What's the Role of Equity in Evaluation Policy?

PREPARED BY: MARY ELLEN WIGGINS AND ALEX SILEO

FEBRUARY 2020

Contents

the
forum
FOR YOUTH INVESTMENT

Acknowledgements

The Forum for Youth Investment is a nonprofit, nonpartisan "action tank" dedicated to helping communities and the nation make sure all young people are Ready by 21®—ready for college, work, and life. Informed by rigorous research and practical experience, the Forum forges innovative ideas, strategies, and partnerships to strengthen solutions for young people and those who care about them. For more information, please visit http://www.forumfyi.org. June 2018 ©The Forum for Youth Investment. All rights reserved.

Introduction

Increasingly, policymakers and practitioners address equity by promoting the use of evidence as an important tool to strengthen opportunity and outcomes for young people. This is a powerful trend, but it needs to be a two-way street. We should also think of equity as a way to strengthen evidence, specifically by taking a hard look at whether and how the processes for generating and using evidence—and for determining what is "sound" evidence in the first place—are themselves equitable.

Evidence is a potent ingredient in shaping all kinds of policy, from legislation and regulations across levels of government to program-level guidance. Particularly at the federal level, evidence has a growing role in determining what programs and services get funded and how work gets done.[1]

As a result, the frameworks that agencies use to ground their work in evidence—such as evaluation policies, principles, and standards—can play a significant role in promoting or hindering equity through the values, parameters, and priorities that they assert.

Without careful attention, evidence frameworks risk reinforcing or even exacerbating inequities potentially embedded in research, policy, and practice—even if that is the opposite of what is intended.

The Forum for Youth Investment (Forum) is building on our evidence-based policy and practice work by applying an equity lens to the processes that underlie how evidence gets generated and used, including answering the following questions:

- Whom is evidence produced for, by, and about?
- What evidence is considered valuable—rigorous, valid, and credible?
- Who has voice during the processes of generating and using evidence?

What Do We Mean by Equity?

The Annie E. Casey Foundation defines the term as "the state, quality or ideal of being just, impartial and fair" that "needs to be thought of as a structural and systemic concept."[2] The American Evaluation Association defines the term as "the condition of fair and just opportunities for all people to participate and thrive in society regardless of individual or group identity or difference."[3]

The Forum believes that, for children and youth, equity is achieved when all young people have a shot at getting the essentials they need to succeed. At an individual level, this requires taking into account the differences in their starting points. At a systems and institutional levels, this means addressing often-entrenched barriers to their success.

Careful consideration of these individual and systemic factors play a role in unpacking what it takes to promote equity in evaluation, both for individual evaluation projects as well as for the policies, norms, and institutions that govern evaluation and other evidence-building.

How Is It Different from Equality?

Equity requires that opportunity be calibrated to need through tailored supports or additional resources. It is distinguishable from equality, which connotes the ideas that supports and resources are distributed evenly. Equity can be both an outcome (where everyone has what they need to thrive) and a process (where those most impacted by inequity are meaningfully involved in the creation and use of evidence, in this case).

These complex questions have many layers and can be approached from a variety of angles. Moving forward, we will release a series of publications that use these questions to position evidence-based policy as a mechanism to actively strengthen equity in youth-serving programs. As a concrete starting point, the Forum will look at how federal youth-serving agencies [i]—including the U.S. Departments of Labor, Education, Health and Human Services, and others—address equity in their evaluation policies and, in instances where official policies are unavailable, evaluation-oriented frameworks and guidelines.[ii]

To support comparison across agencies, this paper focuses specifically on overarching principles or standards in evaluation policies. This publication examines the five most common principles: rigor, relevance or utility, independence, transparency, and ethics.[iii] It explores how equity relates to each principle, provides examples from agency policies, and recommends how federal agencies can incorporate equity into principles and standards moving forward. The Forum's hope is to encourage agencies to consider the role of equity as, in keeping with the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act), they develop or refine their evaluation policies to support their particular missions.

[i] For purposes of this publication, we consider youth-serving federal agencies to be the 21 agencies participating in the Interagency Working Group on Youth Programs, which are listed on youth.gov.

[ii] The Forum recognizes that some agencies have formal evaluation policies and others have evaluation-oriented frameworks or guidelines that might be considered informal policies. Our paper will examine both and will use examples from each of these categories. We recognize that formal and informal policies may have been conceived and implemented with different intents and purposes. Despite these differences, many of the formal and informal evaluation policies share references to common evaluation principles and discuss these principles in highly similar ways and may be informative for purposes of this paper.

[iii] Although other agency policies may use alternative terms, their descriptions of these principles or standards are similar in content to the five most common categories. For purposes of this paper, we do not distinguish between principles and standards. In the case of USAID, which does not articulate principles or standards, we have used their practices as equivalents of principles.

What is evaluation policy, and why does it matter?

Evaluation plays a uniquely important role in evidence-building because, as the Office of Management and Budget (OMB) observes, it "can answer essential questions regarding program effectiveness and efficiency that cannot be answered through performance measurement and monitoring, statistics, or policy analysis."[4] Evaluation policies articulate principles, standards, and/or practices that agency evaluation staff should consider when developing research questions, conducting individual studies, interacting with stakeholders, and communicating results once studies have been completed. As such, they reflect agencies' clearest values and priorities around these evaluation processes and, importantly, they apply to all evaluations the agency conducts or funds.

So how might these high-level principles matter in practice? One instructive example might be home visiting programs.

In 2009, the U.S. Department of Health and Human Services (HHS) launched Home Visiting Evidence of Effectiveness (HomVEE) "to conduct a thorough and transparent review of the home visiting research literature" and assess "the evidence of effectiveness for home visiting models that target families with pregnant women and children from birth to kindergarten entry."[5] These systematic assessments have supported important progress through transparency and rigor (two key evaluation principles), including detailing which models are eligible for HHS Maternal, Infant, and Early Childhood Home Visiting Grants. So what about equity?

The 2011 HomVEE review notes that, despite diversity of race/ethnicity and income in study samples, the sample sizes were generally too small to analyze each subgroup.[6] For evaluations to support all populations, they will need to go further to understand what works best for whom. In response, the agency launched the Maternal and Infant Home Visiting Program Evaluation (MIHOPE) in 2011. The more expansive study was designed from the start to include seven subgroups of interest to the field, such as race and ethnicity, the presence of intimate partner violence, the mother's level of emotional functioning, and other potential demographic risk characteristics.[7]

In 2012, HHS funded the launch of the Home Visiting Applied Research Collaborative (HARC) which, with a new HHS grant from 2017, now works "to strengthen and broaden the impact of home visiting, using innovative research to achieve precision home visiting."[8] HARC notes that while many studies have demonstrated that home visiting participants, on average, experience positive outcomes, not all families are benefiting equally.[9] HARC now is focusing on understanding what works for whom in order to strengthen outcomes for vulnerable families not successfully served by current models.

Framing the research question in this way actively promotes equity by recognizing that different groups of people and families will experience different results and by intentionally seeking knowledge about how to promote success for all groups. Home visiting programs already support more equitable health and education outcomes for many families, which is immensely important. But how might an evaluation policy that explicitly values equity—alongside attributes like rigor and transparency—have helped HHS and the broader field anticipate and confront disparities from the beginning to better reach all communities and populations earlier?

Why now?

Evaluation policies—which several youth-serving agencies already have—will play an expanding role as the Evidence Act requires federal agencies to appoint Evaluation Officers charged with establishing and implementing a policy that:

> ...affirms the agency's commitment to conducting rigorous, relevant evaluations and to using evidence from evaluations to inform policy and practice. The policy will provide the agency's stakeholders with a clear understanding of the expectations related to key principles, such as evaluation rigor, relevance, transparency, independence, and ethics.[12]

The Evidence Act also requires OMB to develop federal program evaluation standards and best practices that can guide program evaluation activities for federal agencies.

The Forum's Managing for Success report noted the importance of evaluation policies in increasing agency capacity to use evidence.[13] This mirrors the recent final report by the Commission on Evidence-Based Policymaking, which informed the Evidence Act.[14] However, neither report explicitly addressed equity as a consideration in evidence-building or in evaluation policy.

We know that equity matters in changing the odds so that all youth can be ready for and succeed in life. Closing equity gaps that persist across many factors—including race and ethnicity, socioeconomic class and income, gender, immigration status, tribal status, and urban/rural geographic divides—is critical to achieving population-level results.[15]

Evaluation and other evidence- and knowledge-generating processes can play a major role in overcoming these equity gaps. At the same time, we have to recognize that using or neglecting an equity lens "shapes worldview and professional practice."[16] If we fail to consider whether and how major frameworks like evaluation policies promote or hinder equity, we risk inherently limiting our ability to reduce inequity or, worse, further entrenching it.

Principle 1: Rigor

Agency evaluation policies acknowledge that rigor has many aspects. Policies for the Administration for Children and Families (ACF), Department of Labor (DOL), and Department of Housing and Urban Development (HUD) all share the following language:

> [r]igor requires ensuring that inferences about cause and effect are well founded (internal validity); requires clarity about the populations, settings, or circumstances to which results can be generalized (external validity); and requires the use of measures that accurately capture the intended information (measurement reliability and validity).[17]

The aspect of rigor that agencies emphasize most may be appropriately matching research methods to the specific research question. For example, both HUD's and ACF's evaluation policies identify the need to use research methods that isolate the effects of programs on outcomes, as well as the need to use other research methods (quasi-experimental or implementation studies, for example) when appropriate for other evaluation questions.

The Department of Education (ED) and the National Science Foundation (NSF) also have released a joint report on common guidelines for education research, which further details the purpose of various types of research studies.[18] ACF and DOL policies also reference the need for each agency to maintain a well-trained evaluation workforce that can answer, with appropriate methodologies, the types of questions agencies are likely to be interested in.

How does equity fit into rigor?

According to these evaluation policies, if agencies choose evaluation partners who are independent of the communities, people, and organizations directly involved in the intervention, then their findings will avoid biases in favor of or against a particular result. If criteria for rigor are too strict or too narrow, it can become unnecessarily difficult for evaluations of programs serving the most vulnerable populations to meet these requirements, including in ways described below. This means that organizations and communities serving the most vulnerable populations may have difficulty finding programs and practices that meet criteria for being "evidence-based," which can be a requirement or incentive for funding. In other cases, such organizations and communities may have difficulty evaluating their own services in ways considered rigorous by funders, which can also have implications for funding. As a result, serving these populations effectively can become more difficult.

The agency examples above use a broad definition of rigor, which promotes equity by recognizing that useful, high-quality evidence can come from different types of research questions and settings that require different methodologies. As discussed further in the paper, these methodologies should be culturally sensitive and valid. In contrast, policy might hinder equity by promoting a single "gold standard" methodology—such as randomized controlled trials (RCTs)—because it provides the greatest level of methodological rigor despite not being appropriate or attainable for all research questions or circumstances. As the Center for the Study of Social Policy points out, the field cannot build an evidence base that will significantly improve outcomes for all—and for the most disadvantaged—children and families if it follows "a one-dimensional hierarchy that ranks programs by how methodologically rigorous their evaluations were."[19] The Government Accountability Office (GAO) weighed in on this issue in 2009 by cautioning against adopting overly narrow criteria for defining rigorous research designs in a study of early evidence-driven work under the Obama Administration.[20]

Designing an evaluation that is both rigorous and appropriate can involve trade-offs given the research questions and on-the-ground factors that affect how a study can be conducted. These realities can have disproportionate effects on the availability of sound evidence for populations most in need of supportive services.[21] The following examples demonstrate conditions that may affect trade-offs in designing evaluations:

- Organizational capacity to conduct an evaluation (such as funding constraints or data access and availability) can limit methodological options for a study design. A community organization that lacks robust administrative data on its programs, potentially due to limited funding for data infrastructure, may not be ready to support an RCT. This means that smaller or lesser-resourced communities and service providers may face particular barriers to conducting RCTs and building rigorous evidence for their work.
- Program practices or circumstances may not be conducive to certain research designs or expectations. For example, a rigorous RCT may not be possible for a program delivered in a residential setting where treatment and control groups cannot be fully separated or for a program serving few participants in a sparsely populated rural community. As a result, programs serving certain vulnerable populations may be more difficult to evaluate using an RCT.
- The most urgent or actionable questions about policy and practice may not lend themselves to the most rigorous methodologies. For example, program managers may gain the most actionable information from a process evaluation showing how an intervention is being implemented across multiple sites. Prioritizing learning for front-line service providers may mean tackling questions where an RCT does not apply.

In all of these cases, recognizing rigor across an array of methodologies and the value of using multiple evaluation methods is an equity issue because it affects the ability to understand and address the needs of diverse populations, programs, settings, and audiences for research.[22]

In addition to study design, what data are captured—and how they are captured—matters for equity. Measures should be culturally sensitive and valid for all of the populations in the study. Agencies also should consider what data is valued enough to be considered necessary for rigorous evaluations. As the American Evaluation Association (AEA) Statement on Cultural Competence notes, "all evaluation reflects culturally influenced norms, values, and ways of knowing."[23] This can translate into assumptions and biases that negatively affect equity, including inadvertently. The Centers for Disease Control and Prevention (CDC), through their adoption of the Joint Committee on Standards for Educational Evaluation's Program Evaluation Standards Statements, is the only agency that explicitly references culture when discussing rigor. Highlighting the need for accuracy in agency evaluations, the Joint Committee's statements note: "Evaluation conclusions and decisions should be explicitly justified in the cultures and contexts where they have consequences."[24]

For example, data that reflects the perspectives and lived experiences of communities or populations served could be disregarded or under-recognized in many methodologies. It is important for evaluators to capture this information in order to reflect the norms, values, and knowledge of people being served. Agencies and organizations capturing such data should use appropriate methods that are rigorous enough to use in studies, are culturally sensitive to the communities being studied, and are empirical enough to validly represent what they intend to measure. They also should consider where a lack of such data may create a gap that reduces the overall validity of findings, particularly in the area of external validity, as defined by both DOL and ACF above.

Finally, findings from the data should be understood in an appropriate cultural context. For example, structural racism or other historical inequities may heavily influence responses to an intervention, but may be difficult for outside evaluators to discern or understand. Involving study participants in interpreting results—as well as in earlier design phases—is an important strategy to provide appropriate context.

Equity gaps will only widen if "the trajectory of evidence" does not give careful attention to how evaluations can effectively address and respond to these communities and circumstances.[25] Broad conceptions of rigor play an important role in how well evaluation policies support and encourage such careful attention.

Principle 2: Relevance or Utility

Agency policies describe the principle of relevance or utility largely in relation to whose interests evaluations should address and, in some cases, what purposes evaluations should inform. For example, USAID notes that:

> ...relevance will be achieved by explicitly linking evaluation questions to specific future decisions to be made by USAID leadership, partner governments, and/or other key stakeholders. Those decisions frequently will be related to how resources should be allocated across and within sectors and thematic areas and/or how implementation should be modified to improve effectiveness.[26]

In this case, the "whose interests evaluations should address" include USAID leadership, partner governments, and other key stakeholders. "What purposes evaluations should inform" include how resources should be allocated and how implementation should be modified to improve effectiveness.

The additional agency examples below show some of the variation across policies:

- The Framework and Guidelines for Evaluation at the Small Business Administration (SBA) highlights the need for staff to "scope and select evaluation questions most closely tied to the goals of the program, the priorities of the Agency, and the intended use by senior leaders."[27]
- The Corporation for National and Community Service (CNCS) offers a more expansive list of stakeholders in its evaluation policy, including "partners such as states, territories, tribes, and grantees; and populations served; researchers; and other stakeholders" alongside "government-wide leadership; agency leadership and staff."[28]
- Both CNCS and DOL emphasize that the agencies "should encourage diversity among those carrying out the evaluations."[29]
- USAID stresses that "the conduct of evaluations will be consistent with institutional aims of local ownership through respectful engagement with all partners, including local beneficiaries and stakeholders, while leveraging and building local evaluation capacity."[30] [iv]
- The CDC adopts the utility standard of the Joint Committee on Standards for Educational Evaluation "to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs."[31]
- ACF encourages evaluation staff to work with program offices for the purpose of influencing their grantmaking activities, stating that "discretionary funding opportunity announcements will require that successful applicants cooperate with any federal evaluations if selected to participate. As legally allowed, programs with waiver authorities should require rigorous evaluations as a condition of waivers. As appropriate, ACF will encourage, incentivize, or require grantees to use existing evidence of effective strategies in designing or selecting service approaches."[32]

How does equity fit into relevance or utility?

Agencies are clearly using a wide variety of language to explain what relevance and utility should mean to their programs, systems, and staff. These policies vary in their treatment of the *who* and *what* questions outlined above. Looking across policies, trends include the following elements.

Whose interests evaluations should address:

- **Broad vs. narrow.** Does a policy refer only narrowly to stakeholders within the agency or the federal government, more broadly across non-federal partners in delivering programs or conducting evaluations, or very broadly across sectors and populations, such as program participants?
- **General vs. specific.** Does a policy describe the "who" in generalized terms, such as "stakeholders," or does it name specific groups whose interests should matter? Specific groups could include Congress; the President; agency leadership and staff; grantees; program participants and customers; states, localities, and tribes; researchers; and others.

[iv] The USAID policy addresses "relevance" and "oriented toward reinforcing local ownership" as separate practices (treated in this paper as equivalent to principles or standards). For purposes of this publication, we are addressing both these practices under the relevance principle for greatest comparability across agency policies.

The principle of relevance can play a strong role in promoting equity—or inequity—based on *to whom* evaluations should be considered relevant. Is it a broad or narrow set of stakeholders who should be valued and consulted throughout the research and evaluation process? Policies that frame relevance or utility broadly by emphasizing the role of non-agency stakeholders may help to promote equity by inviting diverse voices to inform agency planning and processes. This greater diversity may in turn support more equitable systems and outcomes.

In addition to indicating a generally broad or narrow set of stakeholders, it matters whether a policy names specific groups as stakeholders, particularly to capture "on the ground" perspectives from program participants and operational partners. When evaluation policies do not specify stakeholder groups, they risk excluding these groups. If stakeholders are excluded—even inadvertently and with good intentions—policies may fail to promote or even actively hinder equity by perpetuating "blind spots" around biases or needs, by systematically preventing certain groups from raising important issues, and by sending a message that certain groups are not valued as partners in the learning process.[v]

What purposes evaluations should inform: While we saw a trend of agencies citing either distinct or non-distinct purposes, it was not clear that these purposes would impact equity as much as the "who" questions above. For example, non-distinct policies described the purpose of evaluation in generic terms, such as "for decision-making" or "improving effectiveness." Distinct policies named activities such as "resource allocation" and "grantmaking." Though these purposes matter, their effect on equity seemed likely to be shaped by the inclusion or exclusion of different stakeholders. For example, grantmaking involves agency actors; grantees and subgrantees that may include states, tribes, localities, and non-profits; and program delivery staff and participants. As a result, the rest of this section focuses on questions of "who."[vi]

In exploring these questions, the SBA, CDC, and USAID evaluation policies offer informative contrasts:

- **SBA: Narrow and specific.** SBA's framework addresses the principle of relevance around program goals, agency priorities, and the use of findings by senior leaders. It is narrow because it does not explicitly include non-agency stakeholders and specific because it cites senior leaders.[vii]
- **CDC: Broad and nonspecific.** The CDC policy on utility broadly encourages evaluators to "devote attention to the full range of individuals and groups invested in the program and affected by its evaluation," but does not specify groups as likely stakeholders. The policy also states that "evaluation purposes should be identified and continually negotiated based on the needs of stakeholders."[33]
- **USAID: Broad and specific.** On the other hand, the USAID policy emphasizes the role of "all partners," going on to explicitly name local beneficiaries and stakeholders. By being both broad and specific, the USAID framework for determining relevance is more likely to identify needs, question assumptions, and build learning processes that promote equity.

In particular, the USAID policy recognizes the communities or populations being studied as a critical set of stakeholders. In a domestic context, this can mean consulting states, localities, or tribal representatives to inform an agency learning agenda, as both ACF and DOL policies emphasize. It also can mean making sure that providers of and participants in a particular program or service, such as youth living in foster care or participants in a job training program, can help inform research priorities.

---

[v] It is important to recognize that, contributing to the complexity of engagement, each group of stakeholders may be diverse within itself.

[vi] The Forum's assessment of which agency evaluation policies included distinct versus non-distinct purposes appears in the Appendix, along with other factors.

[vii] Importantly, SBA's full Guidelines and Framework document dedicates a section to identifying key stakeholders, defined broadly as "any person or group who has an interest in the program being evaluated or in the results of the evaluation." Moreover, SBA further encourages consideration of participatory evaluation that involves stakeholders in all aspects of the evaluation process, including design, data collection, and analysis. While the high-level principles articulated in the SBA policy lack a clear directive to include non-agency stakeholders in determining relevant research questions, SBA does offer a compelling example of how operational guidance on conducting specific evaluations can encourage appropriate and meaningful participation by external stakeholder groups.

In addition to stakeholder engagement, the DOL policy goes further by recognizing that diversity in the perspectives of researchers and evaluators may help evaluators and the field to clarify and question underlying assumptions or biases. Such assumptions and biases—such as factors related to culture, language, geography, or ability—could negatively affect equity, even inadvertently. Diverse perspectives also may help to identify different variables affecting the outcomes of interest. The CDC addresses this as a utility issue by asserting, "evaluations should clarify and specify the individual and cultural values underpinning purposes, processes, and judgments."[34]

By emphasizing that non-agency stakeholders can meaningfully inform evaluation findings, agency evaluation policies can help guide staff toward more equitable evaluation processes and, ultimately, systems and outcomes.

Principle 3: Independence

Policies consistently emphasize that evaluations must be as objective as possible and unburdened by inappropriate influence from particular stakeholder groups. While evaluation work should be relevant to federal stakeholders (such as Congress, the Administration, or agency leadership) and nonfederal stakeholders (such as businesses, contractors, advocates, or grantees), evaluations also must remain independent of these stakeholders and potential political biases or conflicts of interest. Evaluation policies frequently recommend that agencies choose evaluation partners who do not have any type of vested interest in the evaluation outcome.

Examples of how agencies address independence include the following:

- SBA, similar to some other agencies, requires evaluation by outside parties that do not have a "vested interest in the outcome or will not interpret the results in ways that are self-serving or misleading" and discourages "the appearance of bias to ensure results are properly used."[35]

- Both ACF and DOL promote objectivity through "independence in the design, conduct, and analysis of evaluations,"[36] including by specifying which staff have authority to approve evaluation designs and release findings independent of an agency's political leadership.

- The joint NSF and ED guidelines require external feedback plans for all research types, meaning that evaluation "should be subject to a series of external, critical reviews of its design and activities (including theoretical framework, data collection, analyses, and reporting)" that is "sufficiently independent and rigorous to influence the project's activities and improve the quality of its findings."[37] Feedback mechanisms could include peer review, external review panels or advisory boards, and a third-party evaluator.

How does equity fit into independence?

According to these evaluations, if agencies choose evaluation partners who are independent of the people, communities, organizations directly involved in the intervention, then then their findings will not have biases in favor of or against a particular result. Nonetheless, these partners may still be vulnerable to implicit biases or assumptions that could remain unchallenged without any significant level of stakeholder involvement in the evaluation process. This could potentially weaken both the results of the study as well as the agency's capacity to act on the study's findings. Independence needs to be balanced with appropriate forms and levels of engagement.

For example, agencies could highlight the need for communities to help shape and lead or heavily inform research and evaluation efforts in appropriate ways if they are the ones being researched. This could include helping to formulate key evaluation questions of interest, informing evaluation teams of relevant historical information or cultural norms, and establishing processes for how evaluation teams will interact with members of the community.

There often can be a divide between those conducting evaluations and the service providers or community groups that are being evaluated. As local impact investing organization Chicago Beyond has noted, the research process often has built-in power dynamics that can favor the priorities of researchers or evaluators over the lived experiences of the people and organizations being studied. These dynamics can lead to evaluation results that reinforce agency biases or only partially uncover information key to future decision-making.[38]

Agency evaluation policies should underscore that independence does not mean that stakeholders—communities, organizations, and people—affected by the issues being researched should not have a voice in the evaluation process. The CDC's policy offers a helpful example by affirming the need to "recognize, monitor, and balance the cultural and political interests and needs of individuals and groups."[39] This policy upholds and protects the legitimate roles of different stakeholders in an evaluation while cautioning against undue influence by any particular individual or group.

Principle 4: Transparency

The fourth principle, transparency, refers to how agencies communicate evaluation plans, methods, and findings to interested groups. For example, DOL's policy commits to making evaluation information and findings "broadly available and accessible," including typically releasing results within two months of evaluation completion and archiving data for secondary use by other researchers.[40] ACF's policy echoes these commitments and, as an example of accessibility, explains the policy itself in an approachable YouTube video.[41]

Evaluation policies encourage agencies to publicize information about their evaluation plans before conducting the evaluation and include all findings, regardless of whether they are positive or negative. These evaluation policies encourage agencies to clarify the limits of any evaluation methodology when reporting this information as well. For example, SBA's policy notes that staff should "ensure that the evaluation scope, design, implementation, and results are available for internal and public review, assessment, and critique." [42]

How does equity fit into independence?

Agencies often post evaluation plans and findings online, as some policies explicitly require. This information is theoretically available to all stakeholders, but in practice there is not equity in who can access it, who can understand it, and who can act on it. In terms of equity, what matters is how this information is made transparent and for whom.

The CDC policy, for example, addresses the importance of program stakeholders finding evaluation products, as well as processes, "valuable in meeting their needs."[43] In addition to communicating these findings online and through presentations to academic or policy conferences, evaluators should work with colleagues within their agency, or with external partners, to create actionable tools that allow individuals, organizations, or communities to engage with, easily understand, and act on these findings. While it is important to communicate evaluation findings (regardless of their result) through formal reporting mechanisms, agencies could take this a step further by ensuring that stakeholders, particularly historically marginalized stakeholders, have the capacity to translate these findings into actions that can create positive change.

Evaluators should consider alternative tools or communication vehicles that take the needs of people and organizations being evaluated into account so that they have the power to act on findings and engage the agency in decisions that will influence the policy, program, or practice being evaluated. For example, could potential program participants realistically use the information to make choices, such as whether to participate in or support the program locally? Could busy program operators and managers easily understand what changes are likely or unlikely to strengthen their work?

The Children's Services Council of Broward County has begun to develop new processes to involve youth and other community stakeholders in interpreting research, as well as in combining this research with their own knowledge of the county's history and their lived experiences. The stakeholders then use this information to develop products and tools that they can use to influence local policymakers.44 This example could be instructive for agencies as they determine how to deliver information transparently for all stakeholder groups.

Principle 5: Ethics

Ethics, the final principle, is a far-reaching principle that encourages agencies to conduct evaluations while respecting the dignity and rights of study participants and adhering to professional standards. For example, "ACF-sponsored evaluations will be conducted in an ethical manner and safeguard the dignity, rights, safety, and privacy of participants." [45] This can include maintaining respect for the privacy and confidentiality rights of study participants as well as maintaining respect for any other legal requirements established by Congress or the executive branch.

How does equity fit into ethics?

Equity can be an important lens for ensuring that the dignity and rights of study participants are respected in a sufficiently broad context. For example, evaluators should be aware of how structural racism or other historical inequities could have influenced participants and, subsequently, the interventions under evaluation. While many evaluations attempt to isolate an intervention's impact on specific outcomes, these contextual factors still can affect how a community, or a different population, responds to an intervention. Tragically, there are too many examples of unethical or harmful studies conducted on minority communities—many with government knowledge and approval. Evaluators need to be cognizant of the broader history of evaluation and the more specific history of the particular community of population under study in order to fully respect and preserve the dignity of those affected by their actions. Moreover, ethics are implicated in whether and how the study participants have adequate, appropriate, and accessible opportunities to engage in evaluation processes. Study participants make learning possible by providing information about themselves and their experiences. Through evaluation, their experiences influence critical decisions about whether and how resources and opportunities are made available to specific youth and groups of people, such as for housing, education, job training, child care, healthcare, justice, and other areas.

The dignity, rights, safety, and privacy—among other factors—of study participants and their communities can be directly affected not just by evaluation results but also by what questions are studied in the first place, how findings are interpreted, and how results are made available. Each of these activities is an opportunity to exercise power and influence. Therefore, including or excluding particular stakeholder groups—even inadvertently—can rebalance or redistribute power in how an evaluation is conducted, how results are interpreted, and how findings are acted on. These considerations may be especially relevant for communities and populations that have previously been studied but where, as Chicago Beyond notes, "fruits of those studies have infrequently nourished the neighborhoods where their seeds were planted." [46]

Clearly ethics and equity can be closely linked. However, the fact that equity-driven concerns and practices are not always incorporated into evaluations suggests that lumping these two terms together without explicit equity-focused language may be insufficient. If agency leaders who make evaluation policy believe that ethics encompass equity, they should clearly articulate how equity fits in, to make sure that the policy is clear and promotes accountability.

The Importance of Cultural Competence

Cultural competence has important implications for evaluation policy across each of the five most common principles. As the Equitable Evaluation Initiative describes, it enables organizations—including agencies—"to examine and align their practices and policies within a context of racial and cultural equity and inclusion. The use of such a lens is paramount when evaluating a program whose goals touch on issues of equity or inclusion." [47]

American Evaluation Association defines cultural competence in the following way:

"Cultural competence is not a state at which one arrives; rather, it is a process of learning, unlearning, and relearning. It is a sensibility cultivated throughout a lifetime. Cultural competence requires awareness of self, reflection on one's own cultural position, awareness of others' positions, and the ability to interact genuinely and respectfully with others. Culturally competent evaluators refrain from assuming they fully understand the perspectives of stakeholders whose backgrounds differ from their own.

Cultural competence is defined in relation to a specific context or location, such as geography, nationality, and history. Competence in one context is no assurance of competence in another. The culturally competent evaluator (or evaluation team) must have specific knowledge of the people and place in which the evaluation is being conducted—including local history and culturally determined mores, values, and ways of knowing.

The culturally competent evaluator draws upon a wide range of evaluation theories and methods to design and carry out an evaluation that is optimally matched to the context. In constructing a model or theory of how the evaluation operates, the evaluator reflects the diverse values and perspectives of key stakeholder groups." [48]

Cultural competence and evaluation principles:

This concept is particularly important for each of the evaluation principles. For example, agencies can emphasize rigor in culturally competent ways by valuing data that reflect the perspectives, values, and lived experiences of the diverse communities being served by the agency. Cultural competence matters for understanding how evaluation will and will not be relevant to the needs, assets, and circumstances of different groups or places.

The need for culturally competent practices extends to after an evaluation has been conducted. In terms of transparency, agencies should recognize the need to communicate findings in ways that eliminate potential biases in language by "respectfully and effectively convey[ing] important differences in the worldviews of key stakeholders" and "accurately reflect[ing] how individuals view their own group memberships and creat[ing] nuanced understandings that move beyond simple classifications." [49]

Finally, agencies should recognize that they may have potentially significant power and perceived status over the people, organizations, and communities involved in or affected by their evaluations. Cultural competence is a key component of ethics and respecting the dignity and rights of participants in agency-sponsored evaluations.

Discussion

As the Forum began applying an equity lens to agency evaluation policies and principles, one obvious question seemed to be whether or not equity should stand alone as a discrete principle. However, this question became less meaningful based on our analysis. The Forum has yet to identify any federal agency evaluation policies that explicitly reference equity as a principle and found only one agency evaluation policy (CDC) that referenced equity in any capacity.[viii] Looking more broadly for other words with similar meanings, only one agency (again the CDC) mentioned fairness; four referred to cultural competence, appropriateness, or context; and five referred in some way to diversity (see Appendix for details). As a result, a bigger-picture concern emerged: equity should in some way be apparent as a clear and consistent value in federal evaluation policy, but it isn't. Even where agencies may intend to include equity, or may work intentionally to operationalize equity in evaluations, policies do not clearly and consistently communicate this expectation.

As the discussion above shows, the existing principles have ample room to meaningfully incorporate equity without adding a new, discrete principle. Importantly, some agencies already describe the principles in ways that clearly relate to and align with equity. However, the lack of explicit references to equity—combined with the scarcity of references to similar concepts like fairness, cultural competence, and diversity—raises uncertainty about how agencies approach the issue and how consistently and intentionally agency staff and federally funded evaluators may act equitably under these policies. Silence regarding equity means that, while some agencies or individual evaluators could choose to implement the principles in ways that promote and reinforce equity, it is more difficult for internal and external stakeholders to know whether equity is a priority. Will the current principles effectively guide evaluators to be consistent and intentional in considering equity?

Creating room for equity is not the same as creating a clear expectation of or accountability for equity. At a minimum, agencies should describe the principles in ways that align with equity—such as through broad and specific references to stakeholders and by making sure that historical inequities like structural racism are active ethical considerations—in order to create accountability for equity in concrete ways. There should be no doubt among an agency's stakeholders that equity matters in evaluation and that agencies are accountable for considering equity as they prioritize, design, conduct, and release evaluations.

An easy response is to cite the breadth and diversity of agency missions, both across and within agencies. This is certainly true and could result in equity having different implications in different areas. However, the American Evaluation Association (AEA) offers a telling example about the applicability of equity across areas of evaluation. With the broad mission "to improve evaluation practices and methods, increase evaluation use, promote evaluation as a profession, and support the contribution of evaluation to the generation of theory and knowledge about effective human action," [50] the AEA focuses on evaluation as a field of study rather than on any specific policies, subjects, or outcome areas. Nonetheless, the AEA includes "Common Good and Equity" as one of its five guiding principles, suggesting that it may be presumptuous to dismiss equity as not applicable across the diversity of evaluations that federal youth-serving agencies conduct.

[viii] The CDC's Framework for Evaluation in Public Health, published in 1999, references social equity three times, including as a CDC operating principle and as a potential standard for judging program performance.

Recommendations: So what can the federal government do?

As federal evaluation policy continues to evolve, OMB and each agency have an opportunity to become leaders in making equity a clear and consistent value in evaluation.

OMB's role. The Evidence Act charges OMB to "issue guidance for program evaluation for agencies consistent with widely accepted standards for evaluation" and to "identify best practices for evaluation that would improve Federal program evaluation."[51] Through these actions, OMB will establish the tone, minimum thresholds, and also key considerations and practices worth broad consideration for federal evaluation as a whole. OMB can use these actions to clearly reflect equity as a value that, in keeping with many other types of OMB guidance, agencies can operationalize as appropriate to their missions and circumstances.

Individual agency roles. Agency Evaluation Officers, which OMB guidance M-19-23 mandated be named by July 2019, must establish and implement an evaluation policy that provides "stakeholders with a clear understanding of the expectations related to key principles, such as evaluation rigor, relevance, transparency, independence, and ethics."[52] Importantly, the OMB guidance does not confine agencies to these common five key principles. As Evaluation Officers work across their agencies to formulate or potentially refine evaluation policies, they will inherently be navigating and responding to the particular missions, needs, and cultures of their agencies as a whole, as well as to various bureaus and programs within their agencies. As appropriate to the mission and needs of their agency, Evaluation Officers should act deliberately to clearly articulate equity as a value within their agencies' evaluation policies.

Specific actions. Three options that OMB and agencies could consider to clarify equity as a value are as follows:

1. **Establish a tone** that values equity by referencing it in framing or introduction of the guidance (in OMB's case) or the evaluation policy (for agencies).
2. **Establish minimum expectations** around equity by explicitly including it as a standard or principle or by explicitly addressing it within other standards or principles.
3. **Establish key considerations and practices** that operationalize equity. While OMB is statutorily required to identify best practices, agencies could follow the example of USAID, SBA, and other agencies by creating operational guidelines that can supplement their official policies.

Finally, the federal government is not alone in facing challenges related to equity and evaluation. Just as the sections above highlight opportunities to engage stakeholders and diverse perspectives in executing evaluation policy, agencies and OMB might engage external stakeholders to learn how other entities are addressing equity in policy and practice related to evaluation. AEA is an obvious example, but it is not alone. In addition, the Equitable Evaluation Initiative (EEI) articulates principles specifically to promote equitable evaluation in the nonprofit and philanthropic sectors. For example, EEI holds that evaluative work should be designed and implemented in ways that are multiculturally valid and oriented toward participant ownership.[53]

The Democratizing Evidence in Education initiative, which works to ensure that "diverse stakeholders—people of different civic and professional roles, ages, race, ethnicity, income, and segments of society—have the power and opportunity to participate in producing and using research evidence to improve education for all," offers similar guiding principles. Their four principles focus on inclusive engagement, participation across the research process, responsive research agendas, and responsible evidence production and use.[54]

At the state level, there are fewer examples of evaluation policies. The Minnesota Department of Management and Budget has released an evaluation policy that utilizes the same five principles discussed above. The policy references rigor in a broad way and highlights the need for "evaluation priorities [to] reflect the interests, needs, and sensitivities of the populations served" under relevance.[55] Unlike federal policies, this policy also specifically references the U.S. Department of Health and Human Services' Belmont Report and the need to protect human subjects in research. While other states have examples of research or learning agendas, legislation that specifically defines evidence, and specific statewide strategic goals related to evidence, Minnesota was the only state identified with an evaluation policy.[56]

At a local level, communities are expanding their thinking about who can participate as a researcher, including by engaging young people in conducting research about their own communities. Nonprofits, such as the Brooklyn Bridge Alliance for Youth in Minneapolis,[57] have taken this approach, and cities such as Oakland[58] and Boston[59] also are experimenting with greater community outreach to inform their research and data work. Agencies such as DOL and CNCS have already embraced the notion in their evaluation policies that researchers and evaluators should reflect a diversity of perspectives, competencies, and backgrounds. Engaging more broadly with stakeholders may help to uncover the most feasible and appropriate ways for agencies to incorporate equity into evaluation policy.

Many agencies already have made significant strides to advance evidence-building and use by drafting and adopting evaluation policies that clearly articulate agency values and priorities around evaluation. With the implementation of the Evidence Act, federal policymakers now have an opportunity to update these evaluation policies or adopt them for the first time. Agencies should make clear in their policies that equity is a value by incorporating equity into their existing principles and standards, adding a discrete equity principle, and/or providing an explicit framing statement about the role of equity in evaluation.

Appendix 1: Summary Chart of Agency Evaluation Policies, Frameworks, and Guidelines

The chart on the next page summarizes our analysis of evaluation policies, frameworks, and guidelines for federal youth-serving agencies. We examined all agencies participating in the Interagency Working Group on Youth Programs. The chart states whether or not we found an evaluation policy—including either a formal policy or an informal policy in the form of a framework or other guidelines—for each agency. The chart then summarizes whether or not certain key terms such as "equity," "fairness," or "cultural competence" are mentioned in these documents. While no single term can fully capture the concepts discussed above, a collection of these terms may highlight equity as a key value that should guide agency evaluation activities. Finally, the chart analyzes whether or not agency policies (formal or informal) are conducive to equity in any of the five key principles as discussed above. Importantly, the chart is not meant to be a "check-the-box" exercise. Rather, we provide it for transparency about our own analysis and as a discussion starter for how each agency can better incorporate equity into their evaluation policies in ways that make sense for them.

Agency	Does it have an evaluation policy?	Does the policy mention equity?	Does the policy mention fairness?	Does the policy mention cultural competence?	Does the policy mention diversity?	How does it define rigor (broad or narrow)?A	How does it define stakeholder engagement (broad or narrow)?B	Does it distinguish between different purposes (non-distinct or distinct)?C	Does the policy reference non-agency stakeholders (general or specific)?D
Corporation for National & Community Service	Yes	No	No	No	"diverse...stakeholders" & "diversity among those carrying out evaluations"	Broad	Broad	Non-Distinct	Specific - states, territories, tribes, and grantees; and populations served; researchers; and other stakeholders.
Department of Agriculture - Foreign Agriculture Service	Yes	No	No	No	No	Broad	Broad	Distinct	Specific - program participants and partner governments
Department of Commerce	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Department of Defense - Undersecretary of Defense for Policy	Yes	No	No	"culturally appropriate"	No	Broad	Narrow	Non-Distinct	General
Department of Education - Institute of Education Sciences	Yes	No	No	No	"diverse interests"	Broad	Broad	Non-Distinct	Specific - states, territories, tribes, and grantees
Department of Health & Human Services (HHS) - Administration for Children & Families	Yes	No	No	No	"diverse populations...that ACF programs serve" & "diversity among those carrying out the work"	Broad	Broad	Distinct	Specific - states, territories, tribes, local grantees, populations served, and researchers
Department of Homeland Security	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Department of Housing & Urban Development	Yes	No	No	No	No	Broad	Broad	Non-Distinct	General
Department of Justice	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Department of Labor	Yes	No	No	No	"diverse...stakeholders" & "diversity among those carrying out evaluations"	Broad	Broad	Non-Distinct	Specific - states, territories, tribes, grantees, populations served, and researchers
Department of State	Yes	No	No	No	No	Broad	Broad	Distinct	Specific - universities and colleges, NGO's, and bilateral or multilateral partners
Department of the Interior	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Department of Transportation	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Environmental Protection Agency	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HHS - Centers for Disease Control and Prevention	Yes	No	Yes	"recognize cultural interests and needs" & "cultural contexts of clients"	No	Broad	Broad	Non-Distinct	General
HHS - Substance Abuse & Mental Health Services Administration	Yes	No	No	"cultural competency consultation"	No	Broad	Broad	Non-Distinct	Specific - grantee, scientific community
Millenium Challenge Corporation	Yes	No	No	No	No	Broad	Broad	Distinct	General

National Science Foundation	Joint Statement with ED	No	No	No	No	No	Broad	Narrow	Distinct	General
Office of National Drug Control Policy	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Small Business Administration	Yes	No	No	"cultural.... sensitivities"	No	No	Broad	Narrow	Distinct	General - toward non-agency stakeholders.
Social Security Administration	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
US Agency for International Development	Yes	No	No	No	"diverse set of approaches"	Broad	Broad	Broad	Distinct	Specific – in-county partners, beneficiaries

[A] Narrow means that policies only highlight a defined subset of common methodologies (e.g., RCT and quasi-experimental) as rigorous. Broad means that policies highlight a range of methodologies as having the potential to be rigorous.

[B] Narrow means that policies only cite federal stakeholders when discussing the need for stakeholder engagement to ensure relevant evaluations. Broad means that policies cite non-federal stakeholders as well.

[C] Non-distinct means that polices only cite general purposes such as "for decision-making" or do not reference the purposes of evaluations. Distinct means that policies cite specific purposes such as "resource allocation," "grantmaking," or "improving effectiveness."

[D] General means that policies do not explicitly name non-federal stakeholders. Specific means that policies explicitly mention at least two non-federal stakeholders.

Endnotes

1 Reynolds, K. & Ramakrishnan, K. (2018). Evidence-Based Policymaking at the State Level A Guide for Governors. Washington, D.C.: The Urban Institute. Available at: https://www.urban.org/sites/default/files/publication/99293/evidence-based_policymaking_at_the_state_level.pdf.

2 Annie E. Casey Foundation (2014). Race Equity and Inclusion Action Guide. Baltimore, MD: Annie E. Casey Foundation. Available at: https://www.aecf.org/m/resourcedoc/AECF_EmbracingEquity7Steps-2014.pdf.

3 American Evaluation Association (2019). About AEA. Washington, D.C.: American Evaluation Association. Available at: https://www.eval.org/p/cm/ld/fid=51

4 Office of Management and Budget (2019). Analytical Perspectives: Building and Using Evidence to Improve Government Effectiveness from A Budget for a Better America: Fiscal year 2020 Budget of the U.S. Government. Washington, D.C.: Office of Management and Budget. Available at: https://www.whitehouse.gov/wp-content/uploads/2019/03/ap_6_evidence-fy2020.pdf.

5 U.S. Department of Health and Human Services (2019). Home Visiting Evidence of Effectiveness. Washington, D.C.: U.S. Department of Health and Human Services. Available At: https://homvee.acf.hhs.gov/.

6 Administration for Children and Families (2011). Home Visiting Evidence of Effectiveness Review: Executive Summary. Washington, D.C.: U.S. Department of Health and Human Services. Available at: https://www.acf.hhs.gov/opre/resource/home-visiting-evidence-of-effectiveness-review-executive-summary.

7 Administration for Children and Families (2019). Maternal, Infant, and Early Childhood Home Visiting Evaluation. Washington, D.C.: U.S. Department of Health and Human Services. Available at: https://www.acf.hhs.gov/opre/research/project/maternal-infant-and-early-childhood-home-visiting-evaluation-mihope.

8 Home Visiting Applied Research Collaborative (2019). Innovation towards Precision Home Visiting. Washington, D.C.: Johns Hopkins Bloomberg School of Public Health. Available at: https://www.hvresearch.org/.

9 Home Visiting Applied Research Collaborative (2018). Introduction to Precision Home Visiting. Washington, D.C.: Johns Hopkins Bloomberg School of Public Health. Available at: https://www.hvresearch.org/precision-home-visiting/introduction-to-precision-home-visiting/.

10 U.S. Department of Education and National Science Foundation (2013). Common Guidelines for Education Research and Development. Washington, D.C.: National Science Foundation. Available at: https://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf

11 Small Business Association (2019)

12 Office of Management and Budget (2019). Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance. Washington, D.C.: Office of Management and Budget. Available at: https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf.

13 Thaddeus Ferber and Alex Sileo (2017). Managing for Success. Washington, D.C.: Forum for Youth Investment. Available at: https://forumfyi.org/knowledge-center/managing-for-success/.

14 Commission on Evidence-Based Policymaking (2018). The Promise of Evidence-Based Policymaking. Washington, D.C.: U.S. Census. Available at: https://www.cep.gov/report/cep-final-report.pdf.

15 Farrow, F. & Morrison, S. (2019). Placing Equity Concerns at the Center of Knowledge Development. Washington, D.C.: Center for the Study of Social Policy. Available at: https://cssp.org/resource/equity-at-the-center/.

16 Dean-Coffey, J., Casey, J., & Caldwell, L. D. (2014). Raising the Bar – Integrating Cultural Competence and Equity: Equitable Evaluation. Washington, D.C.: The Foundation Review, 6(2). Available at: https://doi.org/10.9707/1944-5660.1203.

17 Administration for Children and Families (2013). Evaluation Policy. Washington, D.C.: U.S. Department of Health and Human Services. Available at: https://www.acf.hhs.gov/sites/default/files/opre/acf_evaluation_policy_november_2013.pdf.
U.S. Department of Labor (2013). U.S. Department of Labor Evaluation Policy. Washington, D.C.: U.S. Department of Labor. Available at: https://www.dol.gov/agencies/oasp/evaluation/EvaluationPolicy.
U.S. Department of Housing and Urban Development (2016). HUD Program Evaluation Policy—Policy Statement. Washington, D.C.: U.S. Department of Housing and Urban Development. Available at: https://www.govinfo.gov/content/pkg/FR-2016-12-06/pdf/2016-29215.pdf.

[18] U.S. Department of Education and National Science Foundation (2013). Common Guidelines for Education Research and Development. Washington, D.C.: National Science Foundation. Available at: https://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf.

[19] Schorr, L.B. and Farrow, F. (2011). Expanding the Evidence Universe: Doing Better by Knowing More. Washington, D.C.: Center for the Study of Social Policy. Available at: https://cssp.org/wp-content/uploads/2019/04/Expanding-Evidence-the- Evidence-Universe_Doing-Better-by-Knowing-More_December-2011.pdf.

[20] Government Accountability Office (2009). Program Evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions. Washington, D.C.: Government Accountability Office. Available at: https://www.gao.gov/assets/300/298907.pdf.

[21] Whitesell, N. R. (2017). Evidence and equity: Challenges in research design, OPRE Report #2017-76. Washington, DC: U.S. Department of Health and Human Services. Available at: https://www.acf.hhs.gov/sites/default/files/opre/whitesell_opre_brief_clean_508.pdf.

[22] Schorr, L.B. and Farrow, F. (2011). and Schorr, L.B. and Farrow, F. (2014). An Evidence Framework to Improve Results. Washington, D.C.: Center for the Study of Social Policy. Available at: https://cssp.org/wp-content/uploads/2019/07/An-Evidence-Framework-to-Improve-Results-11.14.pdf.

[23] American Evaluation Association (2011). American Evaluation Association Statement on Cultural Competence in Evaluation. Washington, D.C.: American Evaluation Association. Available at: https://www.eval.org/ccstatement.

[24] Joint Committee on Standards for Educational Evaluation (2011). Program Evaluation Standards Statements. Washington, D.C.: Joint Committee on Standards for Education Evaluation. Available at: http://www.jcsee.org/program-evaluation-standards-statements.

[25] Whitesell, N. R. (2017).

[26] U.S. Agency for International Development (2016). Evaluation: Learning from Experience. Washington, D.C.: U.S. Agency for International Development. Available at: https://www.usaid.gov/sites/default/files/documents/1870/USAIDEvaluationPolicy.pdf.

[27] Small Business Administration (2019).

[28] Corporation for National and Community Service (2016). Policies and Procedures. Subject: Evaluation Policy. Washington, D.C.: Corporation for National and Community Service. Available at: https://www.nationalservice.gov/sites/default/files/documents/CNCSEvaluationPolicy.pdf.

[29] Ibid. U.S. Department of Labor (2013). U.S. Department of Labor Evaluation Policy. Washington, D.C.: U.S. Department Labor. Available at: https://www.dol.gov/agencies/oasp/evaluation/EvaluationPolicy.

[30] U.S. Agency for International Development (2016).

[31] Centers for Disease Control and Prevention (2017). Evaluation Standards. Atlanta, GA: Centers for Disease Control and Prevention. Available at: https://www.cdc.gov/eval/standards/index.htm. Joint Committee on Standards for Educational Evaluation (2011).

[32] Administration for Children and Families (2013).

[33] Joint Committee on Standards for Educational Evaluation (2010). Program Evaluation Standards. Washington, D.C.: Available at: https://jcsee.org/program/.

[34] Joint Committee on Standards for Educational Evaluation (2010).

[35] Small Business Administration (2019).

[36] Administration for Children and Families (2013). U.S. Department of Labor (2013).

[37] National Science Foundation (2013). Joint Committee on Standards for Educational Evaluation (2010).

[38] Chicago Beyond (2019). Why am I always being researched? Chicago, IL: Chicago Beyond. Available at: https://chicagobeyond.org/researchequity/.

[39] Joint Committee on Standards for Educational Evaluation (2010).

[40] U.S. Department of Labor (2013).

[41] Administration for Children and Families (2019). ACF Evaluation Policy. Washington, D.C.: U.S. Department of Health and Human Services. Available at: https://www.acf.hhs.gov/opre/resource/acf-evaluation-policy.

[42] Small Business Administration (2019).

[43] Joint Committee on Standards for Educational Evaluation (2010).

[44] Children's Services Council of Broward County (2019). Racial Equity. Broward, FL: Children's Services Council of Broward County. Available at: https://www.cscbroward.org/community-builders/committee/racial-equity.

[45] Administration for Children and Families (2013).

[46] Chicago Beyond (2019).

[47] Dean-Coffey, J., Casey, J., & Caldwell, L. D. (2014).

[48] American Evaluation Association (2011).

[49] American Evaluation Association (2011).

[50] American Evaluation Association (2019). About AEA. Washington, D.C.: American Evaluation Association. Available at: https://www.eval.org/p/cm/ld/fid=4.

[51] Foundations for Evidence-Based Policymaking Act of 2018. Available at: https://www.congress.gov/bill/115th-congress/house-bill/4174

[52] Office of Management and Budget (2019).

[53] Democratizing Evidence in Education. "Statement of Principles." Available at: https://democratizing.edgepartners.org/statement-of-principles

[54] Equitable Evaluation (2019). The EE Framework. Washington, D.C.: Equitable Evaluation. Available at: https://www.equitableeval.org/ee-framework.

[55] Minnesota Department of Management and Budget (2019). Evaluation Policy. Saint Paul, MN: Minnesota Department of Management and Budget. Available at: https://mn.gov/mmb-stat/results-first/evaluation-policy.pdf.

[56] Results for America (2019). Invest In What Works State Standards of Excellence. Washington, D.C.: Results for America. Available at: https://2019state.results4america.org/2019-invest-in-what-works-state-standard-of-excellence.pdf.

[57] Brooklyn Bridge Alliance for Youth (2019). Youth Engagement. Brooklyn Center, MN: Brooklyn Bridge Alliance for Youth. Available at: http://www.brooklynsallianceforyouth.org/youth-engagement.

[58] City of Oakland (2019). Equity in Downtown Oakland. Oakland, CA: City of Oakland. Available at: http://eqtdto.com/#welcome.

[59] Boston Office of New Urban Mechanics (2019). Civic Research Agenda. Boston. MA: Boston Office of New Urban Mechanics. Available at: https://www.boston.gov/departments/new-urban-mechanics/civic-research-agenda.

the forum
FOR YOUTH INVESTMENT

www.forumfyi.org

DC Office: 7064 Eastern Avenue, NW Washington, DC 20012
Michigan Office: 301 W. Michigan Ave., Suite 200, Ypsilanti, Michigan 48197