DRG MISSION USE OF EVIDENCE: LESSONS FROM EVIDENCE UTILIZATION IN USAID DRG PROGRAM DESIGN

Learning, Evaluation, and Research Activity II (LER II)

DECEMBER 2022

This DRG Mission Use of Evidence (MUSE) research was conducted by Mr. Robert Gerstein, consultant at The Cloudburst Group (Cloudburst); Ms. Aleta Starosta, Sr. Evaluation Specialist at Cloudburst; and Dr. Daniel Sabet of USAID's Center for Democracy, Human Rights, and Governance (DRG Center), with support from Ryan Hatano of Cloudburst. The research team would like to thank the USAID staff and implementing partners who shared their experiences and expertise with us.

This document was produced for review by the United States Agency for International Development, Center for Democracy, Human Rights, and Governance (DRG Center) under the Learning, Evaluation, and Research Activity II contract: GS10F0218U/7200AA18M00017.

COVER PHOTO: A young man from Bardales Bueso School in El Ingenio community, Ocotepeque, Honduras demonstrates his mural as part of the "Think Twice" campaign. USAID, through the Honduras Local Governance Activity, helped engage youth in community activities to support efforts to improve social inclusion, in an effort to reduce irregular migration. The mural reads, "If you think something is wrong, think twice. Don't be fooled. If you think it's too good to be true, don't be fooled. Think twice, think more." Photo credit: Honduras Local Governance Activity.

Prepared by:

The Cloudburst Group
8400 Corporate Drive, Suite 550
Landover, MD 20785-2238
Tel: 301-918-4400

CONTENTS

ACRONYMS IV

EXECUTIVE SUMMARY 1

1. INTRODUCTION AND RESEARCH QUESTIONS 4

2. FRAMING THE PROGRAM AND ACTIVITY DESIGN PROCESS 8

3. EVIDENCE UTILIZATION BY TYPE 10

4. FACTORS THAT DETERMINE THE USE OF RESEARCH EVIDENCE 19

5. WHAT ARE THE STRENGTHS AND WEAKNESSES OF HOW SELECTED ACTIVITY DESIGN TEAMS DETERMINE AND PLAN DRG RESEARCH, EVALUATION, AND LEARNING NEEDS AND PLANS? 29

6. HOW COULD RESEARCH EVIDENCE INTEGRATION AND EVALUATION PLANNING BE IMPROVED? 34

7. CONCLUSION 40

ANNEX 1. BIBLIOGRAPHY 42

ANNEX 2. KII LIST 44

ANNEX 3. SURVEY RESULTS 45

ANNEX 4. QUANTITATIVE SURVEY INSTRUMENT 54

ANNEX 5. RETROSPECTIVE QUESTION THEMES 63

ANNEX 6. RFP LANGUAGE REQUIRING OR ATTEMPTING TO REQUIRE EVIDENCE 64

ACRONYMS

ADS	Automated Directives System
CDCS	Country Development Cooperation Strategy
CRB	Contract Review Board
DRG	Democracy, Human Rights, and Governance
E&L	Evidence and learning
FSN	Foreign service national
FSO	Foreign service officer
GLAAS	Global Acquisition and Assistance System
GLP	Global Labor Program
IE	Impact evaluation
IP	Implementing partner
KII	Key informant interview
LER	Learning, Evaluation, and Research
MEL	Monitoring, evaluation, and learning
MRR	Microenterprise results reporting
NWA	Not withstanding authorities
OAA	Office of Acquisition and Assistance
OP	Operational plan
OU	Operating unit
PAD	Project appraisal document
PALT	Procurement action lead time
PDD	Project development document
PE	Performance evaluation
PEA	Political economy analysis
PPL	Bureau for Policy, Planning, and Learning
RFP	Request for proposals
RT	Research team
SMAZ	Strengthening Media for Accountability in Zimbabwe
SOAR	Senior obligation alignment review
SOW	Statement of work
TDY	Temporary duty assignment
USAID	United States Agency for International Development

EXECUTIVE SUMMARY

This study explores how the United States Agency for International Development (USAID) plans for and incorporates research evidence and evaluation into its democracy, human rights, and governance (DRG) programs. As such, the study focuses on both the use and the generation of evidence. On the use side, the goal of this study is to understand the extent to which research evidence is used to inform activity designs, identify obstacles to greater research evidence use, and draw recommendations and conclusions to better integrate research evidence into activity design. On the generation side, the study seeks to understand how activity design teams plan research, evaluation, and learning and ways in which USAID could improve evidence and evaluation planning. This research is intended to inform the DRG Center's strategy to promote the use of evidence and to improve evidence and evaluation planning in the field.

RESEARCH QUESTIONS

The study addresses the following research questions:

1. How is "evidence" understood by activity design teams and how are different forms of evidence used or not used in the design of selected USAID DRG interventions?
2. What factors support or constrain the application of more formal or research evidence in activity design?
3. How could more formal or research evidence be better integrated into project and activity design?
4. What are the strengths and weaknesses of how selected activity design teams determine and plan DRG research, evaluation, and learning needs and plans?
5. How could evidence and evaluation planning be improved?

METHODOLOGY: This investigation uses a mixed-methods design to answer the evidence utilization questions. The research team (RT) conducted a desk review, an online survey offered to previous DRG Center project participants, and key informant interviews (KIIs). The RT) analyzed quantitative survey and qualitative interview data, cross-referencing findings to enhance validity and mitigate the limitations of different data sources.

HOW EVIDENCE IS UNDERSTOOD: In this study, the RT finds that the terms "evidence" or "evidence-based" are understood differently by different USAID staff. This study hopes to increase clarity among USAID staff by distinguishing 1) the decision that the evidence aims to inform (e.g., strategy, diagnosis, prescription, refinement, and targeting) and 2) the type of evidence, which the RT divides into:

- Research evidence: Systematic research testing hypotheses, including impact evaluations (IEs), systematic reviews, and many academic studies.
- Contextual evidence: Research centered on a particular time and place, including most USAID assessments, political economy analyses (PEAs), and local data sources.
- Experiential evidence: Past and personal experience, including input from technical experts, lessons learned exercises, and some performance evaluations (PEs).

FACTORS THAT CONSTRAIN AND SUPPORT THE USE OF RESEARCH EVIDENCE: The use of research evidence is important to USAID's DRG staff; however, time constraints and concerns over

relevance often mean that contextual and experiential evidence are prioritized over research evidence. Additionally, interviewees indicate that research evidence synthesis is not often well timed to meet the needs of USAID's program cycle and activity design schedules. Staff report challenges in interpreting and applying research evidence to support activity design processes. Human resource limitations, including high Mission turnover, undermine a common understanding of the evidence and do not support organizational capacity to focus on research evidence utilization during the design process.

While the constraints are significant, several factors support the use of evidence upon which USAID and the DRG Center can be built. These include 1) an evidence-friendly legal and regulatory environment, 2) recent, substantial increases in the amount and accessibility of research evidence, 3) specialized DRG Center technical experts, 4) evidence champions throughout USAID, and 5) the potential accumulation of knowledge through Foreign Service National (FSN) staff and predecessor awards that generated research evidence.

EVALUATION PLANNING STRENGTHS AND WEAKNESSES: In addition to using evidence to inform the activity design process, USAID has an important role to play in generating research evidence. Planning for external evaluation is often overlooked during the activity design process. Partially as a result, the DRG sector does not meet evaluation spending targets laid out in the Automated Directives System (ADS) and frequently defaults to basic PEs—one-time studies typically involving three to four weeks of largely qualitative fieldwork. This study explores the dynamics that result in limited evaluation planning in activity design, including an input-heavy and time-constrained design process, exacerbated by limited role clarity between the USAID program and technical offices during activity design.

RECOMMENDATIONS: This study recommends that USAID attempt to make existing and future research evidence more accessible and to generate demand for research evidence by fostering a culture of learning. These recommendations build on existing DRG Center initiatives and also present new opportunities and strategic directions, including (when appropriate) ensuring that solicitation documents create an expectation that implementing partner (IP) proposals will include evidence—ideally research evidence—to support their proposed approach. Specifically, this study recommends the following actions that USAID should take to improve research evidence utilization and evaluation planning in its activity design process.

1. **Expand efforts to make evidence accessible:**

   1.1 The DRG Center should strengthen and continue existing Evidence and Learning (E&L) team efforts to make research evidence accessible to USAID's DRG cadre and Mission personnel.

   1.2 The DRG Center should expand efforts to also make research accessible to external audiences such as IPs, other organizations, and interested parties, including through a public-facing website.

   1.3 The DRG Center should ensure that the DRG Center's learning products and evaluations emphasize and strengthen dissemination.

   1.4 The DRG Center should expand the existing pilot evidence review initiative to an "evidence help desk," monitor its utility, and adapt its approach.

1.5 USAID's Bureau for Policy, Planning, and Learning (PPL) should promote the use of Mission monitoring, evaluation, and learning (MEL) platforms (e.g., learning support contracts/mechanisms at the Mission) to conduct evidence reviews as part of their scope.

1.6 The DRG Center should explore means and methods to bolster informal pairing with academics.

2. **Generate demand for research evidence and foster a culture of learning:**

2.1 The DRG Center should encourage its technical staff to serve as promoters of the role of research evidence and evaluation planning in activity design.

2.2 The E&L team should continue to strengthen its research evidence-related training delivery.

2.3 The DRG Center-produced sector-based programmatic guidance materials and training should highlight the value of research evidence.

2.4 Training and other DRG Center outreach efforts should aim to build a culture of learning and shift mindsets.

2.5 The DRG Center should work to establish a "vision of perfect," or an ideal situation of research evidence utilization and evaluation planning, and an action plan to move toward that vision.

2.6 Activity Designers should require (ideally research) evidence in the solicitation process to support any proposed approaches, consistent with existing regulations,

2.7 The DRG Center should continue to support Missions and regional bureaus in building a contextualized research evidentiary base, and Mission learning agendas should adopt this as a priority.

2.8 The DRG Center and the E&L team should continue to support IE and rigorous PE planning, and Missions should prioritize evaluation planning that moves beyond basic PEs.

The recommendations outlined above offer a two-pronged strategy to guide USAID in making it easier to incorporate and plan for evidence while seeking to increase demand for such efforts. The constraints to improved use and generation of evidence are not insurmountable, and with a focus on improving access and demand for research evidence, the RT is confident that USAID's DRG staff can better align real-world design work with the rhetorical and regulatory expectations of an evidence-based, learning organization.

1. INTRODUCTION AND RESEARCH QUESTIONS

There have been several legislative, regulatory, and policy changes requiring the use of evidence in programming (e.g., the Foundations for Evidence-Based Policymaking Act of 2018). The term "evidence" appears 74 times in the ADS Chapter 201 and the ADS repeatedly emphasizes the need to use evidence to inform all levels of decision-making. There is some data to suggest that USAID does a good job of using evidence in informing the design process. A recent non-representative sample of over 600 USAID staff found that most respondents self-reported the use of evidence to inform the design process. As many as 77 percent reported consulting with subject matter experts; 76 percent reviewed evaluations, analysis, monitoring data, and lessons learned from prior activities; and 65 percent consulted relevant data sets (see Table 1).[1] Nonetheless, a smaller percentage (57 percent) report the use of academic literature, and these measurements provide no measure of the intensity of use, potentially overstating the role of evidence. Anecdotally, it seems that a variety of factors, including bureaucratic requirements, personal preferences, and policy steers, risk playing a greater role than evidence.

TABLE 1: PERCENTAGE OF RESPONDENTS SELECTING EACH RESPONSE OPTION, BY OPERATING UNIT (N=637)

RESPONSE	TOTAL # (%) RESPONDENTS (N=637)	# (%) MISSION RESPONDENTS (N=355)
My team adapted or modified an activity design from a similar activity.	304 (47.7%)	179 (50.4%)
My team reviewed evaluations, analyses, monitoring data, and lessons learned from prior activities.	485 (76.1%)	280 (78.9%)
My team conducted site visits and/or otherwise consulted potential beneficiaries and local stakeholders.	395 (62%)	250 (70.4%)
My team reviewed data from relevant national or multinational datasets.	417 (65.5%)	257 (72.4%)
My team referred to findings from peer-reviewed publications, academic literature or publications, and/or think tank reports.	363 (57%)	196 (55.2%)
My team conducted or commissioned new analyses (e.g., geospatial studies, PEAs; cost-benefit or cost-effectiveness analyses, gender analysis, inclusive development analysis, etc.).	341 (53.5%)	204 (57.5%)
My team used evidence or analysis cited in country/regional strategy or project documents.	381 (59.8%)	229 (64.5%)

[1] IT Shows Inc., Deloitte Consulting. (June 2021). USAID Capacity Assessment for Evidence Management and Use: Annexes. USAID.

RESPONSE	TOTAL # (%) RESPONDENTS (N=637)	# (%) MISSION RESPONDENTS (N=355)
My team used toolkits and templates available within or produced by USAID.	350 (54.9%)	183 (51.5%)
My team consulted with USAID subject matter experts.	492 (77.2%)	275 (77.5%)
My team consulted with subject matter experts external to USAID.	351 (55.1%)	191 (53.8%)

As such, this study aims to dig deeper into the role of evidence, particularly research evidence, in the DRG activity design process and to assess the role of evidence in comparison with other factors and demands. The study focuses on both the use and the generation of evidence. On the use side, the goal of this study is to understand how evidence is or is not used to inform activity designs, the obstacles to greater evidence use, and how evidence could be better integrated into activity design. On the generation side, the study seeks to understand how activity design teams plan research, evaluation, and learning and to identify ways in which evidence and evaluation planning could be improved. This research intends to inform the DRG Center's strategy to promote the use of evidence and to improve evidence and evaluation planning in the field.

RESEARCH QUESTIONS

The study aims to address the following questions:

1. How is "evidence" understood by activity design teams and how are different forms of evidence used or not used in the design of selected USAID DRG interventions?
2. What factors support or constrain the application of more formal or research evidence in activity design?
3. How could more formal or research evidence be better integrated into project and activity design?
4. What are the strengths and weaknesses of how selected activity design teams determine and plan DRG research, evaluation, and learning needs and plans?
5. How could evidence and evaluation planning be improved?

METHODOLOGY

This investigation uses a mixed-methods design to answer the evidence utilization questions. The design entails a broad study of evidence utilization in DRG activities and a deeper dive into 12 follow-up DRG activities. The team conducted a desk review, an online survey offered to previous DRG Center project participants, and KIIs. The RT analyzed quantitative survey and qualitative interview data, cross-referencing findings to enhance validity and mitigate the limitations of different data sources.

Desk Review: As part of the background research, the team conducted a desk review of relevant policy and background documents The team also reviewed and coded 39 requests for proposals (RFPs) of DRG

solicitations released between 2018 and 2020 for expectations of evidence use.[2] This included USAID's use of evidence in the solicitation document, an explicit requirement that bidders include evidence in response to a solicitation, and evidence of evaluation planning.

Initial KIIs and Group Discussions (N=22): The authors conducted KIIs and group discussions from November 4–December 22, 2021. Key informants include current and former DRG Center staff and PPL staff. In addition, the RT sought the perspectives of principal investigators, DRG Center learning partners, relevant USAID Mission staff, and IP staff. Accounting for both individual interviews and group discussions, the qualitative data include perspectives shared by 22 individuals in 16 interviews. Please refer to Annex 2. KII List for a summary table of the number of interviews and interviewees by respondent type.

Online Survey (N=83): The RT conducted a quantitative survey with the USAID DRG cadre and support staff that had participated in recent activity designs. The team used ForeginAssistance.gov to identify an initial sample frame of DRG activities commissioned by Missions and USAID/Washington within the last four years. The RT then asked DRG team leads and their colleagues in corresponding Missions to identify technical, program, and contract officers that participated in each of the activity designs. The RT then contacted these individuals to participate in an online survey. Distributed via Google Forms between February 1 and February 18, 2022, the survey included questions about the role of different factors in the design process such as evidence, stakeholder input, and leadership direction. To incentivize responses, the RT offered operating units (OUs) with the highest percent of responses a free "research evidence review" on a topic of their choosing to inform an upcoming design worth up to $12,000.

TABLE 2: DEMOGRAPHIC BREAKDOWN OF MISSION/BUREAU/INDEPENDENT OFFICE SURVEY

GENDER	**Male**		**Female**		**Declined to State**	
	46% (41)		51% (45)		3% (3)	
ACTIVITY DESIGN EXPERIENCE	**1–5 Years**		**6–10 Years**		**10+ Years**	
	26% (23)		31% (27)		43% (38)	
HIRING MECHANISM	**Foreign Service National**		**Foreign Service Office**		**Personal Service Contractor/ Other**	
	57% (51)		28% (25)		12% (12)	
ROLE	**Agreement Officer's Representatives/ Contracting Officer's Representatives**	**DC-Based Support**	**Design Specialist**	**Democracy and Governance Specialist, Team Leader, or Mission-Based Expert**	**Office Director/ Deputy Office Director**	**Office of Acquisition and Assistance (OAA) Agreement Officer/Contracting Officer or Specialist**
	45.5% (40)	6.8% (6)	5.7% (5)	13.6% (12)	20.5% (18)	8% (7)

[2] The identification of these RFPs was carried out by Bryce Watson and Jonathan Rose as part of a study on PEAs in RFPs. Rose, J., and Watson, B. (Forthcoming). The Use of Political Economy Analysis Among USAID Implementing Partners. Thinking and Working Politically Community of Practice.

Of 356 individuals invited to respond to the survey, 83 participated, yielding a response rate of 23 percent. The RT received at least one response for 77 of the 173 activities for which respondents were recruited. Responses covered activities in 53 countries. Table 2 below breaks down quantitative survey respondents by stakeholder group. Annex 3. Survey Results includes a summary of the quantitative data results.

Follow-Up Interviews (N=12): The authors conducted KIIs and group discussions on 12 activity designs presented in Table 3 from March 4–May 22, 2022. The RT selected follow-up interviews based on responses to the quantitative survey and, in three cases, on initial KIIs. The RT intended for follow-up interviews to be selected based on 1) positive deviance, or activity designs that were reported to involve research evidence or research evidence generation; 2) negative deviance, or activity designs that did not use research evidence despite the availability of such evidence; and 3) requirements for the use of evidence in the solicitation. However, due to nonresponse and lack of evidence requirements in the solicitations, the RT accomplished a more limited spectrum of interviews that focused primarily on positive deviants and included three negative deviants.

TABLE 3: SELECTED FOLLOW-UP ACTIVITIES AND PROGRAMS

MISSION/BUREAU/INDEPENDENT OFFICE	ACTIVITY
USAID/DRG Center	Global Labor Program (GLP)
USAID/Indonesia	TOLERANSI (Indonesia Religious Freedom)
USAID/Philippines	Cities for Enhanced Accountability, Governance and Engagement (CHANGE)
USAID/Mexico	ConJustica
USAID/Mexico	Mexico Anti-Corruption Reforms Activity (PRO INTEGRITY)
USAID/Mali	Civic Engagement Program
USAID/Asia Bureau	Asia Religious and Ethnic Freedom
USAID/Malawi	Strengthening Parliament's Role in Malawi's Development
USAID/South Africa	Strengthening Local Government to Improve Gender-Based Violence Response
USAID/Zimbabwe	Strengthening Media for Accountability in Zimbabwe (SMAZ)
USAID/Rwanda	Dufatanye Urumuri Reconciliation Project
USAID/Liberia	Elections and Democracy Activity

IP Interviews (N=2): To better understand how IPs contribute to the use of evidence in the pre-solicitation and solicitation stage of program design, the RT contacted the IPs from six of the positive deviance follow-up interviews, ultimately speaking with only two of them. As such, the IP perspective is underrepresented in the study. The RT asked IPs questions regarding their perspectives on donor expectations, evidence-

based approaches, and how the IPs themselves integrate research evidence into their decision-making and business development.

Data Analysis: The RT used multiple techniques to analyze the data. For the survey data, the team examined descriptive statistics (e.g., means, crosstabs) to obtain statistical profiles of the samples. The team hoped to explore differences by stakeholder group, but the small sample size made this exploration less meaningful. For the qualitative data analysis, the team identified broad themes—both deductively, based on the evidence utilization questions, and inductively, based on interviews—and organized qualitative data by these themes in a spreadsheet. The RT derived findings by comparing data sources within each theme.

Risks and Limitations: There are a few methodological limitations worth noting. In both the survey and the interviews, there is a risk of selection bias. Respondents who complete the survey or agree to conduct the interview are more likely to have a strong opinion or experience with evidence utilization. To mitigate this risk in the quantitative survey, the team sent out personalized emails to respondents asking them to take the survey, followed up three times, and offered an incentive to the Mission or OU with the highest response rate. To mitigate this risk in the follow-up interviews and KIIs, the team conducted repeated personalized follow-ups with identified respondents. However, the response rate for all types of data collection was lower than expected, making selection bias more relevant.

Another limitation was deviation from the proposed follow-up interview selection plan, which was intended to be more robust at the onset of the research and was later revised to better meet the profile of responses that emerged from the online survey.

Other limitations include:

- Deviations from the follow-up interview selection plan (no response on negative deviance, reporting errors in the survey about the use of evidence being required in the RFP).
- Misreporting/misunderstanding types of evidence in the quantitative survey.
- A higher prioritization of research questions on the usage of research evidence over questions on evaluation planning.
- A focus on pre-solicitation, with a limited examination of how evidence is used after award.
- Given recent changes to the ADS to reduce mandatory elements of USAID projects (including their onerous design process), this study did not focus specific lines of inquiry on understanding specifications around project design. Instead, the study took a narrow focus on activity design.

2. FRAMING THE PROGRAM AND ACTIVITY DESIGN PROCESS

At Missions, activities contribute to the development result(s) described in each Mission's Country Development Cooperation Strategy (CDCS) or similar country-level strategy. Activities should also contribute to project results, where applicable.[3] The activity design process is conducted by Missions on a rolling and ad hoc basis as determined by the needs of each Mission and project. Activity design takes

[3] Recent revisions to the ADS made substantial changes to the nature and utility of "projects" at USAID. Previously, projects were mandatory meso-strategies comprised of multiple activities. Substantial policy revisions in 2021 largely removed this meso-strategy layer.

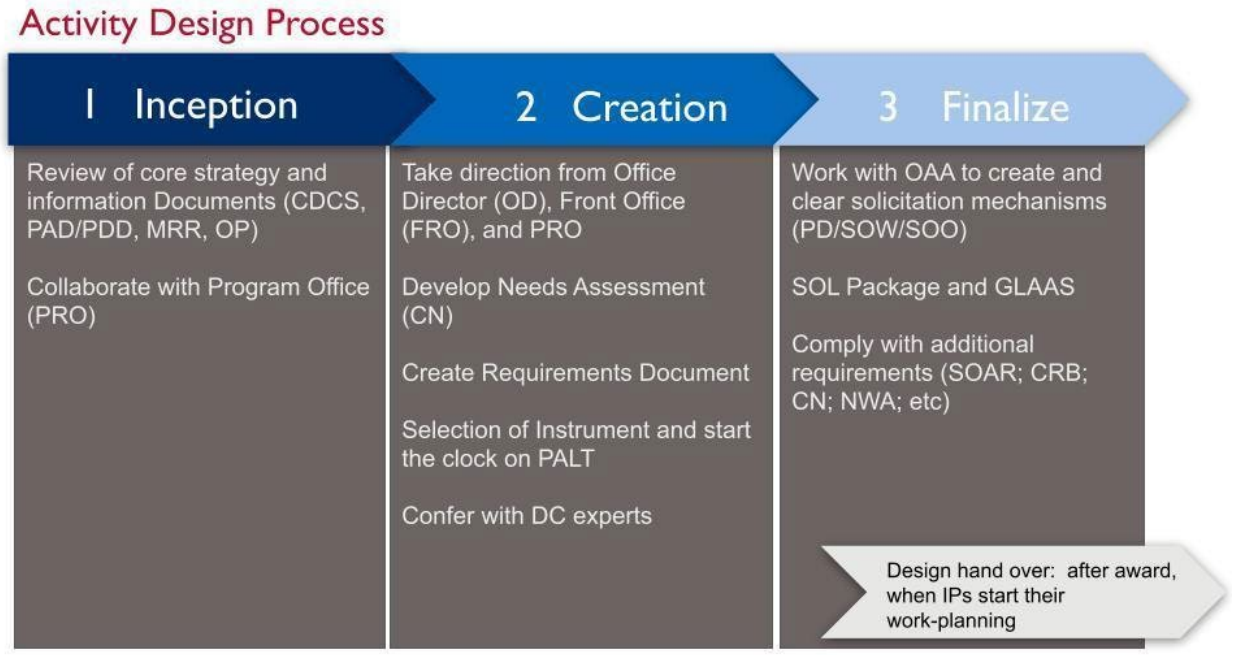
several months to complete, and, depending on the nature of the award, it may take upwards of one to two years.

This study used program cycle elements defined in the ADS, which details USAID's functions, policies, and procedures as well as experiential evidence from internal team members and research participants to outline and validate a summarized activity design process map. The summary process is described in three key phases laid out in Figure 1: inception, creation, and finalization. The process map is aligned with USAID's program cycle but does not strictly adhere to the categorized phases pursuant to ADS 201. The program cycle is USAID's operational model for planning, delivering, assessing, and adapting development programming in a given region or country to advance U.S. foreign policy. The program cycle is codified in ADS 201, which provides the framework from which USAID and its Missions abroad derive procedures for making strategic decisions; focus associated resources; design supportive projects and/or activities to implement these strategic plans; and learn from performance monitoring, evaluation, and other research evidence. ADS 201 provides direction on how and when to use evidence and evaluations to inform USAID's programming.[4]

Multiple ADS chapters describe USAID's approach to activity design from policy, program cycle, and procurement requirements perspectives. An "activity" generally refers to an implementing mechanism that carries out an intervention or set of interventions to advance identified development result(s) in a given country or region. Activities include a wide range of implementing mechanisms, including contracts, cooperative agreements, direct agreements with partner governments, and other mechanisms. Activities also include buy-ins under global agreements (e.g., field support agreements) that generate programmatic results in a specified country or region. Missions and Washington OUs often complement activities with actions undertaken directly by USAID staff such as policy dialogue, stakeholder coordination, or capacity-building.

[4] See ADS 201.3.6.5 Evaluation Requirements; 201.3.2.17 Evaluation During CDCS Implementation; and sections 201.3.6.7 and 201.3.6.8, which also describe evaluation requirements. Additionally, ADS sections that provide evaluation guidance include ADS 201.3.6.2 Evaluation Principles and Standards; 201.3.6.3 Missions and Washington Operating Unit Roles in Evaluation; 201.3.6.4 Types of USAID Evaluations; and the mandatory reference ADS 201saf, Evaluation Triggers.

Figure 1. USAID Activity Design Process



While the ADS is generally referred to and accepted as the statutory process that should be followed, the RT identified during early KIIs that the steps described in the ADS are not interpreted in a strictly linear or static process. Based on Mission orders, operational culture, time constraints, and other practical conditions, the activity design process tends to be somewhat fluid and less standardized (in fact, PPL made recent updates to ADS 201 to reflect this operational dynamic. More in section 4 below). Based on these initial insights, the RT devised a simplified, thought-aligned, activity design process map (above). Using this framework, the study identified when and how activity designers plan for and incorporate evidence into the activity design process. These findings are discussed and analyzed in detail throughout the remainder of this report.

3. EVIDENCE UTILIZATION BY TYPE

How is "evidence" understood by activity design teams and how are different forms of evidence used or not used in the design of selected USAID DRG interventions?

> *'Evidence' doesn't mean 'proof'—it helps to indicate which way we should go (CS 05).*

The terms "evidence" and "evidence-based" mean very different things to different people. For some, "evidence" refers exclusively to experimental evidence produced by IEs; for others, "evidence" is any form of knowledge. USAID's own definition of evidence in Figure 2 is fairly broad and encompasses a wide range of information and sources.

Figure 2. ADS 201.6 Definition of "Evidence"

> "Evidence" is the body of facts or information that serves as the basis for programmatic and strategic decision making in the program cycle. Evidence can be derived from experiential knowledge, assessments, analyses, performance monitoring, evaluations, research, and statistical activities. It can be sourced from within USAID or externally. Evidence should result from systematic and analytic methodologies or from observations that are shared and analyzed. There are four interdependent components of evidence: foundational fact finding and research, policy analysis, program evaluation, and performance measurement. Evidence can be quantitative or qualitative and may come from a variety of sources. Evidence has varying degrees of credibility, and the strongest evidence generally comes from a portfolio of high quality, credible sources rather than a single study.

Evidence can be divided into research, contextual, and experiential evidence with overlap between categories. To clarify the broad concept of "evidence," the RT developed a typology of the three primary types of knowledge that are considered evidence based on the ADS definition and initial interviews with USAID staff. They are:

1. Research evidence, which the RT defines as systematic and analytic methodologies designed to test hypotheses and includes IEs, systematic reviews, and academic studies testing hypotheses.[5]

2. Contextual evidence, which the RT defines as generally non-experimental evidence not necessarily designed to test a specific hypothesis but rather data that centers knowledge unique to a particular time and place. This includes assessments such as PEA or sector-specific assessments, local sources of data, such as government datasets, and some PEs.[6]

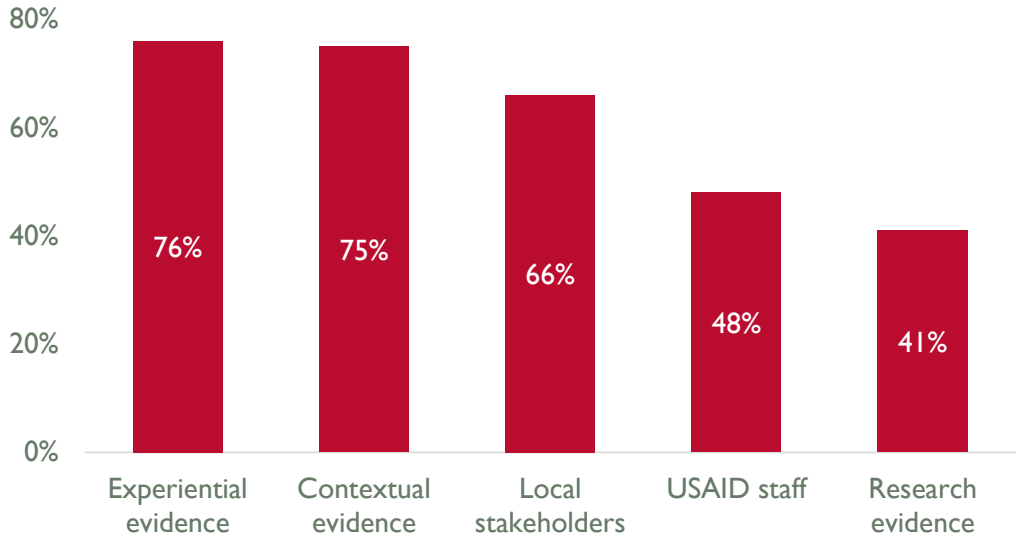
3. Experiential evidence, which the RT defines as informal technical expertise, personal experience, and knowledge. This includes input from technical experts; the experience of USAID staff, IPs, and local experts; and lessons learned exercises and PEs focused on lessons learned.

In the quantitative survey, the RT asked USAID planners and supporting design team staff what factors most influenced activity design, including these three types of evidence, with corresponding definitions. The results, shown in Figure 3, illustrate that experiential and contextual evidence are the most significant influences, followed by input from stakeholders and USAID staff. Research evidence is the least significant (though still relatively common) factor, identified by just over 40 percent of respondents.

[5] See also the USAID Scientific Research Policy, which notes that research is typically hypothesis-driven, testable, and independently replicable (pg. 6).

[6] With all three of these categories, the borders can be somewhat fuzzy. For example, rigorous research evidence conducted in the same context as the activity design would bridge the boundary between the two and could be considered context-specific research evidence. Similarly, there is context- and non-context-specific experiential evidence.

Figure 3. Percentage of Respondents Listing Each Source as a Significant Influence on the Activity Design (n=88)



Evidence should be employed across five activity design decision points. In addition to the type of evidence, the RT also finds variation in when in the design process and to what end evidence is used. Interviews suggest five major decision points where evidence is employed. Evidence can help:

- *Strategize:* Identify the DRG problem USAID should address.
- *Diagnose:* Define the problem and its drivers.
- *Prescribe:* Determine the best programmatic approach to address the problem and its drivers.
- *Inform:* Refine that programmatic approach based on the evidence and for the context and population.
- *Target:* Identify where or on what populations the approach should target and focus.

Given these many potential decision points where evidence could or could not be used to inform decision-making, and given the three types of evidence discussed, the role of evidence varies considerably across designs. The following section ecxplores each type of evidence and how it is understood and utilized by USAID, particularly DRG field officers and program designers, to inform these different decision points.

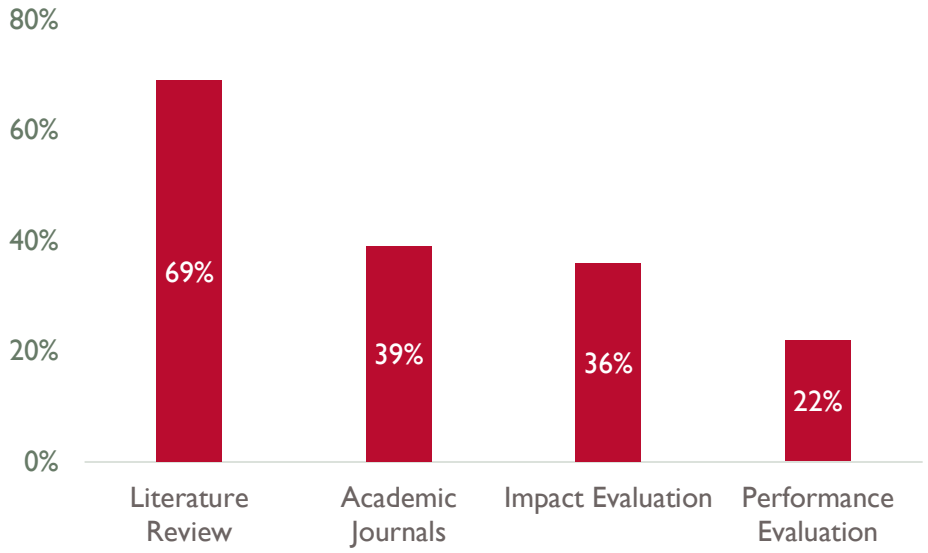
RESEARCH EVIDENCE

Research evidence is the least common form of evidence informing USAID DRG designs. Research evidence—including IEs, systematic reviews, and academic studies testing hypotheses—can be used across the different decision points but is most salient in prescribing and informing; this is to say, answering the "what works" question in selecting and refining a programmatic approach with a proven track record. Where evidence does not exist and in the case of new and innovative pilot programming, USAID has the option (and in some cases, the requirement) to generate its own research evidence through an IE. According to USAID's 2014 Scientific Research Policy:

Impact evaluations and research can form a virtuous cycle: Research priorities help formulate and refine impact evaluation questions so that these can advance the state of knowledge around a particular subject. In turn, impact evaluations ground-truth research findings: they test innovative strategies and approaches in a real-world setting before they are scaled up with USAID funding, and in doing so, reveal new areas of research to be explored.[7]

Despite the potential that research evidence offers, it is the least utilized of the three sources of evidence. Figure 3 above shows that 41 percent of survey respondents claimed to incorporate research evidence into their activity design, and this is likely somewhat overstated.[8] As shown in Figure 4 below, the most common types of research evidence used by these respondents include literature reviews (69 percent of those reporting using research evidence) and academic journals (39 percent). Respondents also used IEs and PEs, but with less frequency. Later sections of this report will explore the various challenges and roadblocks to greater research evidence uptake.

Figure 4. Type of Research Used Among Those Reporting Using Research Evidence (n=36)



Follow-up interviews suggest that research evidence is most useful when USAID is designing new programs without strong priors about what type of intervention is likely to be successful. For example, the DRG Center's Global Labor Program (GLP) design team had an opportunity after a shift in Congressional policies to drastically change the direction of programming for the first time in 60 years.

[7] IEs are also required for any new, untested approach that may be scaled. The ADS 201.3.6.2 Requirement 3 states "Each Mission and Washington OU must conduct an IE, if feasible, of any new, untested approach that is anticipated to be expanded in scale or scope through U.S. Government foreign assistance or other funding sources (i.e., a pilot intervention). OUs should identify pilot interventions during project or activity design and should integrate the IE into their design. If it is not feasible to effectively undertake an IE, the Mission or Washington OU must conduct a performance evaluation and document why an IE was not feasible."

[8] Follow-up interviews suggest this number is likely an overstatement that included contextual evidence that was categorized as research evidence.

If you are going to make a shift, you have to know what you are shifting to (CS 01).

To best take advantage of this window, the GLP team commissioned several pieces of formal research evidence to better understand the more pressing issues currently facing global labor. The commissioned research included an evidence review about global labor collective action; an evidence review on lessons learned from future work and transnational campaigns on migrant labor; an evidence review on adaptations to the platform economy in development countries; a secondary evidence review on the latest experimental/quasi-experimental research on the impact of GLP interventions on core labor standards and the quality of work in terms of formality, employment terms and conditions, wages (e.g., amount over time and stability), and amount of employment; as well as a performance review of the then-current GLP implemented by the Solidarity Center. Leading global labor experts in academia conducted research in partnership with USAID learning partners. As a result of the research evidence, the GLP team released a broad agency announcement that would invite new approaches beyond what had been done for the past sixty years, focused on the role of technology and the gig economy.

Evidence helps ask questions (CS 01).

The GLP team chose the broad agency announcement because it would invite new approaches and new partners to the table, and the research evidence was able to further inform the co-creation process. The research evidence was also used to defend programming choices to Congress.

Nonetheless, as will be discussed in the next question response, other designers were more skeptical about the relevance of research evidence and IEs in particular, citing concerns about generalizability, timeliness, and responsiveness.

We don't use a lot of research. I'll be frank. When Missions link in studies or reports, I tend to regard them as 'academic.' There is a limit to the practicality of the questions. IE is evaluating work after it has happened. The people evaluating are not there in the political dynamics and the dynamics of the country. It is not the same moment in time when we are designing the programming. Something published two years ago might be obsolete the moment it is published… (KII 01).

CONTEXTUAL EVIDENCE

As development practitioners, we need to learn from the context, from the people involved—something that a snapshot in time is not going to get to… (KII 10).

Respondents note that contextual evidence is well integrated into USAID DRG designs. The survey and KII respondents were clear that contextual evidence, particularly PEAs and sector assessments, was a well understood and commonly used type of evidence. As noted in Figure 3 above, 75 percent of survey respondents cited contextual evidence as having a significant influence on survey design. Contextual evidence is appealing because it is viewed as timely, relevant, and relatively affordable compared to research and experiential evidence. As one Democracy and Governance officer told the RT, "We are really good at the contextual analysis"; another confirmed that "PEA is something that we breathe in and out." Contextual evidence is "particularly valuable at the beginning" (IP 02) when critical design choices are being finalized.

The DRG Center has developed several assessment frameworks to guide assessment work. These are largely qualitative exercises conducted by an internal, external, or hybrid team over a three- to four-week period. They include 1) the general DRG systematic assessment framework; 2) other more specific frameworks on civil society, media, countering authoritarianism, women's participation and leadership, human rights, and elections; and 3) a series of primers and handbooks with assessment guidance (e.g., public financial management, misinformation, decentralization, civic education). In one of the case studies in Zimbabwe, the Mission commissioned a formal media assessment to inform the design process (see Box 1). In other cases, the assessment approach is more informal. Interviewed DRG Center technical experts often travel to Missions for a three-week temporary duty assignment (TDY) where they conduct context-focused assessment work as part of a design process. Furthermore, IPs report a more intensive context-focused assessment process as part of their proposal development process.

PEAs deserve particular attention as a common assessment approach. PEAs are at times done as part of the USAID design process, but they are more often done as part of an activity. A recent study of USAID RFPs (i.e., contracts and not assistance) found that 19 of 45 DRG-focused RFPs asked IPs to conduct a PEA.[9] In the sample of case studies, both the Mexico justice and Zimbabwe media activities asked IPs to conduct a PEA as a formative part of the activity. At times, IP-produced contextual evidence is a formative part of the activity—e.g., conducting a programmatic landscape analysis in the first year of the program, as was the case for a USAID/Asia program—though requiring the activity to produce any type of evidence, contextual or otherwise, in the solicitation is rare.

[9] Rose, J., and Watson, B. (Forthcoming). The Use of Political Economy Analysis Among USAID Implementing Partners. Thinking and Working Politically Community of Practice.

Strengthening Media for Accountability in Zimbabwe (SMAZ)

When designing a new media program to strengthen watchdog journalism and media reporting for accountability, USAID/Zimbabwe relied on contextual evidence to inform the program's proof of concept. According to key informants, Zimbabwe is a challenging environment for development work, particularly media, and the Mission strove to balance being innovative and trying new approaches with deploying proven approaches. Contextual evidence is particularly desirable because "What happened three years ago isn't what is happening today" (CS 09), and there is a need to be able to determine what has or has not changed from one program design to the next.

To understand the evolving media landscape, the Mission commissioned a media sector assessment that would inform USAID about the latest developments in Zimbabwe's media sector and its place in ensuring accountability. The assessment relied on desk research and qualitative interviews with more than 70 stakeholders in the media, government, civil society organizations, and donor sectors to answer four questions determined by the Mission. The assessment included recommendations for future programming and an analysis of gender and youth dynamics in the media sector. In addition to the media assessment, the program design team also read academic journals and learned from other USAID Missions that had done a similar intervention.

The assessment was used by both USAID and the future IP when designing the new program: "We relied heavily on the findings of the assessment…The media assessment showed gaps and entry points—it allowed SMAZ to be really comprehensive" (CS 09). The same respondent also mentioned that the assessment findings were particularly easy to integrate into the program design because "the research was already packaged for you," making it faster and easier to utilize than diving into literature (CS 09). The design ultimately led to the SMAZ program, which was awarded in 2021.

EXPERIENTIAL EVIDENCE

> *I bring with me my own experience; I have 30 years of experience with USAID in the DRG sector, including the design and management of DRG activities in other countries, too (CS 03).*

Experiential evidence is the most common evidence reported in USAID DRG designs and, while valuable, it has a higher risk of bias. Experiential evidence—including technical experts, past experience, lessons learned, and less rigorous PEs—is the most common form of evidence reported in the survey informing program design. Over three-quarters of survey respondents reported utilizing experiential evidence in their program designs.

> *We are often not looking to generate new knowledge, but trying to help people understand the experiences in this place (KII 05).*

One of the most important sources of experiential evidence is the USAID FSNs' experiences. As one respondent put it,

Experiential and contextual evidence are used most, and most relevant. We rely on FSN experience and past projects…Do we need to rely more on experiential insight of FSNs? One hundred percent yes. [It's] hard [for USAID] to ask tough questions and get honest answers when talking to [non-governmental organizations] and civil society (KII 01).

Personal experience, whether from FSNs or other USAID staff, has its limitations. One foreign service officer (FSO) commented that FSN knowledge is critical but disadvantaged since it can be siloed or myopic (KII 02). Another respondent stressed that design and evaluation planning is very much based on personalities and their comfort and experience. Many KIIs recognized the importance of designing programs based on more than just the staff's own anecdotal experience. A wide range of studies have recognized that personal experience suffers from several biases stemming from limits to the information that individuals have, how information is filtered, and how one cognitively processes information.[10] From this perspective, personal experience is best used in triangulation with contextual and research evidence, as opposed to being the sole source of evidence-driving design.[11]

USAID also relies on the experience and lessons learned from IPs.

IP proposals are based on experiential learning. You are always getting 'our organization has done X, Y, Z in this country and that is why we think this will work' (KII 04).

Experiential evidence is baked into the annual reporting process and is well understood by both USAID staff and IPs. One USAID interviewee reported requiring a "lessons learned" section in reporting and referring back to these sections as input to follow-on program design (CS 08). This type of evidence is most relevant for long-term DRG programming efforts that rely on similar approaches year after year, such as providing technical support for elections commissions. However, IPs are not always incentivized to honestly discuss what is not working and learn from past mistakes for fear of jeopardizing their reputation. One respondent offered that local partners are more willing to admit when things are not working, as they have a greater incentive to resolve the DRG challenge; they live in the country and need to make it work.

HOW EVIDENCE TYPES WORK TOGETHER

Designs that triangulate among different sources of evidence are viewed as the most successful. KIIs, follow-up interviews, and quantitative survey responses stressed that all three types of evidence have a role in program design. Respondents mentioned combining evidence types during the design phase. For example, when designing the latest iteration of the Countering Trafficking in Persons

[10] See for example, Soyer, E., & Hogarth, R.M. (2015). Fooled by Experience. Harvard Business Review. May, 72–77.
[11] It should be noted that context/ongoing PEA and FSN experiential evidence often go hand in hand to some degree. There is natural reciprocity in many cases, and this helps to strengthen the appeal of these approaches to Mission-based personnel, especially when the alternative is robust research that, on face value, is time-consuming, expensive, and not necessarily context-aware.

program in Cambodia, the program design team incorporated findings from a completed IE, a literature review of academic and gray literature, PEAs, stakeholder consultations, reporting from the previous project, and the expertise of FSNs on the design team to develop the follow-on program. From the IE findings, the team decided to focus on savings group-related interventions; from contextual research, the team determined it would be appropriate to work with the Government of Cambodia on the topic; from the literature review, previous program reports, and past personal experience, the team determined the types of approaches that were most likely to be appropriate and successful.

USAID/South Africa shared a similar experience of combining research evidence with contextual assessments and experiential learning to inform their response to sexual violence in the country. The Mission funded a rigorous research analysis of rape cases over a six-year period that informed the definition of the problem, identification of drivers, and populations to target across different localities. Contextual assessments and documented lessons learned from previous projects showed the limitations of the Mission's central approach of supporting government-supported rape crisis centers. They also identified opportunities for a change of approach in the follow-on, which was eventually pursued.

Table 4 shows a summary of advantages, disadvantages, and best usages for each of the three types of evidence.

TABLE 4: ADVANTAGES, DISADVANTAGES, AND BEST USAGES OF EVIDENCE

EVIDENCE TYPE	ADVANTAGES	DISADVANTAGES	BEST USAGE	DATA QUALITY STANDARDS
Research (e.g., IEs, studies testing hypotheses)	Best source of evidence for knowing what works. Less prone to biases. Lends legitimacy and can help sell ideas to outside audiences.	Requires technical knowledge to interpret. Can be slow to collect and analyze. Can be costly. Can be difficult to generalize findings to new contexts. May be prone to "official" bias/obstruction, especially when subject to host country review boards/regulators.	To identify drivers of a problem and inform approach selection and refinement. Defending subject matter and related needs/gaps to external parties. Measuring in more predictable operative environments.	**Validity:** High **Reliability:** High **Precision:** High **Integrity:** High **Timeliness**: Variable **Relevance:** Variable
Contextual (e.g., PEA, assessments)	Relevant to a specific time and place. Accessible for USAID staff and IPs to collect and learn from.	Less rigorous than research evidence. Cannot test causality and answer the "what works" question.	To define a problem and inform targeting and refinement. When a context is changing rapidly.	**Validity:** Medium **Reliability:** Medium **Precision:** Variable **Integrity:** High **Timeliness:** High **Relevance:** High

EVIDENCE TYPE	ADVANTAGES	DISADVANTAGES	BEST USAGE	DATA QUALITY STANDARDS
Experiential (e.g., IP experience, lessons learned)	Easily available and accessible by USAID staff and IPs. Grounded in the realities of day-to-day work.	Least rigorous. Most prone to biases and errors. Can promote the inertia of past programming.	To inform approach refinement, implementation, and adaptation. To complement research and contextual evidence. For programming on topics in which USAID/IPs have considerable experience. When research and contextual evidence are not available.	**Validity:** Low **Reliability:** Low/Medium **Precision:** Low **Integrity:** Low **Timeliness:** High **Relevance:** Variable

The remainder of the report will focus on research evidence, including the factors that determine the use of research evidence, how to better integrate research evidence into program design, and how to improve the evidence and evaluation planning process.

4. FACTORS THAT DETERMINE THE USE OF RESEARCH EVIDENCE

What factors support or constrain the application of more formal or research evidence in activity design?

Much of the literature on enabling factors and barriers to evidence utilization in activity design centers on the scarcity of effective research translation, perpetuated by the lack of collaboration between researchers and practitioners.[12] To rectify this divide, donors have concentrated resources on initiatives designed to improve researchers' capacity to distill and disseminate evidence for practitioners.[13] However, recent studies suggest that by concentrating on the supply of research dissemination, donors have failed to adequately incentivize the demand for evidence among practitioners.[14] These scholars identify the culture of "evidence complacency," in which evidence is not sought out by practitioners, as a primary barrier to the application of evidence in activity design.[15]

[12] Nutley, S., Davies, H., & Walter, I. (2002). Evidence-based policy and practice: Cross-sector lessons from the UK. Swindon: ESRC UK Centre for Evidence-Based Policy and Practice.
[13] Lavis et al, 2003; Milne, B. J., Lay-Yee, R., McLay, J., Tobias, M., Tuohy, P., Armstrong, A., Davis, P. (2014). A collaborative approach to bridging the research-policy gap through the development of policy advice software. Evidence & Policy: *A Journal of Research, Debate and Practice, 10*(1), 127–136; Nutley, S., Davies, H. T. O., & Walter, I. (2007). Using evidence: How research can improve public services.
[14] Dubois, N. S., Gomez, A., Carlson, S., & Russell, D. (2020). Bridging the research-implementation gap requires engagement from practitioners. Conservation Science and Practice, 2(1), e134; Stewart, R. (2015). A theory of change for capacity building for the use of research evidence by decision-makers in southern Africa. *Evidence & Policy, 11*(4), 547–557.
[15] Sutherland, W. J., & Wordley, C. F. (2017). Evidence complacency hampers conservation. *Nature Ecology & Evolution, 1*(9), 1215–1216.

Building such demand is challenging. Other research has affirmed that "awareness-raising interventions alone have little evidence of success in improving the use of evidence, but building the skills and motivation to use evidence showed promise."[16] Systems designed to promote evidence-based decision-making that fail to center practitioners "risks losing legitimacy and relevance to the very institutions and individuals it aims to serve—the decision-maker."[17] This canon of literature suggests that relationship-building between practitioners and researchers can better facilitate evidence utilization. Moreover, "practitioners will be more interested in connecting when they believe that the scientists will 1) value practitioners' knowledge and 2) efficiently share what they know."[18]

Nonetheless, scholars have suggested other barriers to promoting these "pull activities," or activities that increase the demand for evidence among practitioners. These include a lack of capacity among practitioners for assessing evidence and mistrust toward researchers that disincentivizes relationship-building between practitioners and researchers.[19] Frequent staff turnover, competing interests, communication barriers, and the differences in the pace of development activities and more traditional research activities all present challenges to effective collaboration between researchers and practitioners even after a relationship is formed.[20] Lastly, among implementers, unsupportive organizational cultures and rigid bureaucratic management structures curtail willingness to design activities centered around ex-post evidence and adaptive management because they believe that donors are risk averse and avoid granting awards to projects that require some degree of trial and error.[21]

Findings from this study, including from follow-up interviews, KIIs, and the quantitative survey, generally reflect similar dynamics regarding the factors that support or constrain uptake and usage of research evidence. The RT finds that the constraints outweigh the supporting factors. The following section examines first the obstacles and constraints to greater research evidence utilization, then the supporting factors.

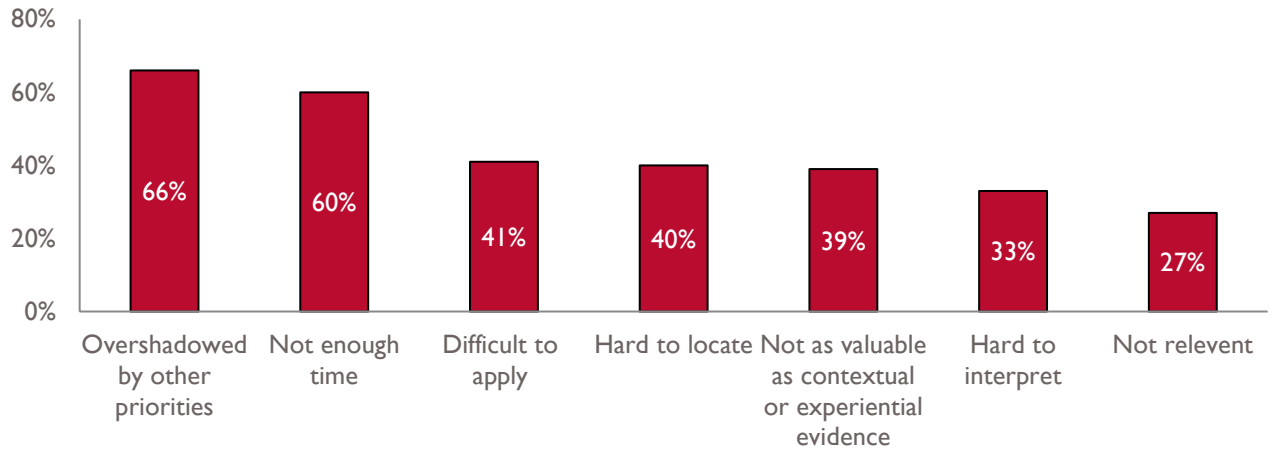
[16] USAID. (July 18, 2018). Request for Task Order Proposal No. 72011518R00007 Judicial Reform in Uzbekistan Program.
[17] Stewart et al., (2019). An integrated model for increasing the use of evidence by decision-makers for improved development. *Development Southern Africa, 35(5).* 616-631
[18] Stewart, R. (2015). A theory of change for capacity building for the use of research evidence by decision-makers in southern Africa. *Evidence & Policy, 11*(4), 547–557.
[19] Ibid; Levine, A. S. (2020). Research impact through matchmaking (RITM): why and how to connect researchers and practitioners. *PS: Political Science & Politics, 53*(2), 265–269.
[20] Brinkerhoff, D. W., Wetterberg, A., & Wibbels, E. (2018). Distance, services, and citizen perceptions of the state in rural Africa. *Governance, 31*(1), 103-124; Dubois, N. S., Gomez, A., Carlson, S., & Russell, D. (2020). Bridging the research-implementation gap requires engagement from practitioners. Conservation Science and Practice, 2(1), e134; Stewart, R. (2015). A theory of change for capacity building for the use of research evidence by decision-makers in southern Africa. *Evidence & Policy, 11*(4), 547–557.
[21] Brinkerhoff, D. W., Wetterberg, A., & Wibbels, E. (2018). Distance, services, and citizen perceptions of the state in rural Africa. *Governance, 31*(1), 103-124; Dubois, N. S., Gomez, A., Carlson, S., & Russell, D. (2020). Bridging the research-implementation gap requires engagement from practitioners. *Conservation Science and Practice, 2*(1), e134.

CONSTRAINTS

(C.1) The design process already entails overcoming many bureaucratic challenges and, given limited time, these take precedence over ensuring evidence and research evidence-based programming.

The RT's survey asked respondents to rate the extent to which various factors were an obstacle to the use of research evidence in the activity design process. The percentage of respondents that rated each factor as either a major or moderate obstacle is presented in Figure 5, which shows that the two most common responses were that research evidence was overshadowed by other priorities and, in a related vein, that there was not enough time to identify, review, and incorporate such evidence.

Figure 5: Percentage Viewing Diverse Factors as a Major or Moderate Obstacle to the Use of Research Evidence in Activity Design (n=77)



There are numerous elements, bureaucratic steps, and clearances required to produce a USAID solicitation. Procurement Action Lead Time (PALT) is the time required by the OAA to undertake various procurement actions. The OAA requires 268 calendar days for competitive contracts and 150 days for competitive cooperative agreements.[22] As such, Missions are frequently in a race to ensure that they can package their needed solicitations in time. The more time needed to accommodate a lengthy PALT, the less time available in planning to adequately and thoughtfully accommodate research evidence in the design itself, not to mention that the PALT may make utilized research evidence out of date by the time an award is up and running. In order to start the clock on a PALT, a design team needs to have their requirements document (e.g., their statement of work [SOW], terms of reference, etc.) complete. Often, completion of the requirements document is subject to substantial commentary from the Mission as well as a rigorous clearance process. This process takes time and is not necessarily accommodated in the PALT. Additionally, the design team must, in tandem with completing the requirements document, build out multiple time-

[22] See ADS 300.3.3 PALTs (p. 14).

consuming extraneous documents required for the procurement packages.[23] According to ADS 300, these include, but are not limited to, upwards of 13 substantial documents, each of which compresses timelines for design teams to accomplish their duties and limits the amount of time available for thoughtfully finding and including research evidence in the design.

As a result, of these many requirements, a lack of time is the second most commonly cited obstacle. Interviewed designers typically carried out a variety of functions (e.g., Contracting Officer's Representative/Agreement Officer's Representative) and undertook design on an irregular basis. As such, despite being an essential duty, program design was typically done on top of well established and already full daily routines (e.g., CS 04). For example, when one design lead who had prioritized research evidence was asked if his Mission ensured adequate time to incorporate research evidence, he noted that he normally works a 10–12 hour day during the week and during the design period he worked Saturdays and Sundays to complete the design (CS 03). Another noted that even though they took the initiative to commission their own research, "since I've been at Post, we've done a couple [research initiatives] whose result we didn't incorporate into our work due to time constraints" (survey response).

(C.2) Evidence is a priority for USAID DRG, but contextual and experiential evidence are prioritized over non-context-specific research evidence.

Another common obstacle identified in the survey in Figure 5 was the prioritization of contextual and experiential evidence. This appears to be a mix of both rational prioritization given limited time among some respondents and a cultural skepticism of the value of research evidence in DRG among others. Interviews suggested two primary critiques of research evidence related to usefulness and relevance.

CONCERNS OVER USEFULNESS: Regarding the former, one interviewee noted that USAID has a lot of control in the health sector, to the point that if the agency invests a certain amount of money, malaria can be reduced by a predictable amount (KII 07). The respondent went on to note that this is somewhat less the case in agriculture, even less so in education, and that USAID has the lowest level of control in DRG. Moreover, he related an anecdote of a USAID staffer who convinced a president to step down and then asked: "how can we talk about evidenced-based programming when person-to-person interactions matter so much?" Another respondent noted that when she sees links to studies or reports in Mission documents, she regards them as "academic" and not embedded in the practical realities of the work (KII 10). Others take a different tack, noting that a study might only tell practitioners what they already know

[23] According to ADS 300, these include but are not limited to upwards of 13 substantial documents including: an approved Activity Approval Memorandum; the justification memo to the contracting officer/agreement officer recommending the choice of instrument; solicitation language designating indicators that the contractor or recipient will be required to collect and report as part of the contractual deliverables or assistance performance goals; independent government cost estimate; proposal submission instructions to the offeror or applicant; technical evaluation or selection criteria (for a competitive action) to be used by the technical evaluation/selection committee; a branding strategy; language on the type of substantial involvement that is anticipated between USAID and the recipient; an approved environmental compliance document such as an initial environmental examination; the inherently governmental and critical functions template; for acquisitions, a written individual acquisition plan or selection assistance plan for assistance; market research documentation; climate risk assessment and climate risk management for USAID projects and activities; and any other documents needed for special clearances (e.g., source nationality/waivers, restricted commodities, limited competition or restricted eligibility, an approved senior obligation alignment review document, and other documents that may be recommended or required).

(KII 4). In short, there remains a perspective among some that research evidence is simply not useful in the DRG sector.

CONCERNS OVER RELEVANCE: Other concerns did not question the value of DRG research evidence per se, but felt that DRG evidence that was not from the same programming context simply was not relevant. One respondent noted that unless an IE is from that country, in that context, and of that same activity, then it is not useful (KII 10). Others noted that "context is king" (KII 5), that "our work is determined by context" (KII 01), and that "Academic research is divorced from the realities on the ground" (CS 09). As noted in Figure 3, contextual evidence was reported to be used far more often than research evidence.

While USAID DRG assessment methodologies do not preclude the incorporation of research evidence, the many formal assessment frameworks, PEAs, and the less formal assessments conducted by DRG Center technical staff on TDYs are overwhelmingly based on understanding the context through a range of interviews with diverse stakeholders related to the planned intervention.

(C.3) While the evidence base in DRG has ballooned in the last two decades, there are still many gaps and challenges IN identifying, interpreting, and applying evidence.

As will be discussed below, there is now a robust DRG evidentiary base that did not exist twenty years ago. Nonetheless, evidence gap maps (alternatively referred to as evidence maps) commissioned by the DRG Center and produced by 3ie demonstrate substantial gaps in the evidentiary base on what works.[24] While some programmatic approaches have generated a large body of research evidence with high-quality systematic reviews, others have produced little to no IEs and much less systematic reviews.

Even where there might be evidence, several survey respondents did consider identifying, understanding, and applying research evidence as a major or moderate obstacle (see Figure 5 above). When asked why research evidence was not used more, one interviewee responded by asking, "how does one figure out what to read, and then who has the time to read all these things, and then even if one is able to read and understand, how does that pertain to what they're working on?" (KII 4). In other words, USAID and IPs need to 1) find and identify relevant research, 2) figure out what to prioritize, 3) find the time to read, 4) actually understand the content and assess its quality and relevance, and 5) apply it.

The fourth step in this process requires a certain level of technical expertise that interviewees acknowledged many staff lack and the fifth step (application) is also particularly challenging. One interviewee noted the wealth of potentially applicable research but questioned how digestible it is and suggested the need for additional steps to turn research into something actionable (KII 5).

(C.4) DRG solicitations often mention evidence but they do not tend to require evidence or research evidence to support the proposed approach, and there is uncertainty on how to responsibly do so.

Best practices in USAID procurement as recommended by the OAA have shifted in recent years to emphasize the "what" rather than the "how." Under this guidance, USAID is to define *what* the DRG problem is to be solved and the outcomes to be achieved, and the IP is responsible for proposing *how*

[24] See for example: 3ie's Strengthening Civil Society Evidence Gap Map.

they will achieve those outcomes, including what programmatic approaches they will take. In effect, this means that while USAID should know and assess the research evidence, it implicitly places the burden for justifying an approach, including providing research evidence on "what works," on the IP. Nonetheless, for this to happen, it remains USAID's responsibility to explicitly require evidence, evidence-based approaches, and research evidence in its solicitations.

To explore this, the RT first asked survey respondents if they required research evidence backing up proposed approaches. Only 10 survey respondents (12 percent) reported requiring research evidence in the solicitation; however, three of these cases were selected for follow-up validation and interviews, and no such requirement was actually included. This suggests that it is both rare and something staff might think they are doing but are, in fact, not doing.

To provide a more objective assessment, the RT reviewed 38 DRG RFPs from 2018–2020 to explore how evidence and research evidence were treated in the proposal. Contracts awarded through such RFPs were examined because, in theory, they were most likely to contain prescriptions from USAID on how DRG problems should be addressed. First, the RT conducted a search of the word "evidence" in section C of the RFP, which lays out the description, SOW, or statement of objectives and turned up 108 mentions in 31 of the 38 RFPs (see Table 5). Seven did not include the word "evidence." In eight RFPs (21 percent), the term was used to describe the DRG problem or to give background information. In 12 RFPs (32 percent), it was used to set an expectation that the proposed approach would be based on evidence, and in 15 RFPs (39 percent) the use of the term was about generating future evidence post-award. As such, while the term "evidence" shows up in most RFPs, it is used to seek an evidence-supported approach in roughly a third of cases.

TABLE 5: INCIDENCE AND USE OF THE WORD "EVIDENCE" IN SECTION C DESCRIPTION/SPECIFICATION OF 38 DRG RFPS, 2018–2020

USE	TOTAL MENTIONS	PERCENT MENTIONS	RFPS WITH AT LEAST ONE MENTION	PERCENT OF TOTAL
To describe the problem or give background	10	9%	8	21%
To set expectations that the proposed approach will be based on evidence	31	29%	12	32%
To reference future evidence generation and use	37	34%	15	39%
Other usage not recreating an expectation of evidence	30	28%	18	47%
Total number of uses of the word "evidence"	108	100%	31	82%
Total RFPs			**38**	**100%**

Regardless of whether or how the term "evidence" was used in the RFPs, the RT examined Sections L (instructions) and M (evaluation factors for award) in greater detail to see if the language employed required evidence in the proposals. As seen in Table 6 below, 26 of the 36 RFPs required or set an expectation that some evidence would be provided by the IP; however, this could be for diverse uses, including defining the problem. Only 13 (36 percent) used language requiring evidence in supporting the proposed approach, and only 1 (3 percent) used language suggesting research evidence was required. In

short, DRG RFPs frequently mention evidence and do at times suggest the need for evidentiary support for a proposed approach, but this is typically not an explicit requirement.

TABLE 6: EVIDENCE REQUIREMENTS IN SECTIONS L AND M OF USAID DRG RFPS FROM 2018–2020

EVIDENCE REQUIREMENTS	FREQUENCY	PERCENTAGE
Number of RFPs that require some evidence	26	72%
Number of RFPs that require evidence supporting the proposed approach, or "what works"	13	36%
Number of RFPs that explicitly require research evidence	1	3%
Total	**36**	**100%**

A deeper examination of the language used in these sections is helpful. The expectation of evidence support was often limited to just an understanding of the DRG problem, which sets a fairly low bar. PEA-friendly language prioritizing contextual evidence was at times used, but this often did not draw a clear connection between the analysis and subsequently proposed approaches. While the term "evidence-based" was often used as a requirement, at times this was not defined, potentially leading to diverse interpretations. From the perspective of ensuring IPs propose an evidence-based approach to support their proposed approach, the most compelling language the RT found was as follows:

> *This document should provide more detail highlighting how the interventions described in the SOW will result in the results USAID wants to achieve, using strong, evidence-based justifications for why the Offeror believes that these are the most effective means to achieve the expected outcomes.*[25]

The language is simple, easy to understand, and clear about what evidence the technical evaluation will be looking for. Nonetheless, the text uses "should" instead of "must," suggesting it is not necessarily a requirement, and the text does not specify research evidence.

These findings are well summarized by an interviewed IP. This group of interviewees reported that they analyze new solicitations for how the terms "evidence" or "evidenced-based" are used in the solicitation. While they note that such terms are common, they report that in many cases, one can distinguish between cutting and pasting standard language and an actual prioritization of evidence. Recognizing that most solicitations are not looking for evidence, they intentionally do not include citations and research evidence in their proposals to save limited space for other content.

While evidence supporting the proposed approach is not typically required and research evidence is rarely required, interviewees suggested several reasons not to include research evidence specifically as a requirement, including 1) inertia—it has not been done traditionally, 2) fear of incentivizing check-the-box gratuitous research evidence citations, 3) the possibility of favoring large IPs who are better positioned to

[25] USAID. (July 18, 2018). Request for Task Order Proposal No. 72011518R00007 Judicial Reform in Uzbekistan Program.

absorb and integrate research evidence as well as learn to use evidence-coded language, and 4) inadequate capacity on the technical evaluation committee to assess presented evidence.

(C.5) The lack of time and challenge of managing the design process are exacerbated by high turnover in Missions and staffing shortages.

Staff turnover is an endemic aspect of programming foreign assistance. The typical post rotation for FSOs is two to four years. The turnover may be more frequent in Missions designated as hardship posts, where FSOs may rotate yearly. Natural turnover is exacerbated by extended vacancies due to medical evacuations and FSO curtailment from a post when FSOs terminate their tour of duty earlier than expected. The COVID-19 pandemic also appears to have magnified this issue; the 2020 USAID year-end fiscal report[26] stated that several USAID Missions "reported that personnel capacity had ceased or was significantly restricted due to the pandemic." Throughout their career, FSO democracy officers work on a wide range of DRG programming across diverse contexts. As such, while they develop substantial technical expertise, FSOs' strongest expertise is typically in USAID processes. More critically, the frequent turnover contributes to brain drain and deficits in knowledge management and organization essential to developing an evidentiary base for programming.

In addition to turnover, there is a shortage of staff. In 2020, the USAID year-end fiscal report indicated that "long gaps in leadership affected the quality of USAID's work…and nearly two-thirds of USAID's Missions reported reduced personnel capacity."[27] Moreover, the 2021 USAID year-end fiscal report noted that, in July 2021, the USAID Administrator testified to Congress that "the funding levels and complexity of our programs has expanded at a rate that significantly outpaces our staffing."[28] To some extent, turnover and limitations to a full complement of staffing at USAID Missions limit their ability to manage the complexities of rigorous research evidence.

SUPPORTS

While the constraints tend to outweigh the incentives for the use of research evidence, there are several important factors that support the use of research evidence and additional factors that could play more of a supporting role. In this section, the RT explores 1) an evidence-friendly legal and regulatory environment, 2) dramatic increases in the amount and accessibility of research evidence, 3) DRG Center technical experts, 4) evidence champions throughout the agency, and 5) the cumulation of knowledge through FSN staff and predecessor awards.

(S.1) The regulatory environment codifies a structure of support for the use of evidence and research evidence in the design process, but it stops short of specific mandates and requirements.

USAID's ADS 201 presents a well-thought-through policy framework for integrating evidence into activity and project design. In fact, the stated purpose of ADS 201 is to further define "the Agency's policies, strategies, and vision…to ensure policy coherence, quality, and technical rigor to support evidence-based decision-making…" To accomplish this, the ADS relies on USAID's collaborating, learning, and adapting

[26] USAID. (2020b). Agency Financial Report Fiscal Year 2020.
[27] USAID. (2020b). Agency Financial Report Fiscal Year 2020.
[28] USAID. (2021). Agency Financial Report Fiscal Year 2021.

approach and its evaluation policy as the standard bearers of its evidence-based decision-making policies. The ADS does not fully articulate how collaborating, learning, and adapting will be used by Missions to integrate evidence-based decision-making in its activity design, and while the sections on evaluation policy are more robust, the ADS stops short of mandating evaluation, with few exceptions that do not innately favor DRG programming evaluation.

Based on survey responses and KIIs, this is not a desirable level of policy mandate. While respondents do not desire an overly burdensome policy mandate, they do want policy that helps steer resources toward evidence-based decision-making.

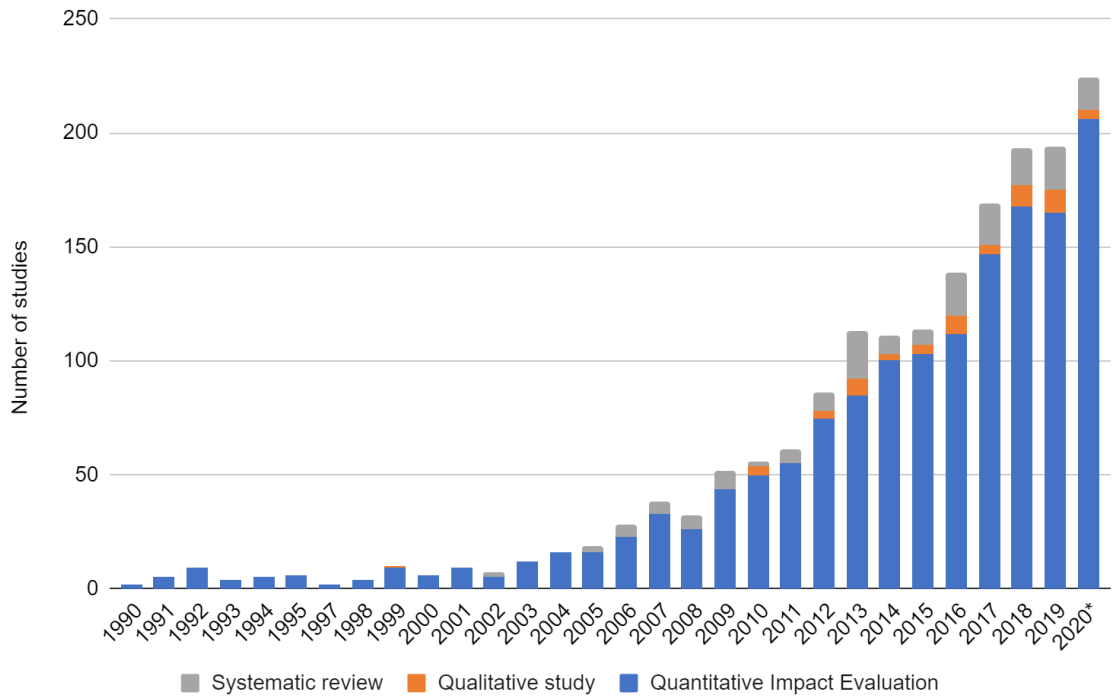
Recent amendments to the ADS further limit mandates that formerly may have triggered DRG activity designers and planners to integrate evidence or evaluation planning in their activities. Formerly, activities were streamlined to align with pre-approved projects under project appraisal documents (PADs). Research evidence was often a featured characteristic of designing and defining a PAD. In some ways, the idea behind the PAD was to use evidence to inform pre-approved activities that were commissioned in a "linear waterfall that takes place in perfect sequence."[29] In reality, this was not really the case, and PADs became a legacy policy superseded by lighter-touch regulations that make projects optional, iterative, and flexible rather than static. In some ways, though, when PADs went away, so too did a natural storehouse of research evidence. While research evidence may be more easily directed toward more strategic endeavors (see recommendation on integrating to CDCS), how best to operationalize for bespoke activity design, prior to award, is a less natural path and full of constraining conditions that limit feasibility.

(S.2) The evidentiary base has grown exponentially in DRG and evidence maps have increased the accessibility of that evidence but the DRG Center does not yet track how this evidence base is used.

3ie and the DRG Center have generated evidence maps, visual depictions of the available evidence (e.g., IEs and systematic reviews) for each DRG program area across a range of interventions and outcomes. The DRG Center also curates a learning harvest, a database of DRG Center-commissioned research and evidence. Figure 6 below provides an overview of the studies contained in the evidence maps by year. These include IEs, systematic reviews, and qualitative studies that can tease out impact. As such, this represents the body of DRG research evidence on the question of "what works." The figure clearly illustrates a dramatic increase in DRG research evidence over the past two decades, from just a handful a year to over 200 a year. In total, there are 1,625 IEs, 63 qualitative studies that meet 3ie's criteria of establishing causal attribution, and 181 systematic reviews contained within the evidence maps. While it might have once been true that there was a lack of research evidence on what worked in the DRG sector, this is clearly no longer the case for the sector as a whole.

[29] ADS 201 Additional Help Document, 2021.

Figure 6: Research Evidence in 3ie Evidence Maps by Year



SOURCE: 3IE DRG EGMS DATABASE

While the DRG Center has promoted both the evidence maps and learning harvest through various platforms, including learning events and conferences and publication in the monthly learning digest, the DRG Center has not yet developed a means to know if and how these sources have informed design processes.

(S.3) DRG Center technical staff have specialized knowledge and mandates, but staying current on the evidence still requires personal initiative.

Several DRG Center staff members have specific technical expertise and long tenures in the agency. For example, at the time of research, the DRG Center had staff with expertise in local governance, parliamentary strengthening, public financial management, elections media strengthening, civic education, anti-corruption, judicial reform, labor, and combatting trafficking in persons, just to name a few. Interviewed technical experts at the center reported supporting Missions in designing 3–7 activities in their field of expertise per year. Their process typically entailed a three-week TDY (or virtual TDY during COVID-19 restrictions) to conduct interviews with stakeholders and work with the Mission to develop the activity design inputs to solicitation documents. Because of their specialization, these individuals reported several instances in which research evidence informed their program design. For example, in the field of local governance, research evidence steered recent activity designs away from decentralization programming in countries where it was unlikely to be effective and away from community-driven development components that were found to be ineffective (KII 13). Nonetheless, even this specialized group recognized that they lack the time and bandwidth to follow and stay current with the research evidence, and, like their peers, they tend to prioritize contextual and experiential evidence.

(S.4) There is variation in the extent to which Mission culture and Mission leadership promote the use of research evidence and the most likely predictor of research evidence appears to be a personal commitment by the designers.

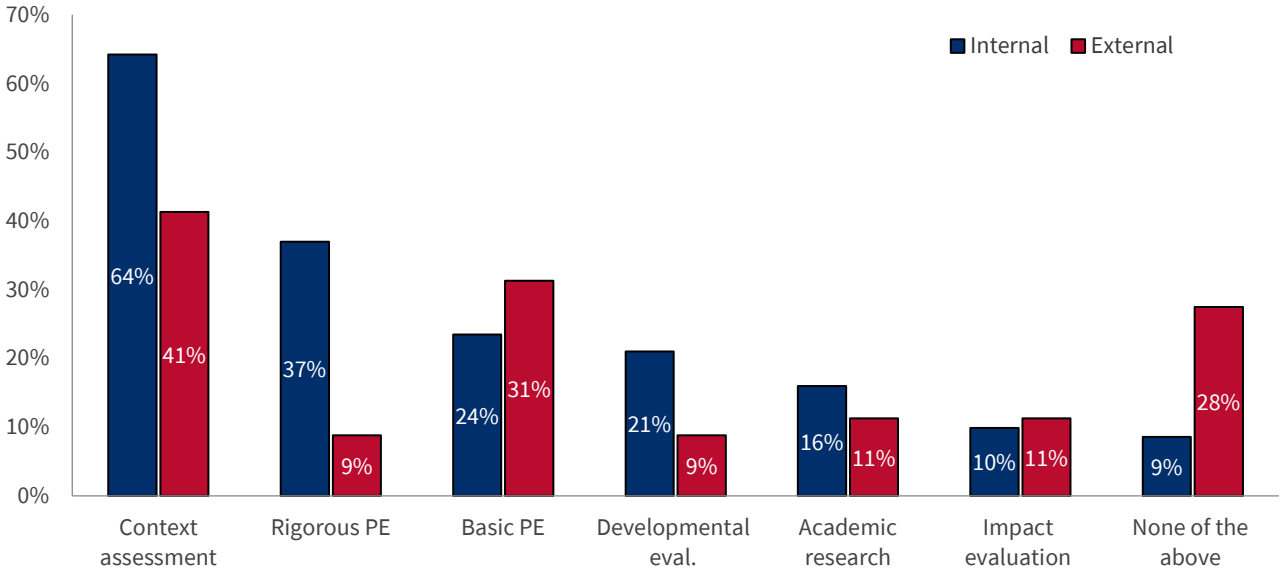
Survey data and follow-up interviews suggest that activity design planners are afforded substantial influence over the direction of the design itself, and local stakeholder input and contextual evidence are valued as key sources of input (each ranked as the three most influential sources of information on an activity design). The survey data suggest that USAID leadership provides a supportive environment for planners to incorporate research evidence into the process. Follow-up interviews suggest that improving uptake and usage of research evidence in activity design could be promoted if leadership (supervisor, sensor staff) were more proactive in helping planners connect to research evidence. Further, case studies suggest that when there is a culture of locally derived evidence (from, for example, local research organizations or civil society organizations), then planners are more likely to be aware of these products and try to incorporate their findings in the activity design (see Mexico, Rwanda, Philippines). Ultimately, when a motivated planner has access to evidence; is supported by their operating environment (OAA and program office encourage research evidence in the design), supervisors, and/or senior staff; and can use that research to back up findings from the contextual analysis, then they are more likely to integrate research evidence into activity design.

In summary, there are substantial constraints to the use of research evidence to inform DRG activity design that lead to the lower-than-expected use of research evidence. Nonetheless, there are also supporting factors, and a robust strategy to address constraints and build on existing supports could increase the use of evidence in USAID activity designs and IP proposals. Recommendations building on the findings in this section are provided below.

5. WHAT ARE THE STRENGTHS AND WEAKNESSES OF HOW SELECTED ACTIVITY DESIGN TEAMS DETERMINE AND PLAN DRG RESEARCH, EVALUATION, AND LEARNING NEEDS AND PLANS?

A variety of sources suggest post-award/assistance learning is a priority for the DRG activity designers. The RT asked survey respondents to identify both internal and external post-award learning planned in their most recent design. As seen in Figure 7, only 9 percent of respondents (7) identified no planned research, and most identified multiple activities. Context assessments, internal PEs including a baseline and endline, and basic external evaluations were the most common forms of post-award learning identified.

Figure 7: Internally and Externally Led Evidence Generation Planned in the Respondent's Most Recent Activity Design (n=88)



As noted above, the study team also explored how evidence was referenced in RFPs and found that USAID DRG activity designers are actually more likely to reference evidence generation post-award than they are to reference it informing the proposed approach. A coder counted the number of times that the word "evidence" was used in Section C Description/Specification of 38 DRG RFPs between 2018–2020. As shown in Table 5 above, out of 108 total uses of the term, 34 percent (37 mentions) were to set an expectation that the project would generate future evidence post-award. This expectation was found in 15 of the 39 RFPs. By contrast, a somewhat smaller 29 percent of mentions (21) set an expectation that the proposed approach would be based on evidence, representing 12 of the 38 RFPs. This finding was corroborated by interviews. For example, one interviewee noted the lack of evidence in a given program area and adopted a strategy of building evidence generation into the award (CS 06).

Nonetheless, external evaluation planning is often not an important part of the activity design process. Among the wide range of post-award learning considered in the study question above, the study team decided to focus on external evaluation planning, given the strong historic emphasis on external evaluation and given that external evaluations are most likely to result in evidence generation about "what works." While Table 7 below suggests internal evaluations are common, they suffer from a clear conflict of interest and are not recommended for high-profile/highly scrutinized activities or when accountability is a goal.[30] Of 35 RFPs examined, the majority (21-60 percent) did not include mention of an external evaluation (see Table 7).

[30] USAID (July 2021) Choosing between a USAID External or Internal Evaluation.

TABLE 7: EXPECTATION OF AN INTERNAL OR EXTERNAL EVALUATION IN 35 RFPS

	INTERNAL	EXTERNAL
No	26	21
Yes	4	8
Possibly	5	6
Total	35	35

Interview evidence suggests that evaluation planning was simply not a high priority, both in Mission-led designs and in designs supported by DRG Center technical staff (KII 1, 11, 13, 14). One DRG Center technical expert noted that MEL-related language is typically just cut-and-paste/stock language. The respondent noted that Missions are typically behind on the drafting and the goal is just to get a scope finalized (KII 13). This concern over the lack of time and more salient requirements that go into a design was echoed by others. For example, another interviewee noted that he did not think he had ever seen evaluation planning done well (KII 1).

While outside of the dates of this study's focus, the important exception to this mentioned by two interviewees was the IE clinics run by the DRG Center from 2013 and 2017, where IE planning was generally part of the solicitation development process. During this period, the DRG Center offered co-funding for Missions to implement IEs and invited Mission staff to a week-long IE clinic to plan an IE. The DRG Center ceased the clinics in 2018 and while it began to support IEs again in 2021, the initiative's momentum was lost. The Center issued a call in 2022 for Missions to express interest in co-funding rigorous evaluations, but only one expression was received by the deadline outside of existing evaluations.

Partially as a result, the DRG sector does not meet evaluation spending targets. According to the USAID Evaluation Registry, from 2016 to 2020, the Agency completed 173 external DRG evaluations valued at over $47 million (see Table 8 below). On the one hand, this is an impressive quantity, but on the other hand, it represents less than 1 percent (0.66 percent) of DRG program expenditures over this period, far less than the previous Agency target of 3 percent during this period, and still less than the recently lowered target of 1 to 3 percent (ADS 201.3.6.5).

TABLE 8: COMPLETED USAID DRG EVALUATIONS (2016–2020)

	PES	IES	TOTAL
No. of DRG evaluations	159	14	173
DRG evaluation budget	$36 million	$11 million	$47 million
DRG program expenditures	-	-	$7,116 million
DRG evaluation budget as a percent of expenditures	-	-	0.66%

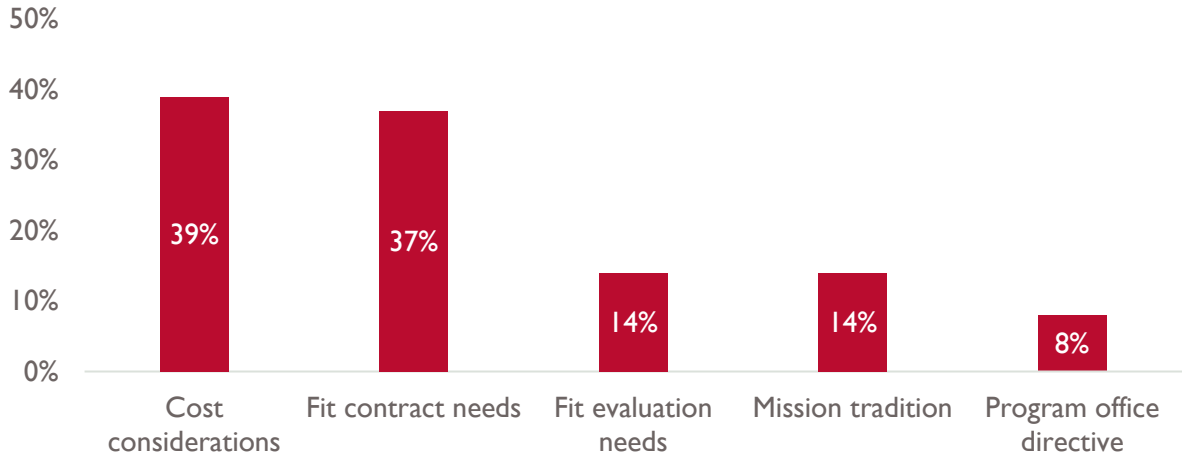
SOURCE: EVALUATION REGISTRY DASHBOARD, FOREIGNASSISTANCE.GOV

The default evaluation is limited to a one-time, basic PE. Table 8 also shows that USAID DRG IEs are still relatively rare. While the evaluation registry does not provide a breakdown of PE types, as suggested in Figure 7 above, basic PEs were the most commonly cited. These are conducted at one point in time, typically involve three weeks of fieldwork, and rely overwhelmingly on KIIs and focus group discussions. Moreover, when evaluations were mentioned in RFPs, Missions seemed to use stock language envisioning a basic PE at mid-term or at the end of implementation. Only in one case did an RFP examined in the review mention a potential IE. While valuable to consolidate experiential evidence, basic PEs are typically unable to measure outcomes or changes in outcomes over time, assess program impact, or otherwise result in research evidence. Given the strong desire to understand outcomes and the range of evaluation options, the preference for this form of evaluation is somewhat surprising.

In the survey, the RT asked respondents who included a basic PE why they chose this form of evaluation. Results are shown in Figure 8 below. The primary explanations included cost followed by a sense that it was the best fit evaluation for their evaluation needs. The cost of IEs and more robust evaluations has been noted as a major obstacle in other sources as well, and several interviewees echoed these cost concerns (CS 02, CS 08).[31] However, several respondents also noted that basic PEs is just what they have always done (15 percent), that they met the evaluation requirement (40 percent), and that it is the best fit for the nature of the agreement/contract (40 percent). In fact, a recent study on DRG evaluation utilization noted a concern that these basic PEs are frequently done as a check-the-box activity to meet a requirement, a sentiment echoed by some interviewees (CS 08).[32]

[31] DRG IE Retrospective.
[32] NORC. (April 2022). DRG-LER I and LER II Research Product Utilization Measurement Analysis. USAID.

Figure 8: Survey Responses to "If you planned for a one-time mid-term or final evaluation rather than an IE or a PE with a baseline and endline, why did you make that determination?" (n=51)



There are a few potential reasons why evaluation planning falls through the cracks. The evidence from this study does not point to a clear cause of evaluation planning limitations, but interviews and the document review do raise potential factors. The first and most important is referenced above. There are so many inputs to the design process, and, given time and bandwidth constraints, evaluation planning decisions are pushed off into the future. Second, evaluation planning appears to fall through a crack between the Program Office and the Technical Office. USAID assigns responsibility for evaluations to the Program Office; however, the Technical Office is the primary user of evaluation learning. In order to plan evaluations that respond to both requirements and learning use needs, Mission staff need to collaborate across this division. Third, in a related vein, there appears to be a crack between steps in the program cycle. Following the development of a five-year CDCS, Missions are required to develop a performance management plan, which entails a high-level five-year evaluation plan. This process varies by Mission but, given the challenges in planning evaluations for the whole Mission for five years, there is a risk that this process, led by the Program Office, amounts to simply identifying required evaluations based on anticipated activity cost. USAID guidance states that the performance management plan will be refined and revised as activities are designed and implemented, processes led by the Technical Office, in later stages of the program cycle;[33] however, as discussed above, this often does not happen.

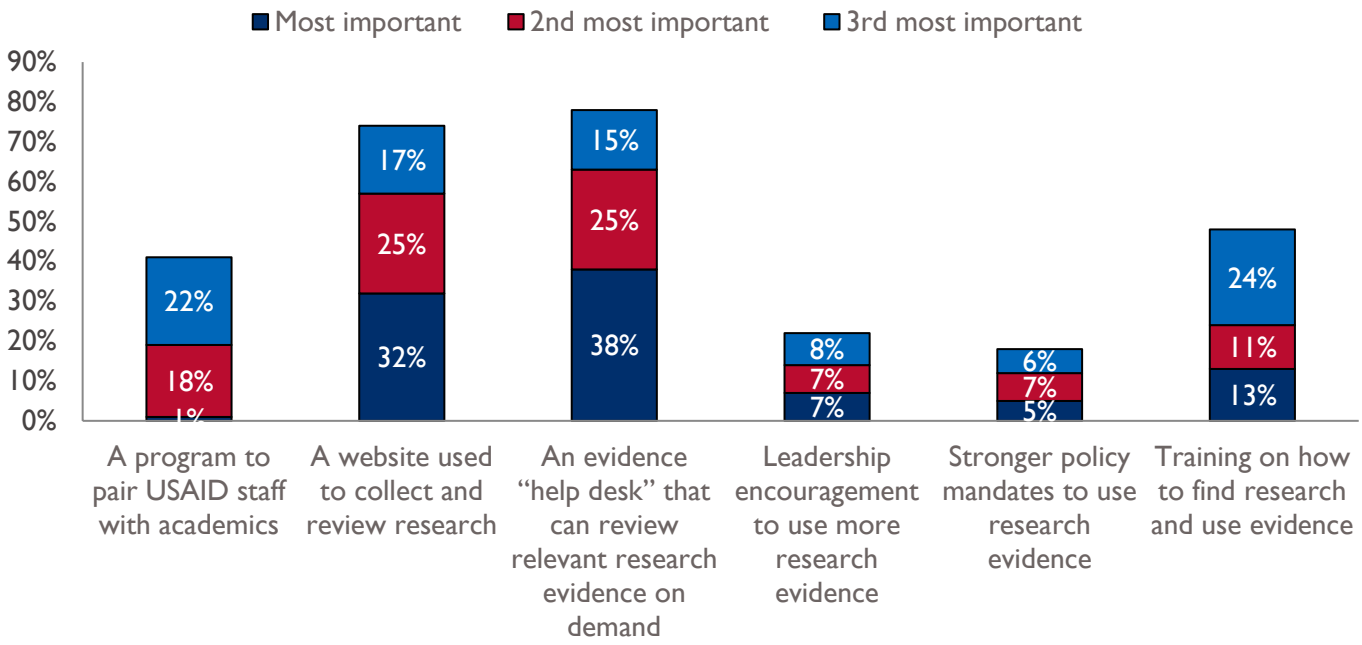
In summary, activity designers recognize the importance of post-award learning and build learning opportunities into their designs. However, in practice, external evaluation planning is a lower-level priority in the activity design process. Partially as a result, the DRG sector does not meet evaluation spending targets and designers tend to default to a basic PE, which, while valuable, does not generate research evidence about what works in DRG. While the RT does not know for sure why evaluation planning falls short, the team identifies potential causes, including time constraints and cracks between the Program Office and Technical Office and between performance management plan evaluation planning and activity evaluation planning.

[33] USAID (Feb 2021) How-To Note: Performance Management Plan.

6. HOW COULD RESEARCH EVIDENCE INTEGRATION AND EVALUATION PLANNING BE IMPROVED?

This section provides and explores recommendations to increase the integration of research evidence and evaluation planning into the design process. To help address research evidence integration, the RT provided survey respondents with a list of initiatives to encourage the use of research evidence and asked respondents to note which they would be highly likely to utilize or support. The most selected initiative was an evidence "help desk" that would review research evidence, followed closely by a website to collect and review research (see Figure 9). Training on how to find and use research evidence and a program to pair staff with academic specialists also received some support. There was only limited support for leadership encouragement to use more research evidence and stronger policy mandates. The paragraphs that follow offer recommendations based on this study's findings. The RT divides those recommendations into two groups: 1) those oriented toward making evidence more accessible and 2) those increasing demand for evidence.

Figure 9: Survey Responses to "Which of the following initiatives to encourage the use of research evidence would you be highly likely to utilize or support?"



1. **Expand efforts to make evidence accessible:** The DRG Center should continue to increase the accessibility and readability of research evidence.

   1.1 **The DRG Center should strengthen and continue existing E&L team efforts to make research evidence accessible.** The DRG Center and the E&L team already have a long history and several initiatives to promote research evidence. This includes commissioning six evidence

maps and associated practitioner briefs, DRG data portraits that summarize key indicators over time for a given country, a monthly learning digest, a section in the quarterly DRG newsletter, weekly Facty Fridays with interesting evidence snippets, the monthly E&L Talk Series, and the Annual Learning Forum. However, several of these initiatives are fairly new or are done irregularly (e.g., E&L Talk Series). As such, the RT recommends that these continue and be strengthened. For example, the E&L team should plan to update evidence maps periodically (e.g., every two years). The E&L team should also experiment with ways to market products to Missions to ensure their relevance and use.

1.2 **Efforts to make research accessible should not be limited to internal audiences but should include IPs and potential IPs and a public-facing website.** Given the important role of IPs in using and generating evidence discussed above, E&L evidence-promotion efforts should be expanded to include this important segment of the USAID community. Some of the activities listed above (e.g., evidence maps, DRG data portraits) are publicly available, some reach limited non-USAID audiences (e.g., learning digest), and some are entirely internal (e.g., E&L Talk Series). The E&L team should explore options to make its efforts more publicly available and not just to select individuals. Key to this effort will be a public-facing website, which the RT understands is moving forward ("DRG Links"). The DRG Center should attempt to replicate the success of similar USAID websites, such as Agrilinks (which has been used for the last ten years to reach the larger agricultural development community), to allow individuals to express interest in USAID activities and evidence and to use as a tool to disseminate evidence and organize webinars.

1.3 **In a similar vein, the DRG Center should ensure that its own learning products and evaluations emphasize and strengthen dissemination.** The RT echoes recommendations in the recent Utilization Measurement Analysis reports to require dissemination plans, tailor dissemination products to intended audiences, and leverage interactive digital tools.[34]

1.4 **The existing pilot evidence review initiative should be expanded to an "evidence help desk," its utility monitored, and its approach adapted.** While the E&L team has commissioned several evidence reviews in the past (e.g., GLP, Armenia Integrity Systems and Rule of Law, Liberia Civic Education), these efforts have influenced only a relatively small percentage of activity designs and have not been scaled up. The E&L team has developed a process and template to conduct bespoke evidence reviews for Missions and OUs to aid in activity design; however, only one has been conducted to date using this new approach. While the evidence maps are an incredible tool to curate and make available research evidence on "what works," this tool still requires USAID efforts to read, understand, and apply. As such, and consistent with the primary recommendation of survey respondents, the DRG Center and E&L team should develop and promote an evidence help desk whereby Missions can commission evidence reviews as part of the activity design process and otherwise receive support in identifying, accessing, and understanding research evidence.

[34] Counterman, M., Conté, S., Starosta, A., Marple-Cantrell, K., Barker, M., & Hatano, R. (2022). Evidence and Learning (E&L) Utilization Measurement Analysis (UMA). USAID; NORC. (April 2022). DRG-LER I and LER II Research Product Utilization Measurement Analysis. USAID.

1.5 **PPL should promote the use of Mission MEL platforms to conduct evidence reviews as part of their scope.** While the E&L team is particularly well situated to offer an evidence help desk, the team has limited bandwidth and buying into the Learning, Evaluation, and Research (LER) II mechanism requires some transaction costs for Missions. Mission MEL platforms, on the other hand, represent an untapped resource to provide these services not just for DRG but for all sectors. PPL has issued extensive guidance on MEL platforms, and while "evidence reviews" are not identified by name in this guidance, they are entirely consistent with the functions of the platform.[35]

1.6 **The DRG Center should experiment with informal pairing with academics.** While pairing the DRG cadre with academics was not a high priority among survey respondents, it does appear to offer an avenue worth experimenting with. There are existing networks that the E&L team could work through, including the Evidence in Governance and Politics network and research 4 impact. As a proof of concept, one option, when appropriate, would be to identify pairings for FSOs recently assigned to the DRG Center who might be asked to provide technical expertise in an area where they have limited experience.

2. **Generate demand for research evidence and foster a culture of learning.** While making research evidence accessible is necessary, it is insufficient. The DRG Center should also expand efforts to increase demand for research evidence, increase its value among the DRG cadre, and foster a culture of learning (2.1, 2.3).

2.1 **Beyond the E&L team, DRG Center technical staff should serve as promoters of the role of research evidence and evaluation planning in activity design.** DRG Center technical staff regularly provide design support to Missions. This study found several instances where their application of research evidence allowed for better-informed designs. Nonetheless, these were the exception rather than the norm, and the DRG Center technical experts interviewed also tended to prioritize contextual and experiential evidence over research evidence and still faced constraints of limited bandwidth, time, and incentive to follow and utilize research evidence. As such, the DRG Center should encourage its technical experts to play a more proactive role in promoting the use of research evidence. This might entail commissioning evidence reviews as part of their standard design process, pairing with researchers, and engaging the E&L team during the design process.

2.2 **The E&L team should strengthen its research evidence-related training delivery.** This should include an expressed objective of increasing the demand for using and generating research evidence. In collaboration with the Bureau for Development, Democracy, and Innovation/Innovation, Technology, and Research Hub, the E&L team piloted a research evidence training in 2021; however, the E&L team has not scheduled subsequent trainings and has only conducted limited follow-up with participants. The team conducts a variety of ad hoc training on IE methods and the use of V-Dem analysis tools and has also generated content for a training on the "learning cycle"; however, this has not been formalized. As such, the team should examine its future intended training schedule and ensure efforts to both build capacity and increase demand are incorporated into that schedule.

[35] USAID. (2020). Illustrative Functions of MEL Platforms (See Section IV Studies—internal USAID access only).

2.3 **Beyond E&L trainings and products, broader DRG Center-produced sector-based programmatic guidance materials and trainings should highlight the value of research evidence.** The DRG Center produces a wealth of primers, assessment tools, guidance documents, and trainings for the DRG cadre. Some of these actively build on existing research evidence on what works (e.g., Civic Education in the Digital Age Primer); however, many others are more separated from the research evidence (e.g., Democratic Decentralization Handbook). New guidance and updates to existing primers and guidance documents should incorporate relevant research evidence. In a related vein, assessment frameworks should also provide guidance to assessment teams for the incorporation of research evidence. For example, the Policy, Coordination, and Integration Team should explore options to increase the role of research evidence in the DRG Strategic Assessment Framework. This might include an analysis of third-party metrics to note changes over time and relative to comparable countries. It might also include the identification and review of key research studies (USAID and non-USAID) completed within the previous five years.

Similarly, trainings should build on the research evidence and direct readers/participants to other resources. The RT recommends that the DRG Center's training support contract include a question about greater incorporation of research evidence as part of their standard review checklist. For example, are there assertions in the training that could be better supported by evidence? Is there existing evidence in this training topic about what works, what does not work, and why that could be better incorporated? DRG Center teams should consider commissioning an evidence review as part of training development.

2.4 **Trainings and other DRG Center outreach efforts should aim to build a culture of learning and shift in mindsets.** For example, training and outreach efforts should address the following misconceptions.

TABLE 9: MYTHS AND MISCONCEPTIONS OF EVIDENCE USE

MYTHS AND MISCONCEPTIONS	COUNTERARGUMENT
I. Contextual and experiential evidence should always be prioritized over research evidence.	Different forms of evidence have different uses in the design process. Research evidence is particularly desirable in identifying drivers of a DRG problem and understanding what works and does not work in addressing the problem. Contextual evidence is particularly valuable in identifying country-specific opportunities and challenges and experiential evidence is best suited to questions related to implementation. Prioritizing one source of evidence over others risks providing USAID with an incomplete picture.
II. There is just not a meaningful research evidence base in DRG like there is in the health and education sectors.	3ie's six DRG evidence maps include 1,625 IEs, 63 qualitative studies aiming to establish causal attribution, and 181 systematic reviews in the DRG sector.[36] While this misconception might have been true twenty years ago, it is no longer true today.

[36] 3ie. (Forthcoming) Chapeau Summary of DRG Sector Evidence. USAID.

MYTHS AND MISCONCEPTIONS	COUNTERARGUMENT
III. Using research evidence is an expensive and untimely undertaking.	While conducting good research can often entail high costs and long timelines, DRG programming can be built on an existing research base that requires a lighter time and cost investment to identify, examine, and consider relevance. Even original research need not be overly expensive or arrive too late to inform decision-making if these risks are planned for and mitigated.
IV. Evaluation and research reduce impact by reducing funding availability for programming.	The ADS notes that 1–3 percent of the Mission's total program funds should be used for evaluation and another 3–10 percent for program monitoring and collaboration, learning, and adaptation. Investing these program funds in MEL initiatives is what lets Missions know if they are spending funds in ways that are likely to produce impacts. The DRG sector is currently not meeting these targets, so it should be spending more, not less.

2.5 **The DRG Center should work to establish a "vision of perfect," or an ideal situation of research evidence utilization and evaluation planning, and an action plan to move toward that vision.** The "vision of perfect" should be multi-faceted and specific, and importantly, it should be created in consultation with those that will help make it a reality: Mission personnel, particularly FSN staff (see discussion above on the role of FSNs).

On the issue of research evidence utilization, this is particularly important given the ambiguity in the ADS about how evidence, and specifically research evidence, should be used in the activity design process. This should be a nuanced vision that accounts for the strengths and weaknesses of research evidence. This study does not suggest a prioritization of research evidence over context evidence or a mandate to always use research evidence would be desirable. One interviewee noted that they have to be humble about what they do not know and adopt a more inquisitive approach to programming (KII 16).

On the issue of evaluation planning, this research has found that the status quo overwhelmingly relies on basic PEs regardless of whether or not that is the best-fit design to address OU learning needs. Changing this equilibrium will be challenging given the time and cost implications of planning and conducting more rigorous research and the pervasive challenge of limited bandwidth to address many competing and important priorities. Fortunately, there is existing evaluation planning guidance for an action plan to build on.[37]

2.6 **Require evidence to support any proposed approaches, and ideally research evidence, in the solicitation process.** Given that IPs have primary responsibility for proposing programmatic approaches, these approaches should be supported by evidence. However, this requires activity design teams to include such evidence as a requirement in solicitation documents, which it currently does not do with regularity. For example, solicitation documents could include:

[37]See ADS 201.3.6.5 Evaluation Requirements; 201.3.2.17 Evaluation During CDCS Implementation; and sections 201.3.6.7 and 201.3.6.8, which also describe evaluation requirements. Additionally, ADS sections that provide evaluation guidance include ADS 201.3.6.2 Evaluation Principles and Standards; 201.3.6.3 Missions and Washington Operating Unit Roles in Evaluation; 201.3.6.4 Types of USAID Evaluations; and the mandatory reference ADS 201saf, Evaluation Triggers.

> *This [technical proposal] must provide evidence, including research evidence, for why the interventions described can be expected to contribute to the outcomes USAID wants to achieve. Research evidence is defined as systematic and analytic methodologies designed to test hypotheses and includes IEs, systematic reviews, and academic studies testing hypotheses.*

This language should go beyond requiring bidders to demonstrate knowledge of a DRG problem and focus on evidence supporting the approach. If programming is to occur in an area with a limited evidentiary base on "what works," then bidders can still use research evidence on the drivers of a DRG problem to justify the approach. Recognizing that interviewees did raise legitimate concerns about incentivizing check-the-box citations and increasing the barriers to successful proposals from small and local partners, research evidence should be thoughtfully included. In addition to being part of the scope of work, the use of evidence should also be included in the technical proposal evaluation criteria, to allow the Technical Evaluation Committee (TEC) the ability to score a proposal based on the inclusion of evidence. If USAID technical evaluation teams feel that they lack the capacity to assess research evidence, experts within or external to USAID can and should be brought in to support the technical evaluation team as non-voting members, as appropriate.

2.7 **The DRG Center should continue to support Missions and Regional Bureaus in building a contextualized research evidentiary base, and Mission learning agendas should adopt this as a priority.** Given the limitations of both uncontextualized research evidence and contextual evidence by itself, building a contextualized research evidence base about drivers of DRG problems and what works in addressing them should be a priority for Missions. Fortunately, partially as a result of the 2018 Evidence Act (Section 312), USAID has added learning agendas as a component of the CDCS process, offering an ideal opportunity for Missions to plan research, evaluation, and evidence generation that builds on the existing evidence base but is specific to their programmatic context. As in other areas, given the longer-term nature of FSN employment and their better understanding of the programmatic context, FSNs should play a key role in this process.

2.8 **The DRG Center and the E&L team should continue to support IE and rigorous PE planning and Missions should prioritize evaluation planning that moves beyond basic PEs.** The RT echoes recommendations in the recent DRG IE retrospective report for the DRG Center to provide co-funding and technical assistance to Missions to support them in conducting rigorous evaluations while incorporating lessons learned from past IEs.[38] This study confirms the retrospective finding that Missions need to be incentivized and supported to go beyond the status quo basic PE approach that currently predominates. Through outreach, trainings, and its technical specialists, the DRG Center should continually promote evaluation planning as an indispensable part of the design process. Moreover, the DRG Center should cultivate champions in the Missions willing to thoughtfully engage in evaluation planning as part of their standard approach to activity design.

[38] Findley, M.G., Starosta, A., Sabet, D. (2022). DRG Impact Evaluation Retrospective: Learning from Three Generations of Impact Evaluations. USAID.

7. CONCLUSION

This study finds that while there is a reasonable foundation for incorporating research evidence into activity design and for planning new evidence generation at the design stage in the DRG sector at USAID, there are several missed opportunities and there is room for considerable improvement.

To begin, the RT finds that the terms "evidence" and "evidence-based" are understood differently by different USAID staff. This study hopes to increase conceptual clarity by distinguishing 1) the type of evidence (research, contextual, experiential) and 2) the decision point that the evidence aims to inform (strategy, diagnosis, prescription, refinement, and targeting). These are presented below in Figure 10. Design teams should employ a template like this to think about how evidence informs their decision-making. While USAID often prioritizes contextual and experiential evidence, research evidence is often able to provide greater confidence in asking questions like: What are the drivers of a DRG problem? What approaches are most effective or ineffective in addressing the problem? Furthermore, some of the more compelling designs examined in this study leveraged multiple sources of evidence across these design decisions.

Figure 10: Template for Connecting Different Evidence Types to Key Design Decisions

The design process has five major decision points that can be informed by the various types of evidence. By completing a template like the one below, planners can plan strategically to find and apply appropriate evidence throughout the process.

Decision point	Type of Evidence (research, contextual, experiential)	How will evidence be collected, and by who?	When and how will findings be incorporated?
Strategize the problem to address			
Diagnose the problem and its causes			
Prescribe the best programmatic approach			
Inform refinement of the approach			
Target who should receive the approach			

Through this study, the RT explores the factors that constrain or support the application of more formal research evidence and finds that the constraints outweigh the supports. On the constraint side, several bureaucratic challenges and requirements take precedence over the incorporation of research evidence. While evidence remains a priority, amidst time constraints and concerns over usefulness and relevance, contextual and experiential evidence are prioritized over research evidence. Beyond these concerns, interviewees also point to gaps in research evidence, challenges in interpreting and applying research evidence, and staffing concerns such as high Mission turnover that undermine a common understanding of the evidence. While the bandwidth constraints are clear, there is a missed opportunity to—when

appropriate—include language encouraging or requiring evidence (and specifically research evidence) in responses to solicitations.

While weaker than the constraints, several factors support the use of evidence upon which USAID and the DRG Center can be built. These include 1) an evidence-friendly legal and regulatory environment, 2) dramatic increases in the amount and accessibility of research evidence, 3) specialized DRG Center technical experts, 4) evidence champions throughout the agency, and 5) the potential cumulation of knowledge through FSN staff and predecessor awards.

In addition to using evidence in the design process, USAID also has an important role to play in generating evidence. This requires planning for learning, evaluation, and evidence generation post-award/contract. While the RT finds that general post-award learning is a priority for DRG activity designers, external evaluation planning specifically is often not an important part of the activity design process. Partially as a result, the DRG sector does not meet evaluation spending targets laid out in the ADS and frequently defaults to a one-time, basic PE. While valuable for answering some questions, these evaluations, which typically involve around three weeks of largely qualitative fieldwork, are not well suited to measuring outcomes and changes in outcomes over time or testing impact. In short, they do not generate research evidence. The RT explores potential reasons why evaluation planning seems to fall through the cracks, including an input-heavy and time-constrained design process and gaps between the USAID program and technical offices.

The RT finds important opportunities to both increase the use of evidence, specifically research evidence, in the activity design process and improve evaluation planning at the design stage. The recommendations are divided into efforts 1) to make evidence more accessible and 2) to generate demand for research evidence and foster a culture of learning. The former set of recommendations include strengthening and building on existing DRG Center initiatives (e.g., existing evaluation planning efforts, DRG learning digests, and other dissemination efforts), but it also entails some new strategic directions. This includes expanding evidence-use promotion to include DRG IPs, creating a DRG Center evidence help desk to conduct evidence reviews as an input to the design process, and expanding the typical role of Mission MEL platforms to conduct evidence reviews. The latter set of recommendations focused on generating demand also includes recommendations to build on existing strengths (e.g., ensuring research evidence is incorporated into DRG Center trainings and guidance materials) and some new directions. The most important of which is, when appropriate, ensuring that solicitation documents create an expectation that IP proposals will include evidence, and ideally research evidence, to support their proposed approach.

Through this two-pronged strategy of both making it easier to incorporate and plan for evidence while seeking to increase demand for such efforts, the RT is confident that the DRG sector can better align real-world design work with the rhetorical and regulatory expectations of an evidence-based learning organization.

ANNEX 1. BIBLIOGRAPHY

Brinkerhoff, D. W., Wetterberg, A., & Wibbels, E. (2018). Distance, services, and citizen perceptions of the state in rural Africa. *Governance, 31*(1), 103-124.

Counterman, M., Conté, S., Starosta, A., Marple-Cantrell, K., Barker, M., & Hatano, R. (2022) Evidence and Learning (E&L) Utilization Measurement Analysis (UMA). USAID.

Dubois, N. S., Gomez, A., Carlson, S., & Russell, D. (2020). Bridging the research-implementation gap requires engagement from practitioners. *Conservation Science and Practice, 2*(1), e134.

Findley, M. G., Starosta, A., & Sabet, D. (2021). DRG Impact Evaluation Retrospective: Learning from Three Generations of Impact Evaluations. USAID.

IT Shows Inc., Deloitte Consulting. (June 2021). USAID Capacity Assessment for Evidence Management and Use: Annexes. USAID.

Levine, A. S. (2020). Research impact through matchmaking (RITM): why and how to connect researchers and practitioners. *PS: Political Science & Politics, 53*(2), 265–269.

Lavis et al, 2003; Milne, B. J., Lay-Yee, R., McLay, J., Tobias, M., Tuohy, P., Armstrong, A., Davis, P. (2014). A collaborative approach to bridging the research-policy gap through the development of policy advice software. Evidence & Policy: *A Journal of Research, Debate and Practice, 10*(1), 127–136

Milne, B. J., Lay-Yee, R., McLay, J., Tobias, M., Tuohy, P., Armstrong, A., Davis, P. (2014). A collaborative approach to bridging the research-policy gap through the development of policy advice software. *Evidence & Policy: A Journal of Research, Debate and Practice, 10*(1), 127–136.

NORC. (April 2022). DRG-LER I and LER II Research Product Utilization Measurement Analysis. USAID.

Nutley, S., Davies, H. T. O., & Walter, I. (2007). Using evidence: How research can improve public services.

Nutley, S., Davies, H., & Walter, I. (2002). Evidence-based policy and practice: Cross-sector lessons from the UK. Swindon: ESRC UK Centre for Evidence-Based Policy and Practice.

Rose, J., & Watson, B. (Forthcoming). The Use of Political Economy Analysis among USAID Implementing Partners. Thinking and Working Politically Community of Practice.

Soyer, E., & Hogarth, R.M. (2015). Fooled by Experience. Harvard Business Review. May, 72-77.

Stewart et al., (2019). An integrated model for increasing the use of evidence by decision-makers for improved development. *Development Southern Africa, 35(5).* 616-631

Stewart, R. (2015). A theory of change for capacity building for the use of research evidence by decision-makers in southern Africa. *Evidence & Policy, 11*(4), 547–557.

Sutherland, W. J., & Wordley, C. F. (2017). Evidence complacency hampers conservation. *Nature Ecology & Evolution, 1*(9), 1215–1216.

USAID. (July 18, 2018). Request for Task Order Proposal No. 72011518R00007 Judicial Reform in Uzbekistan Program.

USAID. (2020a). Illustrative Functions of MEL Platforms (See Section IV Studies—internal USAID access only).

USAID. (2020b). Agency Financial Report Fiscal Year 2020.

USAID. (2021). Agency Financial Report Fiscal Year 2021.

USAID. (Feb 2021). How-To Note: Performance Management Plan.

USAID. (July 2021). Choosing between a USAID External or Internal Evaluation.

ANNEX 2. KII LIST

TABLE 10: KII LIST

POSITION	NUMBER OF INTERVIEWS	NUMBER OF PARTICIPANTS	MALE	FEMALE
USAID/Bureau for Development, Democracy, and Innovation/DRG/ OAA	4	8	4	4
USAID/PPL	3	3	-	-
USAID Mission	16	20	11	9
IPs	2	7	4	3

ANNEX 3. SURVEY RESULTS

Figure 11: Missions Represented in the Quantitative Survey

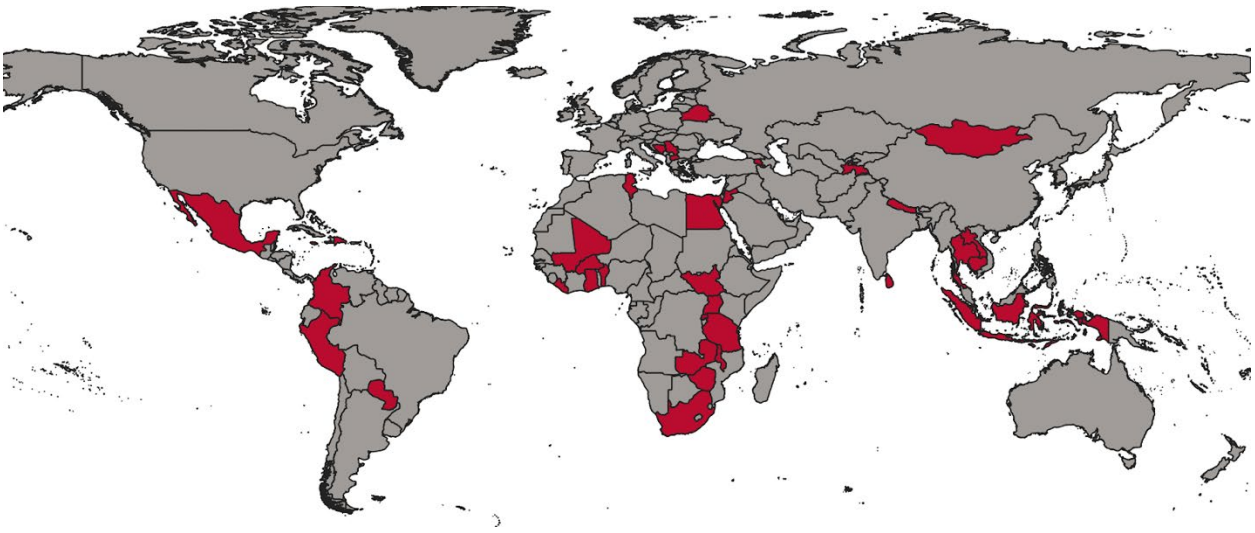
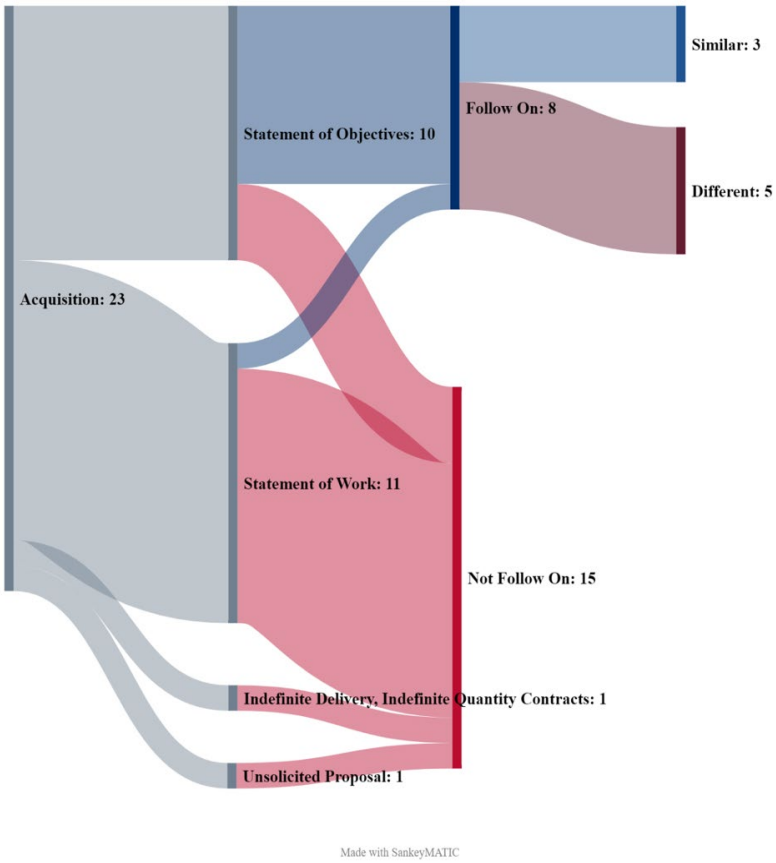


TABLE 11: SURVEY DEMOGRAPHICS N=88

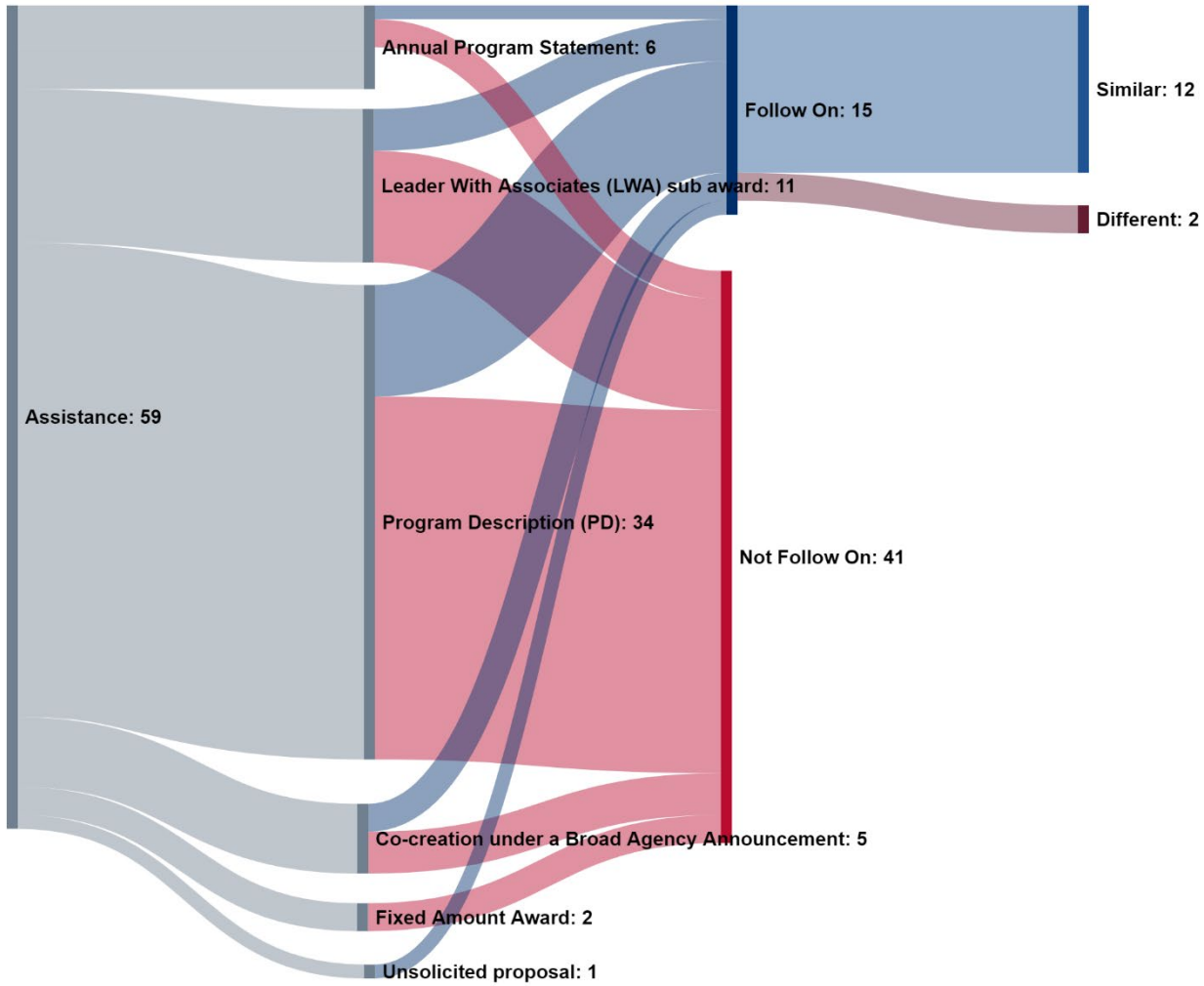
GENDER	**Male**		**Female**		**Declined to State**	
	46% (41)		51% (45)		3% (3)	
EXPERIENCE	**1–5 Years**		**6–10 Years**		**10+ Years**	
	26% (23)		31% (27)		43% (38)	
HIRING MECHANISM	**Foreign Service National**		**Foreign Service Office**		**Personal Service Contractor/ Other**	
	57% (51)		28% (25)		12% (12)	
ROLE	**Agreement Officer's Representatives (AOR)/Contracting Officer's Representatives (COR)**	**DC-based support**	**Design Specialist**	**DG Specialist, team leader, or mission-based expert**	**Office Director /Deputy Office Director**	**OAA Agreement Officer/Contracting Officer or Specialist**
	45.5% (40)	6.8% (6)	5.7% (5)	13.6% (12)	20.5% (18)	8% (7)

Figure 12: Characteristics of Activities Solicited Under Acquisition



Of the 23 of activities that were solicited under acquisition, 10 were a Statement of Objective, 11 were Statements of Work, 1 was Unsolicited, and 1 was an Indefinite Delivery, Indefinite Quantity contract. Additionally, eight were from follow-on activities. Of these eight, three were similar to the predecessor activity and five were different.

Figure 13: Characteristics of Activities Solicited Under Assistance



Of the activities, 59 were solicited under assistance. The most common mechanism was a Program Description (34); the others were Leader with Associate subaward (11), Annual Program Statement, and co-creation under a broad agreement (5). From the 59 activities, 15 were follow-ons to previous programming. Twelve of these follow-on activities were similar to their predecessors.

Figure 14: Percentage of Respondents Listing Each Source as a Significant Influence on the Activity Design

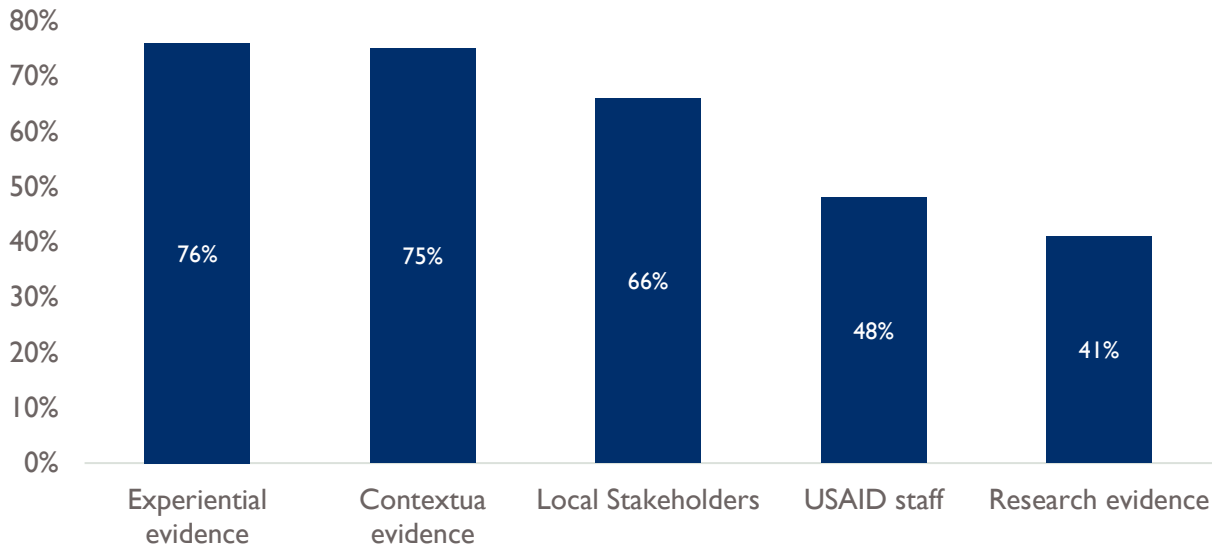


Figure 15: Percentage of Respondents Listing Each Type of Research Evidence Used

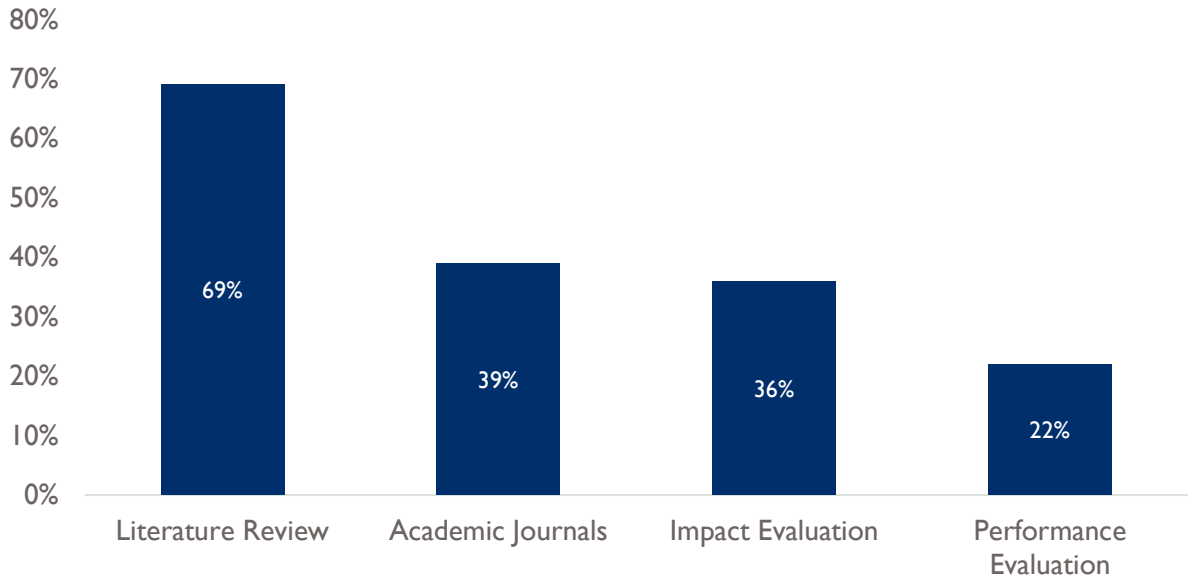


Figure 16: Responses to "Why was research evidence used?"

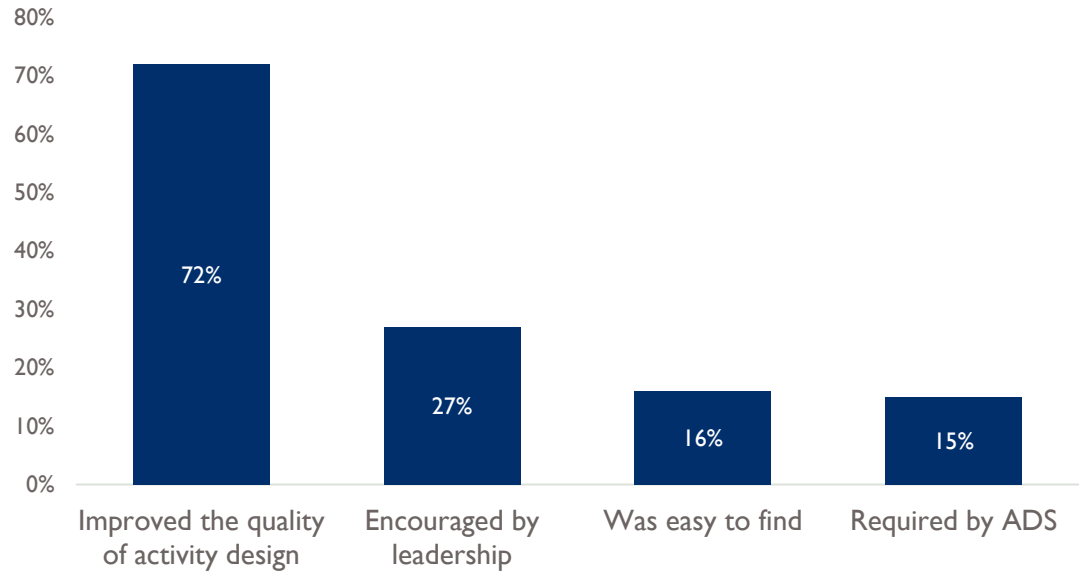


Figure 17: Responses to "Why is research evidence not used more often to inform activity design?"

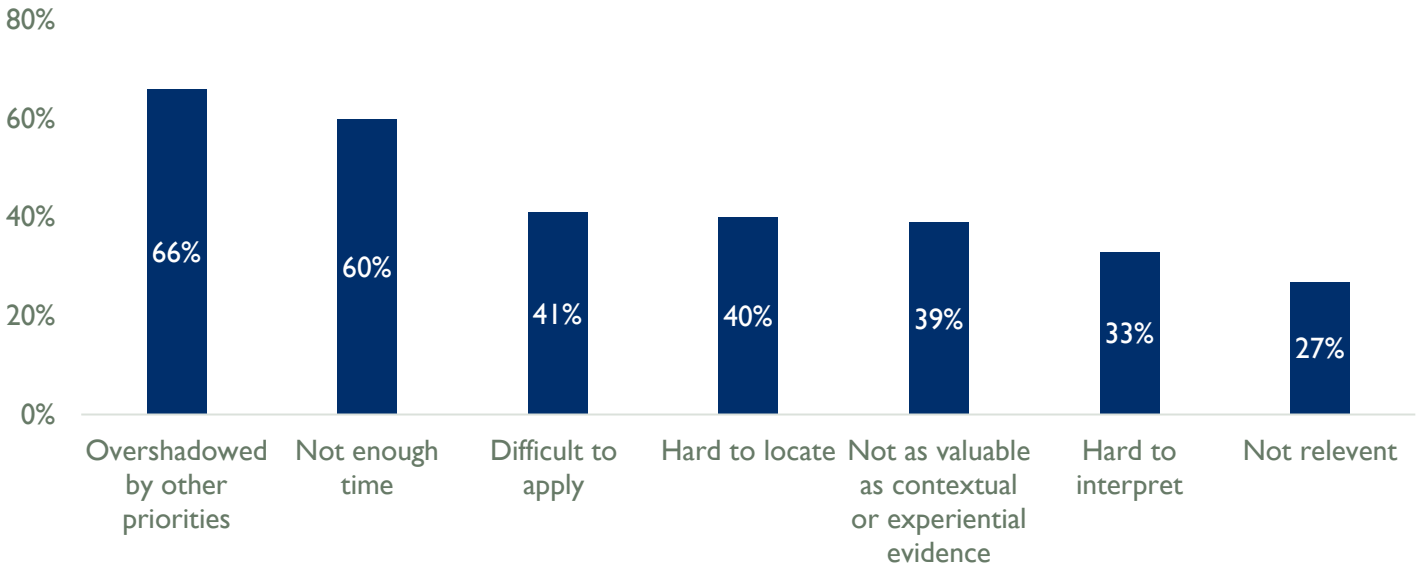


Figure 18: Responses to "In general, would you say that your Mission leadership encourages or discourages the use of research evidence (e.g., impact evaluations, systematic reviews, academic publications)?"

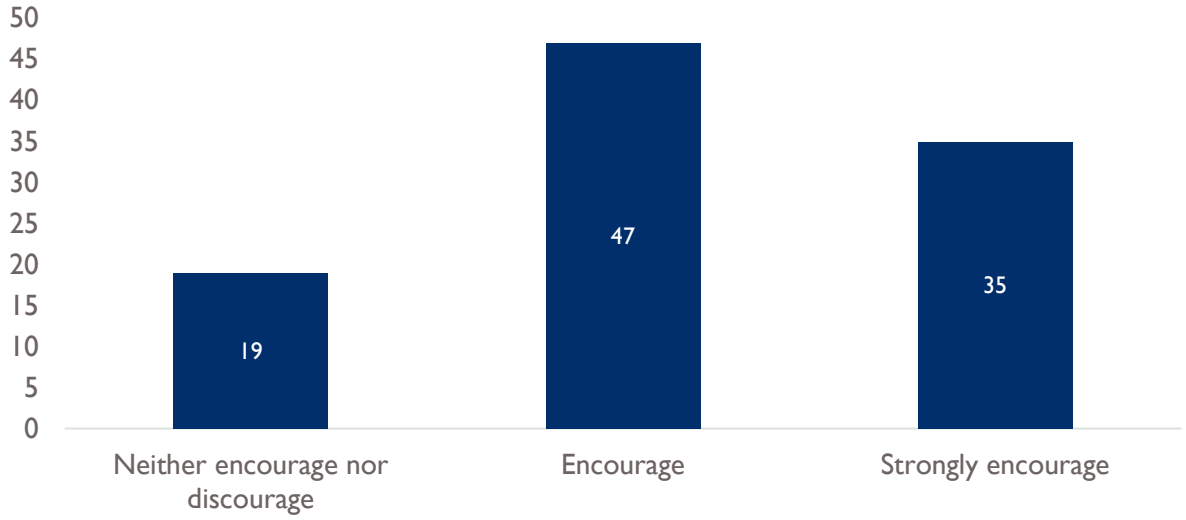


Figure 19: Responses to "In general, would you say that the USAID DDI/DRG and/or USAID/Washington helps to enable the use of research evidence in activity design?"

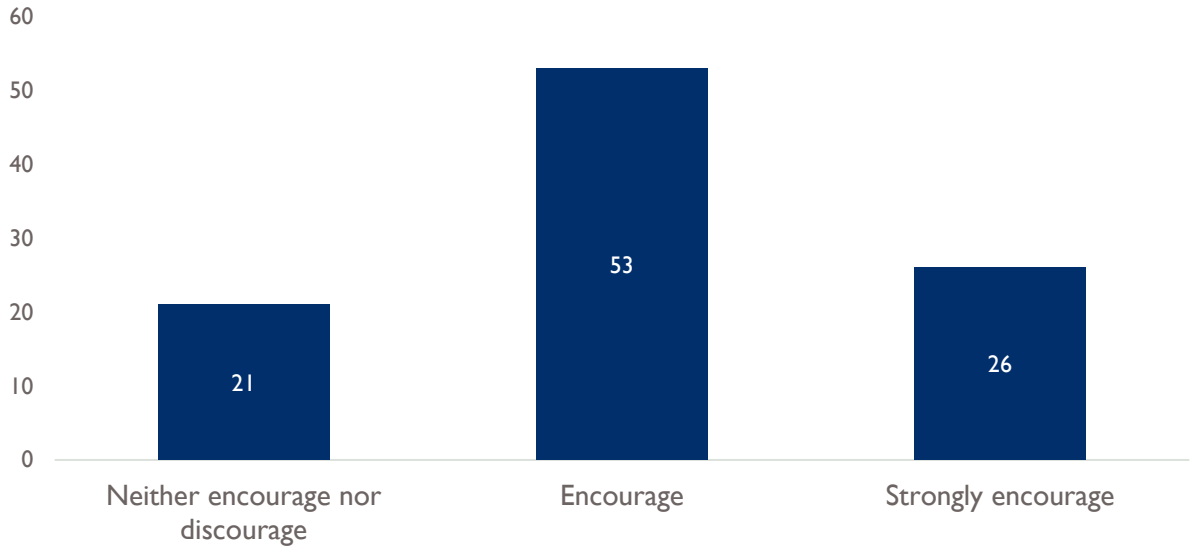


Figure 20: Responses to "During solicitation and before award, did the activity require Offerors to propose an evidence-based approach?"

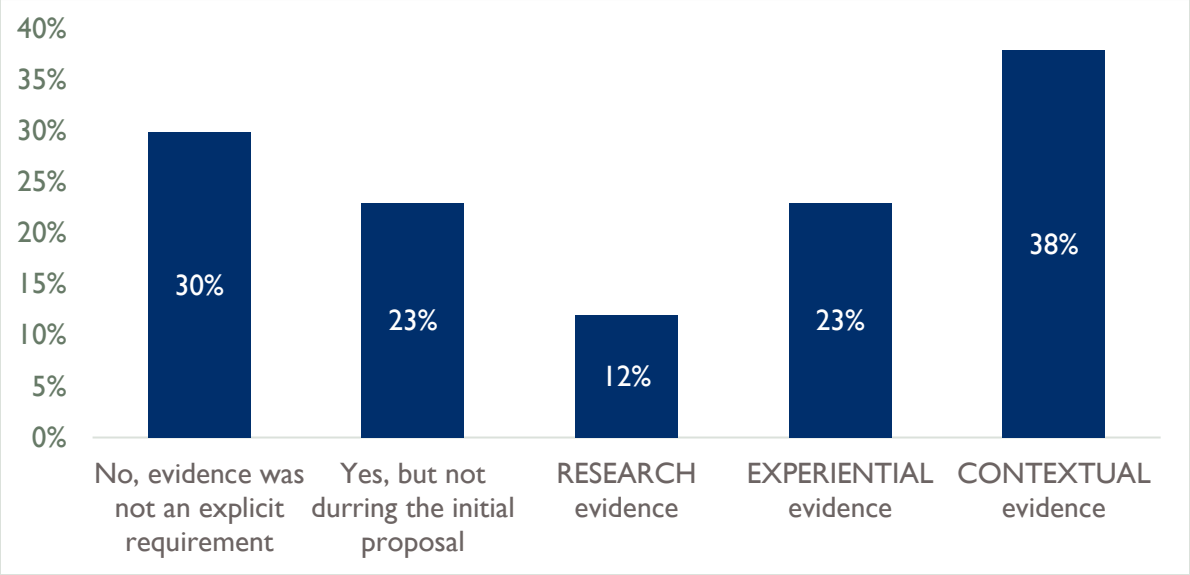


Figure 21: Responses to "What implementing partner-led evidence generation was planned for in your most recent activity design?"

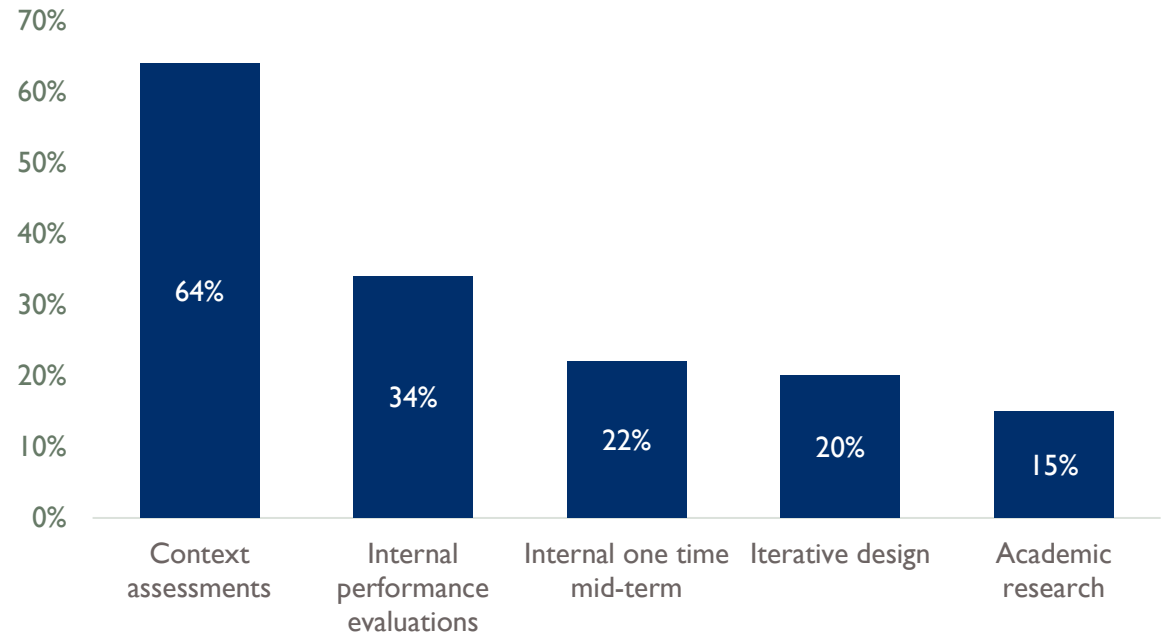


Figure 22: Responses to "What, if any, externally led evidence generation was planned for in your most recent activity design?"

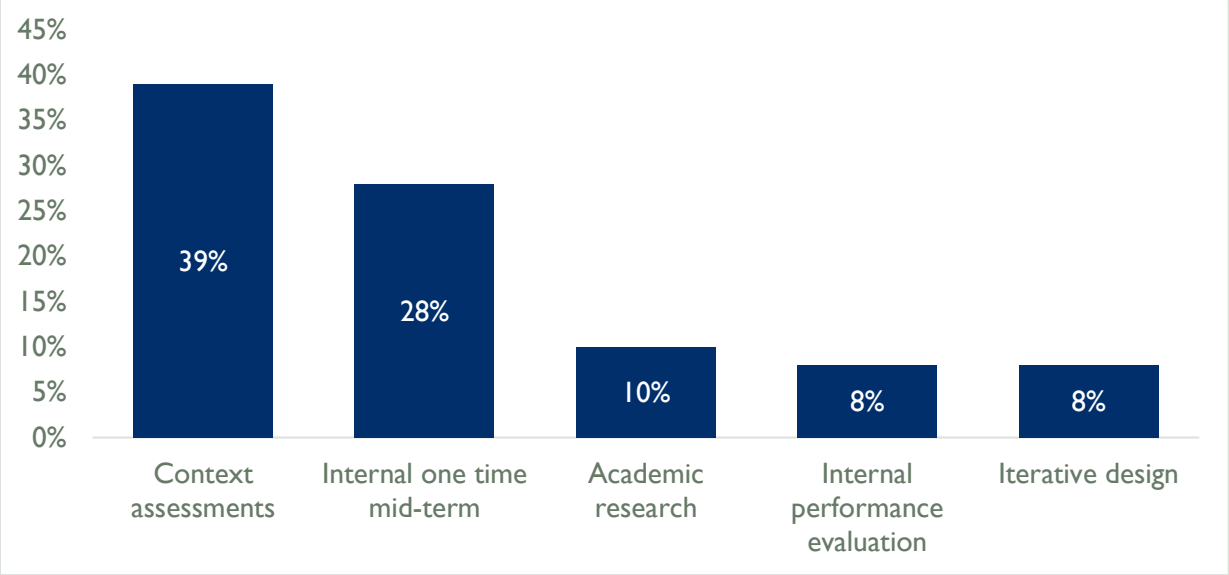


Figure 23: Responses to "Did you consider planning an impact evaluation or performance evaluation with a baseline and endline?"

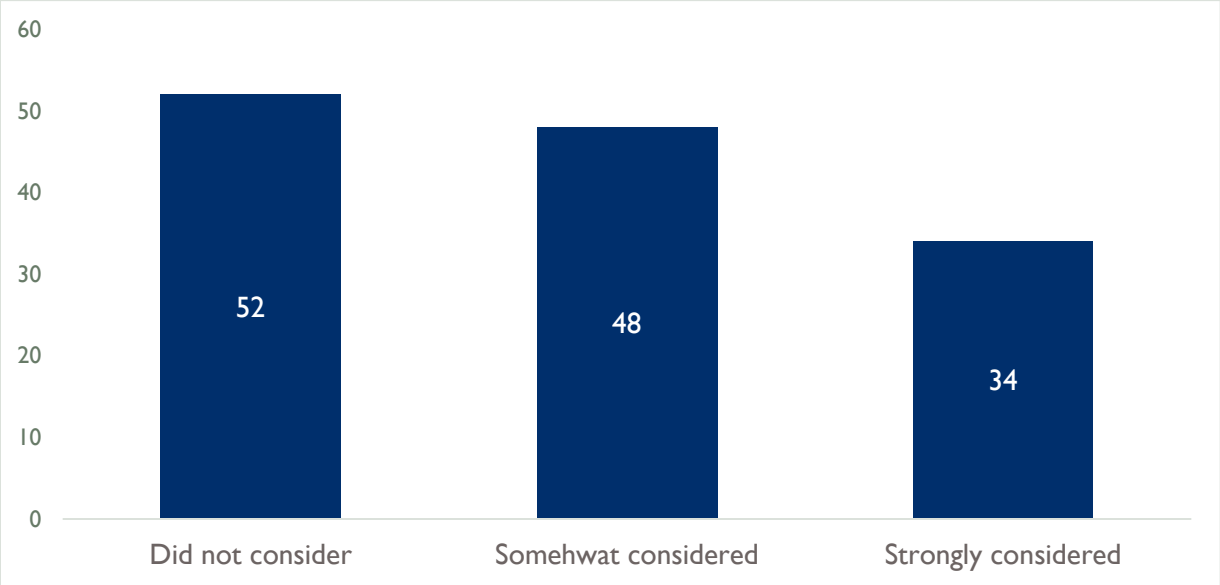
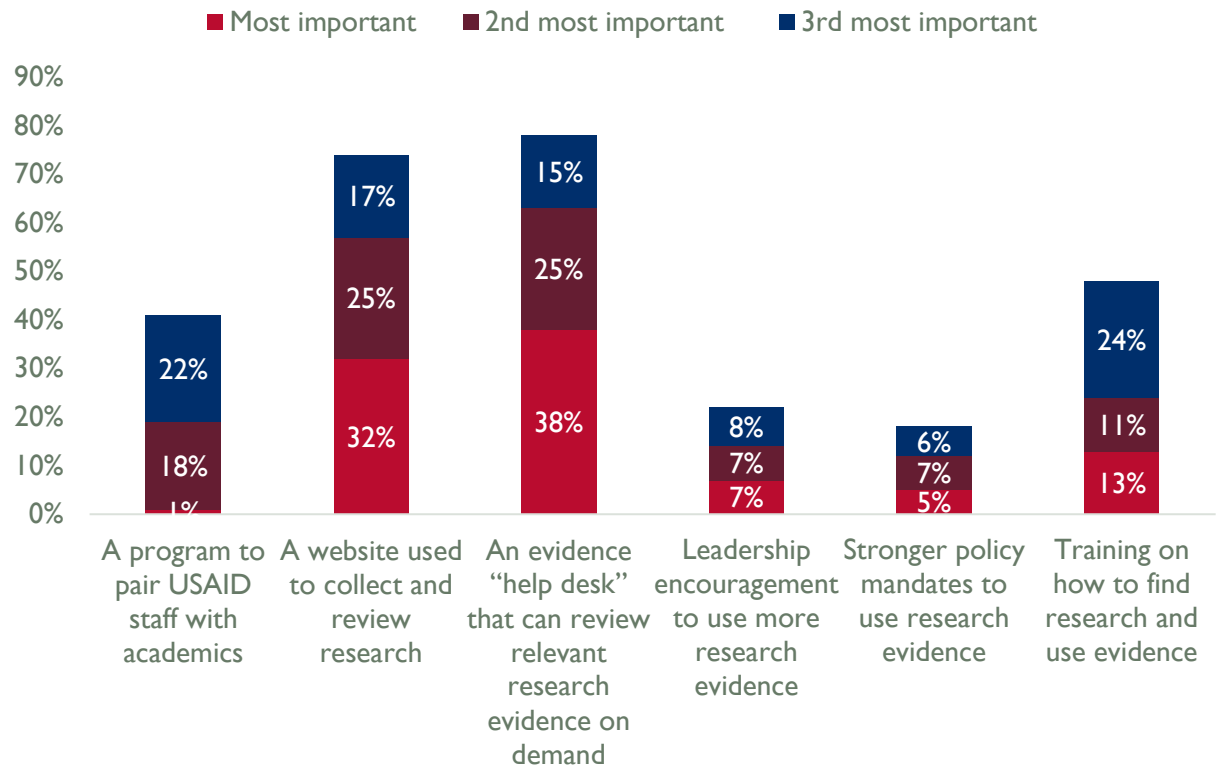


Figure 24: Which of the following initiatives to encourage the use of research evidence would you be highly likely to utilize or support? Number of respondents listing each option as the first, second, or third choice.



ANNEX 4. QUANTITATIVE SURVEY INSTRUMENT

UNDERSTANDING THE ACTIVITY DESIGN PROCESS

CONFIDENTIALITY NOTICE

This survey is being conducted on behalf of USAID's Center for Democracy, Human Rights, and Governance (DRG). This survey is intended for use by and for USAID employees and the global cadre. It will be used to inform future support and result in a report that will be shared with you, disseminated internally, and made publicly available on the Development Experience Clearinghouse. Please do not include any personally identifying information in any open-ended questions.

You are invited to participate in this short 10–15 minute survey on the design process because of your recent experience in designing an activity for USAID. Your responses are confidential and individual responses will not be cited in the report. Your participation is also voluntary. Your participation will help ensure adequate representation of your views in the final results and outcomes. If you agree to participate, you may withdraw your participation in the survey at any time by simply exiting the survey. For more information, please contact Daniel Sabet dsabet@usaid.gov. Please note that your responses will not be submitted until you select the "Submit" button on page 11.

* Required

1. Email* _____

Background and Demographics

2. What best describes your role at USAID?

Mark only one oval.

◯ DC based support for missions
◯ Office Director (OD)
◯ Deputy Office Director
◯ Design Specialist
◯ Agreement Officer's Representatives (AOR)/Contracting Officer's Representatives (COR)
◯ Office of Acquisition and Assistance Agreement Officer/Contracting Officer or Specialist
◯ Other: _____

3. In total, how many years of experience do you have with Program and/or activity Design?

_____

4. What is your hiring mechanism?

Mark only one oval.

◯ Foreign Service National/Cooperating Country National

◯ Foreign Service Officer
◯ Civil Servant
◯ Personal Services Contractor (offshore, local, or Third Country National)
◯ Non-USG partner organization staff
◯ Other: _____

5. Gender

Mark only one oval.

◯ Female
◯ Male
◯ Non-Binary
◯ Decline to State
◯ Other: _____

What activity design have you most recently meaningfully been involved with?

In the following sections we'll ask you some questions about this design process. If you have not been involved in a design process in the last four years, you need not take the survey.

6. Activity Full name:

_____

7. Activity Acronym:

_____

8. What is the name of the Mission or Bureau that hosts the activity?

_____

9. Who is the Implementing Partner? (Please indicate if the IP is a local partner)

_____

10. Was the activity solicited under assistance or acquisition?

Mark only one oval.

◯ Assistance *Skip to question 11*
◯ Acquisition *Skip to question 12*

Skip to question 13

ASSISTANCE

Please indicate the type of assistance mechanism prepared.

11. Please indicate the type of assistance mechanism prepared.

Mark only one oval.

- ◯ Program Description (PD)
- ◯ Leader With Associates (LWA) sub award
- ◯ Co-creation under a Broad Agency Announcement or similar
- ◯ Annual Program Statement
- ◯ Unsolicited proposal
- ◯ Other: _____

Skip to question 13

12. Please indicate the type of acquisition mechanism prepared.

Mark only one oval.

- ◯ Statement of Objectives (SOO) Statement of Work (SOW)
- ◯ Indefinite Delivery, Indefinite Quantity Contracts (IDIQ) Task Order
- ◯ Blanket Purchase Agreement or Purchase Order
- ◯ Unsolicited Proposal
- ◯ Other: _____

13. Was the activity a follow-on activity?

Mark only one oval.

- ◯ Yes *Skip to question 14*
- ◯ No *Skip to question 15*
- ◯ Do not know *Skip to question 15*

Skip to question 15

14. How similar or different would you say the new activity was to the predecessor activity?

Mark only one oval.

- ◯ Very similar
- ◯ Similar
- ◯ Different
- ◯ Very different

USE OF EVIDENCE EXPERIENCE

15. In thinking about this activity design, prior to the solicitation, which of the following had a significant influence on the activity design? (Please select all that apply.)

Check all that apply.

- ☐ Influence of particular USAID staff Local Stakeholder input
- ☐ Experiential evidence (e.g., lessons learned exercises, performance evaluations from past activities, reporting/documentation from previous activities)
- ☐ Contextual evidence (e.g., political economy analysis, assessment work)
- ☐ Research evidence (e.g., impact evaluations, systematic reviews, academic publications)
- ☐ Replication from predecessor award Influence from the host country government
- ☐ Other: _____

16. Please rank the top three most influential sources on the program/activity design

Mark only one oval per row.

	PARTICULAR USAID STAFF, INCLUDING YOURSELF	LOCAL STAKEHOLDER INPUT	EXPERIENTIAL EVIDENCE	CONTEXTUALEVIDENCE	RESEARCH EVIDENCE	REPLICATION FROM PREDECESSOR
Most Influential	○	○	○	○	○	○
Second Most Influential	○	○	○	○	○	○
Third Most Influential	○	○	○	○	○	○

17. What type[s] of research evidence was used in the design of the activity?

Check all that apply.

- ☐ Impact Evaluation (in-country)
- ☐ Impact Evaluation (from another country)
- ☐ Academic journal article[s]
- ☐ Rigorous performance evaluation
- ☐ Literature or evidence review
- ☐ Systemic review
- ☐ Research Evidence was not used to inform the design
- ☐ Other: _____

18. Why did you use research evidence in the design of the activity? (Select all that apply)

Check all that apply.

- ☐ Leadership encouraged the use of research evidence
- ☐ ADS requires it
- ☐ It was easy to find
- ☐ It improved the quality of activity design
- ☐ Mission/office-based incentives or accolades
- ☐ I attended a training that showed me how to use research evidence in activity design
- ☐ Research Evidence was not used to inform the design
- ☐ Other: _____

19. Which of the following USAID staff had a significant influence on the activity design? (Select all that apply)

Check all that apply.

- ☐ Contracting or Agreement Officer
- ☐ Washington DC based staff
- ☐ The Planner and members of the design team
- ☐ Myself (please indicate which role you played in the other box)
- ☐ Front Office leadership
- ☐ Resident Legal Officer
- ☐ Program Office staff
- ☐ Washington DC based staff
- ☐ None of the above
- ☐ Other: _____

20. During solicitation and before award, did the activity require Offerors to propose an evidence-based approach?

Check all that apply.

- ☐ Yes, they had to demonstrate how RESEARCH evidence informed their proposed approaches
- ☐ Yes, they had to demonstrate how EXPERIENTIAL evidence informed their proposed approaches
- ☐ Yes, they had to demonstrate how CONTEXTUAL evidence informed their proposed approaches
- ☐ Yes, but not during the initial proposal. The Offeror could establish a baseline during its start up phase (after award)
- ☐ No, evidence was not an explicit requirement.
- ☐ Other: _____

Evaluation Planning

Next we would like to ask you about any planning you did at the design stage to prepare for future evidence generation or external evaluations

21. What implementing partner-led evidence generation was planned for in your most recent activity design? Please include all that apply. Note: Evaluations conducted by an implementer of their own programs are considered "internal evaluations."

Check all that apply.

- ☐ Context assessments (e.g., Political Economy Assessment, Systems assessments)
- ☐ Iterative design, rapid prototyping, human centered design, etc.
- ☐ Internal one time mid-term OR final performance evaluation
- ☐ Internal performance evaluation including a baseline and endline
- ☐ Internal impact evaluation
- ☐ Academic research
- ☐ None of the above

22. What, if any, externally led evidence generation was planned for in your most recent activity design? Please include all that apply. Note: Evaluations led by third-party evaluators are considered "external evaluations."

Check all that apply.

- ☐ Context assessments (e.g., Political Economy Assessment, Systems assessments)
- ☐ Iterative design, rapid prototyping, human centered design, etc. (e.g., developmental evaluation)
- ☐ External one time mid-term or final performance evaluation
- ☐ External performance evaluation including a baseline and endline
- ☐ External impact evaluation
- ☐ Academic research
- ☐ None of the above

23. Did you consider planning an impact evaluation or performance evaluation with a baseline and endline?

Mark only one oval.

- ◯ Strongly considered
- ◯ Somewhat considered
- ◯ Did not consider
- ◯ Not applicable

24. If you planned for a one time mid-term or final evaluation rather than an impact evaluation or a performance evaluation with a baseline and endline, why did you make that determination?

Check all that apply.

- ☐ Cost considerations
- ☐ It met the evaluation requirements
- ☐ It was the best fit for our evaluation needs
- ☐ It was the best fit for the nature of the agreement/contract
- ☐ It is what we have traditionally done at this Mission
- ☐ Direction from Program Officer and/or Monitoring & Evaluation lead
- ☐ Direction from my supervisor
- ☐ It was already planned for in a higher level strategy, (e.g., Country Development Cooperation Strategy)
- ☐ Not Applicable and/or we did not build in an evaluation requirement
- ☐ Other: _____

General experience with activity design

This section asks questions that are about the activity design process in general. Please respond to these questions based on your cumulative experience with the design process.

25. In general, would you say that your Mission leadership encourages or discourages the use of research evidence (e.g., impact evaluations, systematic reviews, academic publications)?

Mark only one oval.

- ◯ Strongly encourages
- ◯ Encourage
- ◯ Neither encourage nor discourage
- ◯ Discourage
- ◯ Strongly discourage

26. In general, would you say that the USAID DDI/DRG and/or USAID/Washington helps to enable the use of research evidence in activity design?

Mark only one oval.

- ◯ Strongly encourages
- ◯ Encourage
- ◯ Neither encourage nor discourage
- ◯ Discourage
- ◯ Strongly discourage

27. Below are some common reasons USAID staff give as to why research evidence (e.g., impact evaluations, systematic reviews, academic publications) is not used more often to inform activity

design. For activity design in general, across all of your experiences, please rank each obstacle below from 1 to 5 where 1 is NOT AN OBSTACLE at all and 5 is a MAJOR obstacle:

Mark only one oval per row.

	NA	1 (NOT AN OBSTACLE)	2	3	4	5 (MAJOR OBSTACLE)
Known/available research evidence is not relevant to my country/programmatic context	○	○	○	○	○	○
Research evidence is not available when I need it	○	○	○	○	○	○
Research evidence is hard to locate	○	○	○	○	○	○
Research evidence is hard to interpret	○	○	○	○	○	○
The findings from research evidence are difficult to apply	○	○	○	○	○	○
There is not enough time in the activity design cycle to find, read, and absorb research evidence	○	○	○	○	○	○
Research evidence is not as valuable as contextual or experiential evidence	○	○	○	○	○	○
Gathering and using research evidence is overshadowed by other priorities	○	○	○	○	○	○

28. Which of the following initiatives to encourage the use of research evidence would you be highly likely to utilize or support? Below is a list of approaches to encourage the use of research evidence in USAID's activity design processes. Please rank the top three you believe would most likely help design teams complete future DRG activity designs:

Mark only one oval per row.

	A WEBSITE THAT COLLATES RESEARCH EVIDENCE RELEVANT TO OUR WORK.	AN EVIDENCE "HELP DESK" THAT CAN REVIEW RELEVANT RESEARCH EVIDENCE ON DEMAND	A PROGRAM TO PAIR USAID DRG STAFF WITH ACADEMIC SPECIALISTS	TRAINING ON HOW TO FIND AND USE RESEARCH	LEADERSHIP ENCOURAGEMENT TO USE MORE RESEARCH EVIDENCE	STRONGER POLICY MANDATES RESEARCH EVIDENCE	NONE WOULD ENCOURAGE RESEARCH EVIDENCE
Most Likely	◯	◯	◯	◯	◯	◯	◯
Second Most Likely	◯	◯	◯	◯	◯	◯	◯
Third Most Likely	◯	◯	◯	◯	◯	◯	◯

29. In your opinion, how could evidence and evaluation planning be improved? (Please do not include any personally identifiable information in your response)

_____

_

30. If you have any other comments about the use of evidence or evaluation/evidence planning we would love to hear more from you. (Please do not include any personally identifiable information in your response

_____

_

31. In case of any follow up questions, can we contact you for more information?

Mark only one oval.

◯ Yes
◯ No

ANNEX 5. RETROSPECTIVE QUESTION THEMES

INTRODUCTION (5–10 MINS)

- **Purpose:** An evaluation of USAID's use of evidence and research in activity and project design. This interview will shape the design of our quantitative instrument, which will be sent to the DRG cadre. Interviews will inform the selection of case studies. The report will be released next spring.
- **Introductions:** Interviewers.
- **Duration:** Interviews should last about 1 hour.
- **Participation is voluntary:** Questions can be skipped or the interview ended at any time.
- Interviews are auto transcribed; the RT will use the transcripts to flesh out notes. Transcripts will not be shared outside the RT.
- Results of this interview may be used for the evaluation, and the report will be made publicly available online.
- If quotes are used in the report, they will not be attributed to names or position titles.

We tend to see three categories of evidence used at USAID: a) experiential (going from the gut); b) contextual (PEAs, needs assessments, etc.); c) and evaluation/empirical evidence...can you tell us about your experience with these and how they've influenced activity designs you've worked on or know of?

a. Experiential (going from the gut)
b. Contextual (PEAs, needs assessments, etc.)
c. Evaluation/empirical evidence

1. *How has each type of evidence influenced activity designs you've worked on or know of?*
2. *What do you think the role of research evidence should be in the activity design process?*
3. *What are the headaches when searching for relevant evidence?*
4. *What are the headaches when applying relevant evidence?*
5. *When and/or how do Missions incorporate evaluation planning in the activity design process?*
6. *We would love to hear some of the best practices in these areas that the DRG Center should be promoting and replicating as it helps Missions with their activity designs:*

    a. Incorporating evidence
    b. Incorporating research evidence specifically
    c. Evaluation planning

ANNEX 6. RFP LANGUAGE REQUIRING OR ATTEMPTING TO REQUIRE EVIDENCE

SOW LANGUAGE IN SECTIONS L OR M	NOTE
"This document should provide more detail highlighting how the interventions described in the SOW will result in the results USAID wants to achieve, using strong, evidence-based justifications for why the Offeror believes that these are the most effective means to achieve the expected outcomes."	Preferred example of simple, easy-to-understand language. Uses the term "evidence-based" but is also clear about what evidence is sought.
"The Offeror must prepare an evidence-based and results-driven Performance Based Work Statement/Technical Approach that responds to Section C of this RFP. The [Performance Based Work Statement] will: Describe a politically informed, evidence-based technical success approach that addresses each of the three objectives described in Section C…"	The text clearly requires an "evidence-based approach," but there is a risk that the term by itself could be broadly understood and the example above provides a better model.
"The Offeror must provide the analysis and rationale for selecting the targeted organizations as well as propose the selection criteria and a transparent process for selecting what is estimated to be 5–10 mid-level local advocacy NGOs."	Example of requiring evidence in targeting but does not mention evidence supporting the approach.
"The technical approach must articulate the Offerors' understanding of [country] political economy dynamics and the problems [the intervention] seeks to address, and outline possible program pathways, including illustrative interventions and focus areas. Proposals must describe an overall strategy or approach outlining promising entry points and initial programming approaches that demonstrate a nuanced understanding of [country] political economy and propose initial programming."	Example of PEA-focused language that attempts to draw a relationship between the PEA and using context evidence to inform proposed approaches.
"...the soundness of the conceptual approach and the general strategy being proposed for the implementation of each technical component, based on key lessons learned from past and current activities…"	Example of requiring experiential evidence but limited to this form of evidence.
"Demonstrates an in-depth understanding of the development and political contexts impacting domestic revenue mobilization in [country], including the risks to completion of the results."	Limits itself to simply requiring an understanding of the DRG problem. This was the most common evidentiary request.
"The extent to which the Offeror's proposed Performance Work Statement and Technical Approach demonstrates an understanding of current local and regional context and development challenges and convincingly presents a comprehensive, clear, and realistic strategy for achieving the objectives of the SOO."	Potentially problematic language as it only requires "face validity" of the approach rather than evidence.