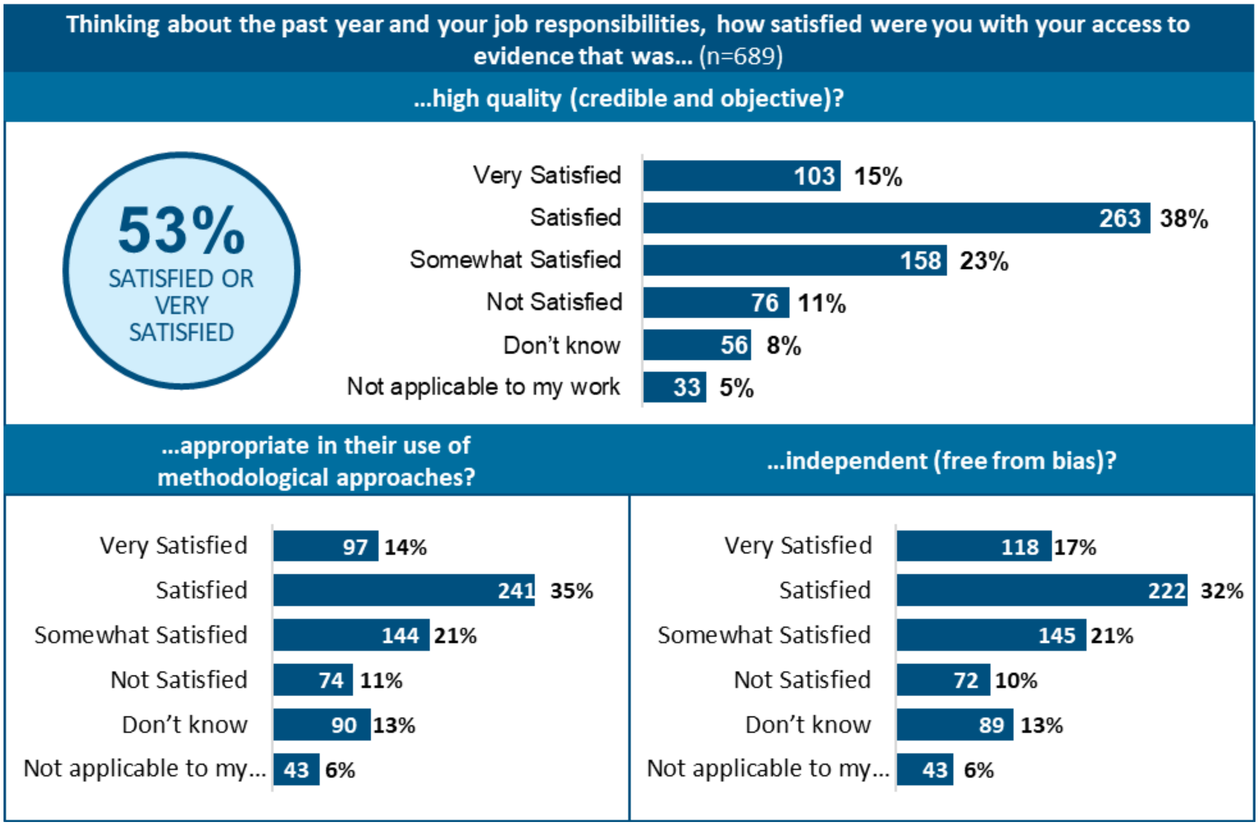
Staff satisfaction with evidence that is:

- Credible and objective
- Appropriate in use of rigorous methods
- Free from bias

This section first presents survey findings on staff-reported levels of satisfaction with evidence that is high quality (credible and objective), appropriate in its use of rigorous methodological approaches, and independent (free from bias). The section concludes with a summary of focus group findings on Quality, Methods, and Independence. The review of selected evidence-related activities presented in Appendix E offers further insights on these topics.

Exhibit 17 presents levels of staff satisfaction with access to evidence over the past year that was high quality (credible and objective), independent (i.e., free from bias), and used appropriate rigorous methodological approaches. Please note that the research team recognizes that we are making an assumption of staff capacity to judge the quality, independence, and rigor of evidence when asking about their satisfaction in this domain. Future iterations of the capacity assessment may dive more deeply into staff knowledge of these features and consider other measures for these domains, including whether mechanisms or protocols are in place that help to ensure the Department uses and produces evidence that is high quality, appropriate in its use of rigorous methodological approaches, and independent.

Exhibit 17: Satisfaction With Access to Evidence that was High Quality, Independent, and Used Appropriate Methods



- **Only half of the respondents reported being satisfied or very satisfied with access to evidence that was high quality, used appropriate methods, and was independent.** Respondents reported similar levels of satisfaction across quality, methods, and independence. The majority of respondents (53 percent) reported that they were satisfied or very satisfied with their access to high-quality evidence over the past year. Just under half of the respondents reported that they were very satisfied or satisfied with their access to evidence that was independent (49 percent) and used appropriate rigorous methodological approaches (49 percent). Approximately a tenth of the respondents reported that they were not satisfied with their access to evidence that was high quality, independent, and used appropriate rigorous methodological approaches.

  As shown in Exhibit 17, respondents could also respond to these questions by noting that they do not know whether they had access to evidence that met these criteria or that access to evidence was not applicable to their work.

- **Eight percent of respondents reported that they do not know whether they had access to high-quality evidence over the past year**, and 13 percent of respondents reported that they do not know whether they had access to evidence that used appropriate rigorous methods or was independent. Between 5 and 6 percent of respondents selected "not applicable to my work" for these three questions. These responses may point to a potential limitation in either the capacity to assess evidence across the domains of quality, methods, and independence (do not know) and/or a culture that facilitates awareness of how evidence is applicable to the day-to-day work of staff ("not applicable to my work").

- **When analyzing responses by agency type, we found similar patterns in respondents' satisfaction with access to evidence that was high quality, used appropriate methods, and independent.** One exception was higher rates among respondents from support agencies, who reported that they were very satisfied (as opposed to satisfied) with their access to evidence across the measures of quality, methods, and independence. Across the different staff levels, the percentage of respondents who were satisfied or very satisfied with their access to evidence that was high quality, used rigorous methods, and was independent was similar to the department-wide findings.
- **SES respondents were more likely to be somewhat satisfied with access to evidence that was high quality (39 percent), used rigorous methods (39 percent), and was independent (33 percent) than other staff levels.** Between 17 and 26 percent of GS-13, -14, and -15–level respondents were somewhat satisfied with their access to evidence that was high quality, used rigorous methods, and was independent. Respondents at the GS-13, -14, and -15 levels were at least twice as likely to report that they did not know about their access to evidence across these domains.

Focus Group Insights on Quality, Methods, and Independence

Focus Groups in 15 DOL agencies (n= 125).
Not a random sample.

While we did not collect detailed information on the quality, methods, and independence of data available to DOL staff during the focus groups, we did gather insights on their access to quality data and factors that affect data use, which encompass characteristics of appropriate methods and independence. (Section 2.3 on Effectiveness also presents some content relevant to this domain.)

STRENGTHS

Participants shared insights on factors that have supported and allowed them to access high quality evidence.

- **Improved data sharing and systems.** Some participants noted improved data sharing between agencies, data systems, and technology has increased access to needed data and ability to analyze these data.
- **Support from the Chief Data Officer.** Participants cited the formalized role of the Department's Chief Data Officer as an important resource to help with data access, including data from other federal agencies.

"We have a data hub that people can click on and see all these visualizations, get answers to questions. And so, I think that's really been a best practice, using these initiatives and also the access within one centralized location."

- **Establishing data quality standards.** Participants also noted that the Department is working to standardize data and data quality standards enterprise-wide, which will provide an existing model for agencies to follow when creating evidence/datasets.
- **Improved staff knowledge and skills.** Participants noted that improving knowledge and skills within their agency/teams has increased the quality and rigor of the research conducted and thus improved access to quality evidence.

CHALLENGES

Participants shared challenges with data accessibility and data quality.

- **Differences in access to and quality of data across agencies.** Some participants shared there are differences in the level of access to and quality of data across agencies. Cited reasons for these differences include challenges with accurate data entry and quality checks, challenges with accessing BLS and Census Data, and agency leadership commitment to developing quality evidence. Participants also pointed to variations in data access and technology across offices within agencies, driven in part by the way programs may be structured.

> *"Some agencies are better at this than others. I think that is a top-down situation for what their leadership think the value the evidence brings."*

- **Challenging user interfaces.** Participants shared the user interface of systems/software that they use to download/extract data can be complicated and challenging. Participants expressed critical needs around user-friendly portals to better access and use data.

> *"The ability to extract the data the way it was extracted before is not there – I have to go to somebody else to help get the data, or make sure it's correct, vs. being able to do it myself."*

- **Need for training on data tools and analysis.** Participants also shared the need for timely training on how to use data tools and conduct data analysis to ensure efficient use and production of evidence. Participants explained that the timing and relevance of training on tools can be an inhibiting factor. While staff members receive training on various tools, they often do not use those tools immediately, leading to learning loss.

SUGGESTED SOLUTIONS

Participants shared some suggested solutions to improve their access to high-quality evidence:

- Investing in user-friendly interfaces/portals
- Easy access to data from other agencies including BLS and U.S. Census Bureau
- Data software, available on demand
- Analysis training, available on demand

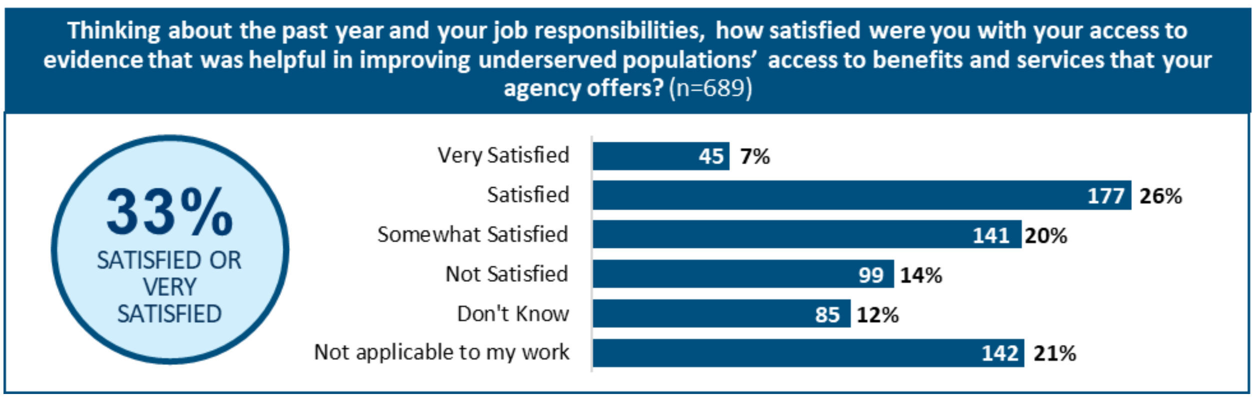
3.2 Equity

Equity:
Staff access to evidence that promotes equity in access to services and benefits

The evidence capacity assessment defines equity as having an appropriate and sufficient degree of evidence activities to address and promote the needs of underserved communities, including access to benefits and services and equitable treatment under policies and programs. Because many other aspects of equity exist, this definition serves only as a starting point for assessing equity's many dimensions in future iterations of the capacity assessment. This section presents survey findings on equity.[13]

Exhibit 18 presents satisfaction levels with access to evidence over the past year that was helpful in improving underserved populations' access to benefits and services offered by the agency.

Exhibit 18: Satisfaction With Access to Evidence That Improves Equity



- **Only one third of respondents are satisfied or very satisfied with their access to evidence that improves underserved populations' access to benefits and services.** Meanwhile, 20 percent reported they were somewhat satisfied, and 14 percent reported they were not satisfied. Examining the responses to this question by agency type, we see that respondents from employment and training and compliance/enforcement agencies reported similar levels of satisfaction as the department-wide findings, with 37 percent and 32 percent of respondents from each agency type respectively reporting they were satisfied or very satisfied. Twenty-two percent of respondents from support agencies reported they were satisfied or very satisfied. Satisfaction with access to evidence that improves underserved populations' access to benefits and services decreased as staff level increased. For example, 35 percent of GS-13s report that they were satisfied or very satisfied compared to 17 percent of SES respondents.
- **Notably, over a fifth of respondents reported that access to evidence that improves underserved populations' access to benefits and services was not applicable to their work.** Reviewing the data by agency type, we see that respondents from support agencies

[13] Focus groups offered limited opportunity to probe issues related to equity. They did involve discussions of staff dissatisfaction with the lack of disaggregation of findings by subpopulations of interest, which has a direct bearing on assessing the equity of investments. (See Section 2.3 on Effectiveness for further discussion of this topic.)

were more likely to report that this aspect of evidence was not applicable to their work (36 percent), followed by those respondents in compliance/enforcement agencies (21 percent), and employment and training agencies (13 percent). The differences in responses to this question by agency type may point to a need to improve how equity is defined and pursued by the full range of DOL agencies.

4. CAPACITY, CONTEXT, AND RESOURCES

4.1 Individual and Team Capacity

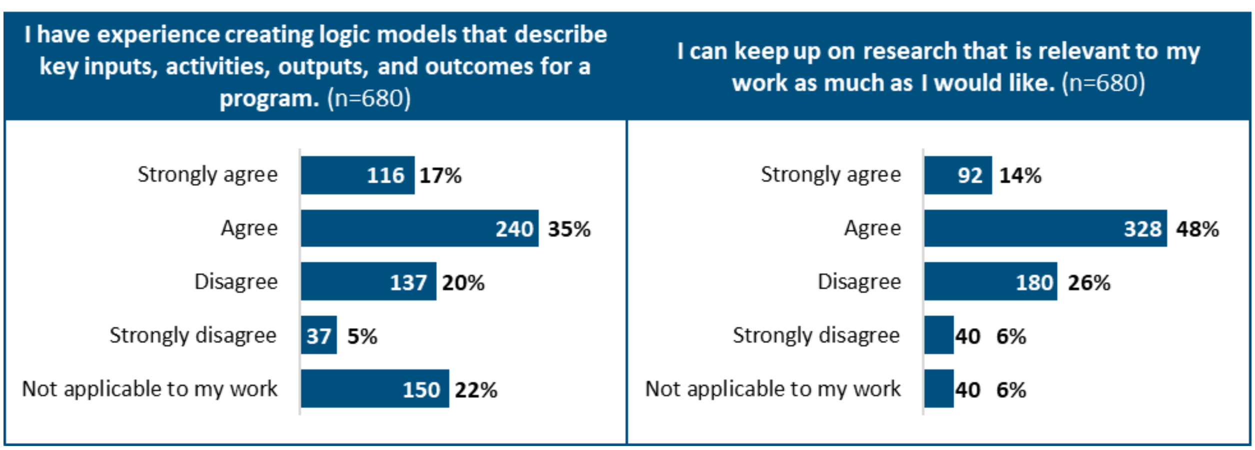
Individual and Team Capacity:

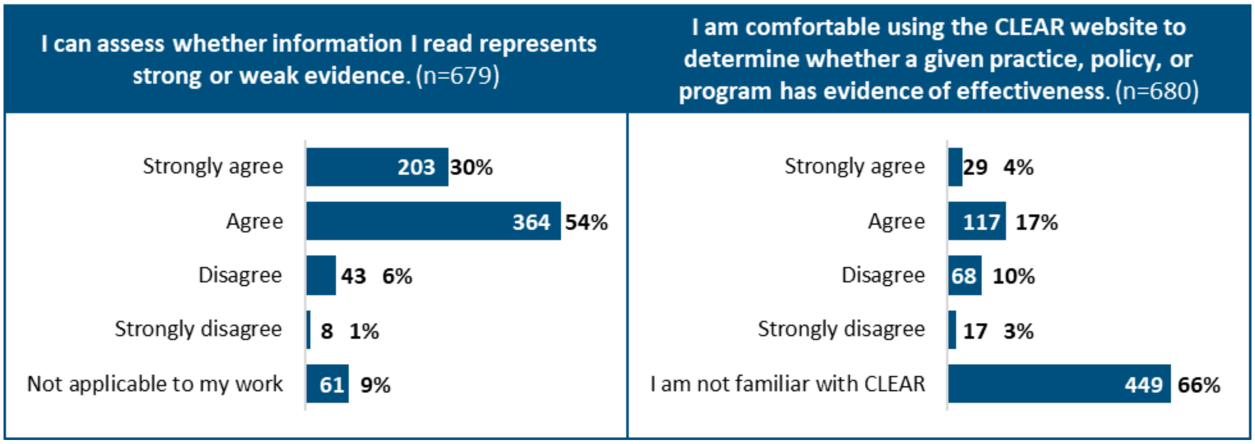
- Staff knowledge, skills, and ability to use and produce evidence
- Staff perceptions of their team's capacity to use and produce evidence

We define capacity as knowledge, skills, and ability to produce and use evidence. To measure individual capacity, we examine individuals' experiences and comfort levels conducting basic, foundational evidence activities. To measure team capacity, we also examine respondents' perceptions of their offices' capacity for using and producing evidence. This section presents a discussion of survey findings on individual staff member experience and familiarity with specific evidence-related tasks as well as staff perceptions of their team's capacity to use and produce evidence. Focus group findings on individual and team capacity are presented at the end of this chapter.

Staff Experience and Comfort Level With Evidence-Related Tasks. Exhibit 19 presents the levels of agreement among survey respondents with statements about their experience creating logic models, keeping up on relevant research, assessing whether information represents strong or weak evidence, and comfort using the DOL Clearinghouse for Labor and Employment Research (CLEAR) website to determine whether a given practice, policy, or program has evidence of effectiveness. We included these questions in the survey as proxies for comfort with foundational evidence activities based on the assumption that without these foundations in place, all other aspects of producing and using evidence will be harder and in recognition that measuring evidence effectiveness in alternate ways will impose too much time burden on respondents. The research team recognizes other sources of evidence in addition to the CLEAR website, but comfort level using the CLEAR website was specifically included as a proxy measure for foundational evidence activities because it is a cross-cutting resource available to all agencies in the Department.

Exhibit 19: Staff Experience and Comfort Level With Evidence-Related Tasks



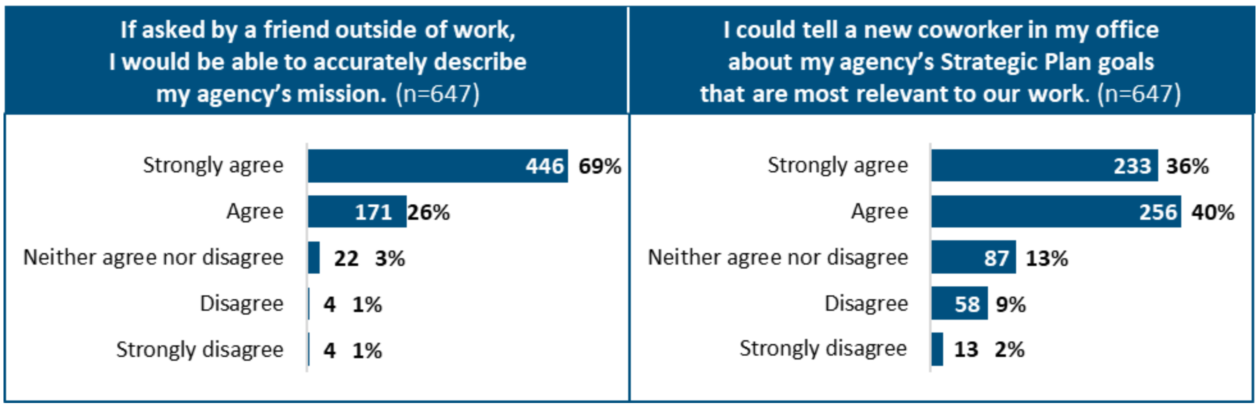


- **The majority of respondents have experience with some evidence-related tasks.** The majority of respondents agreed with statements about having experience creating logic models, keeping up on relevant research, and assessing whether information represents strong or weak evidence. For example, 52 percent of respondents reported that they agree or strongly agree that they have experience creating logic models that describe key inputs, activities, outputs, and outcomes for a program.
- **There are some variations by agency type and staff level**. Sixty-one percent of respondents from employment and training agencies and 63 percent of respondents from support agencies agreed or strongly agreed that they have experience creating logic models. Forty-six percent of respondents from compliance/enforcement agencies agreed or strongly agreed that they have experience creating logic models. Agreement with the logic model statement increased with staff level, with 65 percent of SES respondents agreeing or strongly agreeing with the statement, 62 percent of GS-15, 57 percent of GS-14, and 45 percent of GS-13 respondents agreeing or strongly agreeing with the statement. The number of respondents reporting that creating logic models was not applicable to their work was higher at lower staff levels, with six percent of respondents at the SES level and 28 percent of GS-13 level respondents reporting it is not applicable.
- **Sixty-two percent of respondents agreed or strongly agreed that they can keep up on research that is relevant to their work as much as they would like.** Eighty-four percent of respondents agreed or strongly agreed that they can assess whether information represents strong or weak evidence. We observe similar patterns when looking at responses to these two questions by agency type and staff level.
- In contrast, **two thirds of respondents are unfamiliar with CLEAR.** Sixty-six percent of respondents reported that they are unfamiliar with CLEAR, and 13 percent reported that they are not comfortable using the CLEAR website.
- **Familiarity with CLEAR varies by agency type.** Respondents from employment and training agencies were more likely to report higher comfort using CLEAR, with 34 percent reporting that they agreed or strongly agreed that they are comfortable using the tool, and 52 percent reporting that they were not familiar with CLEAR. By comparison, respondents from compliance and enforcement agencies and support agencies were more likely to report that

they are not familiar with CLEAR, with 73 percent and 62 percent reporting they are not familiar, respectively. Examining responses to this question by staff level points to some interesting differences that show no clear pattern. GS-13s were most likely to report lack of familiarity with CLEAR (75 percent), followed by GS-15s (62 percent). Among both SES and GS-14s, 56 percent of respondents reported lack of familiarity with CLEAR.

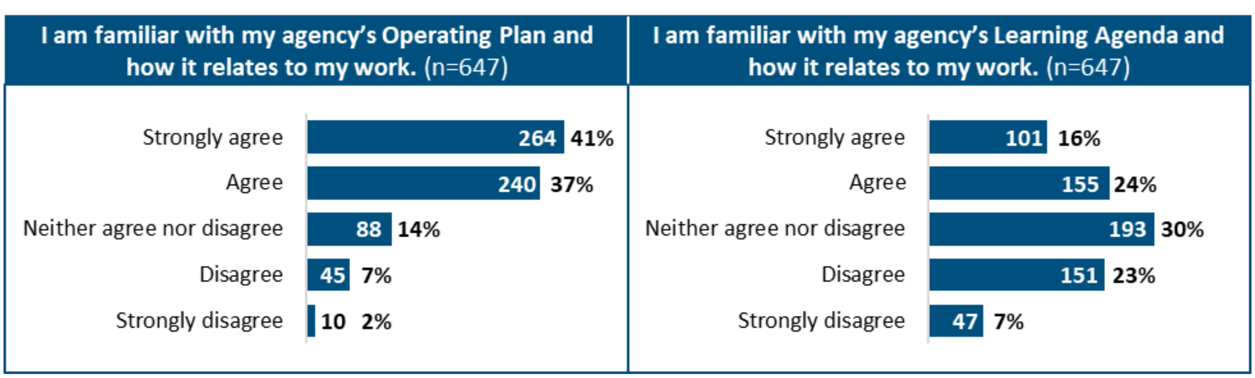
Respondent Familiarity With Agency Drivers of Evidence Use and Production. Exhibits 20 and 21 present the respondents' familiarity with their agency's mission, strategic plan goals, operating plan, and learning agenda, which are key drivers of how an agency uses and produces evidence.

Exhibit 20: Familiarity With Agency Mission and Strategic Goals



- **Respondents generally have a strong understanding of their agency's mission and strategic goals.** As shown in Exhibit 20, almost all respondents report that they could describe their agency's mission (95 percent) and strategic goals (76 percent). We observed similar patterns when looking at responses to these two questions by agency type, but responses to both did show some variation by staff level. Responses from staff at the GS-13 through GS-15 levels were similar to the department-wide findings, but SES respondents were much more likely to strongly agree that they are familiar with their agency's mission (100 percent) and strategic goals (85 percent). No SES respondents disagreed with statements about familiarity with their agency's mission or strategic goals. This finding is not surprising because agency leaders are responsible for ensuring that the agency meets its mission and achieves its strategic goals.

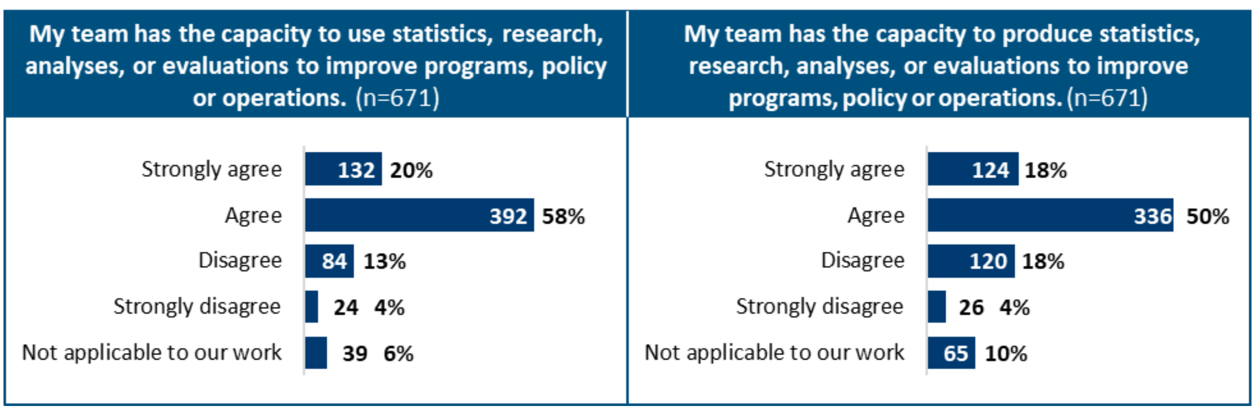
Exhibit 21: Familiarity With Agency Operating Plan and Learning Agenda



- **Respondents are familiar with their agency's operating plan.** Exhibit 21 presents respondent levels of agreement with statements about their familiarity with their agency's operating plan. More than three quarters (78 percent) of respondents strongly agreed or agreed that they are familiar with their agency's operating plan. We observed similar patterns when looking at responses to this question by agency type, but responses by staff level were slightly different. The proportion of respondents that agreed or strongly agreed that they are familiar with their agency's operating plan increases with staff level. Seventy-five percent of GS-13 respondents and 100 percent of SES level respondents reported familiarity with their agency's operating plan.
- **Overall, respondents appear much less familiar with their agency's learning agenda.** Exhibit 21 shows 40 percent reporting that they agreed or strongly agreed that they are familiar with their agency's learning agenda. At the agency level, respondents from employment and training and support agencies were more likely to agree that they are familiar with the learning agenda, with 48 percent and 44 percent, respectively, reporting that they agree or strongly agree. Thirty-five percent of compliance/enforcement agencies agreed or strongly agreed that they are familiar with their agency's learning agenda. As with respondent familiarity with operating plans, the proportion of respondents that agreed or strongly agreed that they are familiar with their agency's learning agenda increased with staff levels. The percentage of respondents agreeing or strongly agreeing that they are familiar with their agency's learning agenda was 32 percent for GS-13, 41 percent for GS-14s, 53 percent for GS-15s, and 85 percent for SES respondents.

Team Capacity to Use and Produce Evidence. Exhibit 22 presents respondent perceptions of their team's capacity to use and produce statistics, research, analyses, and evaluations to improve programs, policy, or operations.

Exhibit 22: Team Capacity to Use and Produce Evidence



- **Most respondents are confident of their team's capacity to use and produce evidence.** The majority of survey respondents reported that their teams have the capacity to use and produce evidence to improve programs, policy, or operations, with respondents appearing slightly more confident in their team's capacity *to use* evidence compared with their team's capacity *to produce* evidence. Seventy-eight percent of respondents agreed or strongly agreed that their team has the capacity to use evidence. At 75 percent, employment and training agencies were the least likely to agree or strongly agree that their team has the capacity to use evidence. Seventy-eight percent of respondents from compliance/enforcement agencies and 84 percent of respondents from support agencies agreed or strongly agreed that their team has the capacity to use evidence. Responses by staff level generally aligned with the department-wide findings.

- **More than two thirds of respondents (68 percent) agreed or strongly agreed that their team has the capacity to *produce* evidence.** Responses by agency type were similar to those observed for team capacity to use evidence, with 63 percent of respondents from employment and training agencies, 70 percent of respondents from compliance/enforcement agencies, and 79 percent of respondents from support agencies agreeing or strongly agreeing that their team has capacity to produce evidence. The proportion of respondents agreeing with this statement increases by staff level. Sixty-six percent of GS-13 staff agreed or strongly agreed that their team has capacity to produce evidence, and 81 percent of SES staff agreed or strongly agreed with the statement.

4.2 Office and Agency Context

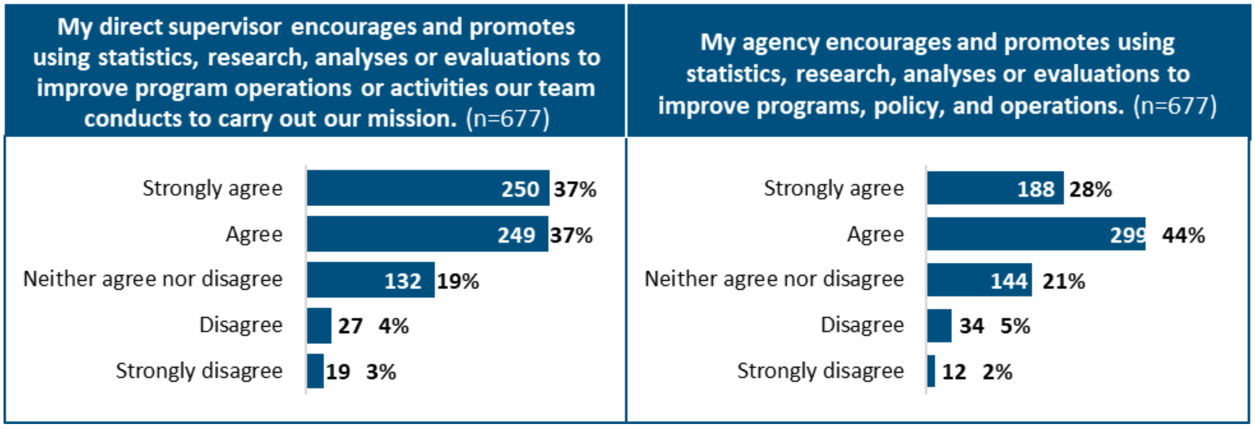
Office and agency context includes office culture and access to adequate resources, which can affect capacity to use and produce evidence. This section presents survey findings on staff perceptions of whether their supervisor and agency encourage using statistics, research, analyses, or evaluations. We then present staff perceptions of whether political support would help to improve their team's ability to use and produce evidence. The section concludes with focus group findings on office and agency capacity and context.

Office and Agency Context:

- Staff perceptions of supervisor and agency support for using and producing evidence
- Staff perceptions of the need for political support to use and produce evidence

Exhibit 23 presents levels of agreement among survey respondents with statements about their supervisor and agency encouraging and promoting the use of statistics, research, analyses, or evaluations to improve programs, policy, and operations as well as other activities conducted to carry out their mission.

Exhibit 23: Supervisor and Agency Encouragement to Use Evidence



- **Majority of respondents feel supported by their agency in using evidence.** Nearly three quarters of survey respondents agreed or strongly agreed that their direct supervisor (74 percent) and their agency (72 percent) encourage and promote using evidence. Around a fifth of respondents held neutral positions on whether their supervisor (19 percent) and agency (21 percent) encourage using evidence. A minority (7 percent) of respondents disagreed or strongly disagreed with statements about their supervisor and agency encouraging and promoting the use of evidence. We observed similar patterns when looking at these findings by agency type, with the exception of respondents from support agencies, who were more likely to strongly agree that their direct supervisors encourage and promote using evidence (47 percent).

- **Respondents at different staff levels responded differently about whether their supervisor and agency encourage using evidence, with levels of agreement increasing by staff level.** At the lower end of the spectrum, 66 percent of GS-13–level respondents agreed or strongly agreed that their supervisor supports evidence use, and 69 percent agreed or strongly agreed that their agency supports evidence use. At the higher end, 88 percent of SES respondents agreed or strongly agreed that their supervisor supports evidence use, and 81 percent agreed or strongly agreed that their agency supports evidence use.

- **A small contingent of respondents signaled the need for political support in using and producing evidence.** The survey also asked respondents to reflect on whether political support would improve their team's ability to use and produce evidence. Eighteen percent of respondents reported that political support would improve their team's ability to *use* evidence, and 16 percent felt that political support would help improve their team's ability to *produce* evidence. Twenty-four percent of respondents from employment and training agencies and support agencies felt that political support would help improve their team's ability to *use* evidence, with 14 percent of respondents from compliance/enforcement agencies noting that they need political support for evidence use. We observed a similar pattern with political support for producing evidence. An examination of both questions by staff level points to similar patterns, with higher proportions (between 20 to 25 percent) of GS-14 and GS-15 staff reporting the need for political support in using and producing evidence, compared with GS-13 and SES respondents (between 6 and 13 percent).

Focus Group Insights on Office and Agency Context and Capacity

Focus group participants offered the following insights on office and agency context, in particular the role of leadership in shaping office culture around using and producing evidence and making investments in staffing and other resources that affect both individual and office capacity to use and produce evidence. The Effectiveness (Section 2.3) and Supports and Training (Section 5) domain sections present related discussions of resources needed to support evidence use and production.

Focus Groups in 15 DOL agencies (n= 125).
Not a random sample.

STRENGTHS AND CHALLENGES

Participants value support from leadership and appreciate when leadership understands the work and time required to use and produce evidence.

- **Leadership sets the tone for evidence use and production.** Participants indicated that support from leadership is important for creating an environment that encourages using and producing evidence and is a necessary ingredient for investments in evidence infrastructure, processes, and staffing. Staff members appreciate when leadership uses evidence to guide decisions, asks for evidence when setting strategy, and asks staff to regularly set aside time to review evidence and data.

"Leadership and support has been positive. (agency name) is talking more about making programs, policies evidence-based. Hearing that from leadership is encouraging and makes it easier to shift resources to establishing that evidence base."

- **Leadership does not always have realistic expectations for evidence-related work.** Participants described how unrealistic expectations from leadership around evidence-related work can pose hurdles to their work. These may be due to a limited understanding of the time needed and tasks involved in using evidence or producing it.

- **Political and programmatic pressures can influence evidence-related work.** Several participants did note experiencing pressure to marshal evidence to support political decisions, which dilutes the importance of evidence-based decision making. Other participants noted that leadership may be disincentivized to conduct rigorous studies of programs or change outcome measures due to the risk of negative findings.

"I don't think leadership necessarily has an understanding of data quality, and it may not be clear to them why things have to be done a certain way, and why that is important. A lot of times poorer quality analysis that can be done quickly will be prioritized over a quality "analysis."

Agency investments in staff capacity and expertise as well as improvements in infrastructure and tools to use and produce evidence are essential in improving individual and office capacity.

- **Investments in new hires are essential.** Noting that it requires significant time and skills to collect, clean, and analyze data, participants explained that a lack of staff capacity and expertise as well as limited bandwidth pose barriers to their agencies' ability to use and produce evidence. Participants also shared challenges related to hiring and retention, such as the timing of hiring new staff members with the political cycle (rather than agency need), slow hiring processes, competition for technical talent from private sector employers, and turnover. Participants shared that leadership investments and new hires with strong technical skills have improved their agencies' capacity and have helped to address these barriers, but more is still needed.

- **Continued commitment to training is needed.** Training helped staff members learn how to better use evidence in their work. For example, one agency noted that a recent training on how to use nontraditional resources (e.g., outside of peer-reviewed journals) was beneficial and useful. Given that staff members across agencies have different levels of experience using and producing evidence, participants cited the need for additional and ongoing trainings on a range of topics, including how to assess the quality of evidence and interpret data. The importance of just-in-time training (that is, training that is available closer in time to when it needs to be applied) was also stressed.

> *"[In]in the past year or 2 [there have been] increased trainings for the … agency. This has helped me think of how to use evidence in my work."*

- **Investments in institutional capacity over contractors is preferred by some.** Participants from some agencies shared that they are overdependent on contractors who have knowledge of agency operations, technology, and data systems. While this facilitates data-driven decision making in the short-term, there is a preference among some staff members to bring institutional knowledge in-house rather than rely heavily on contractors.

- **Improved data infrastructure and tools foster culture for evidence use.** Participants from some agencies also shared that investments in improved data infrastructure as well as tools for reporting and data analysis have helped to improve their agency's culture around evidence use and production.

Agency organizational structure can support or hinder capacity to use and produce evidence.

- **Guidance from leadership can help mitigate organizational silos.** Participants explained that "silos" within an agency can limit access to evidence. Where staff members with specific skills for using and producing evidence are located within the agency can impact who has ready access to evidence. Some participants noted that more clarity from leadership on staff roles and how different offices/teams should interact would help in understanding how to mitigate these silos. Staff from one agency reported that having a specialized team to support data use and analysis that leadership makes clear is available to support the agency has been an asset in using and producing evidence.

- **Cross training and staff details can help in address silos.** Participants noted that cross training and placing staff members in details (such as DOL's ROAD program[14]) within and across agencies offers ways to combat siloed organizational structures and processes and skill gaps.

[14] The DOL Repository of Opportunities, Assignments & Details (ROAD) program was created in 2014 as a short-term rotation assignment and developmental opportunity to increase employee engagement, career development, and cross-training (within or across DOL agencies).

SUGGESTED SOLUTIONS

In discussing findings about office and agency contexts, participants offered the following suggestions for how office and agency leadership can improve individual and office capacity to use and produce evidence:

- Provide staff members with designated time for reviewing and using evidence.
- Offer ongoing training to staffs and leadership on topics, including identifying appropriate sources of information, conducting data analyses, using data software, and interpreting evidence. (Section 5 presents more details on supports and training needs.)
- Continue making investments in hiring new staff members who have specialized, technical skills to use and produce evidence as well as investments in data infrastructure and tools. Create a better talent pipeline out of advanced and professional degree programs by conducting systematic outreach and relationship cultivation with academic institutions.
- Provide clarity on staff roles and how different offices/teams should interact to help mitigate silos. Consider cross training, including continued use of ROAD details within and across the agencies, to help address silos.

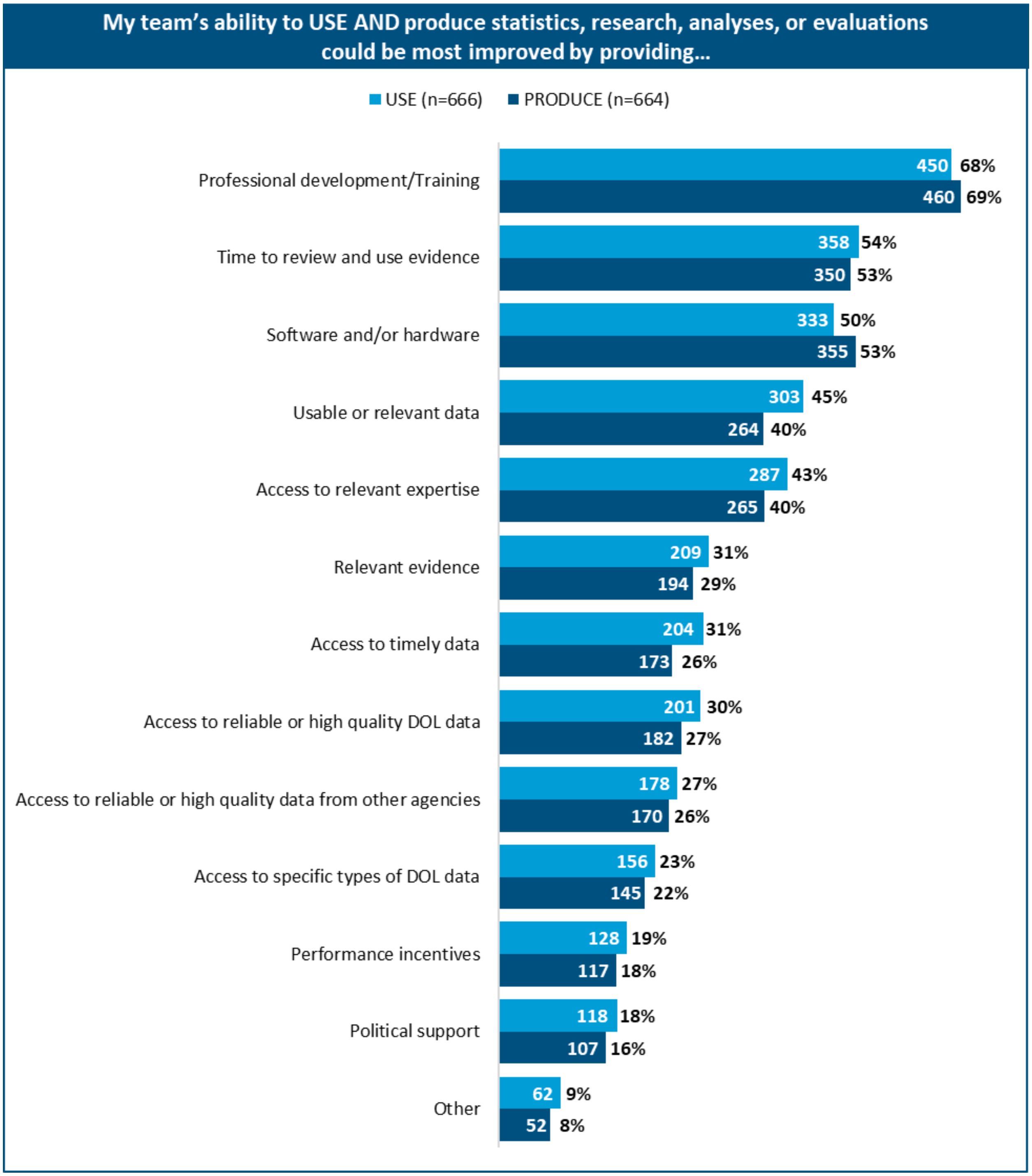
5. OPPORTUNITIES FOR IMPROVEMENT: SUPPORTS AND TRAINING NEEDED

Supports and Training Needed:
Staff needs to improve the capacity to use and produce evidence.

To help DOL improve its capacity and achieve its goals for evidence use and production, the survey asked respondents to identify supports and training that would be most helpful to their teams. This section begins with a discussion of survey findings on supports that could help staff access, use and/or produce evidence, as well as specific training topics in program evaluation and data analytics important to staff members. The section concludes with focus group findings on the supports needed for improving capacity to use and produce evidence. While we recognize that not all domains with identified challenges and needs for support are necessarily within the control of the Department and/or the agencies, we want to share the full suite of insights provided.

- **Respondents reported that professional development and training are the most important support they need to improve use and production of evidence.** We asked respondents *separately* about supports needed to use and produce evidence and allowed respondents to select multiple choices. The five most frequently selected topics were identical across both questions and selected by at least 40 percent of respondents. As Exhibit 24 illustrates, the topics include (1) professional development/training; (2) time to review and use evidence; (3) software and/or hardware; (4) usable or relevant data; and (5) access to relevant expertise. We observed patterns very similar to the department-wide findings when looking at responses by agency type and by staff level. Each staff level selected these same top five topics for using and producing evidence.
- **Some variation by staff level exists on the topics' relative importance**. Most notably, more junior-level staff members requested support on software and/or hardware more frequently than more senior-level staff. For example, to produce evidence, a higher share of GS-13s (54 percent) and GS-14s (58 percent) selected software and/or hardware than GS-15 (44 percent) and SES respondents (38 percent). In addition, GS-15s (52 percent) requested access to relevant expertise to help them use evidence more frequently than GS-13s (40 percent) and GS-14s (44 percent). Respondents at the SES-level were less likely than the other staff levels to report needing more political support or performance incentives to use and produce evidence.

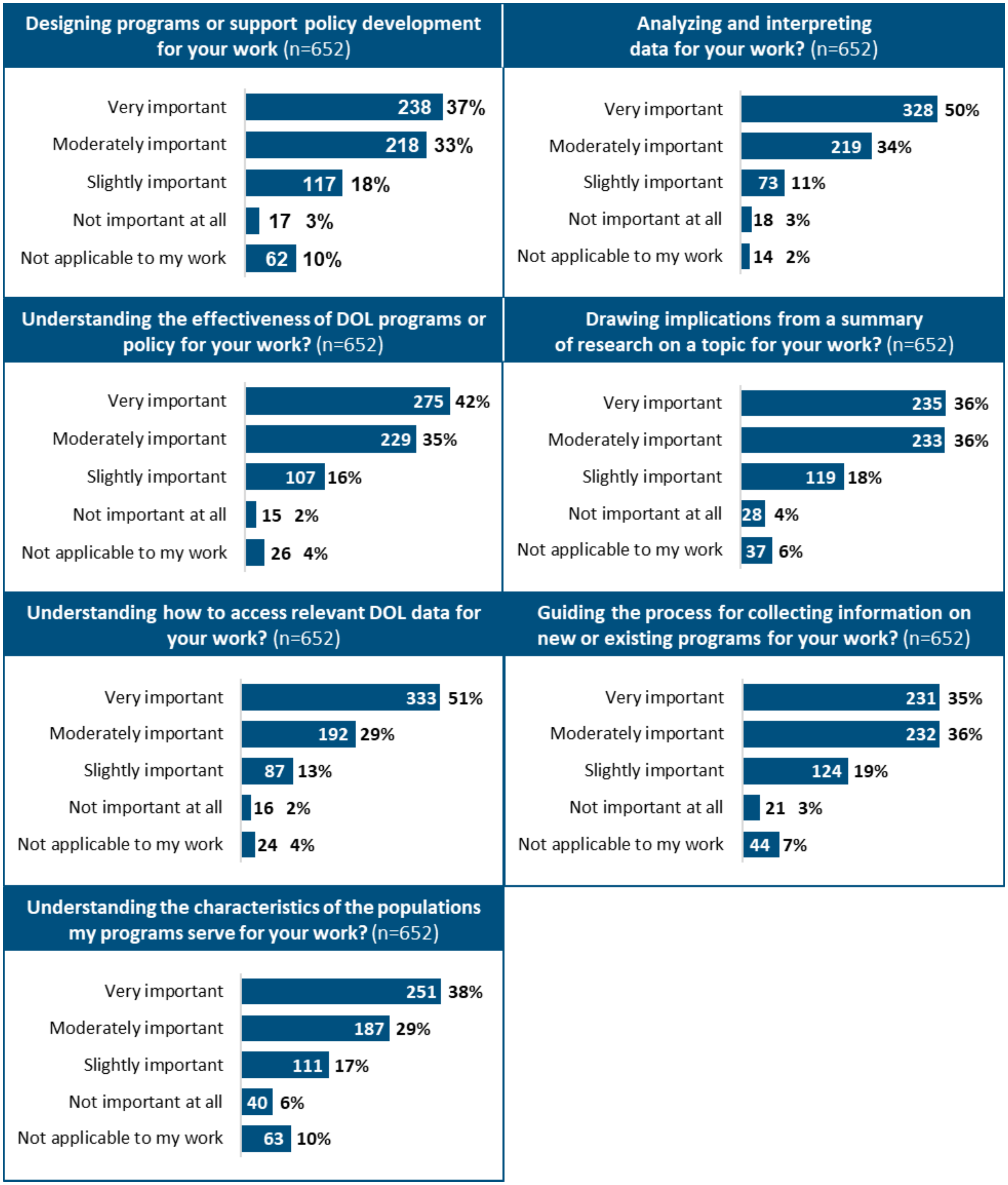
Exhibit 24: Supports Needed to Improve Team USE and PRODUCTION of Evidence



- **Respondents reported that they need training on accessing and using data.** As shown in Exhibit 25, when asked to rate the importance of various training topics, 80 percent or more of respondents rated analyzing and interpreting data and understanding how to access relevant DOL data as either very important or moderately important to do their work.

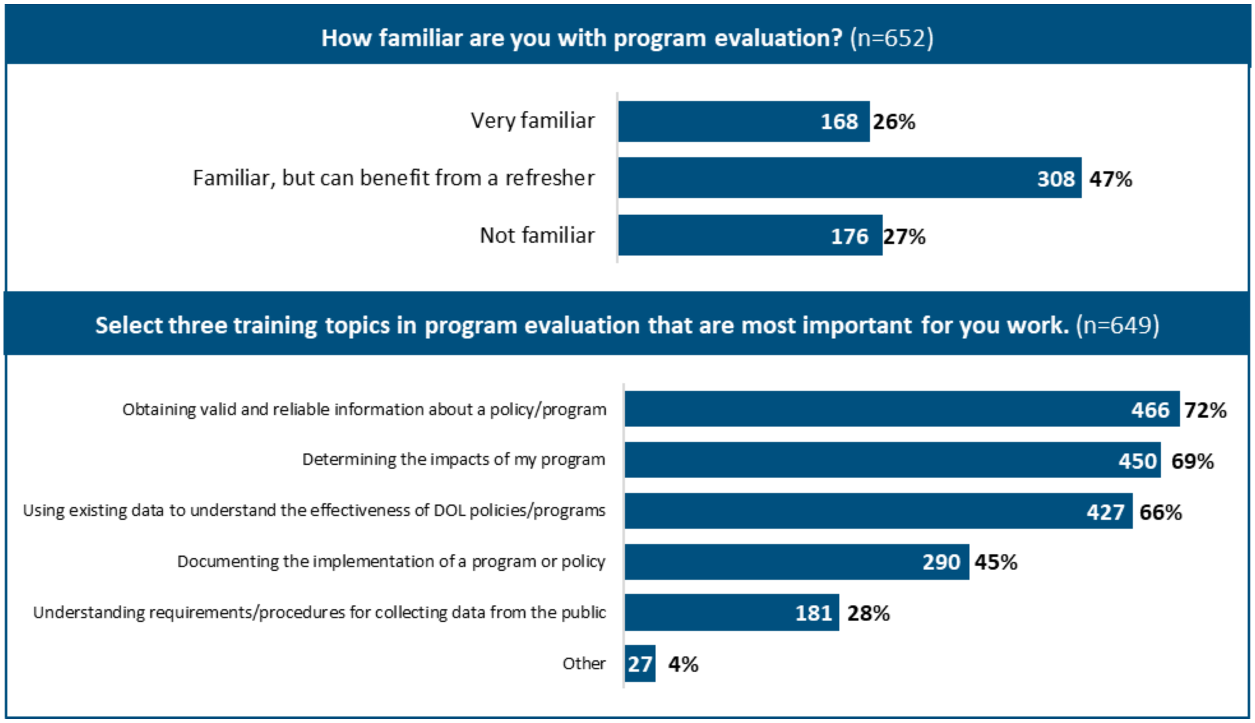
- **Respondents reported high interest in other training topics related to using and producing evidence.** Between 67 and 77 percent of staff reported that the remaining training topics in Exhibit 25 are very important or moderately important to their work.

Exhibit 25: Importance of Training Topics



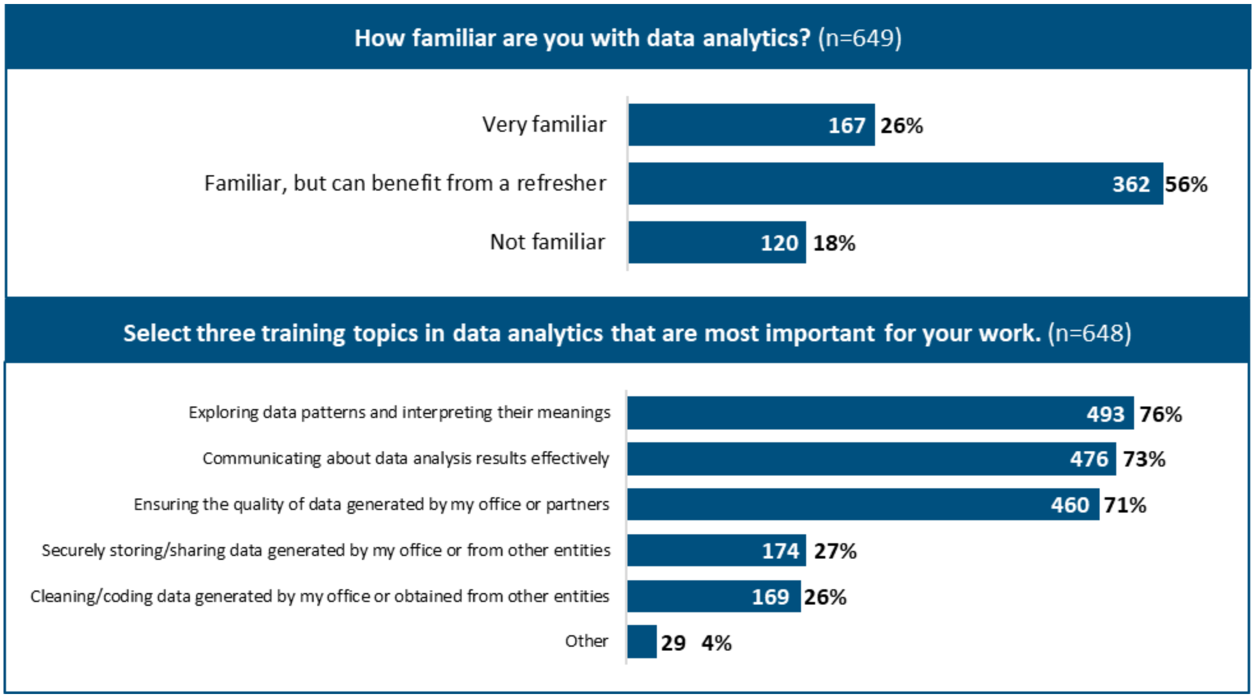
- **Most respondents reported some level of familiarity with program evaluation, but many could use a refresher.** Staff members were asked to rate their level of familiarity with program evaluation. As shown in Exhibit 26, nearly half of respondents (47 percent) reported that they were familiar with the topic but could use a refresher; whereas, the remainder were equally split between being very familiar (26 percent) and not familiar (27 percent) with program evaluation.

Exhibit 26: Program Evaluation Familiarity and Training Topics



- The top three training topics related to program evaluation identified as most important to respondents' work are (1) obtaining valid and reliable information about a policy or program (72 percent); (2) determining impacts of program (69 percent); and (3) using existing data to understand the effectiveness of DOL policies or programs (66 percent). Forty-five percent of respondents selected the topic of documenting the implementation of a program or policy.
- **A similar pattern emerged when respondents were asked to rate their level of familiarity with data analytics (Exhibit 27).** Fifty-six percent reported that they were familiar with this topic but could use a refresher; whereas, 26 percent reported that they were very familiar with data analytics. Eighteen percent of respondents reported that they were not familiar with this topic at all.

Exhibit 27: Data Analytics Familiarity and Training Topics



When asked to select three training topics related to data analytics that are most important to their work, a majority of respondents selected (1) exploring data patterns and interpreting their meaning (76 percent); (2) communicating effectively about data analysis results (73 percent); and (3) ensuring the quality of data generated by my office or partners (71 percent).

Focus Group Insights on Supports and Training Needs

Focus Groups in 15 DOL agencies (n= 125). Not a random sample.

Focus groups generated a rich discussion of the types of supports likely to advance evidence production and use at the agency. Participants shared nuanced insights and specific examples of the types of resources and supports that are likely to be helpful. The most commonly cited additional supports were leadership support; time and staffing resources; training; and technology and data tools to help staff better access and analyze data.

INSIGHTS ON LEADERSHIP SUPPORT

- **More focus on long-term agency goals and vision is valued**. Participants mentioned that a more intentional learning agenda with clearly defined (short-, medium-, and long-term) projects and goals would help them have an increased understanding of their agencies' research priorities and how evidence will inform decisions. That focus would allow them to understand what they need to work toward (i.e., goals that go beyond current administration priorities), as well as what to prioritize in the short-term with the resources they have available.

INSIGHTS ON LEADERSHIP SUPPORT

- **Continued and consistent leadership/political support is important.** Focus group participants noted leadership/political support as one of the most important needs for increasing their agency's capacity to use and produce evidence. Participants expressed that leadership support is necessary for advocating for and securing resources for evidence activities, setting standards for evidence practices at the agency, and demonstrating the importance of using evidence in daily work for agency staff at all levels. This support empowers people to bring evidence into decision-making discussions. (See Section 4 on Agency and Office Context and Capacity for further discussion on the role of political support and politics in evidence use.)

> *"Political support can change from administration to administration–the encouragement of driving our policy decisions on evidence–has been different."*

TIME AND STAFFING RESOURCES

- **Focus on staff bandwidth and specialized skill sets.** While some participants pointed out recent increases in staffing resources, participants across agencies most often noted that additional staff resources (e.g., hiring more staff members, accessing specialized skillsets, better allocation of staff time across tasks) are needed. Reductions in staffing over the years, unfilled administrative and support staff positions, and, in particular, loss of staff members with specialized institutional knowledge and skills have increased responsibilities and workloads of the current staff, leaving less time to focus on using and producing evidence. (The challenges posed by time constraints have been discussed in depth in Section 2.4.)

> *"We completely lack those capabilities [software development]. If we had them, we could evolve and improve over time by using these positive feedback loops where the software engineering staff would be able to respond to the agency and hopefully meet those needs instead of constantly trying to farm out these things in a way that makes them inefficient."*

- **Reexamine staff roles and core functions.** Focus group participants shared that that, to better position future hires in their roles, job descriptions should more accurately represent the types of skills and knowledges that will be used in the positions. To do so, participants shared that staff roles may need to be reviewed/reassessed to accurately document the qualifications needed to perform core functions, including those related to evidence use and production.

- **Dedicate specific time and/or staff specifically for exploring data, including:**
  - further reviewing and analyzing *existing* evidence and data sets;
  - collecting, finding, and downloading *new* or *additional* data;
  - coding data and conducting data entry/field entry/ indexing;
  - linking datasets across sources; and
  - cleaning data and doing quality checks reviewing (quality checking) evidence.

> *"It would be nice…have a couple of statisticians – set them on new datasets and see what we can pull out of there that might be useful – explore new ways to explore the data and present it to policymakers."*

- **Make a designated space for:**
  - practicing new skills,
  - proactively sharing findings that are not just for urgent data requests,
  - scheduling recurring reviews for the usefulness of the current data being collected/used, and
  - supporting external needs/requests.

"We need more capacity to review the measures they have and more willingness to argue the point if something needs to be changed. Be willing to make those changes more often, standardize it."

- **Deepen in-house capacity for data system process improvements.** Have staff (IT staff, software engineers, and web developers) readily available in-house specifically for working with program staff to build or make iterative improvements to data systems in use.

TRAINING NEEDS

Participants agreed that training on program evaluation or data analytics would help them be more successful in their current positions. Most also noted that, if given the time and provided that the topics are relevant to their position, staff members would make it a priority to attend. They indicated that the following features in training are desirable:

- **Actionable and tailored.** Staff members emphasized that training should be actionable and tailored to what leadership feels is required as part of the job and for those skills that are essential to the position.
- **Easy to digest and follow.** Participants expressed the need for trainings with easy-to-follow segments or modules and in-depth step-by-step instructions.
- **Prescheduled and real-time.** The majority of the staff said they preferred the following types of training:
  - **Prescheduled, fixed, type training** as opposed to a self-paced training
  - **Real-time webinars** as opposed to prerecorded webinars
  - **A community of practice type** setting (for sharing rich stories/examples) versus a self-study format
- **Topics of interest.** These included formative evaluation design; qualitative interviewing techniques; survey design; and developing questions for surveys, not just sampling. Additionally, staff members specifically requested training on:
  - evidence-based decision making,
  - how to use and apply evidence daily/in the field,
  - data analytics, including advanced data analytics and using specific data analysis software packages (*R*, Tableau, etc.),
  - how to access and share quality data,
  - computer hardware and software used to house and organize data (navigating software interface),
  - the latest research/data developments, and
  - consensus building and facilitating conversations.

"A training for political appointees for federal appointments to understand data and data requirements."

"If we had some additional team members that would get extra evaluation training on the side, that would be good."

TECHNOLOGY AND DATA TOOL NEEDS

Access to IT Software and Hardware Tools. Participants said improvements to data infrastructure (especially software and hardware tools) are needed to help the staff better organize and sort through existing data. Examples of specific needs in this area of support include:

- **computer hardware and analytical software** appropriate for data analysis,
- **improved user interface** with new and existing software,
- **improved data infrastructure and systems** to support work and allow for more analysis, and
- **data visualization** tools that are accessible for persons with disabilities.

APPENDICES

APPENDIX A: ADDRESSING EVIDENCE ACT AND OMB REQUIREMENTS

The Exhibit below lists Evidence Act and OMB requirements for the capacity assessment, demonstrates how they have been operationalized for the DOL capacity assessment, and notes where relevant findings for each requirement can be found in the final capacity assessment.

Exhibit A-1: Addressing Evidence Act Requirements

Evidence Act PL 115-435, Section 315 C 3	OMB Circular 1	Operationalization in DOL Capacity Assessment	Assessment Domain(s) and Report Chapter
Coverage	What is happening and where is it happening?	Understanding (a) the frequency with which DOL staff use and produce evidence, (b) their satisfaction with having access to evidence that is comprehensive (covers topics they care about), and (c) strengths DOL staff identify. The findings from this domain will complement other DOL evidence-related documents, including Learning Agendas, the Evidence Building Plan, and the Evaluation Plan.	Coverage, Uses; Ch. 2.1 & 2.2
Quality	Are the data used of high quality with respect to utility, objectivity, and integrity?	The approach nests methods and independence as subdomains of quality. Quality is defined as evidence that is credible and objective. Credible evidence is evidence that uses appropriate rigorous methodological approaches and that documents data sources and analysis methods well. Objective evidence is evidence that is independent, either conducted by a third party and/or presented with notes detailing the strengths and weaknesses of the approach.	Quality, Methods, & Independence; Ch. 3.1
Methods	What are the methods being used for these activities? Do these methods incorporate the necessary level of rigor? Are those methods appropriate for the activities to which they are being applied?		Quality, Methods, & Independence; Ch. 3.1
Effectiveness	Are the activities meeting their intended outcomes, including serving the needs of stakeholders and being disseminated?	Examining levels of staff satisfaction with access to evidence that is (a) relevant to outcomes and issues that matter to their office, (b) useful for making program, policy, or operational decisions, and (c) effectively disseminated to internal and external stakeholders.	Effectiveness; Ch. 2.3
Independence	Independence: to what extent are the activities being carried out free from bias and inappropriate influence?	Examined as a subdomain of quality.	Quality, Methods, & Independence; Ch. 3.1
Elements			
(A) a list of the activities and operations of the agency that are currently being evaluated and analyzed;		Compiled by DOL through a separate contractor.	Presented in Appendix G.

Evidence Act PL 115-435, Section 315 C 3	OMB Circular 1	Operationalization in DOL Capacity Assessment	Assessment Domain(s) and Report Chapter
(B) the extent to which the evaluations, research, and analysis efforts and related activities of the agency **support the needs** of various divisions within the agency;		Assessed as a subdomain of effectiveness where staff satisfaction with access to evidence that is useful for making program, policy, or operational decisions is explored.	Effectiveness; Ch. 2.3
(C) the extent to which the evaluation research and analysis efforts and related activities of the agency address an **appropriate balance** between needs related to organizational learning, ongoing program management, performance management, strategic management, interagency and private sector coordination, internal and external oversight, and accountability;		Assessed by staff perception of appropriateness of time spent using and producing evidence needed to carry out the work of their office, given their other responsibilities.	Balance; Ch. 2.4
(D) the extent to which the agency uses methods and combinations of methods that are **appropriate** to agency divisions and the corresponding research questions being addressed, including an appropriate combination of formative and summative evaluation research and analysis approaches;		Covered under discussion of quality.	Quality, Methods, & Independence; Ch. 3.1
(E) the extent to which **evaluation and research capacity is present** within the agency to include personnel and agency processes for planning and implementing evaluation activities, disseminating best practices and findings, and incorporating employee views and feedback; and		Explores individuals' experience and comfort conducting basic evidence activities. Examines individuals' perceptions of team's capacity to access, use, and produce evidence, as well as perceptions of supports and training needed to expand that capacity.	Individual and Team Capacity, Supports, and Training; Ch.4.1 & 5
(F) the extent to which **the agency has the capacity to assist** agency staff and program offices to **develop the capacity** to access and use evaluation research and analysis approaches and data in the day-to-day operations.		Examines office and agency context, including support for evidence access and use among direct supervisor and agency, and perceptions of the degree to which increased political support would improve capacity.	Office and Agency Context; Ch.4.2

DOL's capacity assessment addresses the Evidence Act and OMB requirements listed in Exhibit A-1 for the following types of evidence:

- **Statistics, research, or analyses** such as data analytics, trend analyses, or literature and/or document reviews on specific programs or policy issues.
- **Evaluations** that provide a formal assessment of implementation, impacts, effectiveness, or efficiency of a program, policy, or organization.

APPENDIX B: ASSESSING STATISTICAL CAPACITY AT THE U.S. DEPARTMENT OF LABOR (PREPARED BY BLS)

The Bureau of Labor Statistics (BLS), an agency of the U.S. Department of Labor (DOL), is the principal federal statistical agency responsible for measuring labor market activity, working conditions, price changes, and productivity in the U.S. economy. Its mission is to collect, analyze, and disseminate some of the nation’s most sensitive and important economic data to support public and private decision-making. Founded in 1884, BLS continues to execute its mission with independence from partisan interests while protecting the confidentiality of data providers.

EXISTING CAPACITY

Through the regular production of a wide range of data products, BLS maintains statistical capacity to meet the needs of a diverse set of customers for accurate, objective, relevant, timely, and accessible information and analysis. Data users include the general public, the U.S. Congress, DOL and other federal agencies, state and local governments, business, labor, researchers, educators, students, and more. By continuing to introduce statistical, economic, and technological advancements, BLS is able to improve efficiency and expand its capacity to produce data that are used in support of existing and emerging policies and decisions that affect virtually all Americans, including decisions on fiscal, monetary, and social policy. BLS demographic data, including information on employment, wages, and consumption patterns, are an important input in advancing equity.

BLS is committed to maintaining the highest level of scientific integrity in producing official statistics, while maintaining the confidentiality of respondents and the data they provide. BLS complies with the Statistical Policy Directives and the Standards and Information Quality Guidelines from the Office of Management and Budget (OMB), as well as the National Research Council’s *Principles and Practices for a Federal Statistical Agency*. BLS also conforms to the conceptual framework of the Interagency Council on Statistical Policy’s quality dimensions. BLS data are essential in fostering the President’s priority of providing evidence and supporting evaluation activities, consistent with the Foundations for Evidence-Based Policymaking Act (the Evidence Act). BLS is committed to supporting the President’s Management Agenda of strengthening and empowering the Federal workforce; and delivering excellent, equitable, and secure Federal services and customer experience. Furthermore, by producing gold-standard statistics and analyses, BLS supports the Secretary of Labor’s vision of empowering workers morning, noon, and night.

As a demonstration of capacity, several BLS data series are used in the administration of federal programs. For example, the Internal Revenue Service updates federal income tax provisions based on BLS data, including annual changes to tax brackets based on changes in the Consumer Price Index (CPI). In addition, the Social Security Administration adjusts payments based on changes in the CPI. Select CPIs and Employment Cost Indexes also are used in updates to the Medicare Prospective Payment System, and Consumer Expenditure data are used to adjust cost of living allowances for U.S. military locations. Changes in BLS data have direct effects on overall federal budget expenditures, including federal allocations to state and local jurisdictions. Local Area Unemployment Statistics data are used to allocate federal funds from assistance programs to states and local jurisdictions in such areas as employment, training, public works, and welfare assistance. Businesses use BLS data to make employee wage and benefit decisions, and private citizens make relocation decisions based on unemployment data for states, metro areas, and major cities.

ROLE OF THE STATISTICAL OFFICIAL IN BUILDING CAPACITY

As specified in the Evidence Act, all Chief Financial Officers Act agencies (which includes DOL) must designate a Statistical Official; the Commissioner of Labor Statistics serves as the Statistical Official for DOL. In this role, the Commissioner and BLS staff broaden the impact of the unique knowledge, skills, and practices housed within the statistical function by providing data and analysis to DOL agencies to support their policy initiatives, while at the same time maintaining the confidentiality of statistical data and remaining impartial to policy decisions. Examples of topics about which BLS shares data and analysis with DOL agencies include unemployment, occupational injuries and illnesses, and employee benefits. The Statistical Official role also requires that BLS use its statistical capacity to provide technical assistance and consulting services to assist DOL agencies as they engage in statistical activities, such as data collection and analysis.

An important responsibility of all Statistical Officials is ensuring that all statistics released by the department, including those produced outside of the statistical function, adhere to Federal Information Quality Act standards and statistical best practices. Federal agencies are increasingly called on to release data measuring program performance to promote accountability. Bringing the rigor and expertise developed by BLS and the entire Federal statistical system to these department activities will improve reliability and build public trust.

The Statistical Official advises the entire Department on statistical policy, techniques, and procedures. In so doing, the Statistical Official ensures data relevance (e.g., by validating that data are appropriate, accurate, objective, accessible, useful, understandable, and timely), harnesses existing data (e.g., by identifying data needs and repurposing existing data wherever possible), anticipates future uses (e.g., by building interoperability of data from its inception), and demonstrates responsiveness (e.g., by improving data collection, analysis, and dissemination with ongoing input from users and stakeholders). For example, BLS coordinates with the DOL Chief Evaluation Officer and Chief Data Officer to share information and educate DOL stakeholders about BLS data and publicly-available data sets, including those that could be useful in identifying disparities among demographic groups in employment status, wages/income, consumption patterns, housing, workplace safety, and use of social programs.

Additionally, BLS provides a broad range of information to external stakeholders through the BLS public website and other venues. For example, BLS has considerable data to track black workers and other demographic groups and has shared data with DOL stakeholders on various topics of interest, including wages and benefits by union status, employment of formerly-incarcerated individuals, and household income by race, ethnicity, sex, age, and other variables. Beginning with January data published in February 2022, BLS also will publish monthly and quarterly labor force data for American Indians and Alaska Natives, with historic data available back to 2003. These publicly-available data sets can inform policy decisions to advance the Administration's goals on equity and combatting poverty.

NEW RESPONSIBILITIES TO EXPAND DATA ACCESS

The Evidence Act assigns many new or expanded responsibilities to the newly created Federal Statistical Official role. These trusted intermediaries are charged with supporting expanded access to the government's restricted data assets for evidence building, and to do so while maintaining confidentiality and privacy protections. When fully realized, these provisions will make important strides toward increasing the capacity and use of data for policy and evidence building and advancing Administration priorities, a critical step toward restoring the public's trust in government. BLS and the entire statistical community began working on each of these responsibilities soon after the Evidence Act was enacted, and work continues.

Under the auspices of the Interagency Council on Statistical Policy, BLS is working with other statistical agencies to implement several provisions of the Evidence Act designed to expand research capacity through expanded data access. This work includes three inter-related activities. First, the Presumption of Accessibility provision requires statistical agencies to share data wherever specific laws do not prohibit sharing and allows statistical agencies to request administrative and other data from program offices for exclusively statistical purposes. Second, the Tiered Access provision encourages use of data in settings that align the amount of information disclosed with the minimal amount needed to achieve a project's goals. For example, some projects can be achieved using published tabulations or public-use data, others might be achieved through use of synthetic microdata, perhaps with the possibility of remote access to 'real' data to validate results, while still others might require researchers to directly access confidential microdata in a secure setting. Third, statistical agencies are developing a Standard Application Portal to provide a one-stop-shop for researchers to learn about and apply to use restricted access data. Together, these features are designed to expand the capacity for evidence-building research both within the statistical agencies and among qualified researchers.

Consistent with this expanded access is the need to ensure confidentiality. To this end, the Federal Committee on Statistical Methodology developed a Data Protection Toolkit to share disclosure avoidance techniques among statistical agencies and external researchers. The toolkit is designed to increase statistical capacity by developing tools that can be easily applied to products produced by all departmental business units. In addition to supporting work in the toolkit, BLS is working to expand its own capacity to safeguard sensitive information contained in the agency's datasets. Recently, BLS worked with the National Science Foundation and Penn State University to review and expand formal privacy frameworks and how they might be applied to BLS establishment surveys. As the DOL Statistical Official, the BLS Commissioner and staff can work with business units across the department to identify the tools and training needed to implement these advanced data protection methods.

Finally, to support the Evidence Act's goal of restoring trust in government, statistical agencies are assigned certain fundamental responsibilities, including producing and disseminating relevant and timely statistical information and ensuring protection of data acquired under a pledge of confidentiality or similar statutory requirement. Carrying out these fundamental responsibilities in conformance with OMB's expected implementing regulation requires increased internal statistical agency capacity and technical expertise; BLS and other statistical agencies will be identifying tools and training needed to support this statistical capacity going forward. In addition, the Evidence Act requires DOL and other parent agencies to enable, support, and facilitate statistical agency activities. For BLS and other statistical agencies, this involves partnering with departmental Chief Information Officers, Chief Data Officers, and other officials to ensure that the fundamental responsibilities are being met. Addressing these fundamental responsibilities is an important core statistical agency mission, but is also a necessary precursor to safely expanding data sharing and access to restricted data.

For nearly 140 years, BLS has built and enhanced its statistical capacity to meet stakeholder needs in an ever-changing economy. The Evidence Act provides new challenges for BLS and all statistical agencies, but also provides new opportunities to expand research and data sharing while also enhancing its ability to protect confidential data. This report provides a summary of the efforts BLS has and will be undertaking to meet its expanded responsibilities.

APPENDIX C: SURVEY SAMPLE CRITERIA

IMPAQ/AIR proposed the following sample to distribute the capacity assessment survey to individuals across the Department.

- **Criteria 1**: A sample of GS-13 and above career staff and managers, with the rationale that jobs at this level require understanding and application of analysis, statistics, research, and/or evaluation to inform programs, policies, or operations.
- **Criteria 2**: Exclusion of the following occupational codes, with the rationale that these jobs would not likely have day-to-day work requirements for generating and using evidence. [15]

CODE	OCCUPATION	CODE	OCCUPATION
0080	Security Administration Series	1071	Audiovisual Production
0089	Emergency Management Series	1082	Writing and Editing Series
0201	Human Resource Assistance	1084	Visual Information Series
0203	Human Resource Assistance	1101	General Business and Industry
0243	Human Resource Assistance	1102	Contracting Series
0260	Human Resource Assistances	1160	Financial Analysis Series
0308	Records & Info Management Series	1170	Realty Series
0342	Support Services Admin Series	1176	Building Management Series
0360	Equal Opportunity Compliance	1410	Librarian Series
0510	Accounting Series	1412	Technical Information Services Series
0511	Auditing Series	2003	Supply Management
1001	General Arts & Information Series	2210	IT Management Series
1008	Interior Design Series		

- **Criteria 3:** Exclusion of additional BLS staff after recognizing that BLS would be overrepresented in the sample. A sampling of BLS staff was conducted based on in consultation with BLS, which resulted in a more appropriate sample of 10 percent of the total BLS staff.

The following table shows how many individuals from each agency were surveyed based on the exclusion plan described above.

[15] Future capacity assessment surveys may revise this list of excluded occupation codes.

Agency	Total Sample (Criteria 1, 2, & 3)	% of Total Sample
BLS	82	2.4%
EBSA	416	12.1%
ETA	483	14.0%
ILAB	88	2.6%
MSHA	416	12.1%
OASAM	161	4.7%
OASP	34	1.0%
OCFO	19	0.6%
ODEP	42	1.2%
OFCCP	64	1.9%
OLMS	77	2.2%
OSHA	702	20.4%
OWCP	370	10.7%
VETS	102	3.0%
WB	19	0.6%
WH	371	10.8%
TOTAL	**3,446**	**100%**

APPENDIX D: SURVEY INSTRUMENT

DOL CAPACITY ASSESSMENT SURVEY

[INTRO]

INTRODUCTION

Welcome and thank you for participating in this survey. We are interested in understanding how you use and produce evidence when working on policies, programs, and/or operations. Evidence is data or information, such as statistics, research, analyses, or evaluations.

Your response is voluntary and will be kept confidential. Results will not be reported at the individual level.

We anticipate this survey will take about 10-15 minutes to complete. You may stop this survey when you like, and restart when it is convenient to you. We will remind you from time to time to complete the survey if you have not yet done so.

[Q0]

Your input is important. It will help inform additional supports and resources such as training, process improvements, communications, or studies.

Code	Label	Skip to	Attribute
01	I agree to participate in this survey.	Q1.1	
02	I would like to opt out of this survey.	Close survey	Choose to opt out

[SEC1Contextual Information]

SECTION A: Contextual Information

[QA.1]

Please select your DOL agency:

Code	Label
01	BLS - Bureau of Labor Statistics
02	EBSA - Employee Benefits Security Administration
03	ETA - Employment and Training Administration
04	ILAB - Bureau of International Labor Affairs
05	MSHA - Mine Safety and Health Administration
06	OASAM - Office of the Assistant Secretary for Administration & Management
07	OASP - Office of the Assistant Secretary for Policy
08	OCFO - Office of the Chief Financial Officer
09	ODEP - Office of Disability Employment Policy
10	OFCCP - Office of Federal Contract Compliance Programs
11	OLMS - Office of Labor-Management Standards
12	OSHA - Occupational Safety and Health Administration
13	OWCP - Office of Workers' Compensation Programs

Code	Label
14	VETS - Veterans' Employment & Training Service
15	WB - Women's Bureau
16	WHD - Wage and Hour Division

[QA.2]

Please specify your office within the agency.

Code	Label
01	Please specify:

[QA.3]

Please select your position/level:

Code	Label
01	GS-13
02	GS-14
03	GS-15
04	SES
05	Other. Please specify:
n/a	R skips question

[QA.4]

Which of these describe your major responsibilities/roles? Please select all that apply:

Code	Label
01	Leadership
02	Manager/Supervisor
03	Policy
04	Strategic planning
05	Program management
06	Communications
07	Research
08	Data analysis
09	Agency administration
10	Monitoring/Oversight
11	Field operations
12	Other. Please specify:

[SECBINTRO]

SECTION B: Evidence and Data Use Activities

This section is about how you and your team are using and producing evidence. **Team** is defined as the people you most often work with from your agency/office/division/unit. **Evidence** is data or information, such as statistics, research, analyses, or evaluations.

[QB.1]

Thinking about the past year, in what areas does **your team** typically **use** statistics, research, analyses, or evaluations to inform your work? Please select all that apply.

Strategic Planning and Budget

- ☐ Program strategy and goals
- ☐ Operating plans
- ☐ Agency budget recommendations
- ☐ Research agendas or research questions

Policy

- ☐ Policy development or updates to policy
- ☐ Coordination or communication efforts with stakeholders
- ☐ Response to oversight inquiries (Congressional, OMB, GAO/OIG)

Program Management and Operations

- ☐ Program development or updates to programs
- ☐ Resource allocation
- ☐ Process improvements
- ☐ Service improvements for constituents
- ☐ Corrective action to solve problems
- ☐ Coordination or communication efforts with stakeholders

Code	Label
01	Program strategy and goals
02	Operating plans
03	Agency budget recommendations
04	Research agendas or research questions
05	Policy development or updates to policy
06	Coordination or communication efforts with stakeholders
07	Response to oversight inquiries (Congressional, OMB, GAO/OIG)
08	Program development or updates to programs
09	Resource allocation
10	Process improvements
11	Service improvements to constituents
12	Corrective action to solve problems
13	Coordination or communication efforts with stakeholders
14	Other. Please specify:

[QB.2a-bINTRO]

The next set of questions are about how you and your team **use** different types of evidence. We ask first about you and then your team.

[QB.2a-b]

Thinking about the past year, how often did **you** (not your team):

[QB.2a]

Use **statistics, research, or analyses** such as data analytics, trend analyses, or literature and document reviews on specific programs or policy issues?

Code	Label
01	Frequently
02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.2b]

Use **evaluations** that provided formal assessment of a program, policy, or organization's implementation, impacts, effectiveness, or efficiency?

Code	Label
01	Frequently

02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.3a-b]

Thinking about the past year, roughly how often did ***your team*** (not you):

[QB.3a]

Use **statistics, research, or analyses** such as data analytics, trend analyses, or literature and document reviews on specific programs or policy issues?

Code	Label
01	Frequently
02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.3b]

Use **evaluations** that provided formal assessment of a program, policy, or organization's implementation, impacts, effectiveness or efficiency?

Code	Label
01	Frequently
02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.4a-bINTRO]

The next set of questions are about how you and your team **produce** different types of evidence. We ask first about you and then your team.

[QB.4a-b]

Thinking about the past year, how often did **you** (not your team) help**:**

[QB.4a]

Produce statistics, research, or analyses such as data analytics, trend analyses, or literature and document reviews on specific programs or policy issues?

Code	Label
01	Frequently
02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.4b]

Produce **evaluations** that provided formal assessment of a program, policy, or organization's implementation, impacts, effectiveness, or efficiency?

Code	Label
01	Frequently
02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.5a-b]

Thinking about the past year, roughly how often did ***your team*** (not you):

[QB.5a]

Produce **statistics, research, or analyses** such as data analytics, trend analyses, or literature and document reviews on specific programs or policy issues?

Code	Label
01	Frequently
02	Occasionally
03	Rarely
04	Never
05	Not applicable to my work

[QB.5b]

Produce **evaluations** that provided formal assessment of a program, policy, or organization's implementation, impacts, effectiveness or efficiency?

Code	Label
01	Frequently
02	Occasionally
03	Rarely

04	Never
05	Not applicable to my work

[QB.6A]

Thinking about the past year and your job responsibilities, how satisfied were you with your **access** to evidence that was:

a. **Relevant** to outcomes and issues that mattered to your office?
b. **Comprehensive** in addressing your office's goals and objectives?
c. **Useful** for making program, policy or operational decisions?
d. **Effectively disseminated** to relevant staff and decision makers **inside** the Department?
e. **Effectively disseminated** to relevant stakeholders and decision makers **outside** of the Department?

Code	Label
01	Very Satisfied
02	Satisfied
03	Somewhat Satisfied
04	Not Satisfied
05	Don't know
06	Not applicable to my work

[QB.6B]

Thinking about the past year and your job responsibilities, how satisfied were you with your **access** to evidence that was:

f. Helpful in **improving underserved populations'** access to benefits and services that your agency offers?
g. **High quality** (credible and objective)?
h. **Appropriate** in their use of rigorous methodological approaches?
i. **Independent** (free from bias)?

Code	Label
01	Very Satisfied
02	Satisfied
03	Somewhat Satisfied
04	Not Satisfied
05	Don't know
06	Not applicable to my work

[QB.7]

Please note how much you agree or disagree with the following statement.

My **team** spends the appropriate amount of time using and producing evidence needed to carry out the work of our office given our other responsibilities.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to my work

[SECCINTRO]

SECTION C: Background Information

In this part of the survey, you will be asked to think about your capacity to use and produce evidence. **Capacity** is your personal knowledge, skills, and ability. **Evidence** is data or information, such as statistics, research, analyses, or evaluations. Please note how much you agree or disagree with the following statements.

[QC.1]

I have experience creating logic models that describe key inputs, activities, outputs, and outcomes for a program.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to my work

[QC.2]

I can keep up on research that is relevant to my work as much as I would like.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to my work

[QC.3]

I am comfortable using the Department of Labor Clearinghouse for Labor and Employment Research (CLEAR) website to determine whether a given practice, policy, or program has evidence of effectiveness.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	I am not familiar with CLEAR

[QC.4]

I can assess whether information I read represents strong or weak evidence (i.e., whether I should use it to inform my work).

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to my work

[SECDINTRO]

SECTION D: Office and Agency Context

The next two questions are about the context of evidence use in your office and agency. Please note how much you agree or disagree with the following statements.

[QD.1]

My **direct supervisor** encourages and promotes using statistics, research, analyses or evaluations to improve program operations or activities our team conducts to carry out our mission.

Code	Label
01	Strongly Agree
02	Agree
03	Neither Agree nor Disagree
04	Disagree
05	Strongly Disagree

[QD.2]

My **agency** encourages and promotes using statistics, research, analyses, or evaluations to improve programs, policy, and operations.

Code	Label
01	Strongly Agree
02	Agree
03	Neither Agree nor Disagree
04	Disagree
05	Strongly Disagree

[SECEINTRO]

SECTION E: Office Capacity and Resources

This part of the survey is about the capacity and resources available in your office to use evidence to improve program performance.

You will be asked to think about your team's **capacity** (i.e., knowledge, skills, and ability) in using and producing evidence. **Team** is defined as the people you most often work with from your agency/office/division/unit. **Evidence** is data or information, such as statistics, research, analyses, or evaluations. Please note how much you agree or disagree with the following statements.

[QE.1]

My team has **access** to the statistics, research, analyses, or evaluations it needs to improve programs, policy or operations.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to our work

[QE.2]

My team has the capacity to **use** statistics, research, analyses, or evaluations to improve programs, policy or operations.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to our work

[QE.3]

My team has the capacity to **produce** statistics, research, analyses, or evaluations to improve programs, policy or operations.

Code	Label
01	Strongly Agree
02	Agree
03	Disagree
04	Strongly Disagree
05	Not applicable to our work

[QE.4-5_INTRO]

The next two questions are about resources to improve your team's ability to use and produce evidence.

[QE.4]

My team's ability to **use** statistics, research, analyses, or evaluations could be **most** improved by providing:

Please select all that apply.

- ☐ Professional development/training
- ☐ Software and/or hardware
- ☐ Usable or relevant data
- ☐ Access to reliable or high quality DOL data
- ☐ Access to specific types of DOL data (such as individual or business level data or PII for matching)
- ☐ Access to reliable or high quality data from other agencies
- ☐ Access to timely data
- ☐ Access to relevant expertise
- ☐ Relevant evidence
- ☐ Time to review and use evidence
- ☐ Political support
- ☐ Performance incentives (e.g., performance appraisals, bonuses, or awards)
- ☐ Other. Please specify: ______________________________

Code	Label
01	Professional development/Training
02	Software and/or hardware
03	Usable or relevant data
04	Access to reliable or high quality DOL data
05	Access to specific types of DOL data (such as individual or business level data or PII for matching)
06	Access to reliable or high-quality data from other agencies
07	Access to timely data
08	Access to relevant expertise
09	Relevant evidence
10	Political support
11	Performance incentives (e.g., performance appraisals, bonuses, or awards)
12	Other. Please specify:

[QE.5]

My team's ability to **produce** statistics, research, analyses, or evaluations could be **most** improved by providing:

Please select all that apply

- ☐ Professional development/training
- ☐ Software and/or hardware
- ☐ Usable or relevant data
- ☐ Access to reliable or high quality DOL data

- ☐ Access to specific types of DOL data (such as individual or business level data or PII for matching)
- ☐ Access to reliable or high quality data from other agencies
- ☐ Access to timely data
- ☐ Access to relevant expertise
- ☐ Relevant evidence
- ☐ Time to review and use evidence
- ☐ Political support
- ☐ Performance incentives (e.g., performance appraisals, bonuses, or awards)
- ☐ Other. Please specify: ______________________________

Code	Label
01	Professional development/Training
02	Software and/or hardware
03	Usable or relevant data
04	Access to reliable or high quality DOL data
05	Access to specific types of DOL data (such as individual or business level data or PII for matching)
06	Access to reliable or high-quality data from other agencies
07	Access to timely data
08	Access to relevant expertise
09	Relevant evidence
10	Political support
11	Performance incentives (e.g., performance appraisals, bonuses, or awards)
12	Other. Please specify:

[SECFINTRO]

SECTION F: Training

The next set of questions explore areas and types of training that might be helpful for your work.

[QF.1]

How important is it that you **learn more** about the following topics for you work?

a. **Using evidence** to design programs or support policy development
b. Analyzing and interpreting data
c. Understanding the effectiveness of DOL programs or policy
d. **Drawing implications** from a summary of research studies on a topic
e. Understanding how to **access DOL data** relevant to my work
f. Guiding the **process for collecting information** on new or existing programs
g. Understanding the **characteristics of the populations** that my programs serve (e.g., distribution of demographic characteristics or program eligibility or descriptions of population receiving services)
h. Other. Please specify: (Enter N/A if you unintentionally rated Other.)

Code	Label
01	Very Important
02	Moderately Important
03	Slightly Important
04	Not Important at All
05	Not applicable to our work

[QF.2]

How familiar are you with **program evaluation**?

Program evaluation refers to the planning, execution, and use of research on the implementation and impacts of DOL-funded initiatives and policies.

Code	Label
01	Very Familiar
02	Familiar, but can benefit from a refresher on specific topics
03	Not familiar

[QF.3]

Please select **three** training topics in **program evaluation** that are most important for your work.

Code	Label
01	Obtaining valid and reliable information about a policy or program
02	Documenting the implementation of a program or policy
03	Determining the impacts of my program
04	Using existing data to understand the effectiveness of DOL policies or programs
05	Understanding requirements and procedures for collection data from the public
06	Other. Please specify.

[QF.4]

How familiar are you with **data analytics**?

Data analytics refers to the processing, analyses, and presentation of data.

Code	Label
01	Very Familiar
02	Familiar, but can benefit from a refresher on specific topics
03	Not familiar

[QF.5]

Please select **three** training topics in **data analytics** that are most important for your work.

Code	Label
01	Ensuring the quality of data generated by my office or partners
02	Cleaning and coding data generated by my office or obtained from other entities
03	Securely storing and/or sharing data generated by my office or obtained from other entities
04	Exploring data patterns and interpreting their meanings
05	Communicating about data analysis results effectively
06	Other. Please specify.

[SECGINTRO]

SECTION G: Agency Mission and Goals

This final section asks about the mission and goals of your agency. Please note how much you agree or disagree with the following statements.

[QG.1]

If asked by a friend outside of work, I would be able to accurately describe my agency's **mission**.

Code	Label
01	Strongly Agree
02	Agree
03	Neither Agree nor Disagree
04	Disagree
05	Strongly Disagree

[QG.2]

I could tell a new coworker in my office about my agency's **Strategic Plan goals** that are most relevant to our work.

Code	Label
01	Strongly Agree
02	Agree
03	Neither Agree nor Disagree
04	Disagree
05	Strongly Disagree

[QG.3]

I am familiar with my agency's **Operating Plan** and how it relates to my work.

Code	Label
01	Strongly Agree
02	Agree
03	Neither Agree nor Disagree
04	Disagree
05	Strongly Disagree

[QG.4]

I am familiar with my agency's **Learning Agenda** (i.e., my agency's research priorities) and how it relates to my work.

Code	Label
01	Strongly Agree
02	Agree
03	Neither Agree nor Disagree
04	Disagree
05	Strongly Disagree

[Final Close - Submission Page for Survey Completers]

Thank you for taking time to complete this survey. We appreciate your responses.

Want to know about the latest on all things evidence in the Department?

You can subscribe to the Chief Evaluation Office's new quarterly Building the Evidence Base newsletter by clicking here. You will hear about our latest independent research, data resources, and upcoming events.

Submit Button

[Thank you for Submission Page]

[Submission for Survey Break-offs]

You can complete the survey at a later date by clicking on the link you received via email. We will remind you from time to time about your opportunity to complete the survey later. Remember, the survey will close on March 19th.

APPENDIX E: REVIEW OF SELECT EVIDENCE-RELATED ACTIVITIES

1. METHODOLOGY

Select Evidence Activities. For the in-depth review of selected activities, the team selected three studies from each agency: two evaluations and one foundational fact-finding report, as available. We used different combinations of study types as needed according to each agency's list of evidence-related activities. For example, for agencies with only foundational fact-finding reports, we selected all studies from that category. In addition, the Employment and Training Administration had such a long list of studies (more than 60), that we selected four studies—two evaluations and two foundational fact-finding reports—to more fully represent the breadth of its work. In a few cases, we did not need to select studies, as an agency had only two or three reports available within our targeted date range.

We developed a set of criteria to select individual studies of each type. For evaluations, we selected the two studies with the most rigorous methodologies, in the following priority order: randomized controlled trials (RCT), quasi-experimental impact studies, descriptive outcomes studies, implementation/process studies, and evaluation design option reports. Typically, this criterion was sufficient to complete the selection process. However, for a few agencies (for example, those with more than two RCT studies) we used additional criteria to refine our selection, as described below.

ADDITIONAL SELECTION CRITERIA FOR EVIDENCE-RELATED ACTIVITY REVIEW

The additional criteria were also the starting place for our selection process for foundational research studies, as follows:

1. **Select more recent studies.** We filtered to the two most recent years of publication for that agency (e.g., 2018 and 2019, or 2019 and 2020).
2. **Select studies that have final reports.** We prioritized final reports over interim reports. In addition, we prioritized full reports over shorter summary briefs. However, if a final report was accompanied by an issue brief, we considered both documents as part of the review package. In addition, we prioritized research studies over evaluation design option reports, depending on availability.
3. **Use random selection for the remaining list, as needed**. For the remaining studies, we used a function in Excel to randomly select one study.
4. **Confirm that the selected studies represent different contract vehicles.** For example, agencies often do foundational research that later supports a rigorous evaluation on the same topic. If possible, we preferred that the three selections for each agency represented different areas of investigation. We redid randomization, as needed.
5. **Confirm that the same study is not used for multiple agencies.** This was needed because some studies covered multiple agencies. We redid randomization, as needed.

IN-DEPTH REVIEW ASSESSMENT DOMAIN MEASURE SET

Categories				
High	Medium	Low	Not Applicable	Unknown/No Information

QUALITY

QUALITY (statements below describe the "High" category)		
Integrity	1. Data sources and analysis methods are well documented.	**Category:** **Brief Justification:**
	2. Study notes the strengths and weaknesses of the approach. Outcome evaluations and correlational studies must note that the study is not causal to be placed in the "high" category.	**Category:** **Brief Justification:**
Data Quality	3. **(In the case of original data collection)** There are clear data collection protocols that identify who is collecting what data, when, from whom, and for what purpose (surveys, focus groups, interviews, etc.).	**Category:** **Brief Justification:**
	4. **(Quantitative Studies only: RCTs, quasi-experimental studies, outcomes evaluations and quantitative foundational research)** The analytic dataset is of sufficient quality to support analysis. **Factors to consider:** a. **Missing data.** There is no indication of problematic missing data on key metrics/variables. May keep in the high category if study uses a robust methodology to address missing data, such as random imputation. b. **Sample size.** The size of the analytic dataset is appropriate for the proposed analysis method. Impact evaluations should have a power analysis to justify the sample size. c. **Unit of analysis**. The unit of analysis is appropriate. (Is the data at a sufficiently granular level and from the right groups to answer the research questions?) d. **Accuracy.** There are no strong reasons to be concerned about the accuracy of the recorded data on key variables (e.g., case managers entering in participant data, substantial recall bias or self-reporting bias in a survey). In other words, key variables appear to be valid (measuring what they are intended to measure).	**Category:** **Brief Justification:**

METHODS

METHODS (statements below describe the "High" category)	
1. The research design and method address all the intended research questions.	**Category:** **Brief Justification:**
2. Design and method credibility is strengthened by a peer review/consultation from unbiased experts. Examples are technical working groups (TWGs), third party reviewer, CLEAR database, and quality assurance procedures to monitor survey field workers.	**Category:** **Brief Justification:**

METHODS (statements below describe the "High" category)	
3. **(Quantitative studies only)** The underlying data for the analysis is publicly available or available upon request, so another researcher can reproduce the findings (e.g., public use file, restricted use, full data printed in an appendix) DOL websites with public use data: https://www.bls.gov/developers/ https://www.dol.gov/agencies/oasp/evaluation/publicusedata	**Category:** **Brief Justification:**
4. **(Quantitative RCTs or quasi-experimental impact evaluations only)** For impact studies, studies have high levels of internal validity, based on the factors below. ***Clear Evidence if:*** • **Baseline: (RCT and quasi)** Intervention and comparison group are similar before intervention (including baseline measures of key outcome variables). Any differences are controlled for in the regression. • **Confounds: (RCT and quasi)** There are no confounding factors (no reason to believe an outside factor is driving the results). In other words, except for the intervention, the conditions for the comparison group should be the same as for the intervention group. • **Anticipation: (RCT and quasi)** Sample members' anticipation of the intervention is controlled for, as relevant. • **Attrition: (RCT)** Sample attrition and contamination (control group receiving intervention services, or vice versa) is low. • **Assignment: (RCT)** For RCTs, the probability of assignment into treatment/control is consistent over time. **Overall: (quasi-experimental)** Appropriate methods are used to control for selection into treatment group on observable and unobservable characteristics **Resource: DOL CLEAR Evidence Guidelines**: https://clear.dol.gov/sites/default/files/CLEAR_EvidenceGuidelines_V2.1.pdf	**Category:** **Brief Justification:**
5. **(RCTs and quasi-experimental impact studies only)** Study provides clarity about populations, settings, or circumstances for which results can be generalized (external validity).	**Category:** **Brief Justification:**

INDEPENDENCE

INDEPENDENCE (statements below describe the "High" category)	
1. Evidence activity is designed, conducted, and analyzed by independent researchers (independent of agency and program leadership, program staff, and stakeholders), when feasible.	**Category:** **Brief Justification:**
2. **(Evaluation-only and literature reviews/environmental scans, as relevant)** All results, including favorable, unfavorable, and null findings are shared with relevant audiences**.**	**Category:** **Brief Justification:**

EFFECTIVENESS

EFFECTIVENESS (statements below describe the "High" category)	
1. Studies indicate the ways in which the research questions or analyses are intended to serve the needs of stakeholders (e.g., serves the needs of employers, addresses leadership interest/strategic plan) and the kinds of decisions that the questions or analyses could inform. "High category" if the study clearly describes the policy relevance of the study and/or offers recommendations for program improvement.	**Category:** **Brief Justification:**
2. For publicly available reports, results are accessed or used by others. High category if has more than 1 Google Scholar citation, a conference presentation, or study is publicized by an outside organization's website (outside of DOL or the research organization doing the work). Based on quick Google Scholar and Google search. ("Medium" category has only 1 Google Scholar citation and no other indications of use described above.)	**Category:** **Brief Justification:**
3. Results are shared in a user-friendly (clear and concise) format. Consider the full report package, including any associated issue briefs, podcasts, videos, slides, etc. **"High" does at least 2 of the following.** ("Medium" does only 1 of the following). a. Has a nontechnical executive summary b. Has a nontechnical discussion of results c. Uses graphics to make the material more accessible d. Results are summarized in multiple formats (e.g., issue briefs and main report)	**Category:** **Brief Justification:**

2. FINDINGS

The desk review of publicly available evidence activities aims to understand coverage, quality, methods, effectiveness, and independence. This task leverages an inventory of evidence-related activities provided by CEO and compiled for the interim capacity assessment. Our approach differed by domain. For the coverage, we reviewed the full spectrum of all 313 studies on the data inventory. For quality, methods, effectiveness, and independence, we closely examined 36 selected evaluations and/or foundational fact-finding reports published between October 1, 2017, and September 30, 2020. We selected two evaluations and one foundational research study from each agency, as available, based on the selection criteria described above. We assessed each of the 36 studies on 2 to 4 measures per domain, using a scale of high, medium, low, or unknown based on the measure described in the chart above. While this select review of 36 studies is not fully representative of DOL, it provides a set of indicators for assessing the robustness of studies available to DOL staff. This limited review offers a template that DOL can scale-up to more comprehensively assess their evidence activities in the future.

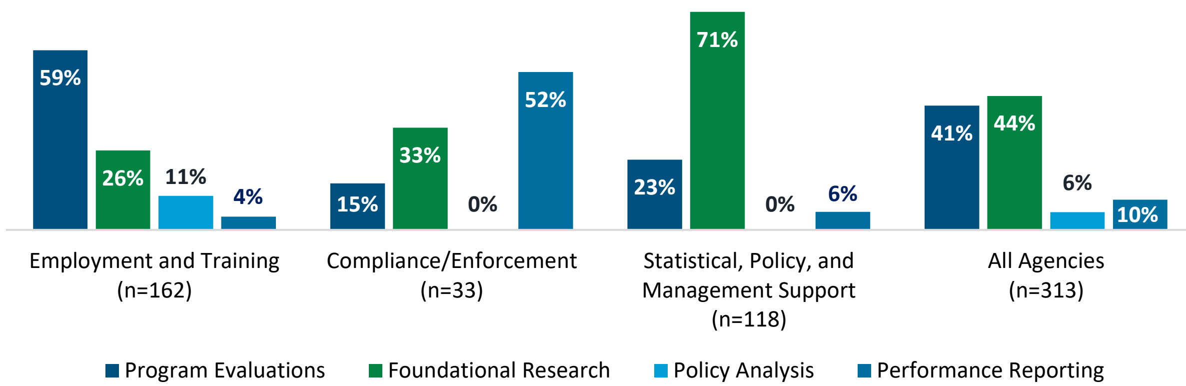
COVERAGE DOMAIN

For the Coverage Domain, we reviewed the distribution of study types across the 313 published studies in the DOL Evidence Review. The study types include program evaluations, foundational research, policy analysis, and performance reporting, as categorized by the agencies. The exhibit below shows the distribution of studies in DOL overall and by agency type. DOL primarily published foundational research (44 percent of studies) and program evaluations (41 percent of studies). The remaining studies included 10 percent performance reporting and 6 percent policy analysis.

There are distinct patterns by agency type. Employment and training agencies focused their evidence-building activities most strongly on program evaluations (59 percent of studies); compliance/enforcement agencies concentrated on performance reporting (52 percent of studies); and support agencies invested most intensely in foundational research (71 percent of studies).

There is also variation in the volume of studies produced by agency type. As shown in Exhibit D-1 below, employment and training agencies and support agencies published 162 studies and 118 studies, respectively, compared with only 33 published studies from compliance/ enforcement agencies. However, compliance/enforcement agencies reported an additional 29 ongoing studies, which will quickly double their body of evidence in the next few years. Employment and training agencies and support agencies reported an additional 71 and 62 ongoing studies, respectively.

Exhibit E-1: Distribution of Studies by Agency Type



Note: Table includes all publicly accessible studies as of October 1, 2017, meaning the assessment is posted online or there are instructions online on how to access it.

Chart definitions:

1. Program Evaluation: Systematic analyses of a program, policy, organization, or component, to assess effectiveness and efficiency, including implementation studies, impact studies, etc.
2. Foundational Research: Any other rigorous research or analysis that does not fall under program evaluation, policy analysis, or performance reporting. Foundational research can include exploratory studies, statistical analysis, qualitative research, etc.
3. Policy Analysis: Analysis of data, such as survey data or program-specific data, to generate and inform policy (e.g., estimating regulatory impacts).
4. Performance Reporting: Reports that provide or summarize agency performance data (e.g., annual reports that provide performance data).

EFFECTIVENESS DOMAIN

Within the Effectiveness Domain, we developed three measures to help make observations on the degree that each of the 36 selected studies is relevant and accessible to stakeholders. We assessed whether the study articulated its policy relevance and/or how its findings serve stakeholders' needs; the user friendliness of the reports; and how frequently studies are accessed**.** The indicators are quite positive. Our findings below also point to some opportunities for DOL to make their studies more user-friendly and more visible to a wider audience**.** Below, we present the results, organized by measure.

Relevance. Most studies clearly articulated their policy relevance: "High" category for 31 of 36 studies (86 percent). By agency type, we assigned this highest category to 13 of 16 studies from the employment and training agencies; 12 of 13 studies from compliance/enforcement agencies; and 6 of 7 studies from the support agencies. We placed four studies in the medium category that left the reader to mostly infer the goals, stakeholders, or policy relevance, and placed one study in the low category where it was particularly vague how one would use the results.

User Friendliness. Nearly all studies had multiple features to share their results in a user-friendly format: "High" category for 34 of 36 studies (94 percent).[16] We placed two studies from support agencies in the low category because they lacked any user-friendly features and were highly academic and technical in nature. We found areas for improvement even in some studies in the high category. Some of these reports had graphics placed at the end of the report instead of embedded with the associated text, and results were discussed quite technically in the executive summary, introduction, or conclusion.

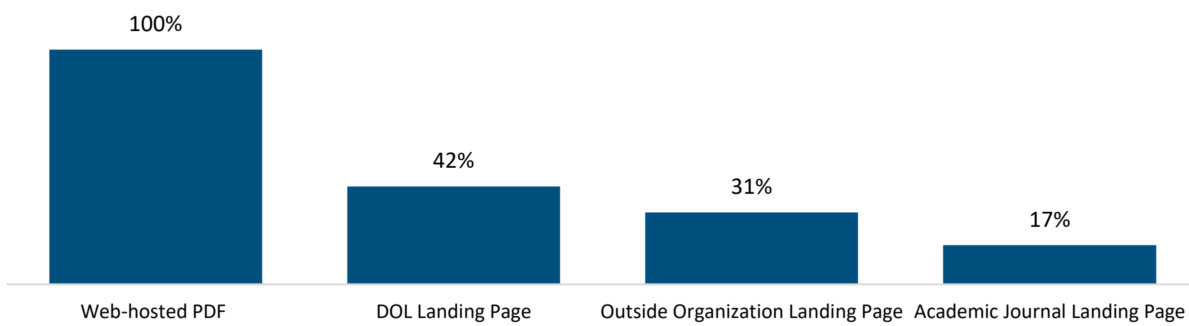
Frequency of Access. We found evidence that the reports were accessed or used by stakeholders for about half of the studies: "High" category for 15 of 36 studies (42 percent). We assigned studies the high category for having Google Scholar citations, conference presentations, or that were advertised by an outside organization's website (outside of DOL or the research organization doing the work).[17] By agency type, we assigned this highest category to: 9 of 16 studies from the employment and training agencies; 3 of 13 studies from compliance/enforcement agencies; and 3 of 7 studies from the support agencies. We categorized studies in the "High" category most consistently for rigorous program evaluations (randomized control trials) and other rigorous academic studies. However, this categorization was true less often for studies from compliance/enforcement agencies. We categorized about half of studies as "Unknown" on this measure: 16 of 36 (44 percent) lacked a Google Scholar entry or evidence of other outside engagement.

Additional Observations. Where studies appear online. In addition to these three measures, we also noted where each of the 36 studies are hosted online. All studies are represented as PDF files on a DOL Web server, which appears in a Google search if one typed the proper keywords (see Exhibit D-2). (This 100 percent retrieval rate is unsurprising because our sampling frame of studies was restricted to those that were publicly available.) However, more opportunities to disseminate DOL studies to external stakeholders may exist, as only 42 percent of the studies have a DOL landing page, 31 percent have a landing page hosted by an outside (not DOL) organization, and 17 percent have an academic journal landing page.

[16] These studies did at least 2 of the following: Had a non-technical executive summary, a non-technical discussion of results, leveraged graphics to make the material more accessible, or summarized their results in multiple formats like an issue brief or podcast.

[17] To assess this measure, we conducted a quick Google search on all studies that we could not find on Google Scholar to look for other outside engagement. Several studies were presented at conferences but have not been cited on Google Scholar.

Exhibit E-2: Web Location of Studies



Bar Chart Definitions: 1) DOL Web-hosted PDF only; 2) DOL landing page, where an agency describes the study and links to any summaries or other versions available; 3) Other organization landing page, where usually the contracting organization or author of the study hosts it on their own website; and 4) Academic journal landing page, where the study has a dedicated webpage built by an academic journal.

QUALITY DOMAIN

For the Quality Domain, we developed four measures to help us make observations about the sufficiency of the study's documentation and the quality of the analytic dataset for each of the 36 selected studies. For documentation, we reviewed how well the study documented its data sources and analysis methods; how thoroughly the authors discussed strengths and weaknesses of their approach; and whether the study had clear data collection protocols, as applicable. Lastly, we considered whether the study's analytic dataset was of sufficient quality to support the analysis. To assess measure 4, we looked for problematic levels of missing data on key variables that may affect the results, if the author had concerns about the accuracy of recorded data, and whether the sample size and unit of analysis was appropriate. These indicators point to strong documentation and robust discussions of the strengths/weaknesses of the research approaches. The indicators also highlight some of the challenges researchers face when building an analytic dataset from DOL administrative data sources or fielding surveys to DOL populations. Below, we present the results, organized by measure.

Documentation. Most studies thoroughly document data sources and analysis methods: "High" category for 31 of 36 studies (86 percent). By agency type, we assigned this highest category to: 15 of 16 studies from the employment and training agency; 10 of 13 studies from compliance/enforcement agency; and 6 of 7 studies from the management agency.

Discussion of Strengths/Weaknesses. Most studies sufficiently discuss the strengths/ weaknesses of their approaches: "High" category for 27 of 36 studies (75 percent). By agency type, we assigned this highest category to 14 of 16 studies from employment and training agencies; 9 of 13 studies from compliance/enforcement agency; and 4 of 7 studies from support agencies. We assigned the "medium" level to 5 studies for the following reasons: 1) It is unclear if we should interpret the regression results as correlational or causal; 2) the authors discussed only strengths but no limitations; 3) the researchers did not explain how weaknesses in their approach could affect the interpretation of results; or 4) researchers did not validate some of the basic assumptions of their econometric model.

Data Collection Protocols. The studies have clear data collection protocols: "High" category for 12 of 13 studies (92 percent).[18] Only 13 studies of our sample of 36 collected original data (e.g., surveys, focus groups), while the rest relied on administrative datasets. Most of these studies (8 of 13) are from employment and training agencies.

Analytic Dataset Quality. About half of studies have an analytic dataset of sufficient quality to fully support their analysis: "High" category for 18 of 36 studies (50 percent). Researchers commonly face data challenges. Most of these studies rely on administrative datasets, often created to serve other stakeholder needs or answer different research questions. Some studies also use surveys, which can suffer from low response rates. By agency type, we assigned the highest category for 9 of 16 studies from the employment and training agencies; 4 of 13 studies from compliance/enforcement agencies; and 5 of 7 studies for support agencies. We assigned about half of studies to the "Medium" category. There are some differences by agency type on the reasons for these lower assessments. For compliance/enforcement agencies, we assigned studies to the "Medium" category primarily due to problematic levels of missing data or reliance on proxy data for key variables. For employment and training agencies, we assigned studies to the "Medium" category due to sample size factors. In most cases, the sample size was too small to measure all the outcomes of interest. In addition, several impact studies lacked a power analysis to validate the sufficiency of their sample size.

METHODS DOMAIN

Within the Methods Domain, we developed three measures to help us make observations about the strength of the methodology of each of the 36 selected studies. The first measure, a general metric, can apply to most studies, if the research design addressed the research questions posed. For studies that attempted to make causal inferences (14 of the 36 studies that used quasi-experimental models and randomized control trials), we also examined whether the basic modeling assumptions were sound (second measure) and whether the study provides clarity on the circumstances to which the impact results can be generalized (third measure). These indicators point to strong study approaches to address the research questions. We also note a few common obstacles researchers face in ensuring the internal validity of causal studies. Below, we present the results, organized by measure.

Addressing Research Question(s). Most studies are designed to adequately address the research questions posed: "High" category for 32 of 36 studies (88 percent). By agency type, we assigned this highest category to: 15 of 16 studies from the employment and training agencies; 10 of 13 studies from compliance/enforcement agencies; and 7 of 7 studies from the support agencies. We assessed three studies as "Medium" that either did not collect data to address one of their research questions, lacked sufficient statistical power to answer one of their research questions, or had an important limitation in their approach that impeded their ability to achieve their research objective.

[18] We assessed one study as "Unknown" that did not provide any information on protocols.

Internal Validity. More than half of studies that make causal inferences have high levels of internal validity/sound modeling assumptions: "High" category for 8 of 14 studies (57 percent). Across agency types, employment and training agencies had the most causal studies. By agency type, we assigned the "High" category to 6 of 8 studies from employment and training agencies, 1 of 3 studies from compliance/enforcement agencies, and 1 of 3 studies from support agencies. We assessed four studies as "Medium" because of likely or definite problems with fundamental modeling assumptions, such as lack of similarity between the intervention and comparison group at baseline, crossover between the intervention and control groups, or failing the parallel trends assumption for some outcomes (Difference-in-Differences model). We also categorized two studies we as "Unknown" for this measure, as there was not sufficient information for an assessment.

External Validity. When it applies, most studies that make causal inferences provide clarity about populations, settings, or circumstances to which results can be generalized: "High" category for 9 of 14 studies (71 percent). By agency type, we assigned this highest category to: 4 of 8 studies from the employment and training agencies; 2 of 3 studies from compliance/enforcement agencies; and 3 of 3 studies from the support agencies. We assessed three studies as "Medium" that made some indirect statements relevant to the generalizability of results but did not discuss the issue directly. We assessed two studies as "No Evidence" that skipped this topic altogether.

ADDITIONAL OBSERVATIONS**.** In addition to these three measures, we also made note of how frequently studies assured the integrity of their methods by making their data available so others could reproduce their results and using an external reviewer. Our analysis is likely an undercount as we were limited to what the authors wrote in the report text and a quick Google search.

Use of External Reviewer. Only 10 of 36 studies (more than 25 percent) mentioned using some type of external reviewer (e.g., Technical Working Groups)**.** By agency type, 6 were from employment and training agencies, 1 was from a compliance/enforcement agency, and 3 were from support agencies.

Reproducing Results. We found evidence of a public use file (or an assertion that the data is available for outside researchers) in only 5 of 32 studies (16 percent), where it could apply.

INDEPENDENCE

Within the Independence Domain, we developed two measures to help us make observations about potential sources of bias for each of the 36 selected studies. First, we checked whether the studies were designed and executed by researchers independent from the agency. Second, we considered whether all analytic results appear to be shared in the report, or if some with unfavorable implications were likely suppressed. These indicators suggest that DOL is producing impartial studies. Future iterations of the capacity assessment could consider the degree to which there are protocols in place to maintain the objectivity of a study and that the Department or its agencies/offices did not influence the findings presented.

Independent Researchers. Most studies were authored by independent researchers (e.g., contractors, academics): 32 of 36 studies (88 percent). Only four studies were performed by an internal researcher, usually an employee of the agency that sponsored the research. By agency type, 1 of these studies was from an employment and training agency, 2 were from support agencies, and 1 was from a compliance agency.

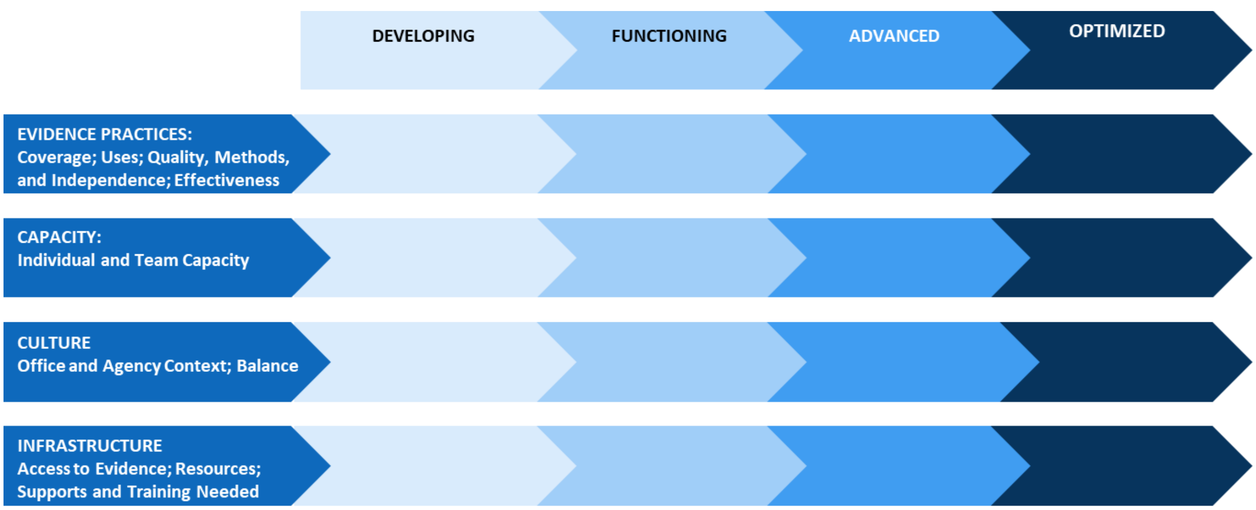
Presents Favorable and Unfavorable Results. All studies which we could reasonably judge appeared to present the full results of their analysis, whether favorable, null, or unfavorable. We do not have access to unpublished or unshared information; we simply considered whether the reports provided a spectrum of results.

APPENDIX F: EVIDENCE MATURITY MODEL FRAMEWORK

To facilitate iterative discussions with agencies, the research team is helping DOL develop a framework for an evidence maturity model suited to DOL's mission and context (Exhibit E-1). This model will serve as a basis for discussions with agency leadership and can also be customized to each agency's context. Organized by assessment domain with four levels of maturity (developing, functioning, advanced, and optimized) this framework was informed by three maturity models developed by other federal agencies.[19,20,21]

In developing the framework, the research team sought to align the requirements of the Evidence Act and the Department's strengths-based assessment approach, which seeks to highlight and build upon the strengths of each agency rather than solely focusing on limitations. For the final capacity assessment, the research team will create departmental descriptions for each level of the four framework domains informed by data collection activities. The research team, in collaboration with DOL, will draw broad conclusions about where the Department stands across the domains based on the findings from the capacity assessment. These conclusions can serve as a starting point for planning actionable next steps for improving capacity at the Department level. The completed departmental maturity model may also function as a tool to help agency leadership document their baseline capacity as well as plan and monitor progress appropriate to each agency's mission, operations, resources, and needs.

Exhibit F-1: Initial Maturity Model Framework for DOL Capacity Assessment



[19] Federal Government Data Maturity Model, https://my.usgs.gov/confluence/download/attachments/624464994/Federal%20Government%20Data%20Maturity%20Model.pdf

[20] Federal Records and Information Management (RIM) Program Maturity Model, https://www.archives.gov/files/records-mgmt/prmd/maturity-model-user-guide.pdf

[21] HRStat Maturity Model, https://www.opm.gov/policy-data-oversight/human-capital-management/hr-stat/hrstat-guidance.pdf

APPENDIX G: ACTIVITIES AND OPERATIONS OF THE DEPARTMENT UNDER EVALUATION OR ANALYSIS

List of ongoing evaluations or analysis as of March 1, 2021.

DOL Agencies	Project Title
All agencies	Evaluation of Behavioral Interventions in Labor Programs
All agencies	Administrative Data Research and Analysis Project (ADRA)
All agencies	Clearinghouse for Labor Evaluation and Research (CLEAR)
All agencies	DOL Capacity Assessment: An Evidence Act Requirement
All agencies	Administrative Data Research and Analysis Project (ADRA): Microsimulation open-source tool with FMLA and ACS Data
EBSA	2018 Form M-1 Bulletin
ETA	Analysis of TAA program data – infographics, community profiles
ETA	ETA Application Fees
ETA	UI Pandemic Unemployment Assistance Risk Modeling
ETA	Unemployment Insurance Work Search Policies and Practices
ETA	American Apprenticeship Initiative (AAI) Evaluation
ETA	America's Promise Job Driven Grant Program Evaluation
ETA	Youth Apprenticeship Readiness Grant Evaluation
ETA	Evaluation Design of Job Corps Experimental Center Cascades
ETA	Information on Subsidized and Transitional Employment Demonstration (STED) Paycheck Plus Site
ETA	Job Corps Evidence-Building Portfolio
ETA	National Agricultural Workers Survey (NAWS)
ETA	Pathways Home
ETA	Performance Partnership Pilots for Disconnected Youth (P3) National Evaluation
ETA	Ready to Work
ETA	Reemployment Services and Eligibility Assessment (RESEA) Research and Implementation Study
ETA	Reentry Employment Opportunities Evaluation
ETA	Labor Market Information on the Native American Population and Workforce
ETA	The Great Recession: Lessons Learned for the Unemployment Insurance System
ETA	Behavioral Interventions Evaluation and Design Options for Long-Term TAA Evaluation
ETA	Unemployment Insurance Deficit Financing study
ETA	WIOA Research Portfolio
ETA, WB	Apprenticeship Impact Evaluation
ETA, WB	State Capacity of Registered Apprenticeships
ETA, WB	Career Pathways Descriptive and Analytical Study

DOL Agencies	Project Title
ETA, WB	National Health Emergency Demonstration Grants to Address the Opioid Crisis: Implementation Evaluation
ETA, WB	Support to Communities Grant Evaluation
ETA, WB	TechHire and Strengthening Working Families Initiative (SWFI)
ILAB	Independent Interim Evaluation of "Supporting Respect for the Working Conditions of Workers in the Agro-Export Sector in Guatemala" Project
ILAB	Independent Interim Evaluation of the ILO Better Work Bangladesh (BWB) Project
ILAB	Independent Interim Evaluation of the ILO Better Work Jordan (BWJ) Project
ILAB	Outcome Evaluation of Project on Increasing Economic and Social Empowerment for Adolescent Girls and Vulnerable Women in Zambia
ILAB	Thematic Performance Evaluation of the Electronic Case Management System (ECMS) Components of 7 ILAB-Funded Projects
ILAB	Adwuma Pa Interim Evaluation
ILAB	Argentina Mult-Project Evaluation
ILAB	Bangladesh CLIMB Final Evaluation
ILAB	Columbia Avanza Final Evaluation
ILAB	FAIR Fish Interim Evaluation
ILAB	Findings on the Worst Forms of Child Labor (TDA)
ILAB	MAP 16 Interim Evaluation
ILAB	Promoting a Better Understanding of Indicators to Address Forced Labor and Labor Trafficking in Peru Interim Evaluation
ILAB	Safe-Seas Interim Evaluation
ILAB	Synthesis Review of OCFT-funded Cocoa and Fishing/Seafood Sector Projects
MSHA	Leveraging MSHA's data as a strategic asset
ODEP	Access to Paid Leave for Family and Medical Reasons Among Workers with Disabilities
ODEP	Center for Advancing Policy on Employment for Youth (CAPE-Youth)
ODEP	Employer Assistance and Resource Network (EARN) on Diversity Inclusion
ODEP	Employment for Persons with a Disability: Analysis of Trends During COVID-19 Pandemic
ODEP	Employment Impacts of COVID-19 on People with Disabilities
ODEP	Partnership for Inclusive Apprenticeship (PIA)
ODEP	Spotlight on Women with Disabilities
ODEP	Youth Supplemental Security Income (SSI) Research Competition
OFCCP	OFCCP Compliance Study
OSHA	Benefits of the OSHA On-Site Consultation Program: An Economic Analysis, Working Paper
OSHA	Employer Adoption of Voluntary Health and Safety Standards
OSHA	Mapping to evaluate OSHA activities

DOL Agencies	Project Title
OSHA	Update to worksheet (Form 33) used by OSHA's On-Site Consultation Program to evaluate employers' safety and health programs.
OWCP	Administrative Data Research and Analysis Project (ADRA): OWCP follow-up analysis on disability management policy change
OWCP	OWCP Annual Report to Congress
OWCP	Workers' Compensation and the Opioid Epidemic: Analysis and Research Design Options
VETS	VETS TAP Apprenticeship Pilot Evaluation
VETS	GrantSolutions
VETS	Homeless Veterans Reintegration Program ("HVRP") Evaluation
VETS	Intergovernmental Personnel Act (IPA) Appointments/Detailees
VETS	New VETS Staff for data analysis and research
VETS	TAP Study: TAP Employment Navigator Formative Evaluation
VETS	TAP Study: TAP Impact Evaluation
VETS, ODEP, ETA	VETS - VA VR&E Apprenticeship Pilot
WB	National Database of Childcare Costs
WB	Opportunity Cost of Caregiving Study
WHD, ETA	Evaluation Logistics and Technical Support