



## **Equitable Evaluation: Incorporating an Equity Lens into Your Evaluation Practice**



*A White Paper by Stacey S. Merola, Ph.D.*

**Guardians of Honor, LLC (GOH)  
1990 K Street NW  
Suite 650  
Washington, DC 20006**

*We would like to thank Tania Jarosewich and Alan Porowski for their helpful comments and suggestions on this paper.*

## Introduction

### What is Equitable Evaluation?

UNICEF (2011) defined equity as: *“all children hav[ing] an opportunity to survive, develop, and reach their full potential, without discrimination, bias or favoritism.”* Societal pressures do not end at childhood and people strive to reach their full potential throughout their life.

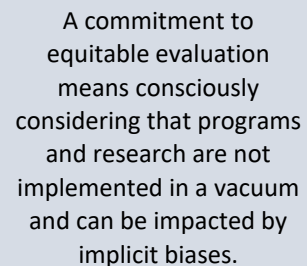
Therefore, this idea broadened to the whole population is that all people should have an opportunity to thrive without discrimination. Undertaking an equitable evaluation means taking a hard look at how these factors are manifesting themselves, and how they might impact study results.

A commitment to equitable evaluation means consciously considering that programs and research are not implemented in a vacuum and can be impacted by both explicit and implicit biases. We work to ensure no explicit bias, but implicit biases can affect the evaluation if we do not attend to certain factors. Relationships, and particularly the power dynamics embedded in these relationships, can influence evaluation results and program implementation. For example, the relationship of the study participants to the evaluator, the relationship of the evaluator to the funder, the relationships among the study participants and the larger community, and the relationships of the study participants to the larger society and history can all impact not only the results of the research, but also the questions that are posed and how

the research is designed. To comprehend these relationships, researchers must gain a holistic view of the

populations being studied from the start of the research, incorporating understanding of historical context and inequitable systems into the interpretation of results, engaging the community in partnership, and making findings transparent to the communities being studied. It also means that there is not a one-size fits all approach to evaluation, but rather that rigor and validity need to be considered in the context of culture, power, and history.

In this paper, GOH discusses a brief history of equitable evaluation, the principles of conducting equitable evaluations, and examples from our own work. GOH has experience leveraging culturally relevant mixed-methods research practices, developing relationships with traditionally underrepresented groups, portfolio analysis, and community-based participatory research, all of which can be incorporated into an equitable evaluation framework.



A commitment to equitable evaluation means consciously considering that programs and research are not implemented in a vacuum and can be impacted by implicit biases.

## A Brief History of Equitable Evaluation

### An Idea That Has Finally Found Its Moment

Though there have been calls over 40 years to move away from the strict empiricism that formed the basis of evaluation research (see Ball's 1977 argument to study processes and relationships as an example) and social science research more broadly, it was not until the turn of the 21<sup>st</sup> century that movement to view program funding and evaluation of investments from the perspective of promoting equity gained momentum. The last decade has been one of rapid creation and adaptation of frameworks to promote equitable programming and evaluation practices within the philanthropic and nonprofit realms. Most recently these frameworks have begun to spread into the public sector.

### Organizations Stand to Support Equitable Evaluation

In 2011, the American Evaluation Association (AEA) released a statement on Cultural Competence in Evaluation applicable to evaluations conducted in the U.S., which was the culmination of six years of work by the Cultural Competence in Evaluation Task Force and close to 12 years of thought and planning on how to address the needs of evaluators working across cultural contexts (AEA, 2011). AEA took a public stance via this statement that evaluation team members must demonstrate cultural competence, and that valid evaluation results consider cultural context. AEA (2011) defines cultural competence as a world view or

A culturally competent evaluator is prepared to engage with diverse segments of communities to include cultural and contextual dimensions important to the evaluation. Culturally competent evaluators respect the cultures represented in the evaluation throughout the process.

-American Evaluation Association (2011)

perspective on culture, rather than a specific set of skills, and that “a culturally competent evaluator is prepared to engage with diverse segments of communities to include cultural and contextual dimensions important to the evaluation. Culturally competent evaluators respect the cultures represented in the evaluation throughout the process”(p.1). Additionally, AEA (2011) posits that a lack of cultural competence threatens the validity of evaluations in that, “inaccurate or incomplete understandings of culture introduce systematic error that threatens validity” (p.5). In 2018 the AEA released a new version of their guiding principles which included the provision that “evaluators strive to contribute to the common good and advancement of an equitable and just society” (AEA, 2018).

In 2014, the Annie E. Casey Foundation published guidelines for promoting racial equity both in evaluation practices as well as the funding decisions of foundations. One of the steps is conducting a “racial equity impact assessment for all policies and decision making” (The Annie E. Casey Foundation, 2014).

In 2016, the National Committee for Responsive Philanthropy (NCRP) released a 10-year framework for strategies to promote the goal of a more equitable and just society by leveraging the resources of the philanthropic and nonprofit sectors (National Committee for Responsive Philanthropy, 2016).

One of the tools employed by the NCRP to achieve this goal is the evaluation of how well the policies and practices of foundations are benefitting those with the “least wealth, opportunity, and power”

(National Committee for Responsive Philanthropy, 2016).

Building on these trends, the Equitable Evaluation Initiative (EEI), formed in 2018 and officially launched in 2019, offers a vision in which “evaluative practice works toward creating a world in which we all thrive and where the multiple truths of the human experience are valued and valid” (Equitable Evaluation Initiative, n.d.). The endeavor focuses on building a pool of evaluation practitioners dedicated to the principles of equitable evaluation. Similarly, the Annie E. Casey foundation has created the Expanding the Bench initiative which seeks to increase both the numbers of evaluators from historically underrepresented groups and the numbers of practitioners of culturally responsive and equitable evaluation (CREE) practices (Expanding the Bench, 2021).

Much federal data is not disaggregated by demographics such as race, gender, and disability status. One implication of this challenge is that a lack of measurement leads to a paucity of information related to the impacts of policies on specific groups.

a paucity of information related to the impacts of policies on specific groups. For example, recent news reports have uncovered that since the IRS does not collect information on race, it has only recently come to light how much tax laws have traditionally disadvantaged people of color (Faler, 2021).


The Interagency Equitable Data Working Group is identifying gaps in equitable data collection and reporting through a case study approach. The case studies focus on questions and topics that have been difficult to address due to lack of appropriate data. For example, “What disaggregated data is needed to measure whether the CARES Act benefits and American Rescue Plan benefits were equitably distributed?” (Nelson and Wardell, 2021). By addressing these questions in a holistic way, the group will develop recommendations for potential changes to legislation and agency procedures.

Nelson and Wardell (2021) acknowledge the challenges of reporting detailed information about groups while protecting their

privacy. Researchers must design their studies to ensure adequate and accurate representation across groups to allow for a meaningful disaggregation and avoid the pitfalls of small group sizes. Traditionally researchers have done this by oversampling from segments of the population that would likely make up only a small percentage of a traditional random sample. A good example of what can happen when this is not done was a 2016 survey conducted by the Washington Post. In 2016, survey results reported by the Washington Post and extensively disseminated by other media outlets indicated that 9 in 10 Native Americans were not offended by the former name for the Washington football team, the Redskins. This

### Greater Representation and Understanding Through Data Disaggregation Policy

On his first day of office, January 20, 2021, President Joe Biden signed an executive order initiating a policy that the federal government will pursue equity for all (The White House, 2021). The order called for an assessment of internal agency policies to promote equity and greater engagement with underserved communities and established an Interagency Working Group on Equitable Data. A challenge facing this working group is that much federal data is not disaggregated by demographics such as race, ethnicity, gender, LGBTQ+, veteran, and disability status. One implication of this challenge is that a lack of measurement leads to



Washington Post survey was problematic since the sampling frame did not consist exclusively of Native people, but the whole population of the United States. This is akin to wanting to do a survey of teachers, and then including people from all occupations and hoping some teachers will be in the sample to respond to your survey. In the case of the Washington Post, an ongoing survey of the general population (with a 10% response rate) without any stratification or over sampling of Native Americans was used as a vehicle for their questions. Respondents who self-identified as Native American were then given some additional questions to answer regarding the Redskins' name, an ordering that could have in itself biased the results depending on what was asked on the main part of the survey. The researchers aggregated the results from respondents who self-identified as Native American over a period of weeks and weighted the responses based on U.S. Census data. After aggregating, the final sample was 504 Native Americans from all 50 states, meaning that about 10 Native Americans were representing all Native people in any given state.

Not surprisingly, the survey caused quite a bit of controversy. Though technically the results were disaggregated, there were questions about the representativeness of the Native American proportion of the sample, not to the *national* population, but to the *Native American* population. Though something of a moot point since the name has been changed, this survey may have been less controversial if the effort

had been made to recruit Native Americans directly either as part of a stratified random sample or a stand-alone survey of Native Americans. Recruitment would have required a good-faith effort to engage the Native American community early in the study design, which would have also lent credibility to the study.

One must also have valid measures that allow for meaningful disaggregation. A thoughtful reevaluation of the construct validity of existing and proposed measures should be undertaken when commencing research. One historically contentious example of a need to reevaluate measures is the use of racial/ethnic variables in medical research. Ioannidis, Powe, and Yancy (2021) provide a cogent discussion of why it is important to thoughtfully consider the inclusion of race/ethnicity variables in research and if they are included, to consider the implications of choosing a specific reference group. Ioannidis et. al (2021) recommend that researchers review prior work including race/ethnicity as a measure to ascertain whether anything new can be added to the existing corpus of knowledge, and, if it is determined that an additional study including race/ethnicity adds value to the field, to carefully consider: 1) whether there are any mediating biological or sociological factors that should also be included, 2) whether in comparative analyses White is an appropriate reference group, and 3) whether the research will have positive or deleterious clinical or social consequences.



## Key Principles

Equitable evaluation is not one specific research method or model, it is an approach that can be, and should be, incorporated into a variety of research designs, tools, measurement, processes, and analysis methods (Equitable Evaluation Initiative, n.d.). A key principle of equitable evaluation is that an evaluation should be tailored to local context, incorporating local culture, history, power dynamics, resource distribution, etc. Community support and understanding of the need for evaluation should also be cultivated, so evaluation moves from something performed exclusively for and by outsiders, to something the community actively participates in and brings them value, such as through Community-Based Participatory Research (CBPR), which engages a community as an equal partner with evaluators to respond to community needs.<sup>1</sup> To be most effective, this tailoring should begin during the initial stages of the evaluation and continue through to reporting of the research findings.

In 2011, Bamberger and Segone provided guidelines for the development of equitable evaluations, suggesting refinements to existing evaluation methods rather than calling for the development of completely new methods. They

also called for empowering traditionally disadvantaged groups through the evaluation process. The authors specified that an evaluation of complex equity-focused policies might require more facets than the evaluation of a single project or program. Some additional, primarily qualitative, approaches recommended by Bamberger and Segone (2011) to understand the social systems in which programs are being

Evaluations should be tailored to local context, incorporating local culture, history, power dynamics, resource distribution, etc. Community support and understanding of the need for evaluation should also be cultivated, so evaluation moves from something performed exclusively for and by outsiders, to something the community actively participates in and brings them value.

implemented (i.e., cultural, historical, community relationships, etc.) include: 1) Systems Dynamics Approach, 2) Soft Systems Methodology<sup>2</sup>, 3) Cultural-Historical Activity Theory<sup>3</sup>, 4) Unpacking Complex Policies, 5)

<sup>1</sup> The National Institute on Minority Health and Health Disparities has a program focused on CBPR and provides a nice overview of this approach, <https://www.nimhd.nih.gov/programs/extramural/community-based-participatory.html>

<sup>2</sup> From page 64 of Bamberger and Segone (2011): *Soft Systems Methodology focuses on the multiple perspectives of a particular situation. The first step is to provide a “rich picture” of the situation and then to provide a “root definition” (the essential elements) of the situation in terms of the beneficiaries; other actors; the transformation process (of inputs into outputs); the world-views of the main actors; the system*

*owners (who have veto power over the system); and environmental constraints.*

<sup>3</sup> From page 65 of Bamberger and Segone (2011):

*Key elements of the Cultural-Historical Activity Theory approach are that: systems have a defined purpose; they are multi-voiced (different actors have different perspectives); systems are historical and draw strongly from the past; changes in a system are produced largely by contradictions which generate tensions and often conflict; and contradictions provide the primary means by which actors learn and changes take place. The changes can produce further*

Pipeline Designs, 6) Policy Gap Analysis, 7) Concept Mapping, and 8) Portfolio Analysis.

Bamberger and Segone's (2011) guidelines for empowering disadvantaged groups recommend that equitable evaluation should be highly ethical, culturally sensitive, and ensure that marginalized groups are involved with the evaluation design process from the inception of the project. The Annie E. Casey Foundation (2014) expanded on these ideas with seven additional steps for advancing racial equity:

- 1) Establish an understanding of race equity and inclusion principles.
- 2) Engage affected populations and stakeholders.
- 3) Gather and analyze disaggregated data.
- 4) Conduct a systems analysis of root causes and inequities.
- 5) Identify strategies and target resources to address root causes of inequities.
- 6) Conduct race equity impact assessment for all policies and decision making.
- 7) Continuously evaluate effectiveness and adapt strategies.<sup>4</sup>

As implied by the strategies above (and as explicitly recommended by Bamberger and Segone, 2011), a mixed-methods research design is well suited to incorporating diverse perspectives and studying the mechanisms that lead to specific outcomes. Johnson and Onwuegbuzie (2004) define mixed-methods research as "class of research where the researcher mixes or combines quantitative and

qualitative research techniques, methods, approaches, concepts or language into a single study" (p.17). A mixed-methods research design incorporates both quantitative and qualitative data, providing the benefits of being able to conduct statistical analyses and provide generalizations based on the quantitative data, with the insights that can be gathered from the analysis of in-depth, qualitative data. An added benefit of mixed-methods research is that it allows for triangulation, or the ability to assess and confirm findings from different perspectives due to the multiple data sources. Bamberger and Segone (2011) pose that triangulation is a particularly important component of equitable evaluation in that it allows for the synthesis of different perspectives.

There are numerous configurations of mixed-methods research designs, with variations in the timing of the qualitative and quantitative components and ranging in the proportions of qualitative and quantitative data collected. Johnson and Onwuegbuzie (2004) provide nine mixed-methods designs and note that there can easily be more. Johnson and Onwuegbuzie (2004) recommend that "researchers should mindfully create designs that effectively answer their research questions" (p.20). This flexibility and pragmatism help make these approaches ideal for the study of equity given the wide range of processes and factors that need to be considered and will vary based on environment.

---

*contradictions, so processes of change are often cyclical.*

<sup>4</sup> A side-by-side comparison of recommendations from AEA (2011), the Annie E. Casey Foundation

(2014), and Bamberger and Segone (2011) is presented in the Appendix A.

## Recommendations

Based on the foundations of equitable evaluation described above and on our own research, GOH suggests the following considerations for conducting research with an equitable lens:



***Incorporating an emphasis on equity into research requires altering course from “business-as-usual” practices.***

Successfully implementing these research designs means allowing extra time to build trust and determine questions that are relevant to the community (e.g., participatory research), as well as methods of providing the information back to the communities in useful ways, to ensure the communities have directly benefited at the completion of the evaluation, and that the exchange of information has not solely flowed in one direction from the community to the evaluator.



***Consider carefully which groups are being included in the research and which are being excluded.*** If what we measure is what we know about, one needs to consider what is excluded as well as what is included, and whether the groups or information that are excluded really are not necessary or are being excluded for other reasons. One well-documented example is the gender bias existent in traditional health research where women were excluded from clinical trials and their health concerns understudied (see Holdcroft, 2007 for an overview). Other examples are the overlooking of the impacts of programs on disabled people (see Rios, Magasi, Novak, and Harniss (2016) for a discussion of the exclusion of disabled people from health research) and rural populations (see Hayes (2021) for a discussion of a forthcoming study from Cornell University on the impacts of government policies on “left behind” rural communities).



***Incorporate a plan to disaggregate data into the initial stages of the research design.*** Sampling frames and designs must be thoughtfully created to have both a valid and adequate sample upon which to draw conclusions from disaggregation. This may mean using several techniques, including oversampling some segments of the population; stratified random sampling; and purposeful sampling, a technique commonly used in qualitative research to ensure “information rich” participants are included in the research (Palinkas, Horwitz, Green, Wisdom, Duan, and Hoagwood, 2015). Multiple sampling frames may be needed to have valid subsamples (e.g., having to sample from multiple Tribal Nations to get an adequate, representative sample of Native Americans).



***Consider moderating and mediating processes.*** Evaluations should be approached with a curiosity regarding underlying processes existing beyond the intervention itself and questions related to program implementation. One way to gain a fuller picture of underlying processes is

Evaluations should be approached with a curiosity regarding underlying processes existing beyond the intervention itself and questions related to program implementation.



If what we measure is what we know about, one needs to consider what is excluded as well as what is included, and whether the groups or information that are excluded really are not necessary or are being excluded for other reasons.

to consider potential moderators and mediators of the effect. Moderators are the context under which an effect occurs and will impact the direction or strength of an effect, while a mediator is an in-between step necessary for achieving an effect. A consideration of moderators and mediators in statistical modeling allows one to achieve a fuller picture of why, how, and to what extent an intervention is achieving an effect. Mediators and moderators could be mapped out in the initial logic model created for the program, and then tested. A consideration of moderators would be helpful to include as part of an equitable evaluation in that gender, race, and socioeconomic status are common ones.



***Qualitative data must be an integrated part of the mixed research design, not just an add on piece.*** Mixed-methods designs can be a powerful tool in the equitable evaluation toolkit. To take full advantage of the strengths of this design, qualitative and quantitative data need to be incorporated at each step in the research process. Qualitative data cannot simply provide examples or quotes for trends seen in the quantitative data. Consider all data as having equal weight, with qualitative data illuminating underlying trends, mediators, moderators, histories, beliefs, and power dynamics to better understand the processes and mechanics behind quantitative trends and findings.



***Undertake a thoughtful approach to measurement and construct validity.*** A thoughtful evaluation of the construct validity of existing and proposed measures should be undertaken as part of the research design. Consideration should be given to whether the measure is adding value to the research, and whether the measure is capturing the concept of interest, or if there a mediating process that should be included in the study. Also consider if the appropriate reference group is being used for comparisons. The ordering of categories in categorical variables can also have an impact on how people respond, as can placement of a question in a survey.

## Examples From Our Experience

Throughout its history Guardians of Honor, LLC (GOH) has been dedicated to giving voice and respect to underserved communities and has a sizeable portfolio of culturally relevant evaluations across its 24-year period of performance. GOH has contributed to equitable evaluation networks such as Expanding the Bench initiative, whose mission is to support diverse evaluators, and funders of evaluation to value, practice, and promote a culturally responsive and equitable evaluation (CREE) ecosystem (see [About Expanding the Bench – Expanding the Bench](#) for more information).

GOH’s commitment to equitable evaluation strategies and approaches ensure that our research and evaluation activities promote diversity, equity, and inclusion – such as mixed methods and Community-Based Participatory Research designs (CBPR). We adhere to the American Evaluation Association’s guidelines for contributing to the common good and equity through 1) balancing the needs of evaluation stakeholders while maintaining the rigor of the study; 2) fostering transparency and equitable access to information through the active sharing of data and findings; and 3) monitoring and mitigating power imbalances or threats to the common good. We also incorporate the Expanding the Bench, CREE principle of “integrating diversity, inclusion, and equity into all phases of evaluation” (Expanding the Bench, 2021). Paramount is establishing a culture of inclusivity and reciprocity with communities whereby they can clearly see that their voices and inputs are reflected in our approaches and findings; and individuals and communities benefit from the evaluation and findings.

An example of our approach was in community conversations in Flint, Michigan for AmeriCorps


to better understand the role of volunteerism in responding to disasters. The project began with a focus on AmeriCorps research questions, with questions of direct interest to the community included as the research progressed. The preliminary reports focused discussion at a series of targeted community conversations and town hall-style meetings during which researchers and community members discussed the findings. The Flint community conversations on civic health and volunteering facilitated open dialogue and provided important insights into how to promote and measure civic engagement and about the characteristics of successful community coalitions in response to disasters.

Additional qualitative research has consisted of listening sessions and portfolio analysis. In Fall 2020, GOH supported a series of virtual Tribal Consultations for the Office of Head Start (OHS), meetings with tribal leaders similar to listening sessions.

During these consultations, tribal leaders shared testimonies with OHS staff members. OHS used this feedback to learn what was working, and not working, with the Head Start programs serving tribal communities.

The challenge in the 2020 consultations was how to incorporate cultural norms and traditions due to social distance restrictions. In many ways, virtual platforms are in direct opposition to tribal norms. Virtual platforms are not conducive to open dialogue since the meeting view can limit the number of participants seen on a screen at any given time, and the show producer and facilitator take on an important role – whereas typical tribal consultations are led by the tribal leaders. GOH staff worked directly with tribal leaders to develop the agenda, facilitation techniques, and virtual platform configuration needed to foster trust, transparency, and open dialogue.

Paramount is establishing a culture of inclusivity and reciprocity with target communities whereby they can clearly see that their voices and inputs are reflected in our approaches and findings; and individuals and communities benefit.



Since 2004, GOH has supported a vast array of National Science Foundation (NSF) broadening participation initiatives using the types of portfolio analysis techniques recommended by Bamburger and Segone (2011) to assess impact on equity of initiatives with many individual projects and programs. We conducted a meta-evaluation of the NSF Centers of Research Excellence in Science and Technology (CREST) program, program analysis of the NSF Historically Black Colleges and Universities Undergraduate Program (HBCU UP) program, and served on the leadership team for the Louis Stokes Alliances for Minority Participation (LSAMP) Midwest Center of Excellence. On behalf of the NSF, we convened an NSF working group on Program Evaluation for Broadening Participation Programs and provided technical writing for the public dissemination of the conference proceedings report. The focus of this report is to articulate culturally sensitive evaluation methodologies and best practices. We also conducted portfolio analyses to identify investments and impacts of the NSF broadening participation investment portfolio, including all programs in the Human Resources Division. We utilized quantitative, qualitative, and bibliometric analysis to identify grant project outcomes and report on broader impacts of various grant projects and programs.

Finally, we have designed and implemented mixed-methods evaluations with diverse communities in which we triangulate across data sources to provide insights about the processes at

work. Dr. Stacey Merola has been the external evaluator on seven STEM education research projects for The Y in Central Maryland funded through the U.S. Department of Education (ED) 21<sup>st</sup> Century Community Learning Centers (21<sup>st</sup> CCLC) program. The 21<sup>st</sup> CCLC program supports the creation of community learning centers that provide academic enrichment opportunities during non-school hours for children, particularly students who attend high-poverty and low-performing schools, with an articulated emphasis on hands-on and experiential STEM learning and socio-emotional well-being. We work collaboratively with the Y to collect relevant data and information to assess the effectiveness of programs; and glean insights into what works, for whom, and in what context. Evaluation activities meet the requirements of the students, schools, and Y staff without sacrificing methodological rigor. Working in collaboration with the Y we have provided trainings to both Y staff and parents, in the importance of evaluation for the success of the programs and how accurate and timely data collection is essential. We have also provided technical assistance in formulating measures and goals. We have adapted data collection instruments for a range of academic abilities and English-language proficiencies. Data from quantitative and qualitative sources are triangulated to provide a full picture of the effects of the program on student achievement, school attendance and socio-emotional learning. Fidelity of implementation is also assessed.



## Summary

The equitable evaluation movement has made great strides over the last decade, culminating in an executive order calling for federal agencies to implement equitable data policies. Rather than a specific research method, equitable evaluation means recognizing that a one-size fits all approach to research will not be adequate. It means considering equity at every part of the research process – planning, analyses, and reporting of results, as well as the research products that are created. This paper presents the following recommendations for incorporating an equitable framework into your research:


- Incorporating an emphasis on equity into research requires altering course from “business-as-usual” practices.
- A plan to disaggregate data must be incorporated into the initial stages of the research design.
- Qualitative data must be an integrated part of the mixed research design, not just an add on piece.
- Consider moderating and mediating processes.
- Undertake a thoughtful approach to construct validity.
- Consider carefully which groups are being included in the research and which are being excluded.

GOH has experience conducting equitable evaluations for a range of clients across its 24 years in business. For more information, contact Stacey Merola at [smerola@gohnow.com](mailto:smerola@gohnow.com).

## References

- American Evaluation Association. (2011). *American Evaluation Association public statement on cultural competence in evaluation*. Fairhaven, MA: Author. Retrieved from [aea.cultural.competence.statement.pdf \(eval.org\)](http://aea.cultural.competence.statement.pdf)
- American Evaluation Association. (2018). *Guiding principles for evaluators*. Washington, DC: Author. Retrieved from <https://www.eval.org/p/cm/ld/fid=51>
- Ball, R. (1977). *Equitable Evaluation through Investigative Sociology*. *Sociological Focus*. 10 (1), 1-14.
- Bamberger, M., & Segone, M. (2011). *How to design and manage equity-focused evaluations*. New York, NY: UNICEF. Retrieved from [https://evalpartners.org/sites/default/files/EWP5\\_Equity\\_focused\\_evaluations.pdf](https://evalpartners.org/sites/default/files/EWP5_Equity_focused_evaluations.pdf)
- Equitable Evaluation Initiative. (n.d.) *About Equitable Evaluation Initiative*. Retrieved from <https://www.equitableeval.org/about>.
- Expanding the Bench (2021). *About Expanding the Bench*. Retrieved from [About Expanding the Bench – Expanding the Bench](#).
- Faler, B. (2021) [Taxes may not be colorblind, and critics say more data could prove it - POLITICO](#)  
Retrieved from [Taxes may not be colorblind, and critics say more data could prove it - POLITICO](#)
- Hayes, M. (2021). *Study to explore rural communities and inequality in the US*, *Cornell Chronicle*, Retrieved from [Study to explore rural communities and inequality in the US | Cornell Chronicle](#).
- Holdcroft, A. (2007). *Gender bias in research: How does it affect evidence-based medicine?* *Journal of the Royal Society of Medicine*. Jan; 100(1): 2–3. Retrieved from [Gender bias in research: how does it affect evidence based medicine? \(nih.gov\)](#)
- Ionnidis, J.P., Powe, N.R., and Yancy, C. (2021). *Recalibrating the Use of Race in Medical Research*. *JAMA* February 16, 2021 Vol.325, (7), Retrieved from [Recalibrating the Use of Race in Medical Research | Health Disparities | JAMA | JAMA Network](#).
- Johnson, R.B., and Onwuegbuzie, A.J. (2004). *Mixed Methods Research: A Research Paradigm Whose Time Has Come*. *Educational Researcher*, Vol. 33, No. 7, pp. 14–26.
- National Committee for Responsive Philanthropy. (2016). *Strategic Framework*. Retrieved from [Strategic Framework - National Committee For Responsive Philanthropy \(ncrp.org\)](#).
- Nelson, A., and Wardell III, C. (2021). *An Update from the Equitable Data Working Group*. The White House. Retrieved from <https://www.whitehouse.gov/briefing-room/blog/2021/07/27/an-update-from-the-equitable-data-working-group/>.
- Palinkas, L., Horwitz, S., Green, C.A., Wisdom, J.P., Duan, N., and Hoagwood, K. (2015). *Purposeful sampling for qualitative data collection and analysis mixed-method implementation research*. *Adm Policy Ment Health*. Sep; 42(5): 533-544.
- Rios, D., Magasi, S., Novak, C., and Harniss, M. (2016). *Conducting Accessible Research: Including People with Disabilities in Public Health, Epidemiological, and Outcome Studies*. *American Journal of Public Health*, December; 106(12): 2137-2144. Retrieved from [Conducting Accessible Research: Including People With Disabilities in Public Health, Epidemiological, and Outcomes Studies \(nih.gov\)](#).





The Annie E. Casey Foundation. (2014). *Race Equity Inclusion and Action Guide*. Retrieved from <https://www.aecf.org/resources/race-equity-and-inclusion-action-guide>.

The White House. (2021). [\*Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government\*](#) | The White House.

UNICEF. Iraq Co (2011) *Concept, Methodology and Next Steps for Child-Centred Equity Analysis*. Retrieved from [Child-Centred Equity Analysis \(unicef.org\)](#).

## Appendix A. Side-by-Side Comparison of Recommendations from Prior Literature

American Evaluation Association (2011) <i>Essential Practices for Cultural Competence</i>	Annie E. Casey Foundation (2014) <i>Seven Steps for Advancing Racial Equity</i>	Bamberger and Segone (2011) <i>Empowering Worst-Off Groups, including Children, through Equity-Focused Evaluation Processes</i>
<ol style="list-style-type: none"> <li>1. Acknowledge the complexity of cultural identity.</li> <li>2. Recognize the dynamics of power</li> <li>3. Recognize and eliminate bias in language.</li> <li>4. Employ culturally appropriate methods.</li> </ol>	<ol style="list-style-type: none"> <li>1. Establish an understanding of race equity and inclusion principles.</li> <li>2. Engage affected populations and stakeholders.</li> <li>3. Gather and analyze disaggregated data.</li> <li>4. Conduct systems analysis of root causes and inequities.</li> <li>5. Identify strategies and target resources to address root causes of inequities.</li> <li>6. Conduct race equity impact assessment for all policies and decision making.</li> <li>7. Continuously evaluation effectiveness and adapt strategies.</li> </ol>	<ol style="list-style-type: none"> <li>1. Equity-focused evaluation should be culturally sensitive and pay high attention to ethics</li> <li>2. Equity-focused evaluation should use participatory and/or empowerment evaluation processes to ensure worst-off groups are involved and/or co-leading the Equity-focused evaluation process starting at the design phase.</li> <li>3. Equity-focused evaluations should also involve children as appropriate, since children are also among the worst-off groups.</li> </ol>