INTRODUCTION

With the enactment of new legislation such as the Foreign Aid Transparency and Accountability Act[1] and the Foundations for Evidence-Based Policymaking Act[2], there is a renewed focus on the importance of conducting evaluations within the United States Government (USG). The Department of State (DoS) has revised its Evaluation Policy (18 FAM 300[3]) to emphasize the utilization of monitoring and evaluation (M&E) data in program design efforts, with each DoS Bureau being required to conduct its program design, M&E, and learning activities in compliance with the new policy. The Bureau of Educational and Cultural Affairs (ECA) has refined its practices to create a comprehensive M&E system that will meet the demands of the revised DoS Evaluation Policy and ensure ECA program teams and senior leadership benefit from timely, credible evidence that can be utilized for evidence-based decision-making.

THE BUREAU OF EDUCATIONAL AND CULTURAL AFFAIRS

The ECA Bureau builds friendly, peaceful relations between the people of the United States and the people of other countries through academic, cultural, sports, and professional exchanges, as well as public-private partnerships. Given the global nature of the Bureau's activities as well as the number and types of stakeholders that are engaged in programming, the nature of the outcomes and impact are contextually dynamic and complex in nature.

The ECA Bureau's Monitoring Evaluation Learning and Innovation (MELI) Unit is the primary entity tasked with conducting M&E to assess the effectiveness of ECA programs. As the first M&E entity established in DoS, ECA's MELI Unit has been at the forefront of the Department's M&E efforts since 1999. To date, ECA evaluations have primarily utilized mixed methods to obtain data on ECA's long-term results to inform ECA program teams and senior leadership of program performance and areas for improvement, as well as the merit and worth[4] of the Bureau's programs.

[1] https://www.congress.gov/bill/114th-congress/house-bill/3766/text
[2] https://www.congress.gov/bill/115th-congress/house-bill/4174
[3] https://fam.state.gov/fam/18fam/18fam030104.html
[4] **Merit** refers to the *intrinsic* value of a program; how effective a program is in meeting the needs of those it is intended to help. **Worth** refers to *extrinsic* value to those outside the program such as the larger community or society (for example, employers being happy to have former ECA exchange participants they can hire who have relevant skills). (Mertens & Wilson, p.7)

The MELI Unit takes seriously the renewed emphasis on strengthening M&E practices within the U.S. Government (USG), and has set out to better define its stance on M&E. Recognizing that decisions about how an evaluation will be conducted begin not with the design of the evaluation and which methodology will be used, but rather with the background and worldviews of the evaluator and how these elements impact their thinking on evaluations, the ECA MELI Unit has rooted its practices in evaluation theory. This paper outlines the foundational philosophy undergirding the processes for how and in what ways the ECA MELI Unit approaches and conducts its M&E activities.

EVALUATION THEORY AND PARADIGMS

As evaluation theorist William Shadish stated: "All approaches to evaluation involve tradeoffs among the many goals we try to maximize in evaluation… so we need to have conceptual tools to help us understand those tradeoffs" (Shadish, p. 8). Evaluation theory and the overarching evaluation paradigms provide the foundation for understanding the range of evaluation designs, methods, and choices available to evaluators, and can guide evaluators to make thoughtful decisions in selecting from among the many options available.

EVALUATION THEORY

In William Shadish's 1997 presidential address during the American Evaluation Association (AEA) annual conference he stated "evaluation theory is who we are in the sense that it is the knowledge base that defines the profession. It is what we know that other professions don't know. To be an evaluator, you need to know that [evaluation] knowledge base that makes the field unique" (Shadish, p.5). While Shadish argues the importance of evaluation theory, there are multiple classifications of evaluation theory that exist. Shadish, Thomas Cook, and Laura Leviton have determined that evaluation theories should include five components: practice, knowledge, value, use, and social programming resulting in "the fundamental purpose of program evaluation theory [being] to specify feasible practices that evaluators can use to construct knowledge of the value of social programs that can be used to ameliorate the social problems to which programs are relevant" (Shadish, et. al., p.36). Marvin Alkin and Christina Christie have outlined evaluation theory as including three main dimensions (use, methods, and valuing), which they depict in the form of a tree with each of the dimensions serving as branches which match to their respective roots (social accountability, social inquiry, and epistemology) (Alkin & Christie, p.12). Regardless of the specific definition offered, evaluation theory is meant to specify what a good evaluation is and guide evaluators in how evaluation should be conducted.

While not explicitly utilized in daily decision-making, knowledge of evaluation theory interacts with other dimensions of evaluation to assist evaluators in framing how evaluation is practiced. Specifically, "[evaluation design] should interact with theory by forcing reconsideration of: (1) the kind of evaluation that may be feasible based on what has been learned about the program context and especially its history; (2) the types of evaluation questions that will be possible to answer; and (3) the methods – individual or combined – that are appropriate" (Chelimsky, p.

93). Evaluation theory reinforces good practice and assists in the planning and execution phases of an evaluation.

However, evaluation theory alone does not drive evaluation practice. Evaluation is informed by numerous dimensions (including an evaluator's background, experience level, and the context in which they work) all of which combine to shape how an evaluator makes interpretive judgments about programs and policies being evaluated (Chouinard, et. al., p.493). Social Science Inquiry Paradigms are made up of multiple theories (evaluation theory, social science theory and research, and program theory) as well as four sets of philosophical assumptions that address what is ethical (axiology), what is reality (ontology), what is knowledge (epistemology), and what is considered to be appropriate systematic approaches to gathering information (methodology) (Mertens and Wilson, p.38). Evaluation theories are broad constructs inclusive of these elements that work in concert to form how an evaluator conducts, or should conduct, evaluations.

Evaluation Paradigms

Four major social science paradigms include: Postpositivist, Pragmatic, Constructivist, and Transformative. Each paradigm is constructed of varying theories and assumptions. The foundation for an evaluator's philosophical assumptions is what determines which paradigm will be most closely followed (or aligned with) and, consequently, how evaluations will be conducted. The four philosophical assumptions are:

- Ontology – The nature of reality. An evaluator's ontological assumptions center around there being one reality that can be discovered versus multiple realities that differ depending on the experiences and conditions of the stakeholders in the specific context (Mertens and Wilson, p.38).
- Epistemology – The nature of knowledge and the relationship between the evaluator and stakeholders in an evaluation. Epistemology shapes the researcher's conceptualization of the participant during the data collection and analysis phases of an evaluation and determines how an evaluator communicates with his or her audience, affecting the nature of reporting. Differing epistemological positions will be drawn to different methodologies and different variants of the same methodology (Carter & Little, p.6).
- Methodology – The nature of gathering credible and actionable evidence. "Methodology influences the objectives, [evaluation] questions and design, and thus have a profound effect on the implementation of methods used to conduct an evaluation" (Carter & Little, p.13).
- Axiologoy – The nature of ethics. Axiology is shaped by the society, culture, religion, beliefs, and values which have developed the principles and ideas you believe (Mertens and Wilson, p.38). "Evaluators' personal sense of what is ethical and correct shape a fair amount of the way in which many go about their practice and subsequently arrive at conclusions about what they have been charged to evaluate" (Alkin et. al., p.37).

FOUR EVALUATION PARADIGMS

It is critical that evaluators and evaluation stakeholders understand the four paradigms and how they differ. There are a range of methods and perspectives that form a large toolkit of evaluation approaches that are deemed acceptable in the evaluation field. The knowledge of the similarities and differences of the paradigms assist evaluation stakeholders (and evaluators themselves) to better understand the foundation of evaluation practice and properly place assumptions and/or criticisms about an evaluator's philosophy and how they choose to conduct evaluation in the proper context (Donaldson & Lipsey, p.61). This section briefly describes the evaluation paradigms in relation to the philosophical assumptions and the role of the evaluator.

Postpositivist Paradigm/Methods Branch. The field of evaluation emerged during the Johnson Administration's Great Society programs in the 1960s when Congress expressed interest in understanding if these new programs were causing the expected effects while avoiding negative unintended consequences (Shadish, et. al., p.22). The early knowledge base of evaluation borrowed heavily from scientific experiments, which saw a predominance of quantitative approaches to data collection, primarily randomized control trials (RCT, which is the random selection of subjects and random assignment to be a part of the program or not) and quasi-experimental designs (which lack the random assignment of an RCT, but still provide a non-treatment group to compare results to) and sophisticated statistical analyses. Postpositivists hold the "belief that one reality exists and that it is independent of the observer, and that the evaluator's distance from the object of study contributes to reducing bias in research" (Mertens & Wilson, p.55). A "valid and rigorous" research design is what drives those guided by the Postpositivist Paradigm, which leads them to view the use of quantitative methods as superior to other methodological options.

Pragmatic Paradigm/Use Branch. Subscribers to the Pragmatic Paradigm focus on the need to be responsive to stakeholders/decision-makers to foster use of evaluation findings and recommendations. The main axiological assumption underlying the work of Pragmatists holds that the value of something is a function of its consequences, while the epistemological standpoint is that the relationship of the evaluator to the participants and the level and type of interaction between them depends on "how well that relationship allows you to achieve your purpose in the evaluation" (Mertens and Wilson, p.86) – with the purpose being the utilization of the findings and recommendations. The ontological viewpoint of pragmatists is that there is no absolute truth concerning reality as there are multiple explanations of reality and, at any given time, one explanation of reality may be truer than another (Christie & Fleischer, p34). Where postpositivists prefer to use quantitative methods, evaluators associated with the Pragmatic Paradigm match their actions to the underlying purpose of the study – with mixed methods (both qualitative and quantitative methods) being the most common type utilized.

Constructivist Paradigm/Values Branch. Michael Scriven, one of evaluation's earliest theorists, places an emphasis on the role of valuing in evaluation stating: "bad is bad and good is good and it is the job of evaluators to decide which is which. Evaluation research must produce as a conclusion…a judgement of value, worth, or merit" (Scriven, p.74). This concern with valuing is primary to those that adhere to the Constructivist Paradigm. "Constructivists believe that there

are multiple, socially constructed realities uncovered through reflection upon their experiences and in interaction with others" (Mertens & Wilson, p.132). It is the role of the evaluator to uncover these realities through dialogue and close and continuing contact with the participants in an evaluation. Methodologically, evaluators on the Values Branch predominantly use qualitative methods to capture the diversity of stakeholders' experiences and the implications of different values for determining how to interpret those findings (Mertens & Wilson, p.126).

Transformative Paradigm/Social Justice Branch. The Transformative Paradigm places addressing inequities and promoting social justice as the primary principle guiding evaluators (Mertens & Wilson, p.157). The context of power and privilege and how power relations are shaped in cultural, social, and historical contexts are the lens through which a Social Justice Branch evaluator sees the world (Mertens & Wilson, p.157). To uncover these power dynamics and the contexts through which they are shaped, evaluators are expected to build trust and develop close and collaborative relationships with participants and other community stakeholders. In this Paradigm, the focus is on engaging all stakeholders – particularly those who have traditionally been excluded from decision-making. While no single methodology is associated with the Transformative Paradigm, mixed methods are most commonly utilized.

The chart below displays information on each of the four paradigms[5]:

	Postpositivist Paradigm (Methods)	**Pragmatic Paradigm (Use)**	**Constructivist Paradigm (Values)**	**Transformative Paradigm (Social Justice)**
Description (Primary focus)	The use of statistical methods of inquiry (quantitative designs and data)	Data that are found to be useful for stakeholders; advocates for mixed methods	Identify multiple values and perspectives through qualitative methods	Take viewpoint of marginalized groups and challenge power structures to further social justice
Axiology	Beneficence, respect, and justice – through the conduct of valid scientific procedures	Gain knowledge to show that results answer the question being asked	Evaluator is aware of others' values as well as their own	Cultural respect, promotion of human rights, addressing inequities
Ontology	One reality exists and it is independent of the observer	Multiple explanations of reality exist and all individuals have their own unique interpretation	Multiple, socially constructed realities uncovered through reflection and interaction	Different perspectives based on diverse values and life experiences; best understood in terms of power relations

[5] This chart has been adapted from numerous sources cited throughout this paper

	Postpositivist Paradigm (Methods)	Pragmatic Paradigm (Use)	Constructivist Paradigm (Values)	Transformative Paradigm (Social Justice)
Epistemology	Evaluator remains objective through minimal interaction during the study	Relationships determined by what the evaluator deems as appropriate to the study	Meaningful dialogue and reflection used to create knowledge	Knowledge constructed within a context of power and privilege
Methodology	Experimental and Quasi-Experimental designs	Method matches the purpose of the study – often mixed methods	Primarily qualitative; participatory	Method matches best way to facilitate findings to enhance social justice
Evaluator Role	Evaluator as detached, objective, neutral party	Social relations manager to facilitate use	Communicator who engages in meaningful dialogue	Relationship-builder based on trust and cultural respect

While the paradigms provide a framework for how to discuss and think about the field of evaluation, it should be noted that no matter the paradigm that an evaluator places themself in, "all evaluators consider methods, valuing, and use as part of their conceptualization about what evaluation is and how it should be conducted. For instance, although methods-based evaluators rely heavily on the scientific process to inform their final decision about merit and worth, that is not to say that they do not value stakeholder input en route to that decision" (Alkin, et. al., p.38). Thus, if an evaluator is aligned with the Pragmatic Paradigm, it does not mean that rigorous methods will not be utilized in the evaluation (Postpositivist Paradigm) or that the mechanisms for valuing (Constructivist Paradigm) will not be properly considered. It simply means that the prioritization for an evaluator that is part of the Pragmatic Paradigm will be on the use of evaluation findings, as well as influencing the attitudes and behaviors of evaluation stakeholders.

ECA MELI UNIT – PRAGMATIC PARADIGM

The ECA MELI Unit's values are most closely aligned with the Pragmatic Paradigm (the Use Branch) of evaluation. This means that the Unit's procedures are centered on the involvement of stakeholders (primarily staff of the Program Office(s) who requested the evaluation, but also ECA senior leadership) throughout the process with the goal of ensuring responsiveness to their needs and requests. Even though the ECA MELI Unit's practices are aligned with the Pragmatic Paradigm, using rigorous data collection and analysis methods as well as placing value on those findings is still important; however, the primary priority is to ensure that the data, findings, and recommendations from evaluations conducted are utilized. Before discussing use, this paper will explore the topics of valuing, methods, and generating credible evidence – and the ECA MELI Unit's position regarding these core aspects of evaluation.

VALUING

"To understand valuing in evaluation is to understand the methods by which we assist our natural abilities to judge the value of alternatives" (Julnes, p.4). Valuing, along with methods and generating use of evaluation results, together form the core aspects of evaluation. Valuing is concerned with analyzing data collected within a particular context (through establishing criteria of merit and constructing standards), synthesizing the elements (through measuring the performance of the program and comparing to the constructed standards), and integrating the data collected against the standards and criteria to form a value judgment of a program's merit, worth, and/or areas for program improvement (Julnes, p.6).

There are two types of approaches to valuing: descriptive and prescriptive. Prescriptive valuing promotes a particular set of values. For instance, those evaluators on the Social Justice Branch (Transformative Paradigm) may approach their evaluations through a Critical Race Theory lens to assist in the determination of power issues present in the context of the program being evaluated. The prescriptive approach to valuing holds that an evaluator's particular value should have priority in an evaluation. A descriptive approach to valuing "selects criteria of merit from descriptions of values held by stakeholders. This approach identifies relevant stakeholder values, uses these to construct criteria and standards, and gathers and reports evaluative data in terms of those criteria" (Scriven, p.97).

Evaluation theorists Marvin Alkin, Anne Vo, and Christina Christie created a typology of evaluator valuing roles to describe three ways in which valuing is done – by stakeholders alone, by evaluators only, or by stakeholders and evaluators working together (Alkin, et. al., p.38). As ECA MELI Unit procedures emphasize a collaborative approach and working closely with stakeholders throughout the entire evaluation process, the valuing typology most often used is that of evaluators and stakeholders together.

Each evaluation commissioned or conducted by the ECA MELI Unit relies on stakeholder input and a firm grasp of the program and its operating context. This knowledge lends itself to the evaluator being able to make a judgement of the program's merit or worth through joint understandings of all value positions.

CREDIBLE EVIDENCE

Another important aspect in evaluation and research that merits discussion is that of credible evidence as it is the foundation on which knowledge is created and decisions are made. Producing credible evidence is important to evaluators in each of the four paradigms, though each prioritize different aspects of credible evidence and how it is created. This section will explore how credible evidence is defined and derived, and what aspects of credible evidence the ECA MELI Unit values.

DEFINITION OF CREDIBLE EVIDENCE

There is no one definition of credible evidence or a consensus on what counts as credible evidence. The Office of Management and Budget defines credible evidence as "objective and of

sufficient quality, utility, and integrity" with each USG agency being required to establish procedures to ensure the objectivity, utility, and integrity of information provided (OMB Circular A-11). Evaluation scholar Thomas Schwandt defines credibility as "a matter of whether the evidence is true, and that judgment is subject to interpretation by the relevant reference group charged with making that judgment" (Schwandt, p.73), suggesting that context and an evaluation's stakeholders play defining roles in determining what constitutes credible evidence.

Surprisingly, research "…tells us that facts are only a small part of how credibility judgments are made" (Miller, p.41). An evaluation's stakeholders will draw upon a multitude of aspects of information provided to determine whether or not it is credible, such as:

- the validity of methods utilized to collect the evidence;
- how relevant, useful, up-to-date, and balanced the evidence is;
- how strongly the evidence in question points to the conclusion being considered; and
- characteristics of who produced and received the information – particularly surrounding perceived similarity of a source's beliefs as "clients are prone to believe information produced by sources who are trusted and who possess compatible worldviews, as well as the perceived lack of bias and the impartiality of context produced by a source" (Miller, p.48).

CREDIBLE EVIDENCE AS VIEWED FROM EACH PARADIGM

Evaluation theories offer guidance in terms of how to gather credible and actionable evidence in contemporary practice (Donaldson, p.7). Each paradigm places great emphasis on providing credible evidence, however the lens through which evaluators in each paradigm would view credible evidence is different.

- Postpositivists emphasize that it is the rigor of the methods by which the evidence is collected that determines the credibility of evidence – using a narrow definition of rigor wherein RCTs or quasi-experimental designs reign supreme.
- Those on the Use Branch (Pragmatic Paradigm) focus on the actionability of the evidence, defined as "the perceived degree to which evidence is suitable as a guide to possible action", with the views of how stakeholders (specifically clients and/or decision-makers) would deem evidence as being credible.
- Constructivist evaluators place importance on who the data are collected from; holding that evidence is credible if it is gathered from and takes into account the viewpoints of the widest range of stakeholders possible.
- Like Constructivists, Transformative Paradigm evaluators consider evidence as being credible if the data represent multiple value perspectives, though the focus of the evidence should promote social justice through highlighting inequities or unequal power structures amongst the stakeholders.

GENERATING CREDIBLE EVIDENCE (METHODS)

Context is an important factor in generating credible evidence for an evaluation. The question of how credible evidence is gathered was at the heart of a debate for decades in the evaluation and research fields (Donaldson, p.7) centering around the question of whether or not

qualitative methods were rigorous enough (as compared to quantitative methods, primarily RCTs) to generate credible evidence. Fortunately, these debates have subsided, but the topic still deserves a brief mention.

Particularly in regard to those evaluators in the Pragmatic Paradigm, context plays a vital role in determining whether or not evidence will be seen as credible within a particular environment. Methods tend to be valuable for certain purposes in certain circumstances with "no method better than another in the absolute, but [some] being more or less relevant depending on the context" (Tourmen, p. 28). There are a number of contexts that factor into the decision-making process and play a role in influencing the evaluation design and method choice: the (1) decision context, (2) evaluation context, and (3) the program context.

The decision context refers to the importance of knowing who the decision-makers are who need the evidence, and the standards of rigor expected from them – which then leads the evaluator to choose the method aligned with these aspects for the evaluation. The decision context includes several dimensions: the types of decisions and information needs of stakeholders, the multiple questions and tasks the evaluation is expected to address, the multiple levels of conclusions and decisions that need to be made, and the form in which the decision-makers would like to see the conclusions (Julnes & Rog, p.223). The dimensions of the decision context highlight the importance of working closely with stakeholders to determine their needs and expectations and can influence how the evidence of the evaluation should be collected. With the decision context, the evaluation design and methods utilized to gather credible data are driven by the stakeholder and do not favor a pre-determined set of methods. "The basis on which we substantiate the use of any method to generate evidence is not a hierarchy of method – with RCTs at the highest level and expert opinion at the lowest level – but a judgment of aptness of a given method to generate the kind of information needed to produce the evidence needed to answer the question under investigation" (Schwandt, p.264).

The evaluation context refers to the constraints placed on evaluators. Most commonly this refers to three constraints: time (how quickly a decision-maker needs the results), budget, and data (lack of data altogether or existing data that are not valid or reliable). Regarding the constraints of budget and time, a mixed methods approach to gathering credible evidence is seen as a more cost-effective method (Julnes & Rog, p.238) as compared to RCTs, and often takes less time to implement.

Finally, the program context heavily "influences method choice. The more complex the systems relationship in the domain being studied, the more important it is that multiple sources of evidence support the resulting conclusions" (Julnes & Rog, p.253). One of the realities acknowledged with broad-based initiatives, such as those that attempt to change systems or environments, or where there is a significant lag between the onset of a program and changes in measured outcomes, the "more difficult it is to rule out alternative causal explanations" (Julnes & Rog, p.238) as RCTs are designed to do. A qualitative or mixed methodology provides the contextualization and understanding of personal experience in which quantitative methods are deemed to be "sorely lacking in explanatory or educative power" (Simons, p.410).

CREDIBLE EVIDENCE IN THE ECA BUREAU

The ECA MELI Unit is concerned with generating evidence that is not only credible but actionable as well, taking into account the information stakeholders have requested, any evaluation context constraints, and the real-world factors of ECA programming. As the name suggests, randomized control trials require random assignment of participants to two different groups (one group that is part of the program, the other not) to determine whether the program caused the outcomes. In relation to a study abroad program, "such a design may not be feasible given the naturalistic setting of study abroad programs and the difficulty in randomly assigning students to both treatment and control groups" (Savicki, p.123).

For ECA-funded exchange programs, the selection process for potential exchange participants is rigorously managed, with neither program staff nor senior leadership wanting to sacrifice the selection of high-quality candidates for the sake of an evaluation being seen as more rigorous. The ECA MELI Unit views the balance between program needs and evaluation methods in the way that prominent evaluation theorist Carol Weiss has suggested: "When evaluation needs and program needs conflict, the program gets priority. Evaluators must compromise their methodological standards when it suits the program, and often forgo questions that program personnel and managers do not see as legitimate" (Weiss, p183).

Because the goal of ECA exchange programs is to create networks of people designed to increase mutual understanding and build peaceful relations, the programmatic context is considered long-term and complex in nature ('complex' meaning there are a number of external factors playing a role in potential outcomes). While the MELI Unit is amenable to conducting evaluations using any and all methodologies, because of the aforementioned aspects of ECA programs and context, the ECA MELI Unit generally conducts evaluations utilizing mixed methods. "[RCTs] can only address a small number of presumed mechanisms and can offer limited insight on how mechanisms interact, [while utilizing a mixed methods approach] can provide the needed insights about complex mechanisms and…the rich understanding needed" (Julnes & Rog, p.238).

This is not to say that RCTs have no value in the international or diplomatic arenas, only that qualitative methods (through the collection of stories, personal narratives, and contextual factors of ECA programs) provide an important source of information that create evidence that is actionable for ECA stakeholders. All evaluations conducted or commissioned by the ECA MELI Unit must place equal value on evidence being credible, actionable, and gathered through appropriate methods; as former Program Director for the U.S. Government Accountability Office and evaluation theorist Eleanor Chelimsky stated: "what [evaluators] should aim for is the greatest possible validity in our methods combined with the greatest possible persuasiveness in our explanations and accounts" (Chelimsky, p.187).

USE IN ECA EVALUATIONS

The ECA MELI Unit follows the requirements of the OMB closely by conducting evaluations that use the most rigorous methods that are appropriate to the evaluation and are feasible within

budget, time, data, and other constraints (OMB Circular A-11). The MELI Unit sees its role as ensuring all ECA evaluations are conducted with the most rigorous methods that fit the context, while also building the evaluation capacity of stakeholders and fostering use of evaluation results by including stakeholders throughout the evaluation process.

There are several different types of use, though the ECA MELI Unit is concerned with two: instrumental use (direct, visible action taken based on evaluation findings) and enlightenment use (contribute to general understanding, which does not necessarily lead to change in overt behavior in the short-term) (Kirkhart, p.9). "Helpful activities for instrumental use include: identifying users early in the evaluation; having frequent contact with users, especially during question formation; studying things that users can control; providing interim results; translating findings into actions; and disseminating results…" (Shadish et. al., p55). The MELI Unit has created its standard operating procedures to foster instrumental and conceptual use through these research-informed strategies of conducting evaluations utilizing different models, involving stakeholders throughout the evaluation process, and facilitating the implementation of recommendations stemming from an evaluation.

CONDUCTING THE EVALUATION: INVOLVEMENT OF STAKEHOLDERS

Evaluators adhering to the tenets of the Use Branch often involve stakeholders in the evaluation. The ECA MELI Unit encourages ECA stakeholders to participate throughout the entire evaluation process. The MELI Unit solicits requests for evaluations twice each fiscal year through an email sent to all ECA Managers and announcements in all-staff meetings. In the past, the MELI Unit selected the programs to be evaluated without the involvement of Program Officers, which led to the findings and recommendations from the reports not being widely used. This new process provides the MELI Unit with the buy-in and support of the program staff that manage the programs.

Once the programs to be evaluated are selected, the MELI Unit creates a User Group to assist in writing the evaluation scope of work. The User Group includes the Program Officer(s) directly responsible for the program being evaluated and the senior leadership for that particular Neighborhood (to include the Division Chief, Managing Director, and Senior Advisor). The MELI Unit then meets with the User Group to obtain more information as to the nature of the program – including overarching questions that need to be answered, sensitivities of the program itself and the context in which it is implemented, and intended use of the results of the evaluation. These interactions and relationships matter as credibility is built "through evaluative relationships enacted with norms of trust, acceptance, and respect and through evaluative communications that are inclusive and dialogic" (Greene, p.210).

Creating an effective communication system to disseminate information to the User Group throughout the evaluation is critical to ensuring that evaluation results are utilized. To do this, the MELI Unit uses a collaborative approach to all evaluations. Throughout the evaluation, the MELI Unit communicates regularly with the ECA Program Officer(s) managing the program.

When evaluation products are delivered, the MELI Unit schedules a formal meeting with the User Group to explain the product, discuss any decision points that need to be made, and provide recommendations from a technical standpoint as to what the User Group should take into consideration when reviewing the product. Utilizing a collaborative approach results in stronger evaluation designs, enhanced data collection and analysis, and results that the stakeholders understand and use (Rodriguez-Campos, p.11).

For those stakeholders that are interested and able to dedicate the time required, the MELI Unit is receptive to using participatory methods. Participatory evaluation differs from collaborative evaluation in that the evaluator shares control of the evaluation with the stakeholders rather than only involving them at certain points in time during the evaluation. "Participatory evaluation enhances evaluation use by increasing the depth and range of participation.... [it] creates a synergistic effect between the evaluation team and the core stakeholders and contributes to reflective and mutually beneficial learning" (Zukoski & Bosserman, p.48). Due to the time commitment required, this approach to evaluation is not utilized often, but it is an option that is encouraged when possible.

CONDUCTING THE EVALUATION: INTERNAL EVALUATIONS

The MELI Unit conducts a variety of evaluations, including external (conducted solely by evaluators external to ECA), hybrid (a combination of an external evaluation team and the ECA MELI Unit), and internal evaluations (conducted solely by evaluators from the MELI Unit). Regardless of the type of evaluation, the MELI Unit involves ECA stakeholders throughout the evaluation process.

Internal evaluations have been criticized due to the perceived lack of objectivity; because the evaluator is a part of the institution implementing the program, it is assumed that there may be pressure or additional bias in the process and that perhaps the findings and conclusions of the evaluation may be more positive than they would have been had an external evaluation team conducted the evaluation. However, interactions and relationships can matter for credibility: "when an evaluator has direct interaction with individuals who are recipients of evaluative information, the evaluator's level of diplomacy, charm, goodwill, and likeability may decrease or increase experienced credibility of him or her. People are biased toward extending credibility to those whom they find likeable" (Miller, p.45). When conducted in a way that builds trust and respect, internal evaluations can produce credible evidence that stakeholders are more likely to utilize to improve their programs.

OMB regulations and 18 FAM 300 approve of internal and hybrid evaluations as long as "there is an appropriate level of independence by experts external to the program either inside or outside an agency" (OMB Circular A-11). The MELI Unit is located in the Policy Neighborhood, a cross-cutting entity in the Bureau that manages only a small fraction of the programs in ECA. With a high proportion of programmatic activity taking place external to Policy, the MELI Unit is adhering to the requirements set forth in the guidance. 18 FAM 300 also outlines the importance of the internal evaluation team needing to have the skills necessary to successfully conduct an evaluation. The ECA MELI Unit takes this guidance seriously. Each member of the

MELI Unit has the requisite M&E experience and has signed a code of conduct acknowledging their familiarity with the AEA Guiding Principles[6] and AEA Competencies[7], and their intent to adhere to the ethics and standards placed on all evaluators.

When the ECA MELI Unit conducts hybrid and internal evaluations of its programs, it is recognized that, rather than being a hindrance, the proximity to the program provides better insight into the political realities of the programs and organizational context, and what information stakeholders will find most useful. Evaluators who work closely with those managing programs "…are well positioned to do evaluations that may have a higher rate of impact on organizational decisions" (Rossi, p.394).

ACTION PLANS (INSTRUMENTAL USE)

As part of ECA's commitment to 18 FAM 300, the MELI Unit has begun using Action Plans to guide implementation of recommendations provided in evaluation reports. Once an evaluation is completed, the ECA MELI Unit creates an Action Plan that outlines the recommendations made in the evaluation report. A meeting with the User Group is then held to determine which recommendations are deemed actionable, who has responsibility for implementing the recommendation, and the action to be taken and date by which the responsible party will carry it out. The MELI Unit will maintain responsibility for inquiring as to the status of the recommendations until all are considered completed as agreed. Through the Action Plan process, the ECA MELI Unit directly fosters instrumental use of evaluations.

DISSEMINATION OF RESULTS (ENLIGHTENMENT USE)

Evaluators should not restrict themselves to seeking direct, instrumental use, but should plan for findings that may facilitate enlightenment. An "enlightenment function that serves a broad political audience, including the public, may create long-term effects for evaluation" (Mark & Henry, p.83). Each evaluation that the ECA MELI Unit commissions or conducts is posted in its entirety online – making the report accessible to anyone with an internet connection. By doing this, the MELI Unit is promoting enlightenment use in the sense that the information contained in the reports may contribute to future internal decisions or decisions in other settings.

CONCLUSION

All approaches to evaluation involve tradeoffs that require evaluators to prioritize how an evaluation will be conducted. The ECA MELI Unit prioritizes the use of evaluations and has built its practices and procedures around maximizing the utilization of evaluation findings and recommendations. The MELI Unit will continue to learn, adapt, and be responsive to evaluation-related legislation and DoS policy, while conducting evaluations that are rooted in the Pragmatic Paradigm.

[6] https://www.eval.org/p/cm/ld/fid=51

[7] https://www.eval.org/page/competencies

SOURCES

Alkin, M. C. & Christie, C.A. (2013). In Alkin, M., *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). Thousand Oaks, CA: Sage.

Alkin, M. C., Vo, A. T., & Christie, C.A. (2012). The Evaluator's Role in Valuing: Who and With Whom. In G. Julnes (Ed), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, 133*, 29-41.

Carter, S.M., & Little, M. Justifying knowledge, justifying method, taking action: epistemologies, methodologies, and methods in qualitative research. Qualitative Health Research 2007; 17(10): 1316-1328. doi: 10.1177/1049732607306927

Chelimsky, E. (2013). Balancing Evaluation Theory and Practice in the Real World. *American Journal of Evaluation*, 34(1), 91-98.

Chelimsky, E. (2015). Credibility, Policy Use, and the Evaluation Synthesis. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 177-200). Los Angeles: Sage.

Chouinard, J.A., Boyce, A., Hicks, J., Jones, J., Long, J., Pitts, R., Stockdale, M. (2017). Navigating Theory and Practice Through Evaluation Fieldwork: Experience of Novice Evaluation Practitioners. *American Journal of Evaluation*, 38(4), 493-506.

Christie. C. & Fleischer, D. (2015). Social Inquiry Paradigms as a Frame for the Debate on Credible Evidence. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 3-26). Los Angeles: Sage.

Donaldson, S. I. (2015). Examining the Backbone of Contemporary Evaluation Practice: Credible and Actionable Evidence. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 3-26). Los Angeles: Sage.

Donaldson, S. I., Lipsey, M.W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J.Greene, & M. Mark (Eds.), *The Handbook of Evaluation: Policies, Programs, and Practices* (pp. 56-75). London: Sage.

Engberg, M. E. & Davidson, L. M. (2015). Quantitative Approaches to Study Abroad Assessment. In Savicki, V. and Brewer, E. (eds.), *Assessing Study Abroad: Theory, Tools, and Practice*. (pp. 122-141). Sterling: Stylus.

Greene, Jennifer. (2015). How Evidence Earns Credibility in Evaluation. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 221-258). Los Angeles: Sage.

Julnes, G. (2012). Managing Valuation. *New Directions for Evaluation*, 133, 3-14.

Julnes, G. & Rog, D. (2015). Actionable Evidence in Context: Contextual Influences on Adequacy and Appropriateness of Method Choice in Evaluation. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 221-258). Los Angeles: Sage.

Kirkart, Karen E. (2000). Reconceptualizing Evaluation Use: An Integrated Theory of Influence. *New Directions for Evaluation, 88*, 5-21.

Mark, M. & Henry, G. (1998). Social Programming and Policy-Making: A Realist Perspective. *New Directions for Evaluation*, 78, 73-88.

Mertens, D. M. & Wilson, A. T. (2019). *Program Evaluation Theory and Practice: A Comprehensive Guide* (2nd Ed.). New York: Guildford Press.

Miller, R.L. (2015). How People Judge the Credibility of Information: Lessons for Evaluation from Cognitive and Information Sciences. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 39-61). Los Angeles: Sage.

Office of Management and Budget, Executive Office of the President, OMB Circular A-11 Section 200.14, Management of Federal Information Resources (2017)

Rodriguez-Campos, L. Essentials of Collaborative Evaluation. In Fetterman, D. m>, Rodriguez-Campos, L, & Zukoski, A. (Eds.), *Collaborative, Participatory, and Empowerment Evaluation.* (pp. 10-20). New York: Guilford.

Rossi, P. H. (1991). Comprehensive, Tailored, Theory-Driven Evaluations – A Smorgasbord of Options. In Shadish, W. R, Cook, T. D., & Leviton, L. C. (Eds.), *Foundations of Program Evaluation: Theories of Practice.* (pp. 377-437). London: Sage.

Schwandt, Thomas. (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Stanford: Stanford Business Books.

Schwandt, Thomas. (2015). Credible Evidence of Effectiveness: Necessary but Not Sufficient. In Donaldson, S. I., Christie, C. A., and Mark, M. M. (Eds.), *Credible and Actionable Evidence: The Foundations for Rigorous and Influential Evaluations*, (pp. 259-273). Los Angeles: Sage.

Scriven, Michael. (1991). The Science of Valuing. In Shadish, W. R, Cook, T. D., & Leviton, L. C. (Eds.), *Foundations of Program Evaluation: Theories of Practice.* (pp. 73-118). London: Sage.

Shadish, W.R. (1998). Evaluation Theory is Who We Are. *American Journal of Evaluation*, 19(1), 1-19.

Shadish, W. R, Cook, T. D., & Leviton, L. C. (1991). *Foundations of Program Evaluation: Theories of Practice*. London: Sage.

Simons, Helen. (2004). Utilizing Evaluation Evidence to Enhance Professional Practice. *Evaluation*, 10(4), 410-429.

Tourmen, Claire. (2009). Evaluators' Decision Making: The Relationship Between Theory, Practice, and Experience. *American Journal of Evaluation*, 30(1), 7-30.

Weiss, Carol. (1991). Linking Evaluation to Policy Research. In Shadish, W. R, Cook, T. D., & Leviton, L. C. (Eds.), *Foundations of Program Evaluation: Theories of Practice.* (pp. 179-224). London: Sage.

Zukoski, A. & Bosserman, C. (2018). Essentials of Participatory Evaluation. In Fetterman, D. M., Rodriguez-Campos, L, & Zukoski, A. (Eds.), *Collaborative, Participatory, and Empowerment Evaluation.* (pp. 48-56). New York: Guilford.