

Better STEM Outcomes: Developing an Accountability System for Broadening Participation

*Report on an NSF-funded Workshop**

Daryl Chubin, Ira Harkavy, and Louis Martin-Vega

March 2017

“If ability, and not the circumstances of family fortune, determines who shall receive higher education in science, then we shall be assured of constantly improving quality at every level of scientific activity.”
—Vannevar Bush, 1945

“... to ‘cultivate a world-class, broadly inclusive science and engineering work force and expand the scientific literacy of all citizens’ is crucial to the future of science and democracy itself.”
—David B. Spencer and Sharon Dawes, 2009

“We know that diversity fuels innovation, so having multiple perspectives and people of different backgrounds, that’s what’s going to make organizations successful.”
—Peter Finn in Williams, 2016

“... it is not only in the students’ interest but that of our institutions and our economy to double down on our efforts to enroll and graduate students who have too often been overlooked in the past. We can no longer pay lip service to closing attainment gaps, nor can we continue to rely solely on poorly funded community colleges and open-access urban institutions to carry the load. All of our institutions, including four-year colleges and research universities, have to do their share to educate and graduate more students of color.”
—Joe Garcia and William Serrata, 2016

“The U.S. science and engineering workforce can thrive if women, blacks, Hispanics, and people with disabilities are represented in percentages comparable to their representation in the U.S. population. According to the National Center for Science and Engineering Statistics, we have a long way to go to reach that goal. We can achieve national STEM diversity and its benefits to our Nation if we commit to national STEM inclusion.”
—France A. Córdova, 2016

*This material is based upon a workshop supported by the National Science Foundation (NSF) under Grant No. OIA-1639188, awarded May 27, 2016 to Ira Harkavy (Principal Investigator) and Louis Martin-Vega (Co-Principal Investigator). Any opinions, findings, and conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of NSF.

Acknowledgments

The “Workshop on Assessing Performance and Developing an Accountability System for Broadening Participation” and the writing of this report were made possible through support from the National Science Foundation, under Grant No. OIA-163918.

We are very grateful to the panelists and participants of the workshop for their insightful remarks that led to the findings and discussion in this report. Participants are listed in Appendix A. We want to thank Joan Burrelli, NSF Science Resources Analyst, for her notes to Ira Harkavy, in his role as Chair of the Committee in Equal Opportunities in Science and Engineering (personal communication, Oct. 24, 2016), from which we drew for this report. We also wish to thank the other NSF staff members who contributed to the workshop.

We would like to recognize our colleagues at the Netter Center for Community Partnerships at the University of Pennsylvania who helped with the planning and organization of the workshop and the development of this report. We particularly acknowledge the significant contributions made by Joanna Chae, Director of the Moelis Access Science Program, and Rita Hodges, Assistant Director, as well as the support of Tina Ciocco, the Netter Center’s Administrative Coordinator.

Table of Contents

Executive Summary	iii
Introduction	1
Workshop Description	3
Framing an Accountability System for Broadening Participation	12
NSF's Pivotal Role	18
Recommendations and Conclusions	21
References	23
Appendices	27

Executive Summary

In October 2016, a workshop was held at the National Science Foundation (NSF) to create the framework for developing and implementing an accountability system for broadening participation in science, technology, engineering, and mathematics (STEM). Developing such an accountability system has become a priority among federal research and development (R&D) agencies. The workshop was designed to help develop an accountability system at three levels: funding agencies, individual grantees, and institutions receiving awards (particularly institutions of higher education).

The “Workshop on Assessing Performance and Developing an Accountability System for Broadening Participation” engaged 50 educators, administrators, and evaluators from all corners of the U.S. science and engineering (S&E) enterprise in a two-day conversation framed to support the successful implementation of the bold initiative to broaden participation of underrepresented groups in STEM, as outlined in CEOSE’s 2011-2012 and 2013-2014 Biennial Reports to Congress.

There was general agreement that the current approach to broadening participation must be reconsidered and reconfigured to recruit and nurture talent along many pathways, beginning in pre-Kindergarten (pre-K) and continuing into early careers. There was also agreement that a valid and constructive accountability system must be created with the explicit understanding that the pursuit of science and engineering occurs in, and is affected by, various institutional settings. Importantly, workshop participants emphasized that if an accountability system is to change institutional, disciplinary, and professional culture for the better, it must be embedded, embraced (not imposed), and practiced willingly.

This report first describes the workshop proceedings. Next, it turns to what is needed to develop an accountability framework, including (a) the principles guiding an accountability system that would encourage and reward broadening participation; (b) the roles of different sectors as stewards of accountability, with particular emphasis on the leadership role of higher education institutions; and (c) the pivotal role that NSF should play as lead catalyst among federal agencies. The report concludes with recommended actions that organizations could take to institute an accountability system to broaden participation in the S&E enterprise.

Workshop Description

The agenda was organized around three thematic panels (each with three speakers and a facilitator), breakout small group discussions following each panel, and a report-out to all the participants. The three panel themes were: Exemplary Programs, Metrics and Measurement, and Accountability Systems.

The three programs highlighted during the opening workshop panel—Participation and Advancement of Women in Academic Science and Engineering Careers (ADVANCE), National Center for Women in Technology (NCWIT), and the Louis Stokes Alliances for Minority Participation (LSAMP)—are instructive for their focus on institutional change.

While institution-level change is an overarching goal, there are also more micro-units of change, beginning with student participation and student learning. Broadening participation measures should not be based on numbers alone—students’ “stories matter.” These stories of overcoming institutional barriers to pursue STEM education and careers should be calibrated against time and data expectations.

The panel on metrics and measurement served as a bridge to conceptualizing an accountability system with broadening participation at its core. Data should enable us to make decisions and take actions. Data should also be regularly collected to measure progress, and such efforts should be made to identify what is facilitating and preventing success. Metrics should guide implementation that is both flexible and effective.

The metrics panel offered three key ideas: collect measures that drive change and improvement that could be coordinated in a strategic roadmap; exchange knowledge across STEM disciplinary and organizational boundaries; and adapt successful practices for use locally.

The accountability panel spoke both to STEM as it is and visions of what it could be. The building blocks of an accountability system were presented as precursors emanating from participants’ values, resulting in outcomes that first “do no harm,” and ultimately improving the participation of all individuals and institutions.

Workshop participants identified the most important aspects of an accountability system as a set of clear goals, assumptions, definitions, metrics, and a strategy for change. The general sense was that the ultimate goal is to “democratize STEM” such that barriers to full participation by all groups are reduced and that there is meaningful participation by all. Several barriers to full participation cited at the workshop included: persistent inequality across a variety of dimensions, false expectations around who can “do” science, lack of mentors and support networks, high teaching loads for STEM faculty, widespread lack of awareness of STEM programs, and bias—explicit and implicit.

Framing an Accountability System for Broadening Participation

The workshop identified four broadening participation needs that an accountability framework could accommodate merely by applying additional performance criteria to those currently employed by R&D agencies:

- Leverage resources and measure successful partnerships;
- Widen participation in programs or projects to include groups underrepresented in STEM;
- Broaden participation among panelists in the peer review process; and
- Exercise leadership by amending program announcements to emphasize PI creativity in broadening participation and instructing merit review panels on the value of impactful broadening participation activities.

Principles of a sound accountability system for broadening participation were identified. They included the following:

- The system should take account of local conditions, context and history. All institutions should not be held to the same standards. Rather, the accountability framework should take into account “multicultural validity”—social or cultural factors that may influence a particular assessment or measure.
- Successful efforts should have a research foundation and an emphasis on accountability from the beginning. Programs and projects need to have the evaluation structured into the research, not added as an afterthought. Multidimensional metrics, both quantitative and qualitative, are needed to ensure understanding and meaning beyond metrics alone. Measurement needs to be made at regular intervals, which includes monitoring contextual factors.
- A successful evaluation system will involve periodic feedback that can be used to modify practice. Evaluation should guide the evolution of programs and projects. Evaluation benefits from participatory methods that engage the groups affected, that involve learning from failure, and that are flexible.
- Learning from and through implementation of programs is crucial to evaluation as well as to the evolution and replication of successful approaches to broadening participation.
- Connecting organizations within and across sectors can heighten the impact of accountability systems. These must span education-to-workforce boundaries to form structural pathways. Such programs require incentives for collaborating organizations that increase participation system-wide.

For accountability systems to take shape and thrive, stakeholders must become stewards of such a system for broadening participation. These stewards can be found in all sectors and types of organizations, particularly institutions of higher education, the federal government, corporations and small businesses, and nonprofit organizations. Each has a particular role, but the key is working in concert. Importantly, institutions of higher education, on the front line of STEM research and education, must lead the way.

NSF’s Pivotal Role

NSF was identified as the primary catalyst for incentivizing principal investigators and institutions of higher education to move with urgency toward the goal of broadening participation. Several suggestions were targeted specifically by the workshop to NSF (and could be generalized to all funding agencies that value broadening participation efforts):

1. Hold universities more directly accountable for broadening participation. Broadening participation could be a funding criterion in all program solicitations, and information on institutional performance in this area could be required. Higher education grantees should be held accountable for: rigorous evaluation, details on who participates in each project and why, closing disparities in participation, raising degree completion rates, forming partnerships with local K-12 schools and other community-based organizations, and clarity on the consequences of the project/program/center success for various publics.
2. Convene grantees in national forums to discuss how to change the mindset and culture of PIs and institutions of higher education to embrace inclusive STEM learning and research environments. Projecting broadening participation as central to the future of science and

engineering through national forums would influence individual campuses, which could host forums of their own. Discussions of how to reconcile the conflict between short-term grants and measuring change over time would be fruitful.

3. Reward institutions for positive past and present contributions to broadening participation. In the review process, reviewers could be encouraged to discuss the strengths and weaknesses of institutional performance in broadening participation when evaluating the resources available to the PI. The institution could also aggregate data from grants that document PIs' broadening participation efforts. Combining data from distinct NSF-funded projects, as well as those from different agency programs, can increase understanding of the scale and duration of effort as well as the number of students impacted.
4. Provide multiple levels of training—in-house with NSF program officers, and in the field with PIs and program evaluators. NSF could offer training through a technical assistance corps for those who need to know how to plan, implement, and evaluate broadening participation initiatives and projects. Access to tools should be provided to all involved in the grant-making process.
5. While the NSF merit review system is sound, criteria could be strengthened as regards broadening participation. If intellectual merit is conceptualized to *include* broadening participation, the connection between inclusion and better science would be made clear. Given the impact that diversity has on intellectual creativity and the production of better science and engineering, conceptualizing broadening participation as a component of intellectual merit makes particularly good sense.

Recommendations and Conclusions

The action steps below distill what the workshop (and the literature relevant to broadening participation) recommends. This list operationalizes ten actions to institute an accountability system that strengthens the S&E enterprise. Any organization dedicated to broadening participation in STEM could undertake these actions in concert with its stakeholders and partners.

Actions to Institute an Organizational Accountability System

1. Conduct a self-study that takes stock of your organization's current broadening participation portfolio and climate.
2. Construct a timeline (near- and long-term) for achieving broadening participation outcomes articulated by your theory of change consistent with the institutional mission and strategic plan.
3. Identify data and measures that are required—either extant or to be created—to gauge progress organization-wide (and within operating units) toward your broadening participation outcomes.
4. Engage stakeholders to define a common agenda and recruit partners to work toward agreed-upon outcomes, disaggregated by demographic, educational, and careers stages as much as possible.
5. Communicate gains and setbacks with national as well as local stakeholders through a variety of media, sharing information to reach out for new partners and ideas.

6. Update and revise plans and practices as new knowledge and experience (e.g., program cost, workforce turnover) reshape your organization's thinking about how to experiment, accelerate, and expand broadening participation outcomes.
7. Incorporate what has been learned from ongoing longitudinal assessments of your organization's broadening participation programs.
8. Re-examine how the roles of government, institutions of higher education, the private sector, and nonprofits are expediting or inhibiting outcomes that your organization values and report on the contributions of each role-player to broadening participation.
9. Appraise the performance of your organization in taking steps toward increasing accountability and institutionalizing a democratized science and engineering system.
10. Be ready to begin again, as accountability for broadening participation is a recursive, iterative, and ongoing process.

The members of the October 2016 workshop looked to universities united with sponsors, notably NSF and other federal R&D agencies, to galvanize their myriad partners and function as the principal change agents for broadening participation in the S&E enterprise. This would begin to fulfill the workshop's call to democratize science and engineering. The workshop participants highlighted the need to continue the conversation on accountability for broadening participation, and to engage all stakeholders in developing and implementing an accountability system that serves all.

Introduction

In October 2016, a workshop was held at the National Science Foundation (NSF) to create the framework for developing and implementing an accountability system for broadening participation in science, technology, engineering, and mathematics (STEM). Developing such an accountability system has become a priority among federal research and development (R&D) agencies. Principal Investigator (PI) Ira Harkavy and Co-PI Louis Martin-Vega (2016)¹ designed the workshop to help develop an accountability system at three levels: funding agencies, individual grantees, and institutions receiving awards (particularly institutions of higher education).

The “Workshop on Assessing Performance and Developing an Accountability System for Broadening Participation” engaged 50 educators, administrators, and evaluators from all corners of the U.S. science and engineering (S&E) enterprise (see **Appendix A**—Workshop Participants). The workshop was framed in the PI’s words, “to support the successful implementation of the bold initiative to broaden participation of underrepresented groups in STEM,” as outlined in CEOSE’s 2011-2012 and 2013-2014 Biennial Reports to Congress.² Since its inception, the Committee on Equal Opportunity in Science and Engineering (CEOSE) has championed broadening participation in STEM fields for women, underrepresented minorities, and persons with disabilities.³ How to broaden participation in the S&E enterprise challenges all stakeholders in the STEM community⁴ to identify, attract, mentor, and support students to complete degrees and seek careers in STEM-based professions.

As acknowledged at the October workshop, the STEM community is accountable to a system of incentives and rewards—mainly grants and publications. This system by and large places insufficient value on the human resources that embody skills, generate knowledge, and contribute mightily to a vibrant society. A key question confronting the workshop, then, was how to make broadening participation a more significant and serious priority.

If there was unanimity on anything at the workshop, it is that the current approach to broadening participation must be reconsidered and reconfigured to recruit and nurture talent along many pathways, beginning in pre-Kindergarten (pre-K) and continuing into early careers. This would ensure that all stakeholders in all sectors of U.S. society—not just educational institutions, spon-

¹ The PI also serves as Chair of the Committee on Equal Opportunity in Science and Engineering (CEOSE); the Co-PI is CEOSE Vice Chair.

² Throughout this report, only quotes from participants will be so noted. Other points of emphasis will be *italicized*.

³ For CEOSE’s 2011-2012 and 2013-2014 Biennial Reports to Congress, see

https://www.nsf.gov/od/oia/activities/ceose/reports/Full_2011-2012_CEOSE_Report_to_Congress_Final_03-04-2014.pdf and https://www.nsf.gov/od/oia/activities/ceose/documents/2013-2014%20CEOSE%20Biennial%20Report%20to%20Congress_Final%20Version_09-08-2015.pdf

CEOSE responds to a 35-year-old congressional mandate (Pub. L. 96–516, § 32, Dec. 12, 1980, amended in 1985 and 2002) for NSF to implement policies and programs that promote full participation by women, underrepresented minorities, and persons with disabilities at all stages—education to workforce—of America’s STEM enterprise. Underrepresented minorities are African Americans, Hispanic or Latino Americans, Native Americans, Alaska Natives, Native Hawaiians, and Pacific Islanders.

⁴ While STEM is a collection of disciplines, the community extends beyond the educational sphere to other institutions (such as corporate laboratories and a host of nonprofit organizations) and is more inclusive of the STEM workforce. Although not cited at the workshop, demarcating the “STEM workforce” is itself an issue (NSB, 2015).

sors, and employers—are responsible for nurturing new scientists and engineers among women, underrepresented minorities, and persons of disabilities as an intentional outcome of many interventions.⁵ The next generation of scientists and engineers should not be seen as a mere *byproduct* of a knowledge and innovation system. Workshop participants emphasized that such a reconfiguration would result in *democratizing science and engineering*. As one participant succinctly stated, “This not only broadens participation, but produces *better science*.”⁶

Further convergence at the workshop was on this proposition: a valid and constructive accountability system must be created with the explicit understanding that the pursuit of science and engineering occurs in, and is affected by, various institutional settings. Accounting for the differences that institutions of higher education make in the lives of their students requires distinctive baselines, timetables, and investments for calculating desired outcomes, such as the number of STEM majors and graduates, PhDs awarded, and new faculty hired, as well as other indicators of progress towards a democratic STEM accountability system. Without such specificity and nuance about institutional history and mission, a one-size-fits-all approach can develop that is inappropriate and perhaps damaging to different types of higher education institutions. *If an accountability system is to change institutional, disciplinary, and professional culture for the better, it must be embedded, embraced (not imposed), and practiced willingly.*

This report summarizes the workshop's deliberations, then proposes action steps.⁷ First, it describes the proceedings as a series of panel presentations and discussions among all participants and in small groups. Next, it turns to what is needed to develop an accountability framework, including (a) the principles guiding an accountability system that would encourage and reward broadening participation; (b) the roles of different sectors as stewards of accountability, with particular emphasis on the leadership role of higher education institutions; and (c) the pivotal role that NSF should play as lead catalyst among federal agencies. The report concludes with recommended actions that organizations could take to institute an accountability system to broaden participation in the S&E enterprise. Participants described the workshop as a rallying-point, not an end-product, to compel action.

⁵ A word about “stakeholders,” since it is used throughout this report and heard regularly in Washington, DC, policy circles: a stakeholder has an interest in an organization’s business, that is, a stake in the outcome of its activities.

⁶ NSF’s Advisory Committee for GPRA Performance Assessment emphasized that the goal of broadening participation is not only an issue of fairness and equal opportunity, but is also the means of bringing diversity and intellectual breadth to the transformation of science itself (Spencer and Dawes, 2009). Another NSF-funded workshop resulted in a white paper that provides a detailed discussion of how broadening participation and inclusion, particularly through higher education-community engagement, contributes to better science. The authors emphasize, “Higher education-community engagement focused on locally manifested universal problems is an effective strategy for realizing full inclusion and for producing better science and a better society” (Harkavy, Cantor, and Burnett, 2015, p. 1).

⁷ The report refrains from recounting the voluminous data on participation in S&E. Instead, it points to illustrations and sources, notably the resources of NSF’s National Center for Science and Engineering Statistics (<https://www.nsf.gov/statistics/>) that paint a more complete picture by highlighting social science research and accounts that have appeared in more popular (and even social media) outlets accessed by readers attentive to issues of structural inequality and impediments to career opportunity in the U.S.

Workshop Description

The October 2016 workshop built on momentum from particularly the last five years of CEOSE's work and reports calling for a bold initiative to broaden participation of underrepresented groups in STEM (2013, 2015).⁸ Developing metrics and an accountability system for broadening participation is an intellectually challenging endeavor that necessitates multiple stakeholders and a comprehensive, systematic approach. The workshop, therefore, convened CEOSE members and experts from federal agencies, institutions of higher education, foundations, nonprofits, and for-profit organizations to discuss, brainstorm, analyze, and plan how to build an accountability system for broadening participation.⁹

While the charge to the workshop was directed to federal agencies in general, participants immediately turned their attention to NSF, which they regarded as the lead catalytic agency for broadening participation primarily through its university-based grantees.¹⁰ Moreover, NSF was seen as crucial for establishing an effective, cutting-edge performance assessment and accountability system. The conversation was wide-ranging, but used NSF as a familiar and central frame of reference.

The agenda was organized around three thematic panels (each with three speakers and a facilitator), breakout small group discussions following each panel, and a report-out to all the participants (see **Appendix B**—Workshop Agenda). The three panel themes were: Exemplary Programs, Metrics and Measurement, and Accountability Systems (see Panelists' Slides archived at <https://upenn.box.com/v/AccountabilityWorkshop>).

The breakout groups were asked to reflect on three identical questions: what was learned, what remains to be learned, and what are appropriate next steps. In short, the charge was how to get “from here to there”—from the current accountability for broadening participation to a more robust and effective broadening participation accountability system in the future.

Emerging at the outset of the workshop were several fundamental issues to be addressed in developing an accountability system for full participation in STEM. Foremost among them is the symmetrical character of the system developed and implemented. All stakeholders—sponsors

⁸ The latest CEOSE reports are credited with yielding what became (in September 2016) NSF INCLUDES (Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science) (https://www.nsf.gov/news/news_summ.jsp?cntn_id=189706). In the words of the NSF solicitation, “The grand challenge of broadening participation in STEM is to transform the STEM enterprise at all levels in order to fully engage the nation’s talent for the ultimate improvement of the STEM enterprise The goal is to achieve national level impact and progress toward STEM inclusion.” (NSF INCLUDES, 16-544). For a discussion of the impact of CEOSE Reports on INCLUDES, see <http://www.sciencemag.org/news/2016/02/nsf-launches-long-awaited-diversity-initiative>.

⁹ Many of the workshop participants were intimately familiar with the federal funding system—from the inside and the outside—that they sought to improve. They included a half-dozen NSF staff present at the workshop plus representatives from two other federal agencies—the U.S. Aid to International Development, and the Environmental Protection Agency.

¹⁰ NSF’s portfolio currently features 17 “dedicated” broadening participation programs and 11 others that emphasize participation as a distinguishing feature (https://www.nsf.gov/od/broadeningparticipation/bp_portfolio_dynamic.jsp).

(federal, nonprofit, and for-profit alike), grantees (institutions and individual), and advisory policy bodies (e.g., CEOSE) are accountable to one another for outcomes and the processes that justify them. This is where the two NSF funding criteria are so vital.¹¹ “Intellectual merit,” as was observed by workshop participants, should include the composition of the S&E workforce as a core component of good science and engineering. Who does science and engineering shapes what and how it gets done. The second NSF funding criterion is “broader impacts,” which is where broadening participation currently resides.¹² Many at the workshop claimed that the placement of broadening participation as one of a number of impacts devalues its significance. Given that broadening participation increases intellectual creativity through the diversity of perspectives and ideas, workshop participants felt broadening participation should instead be a component of intellectual merit.

The workshop surfaced these issues with illustrative empirical evidence—exemplars of programs and policies that have operated in different settings to boost participation of underrepresented STEM students and faculty.¹³ These programs demonstrate that progress can be made within the system as it is—the baseline condition—but a better system would have enabled even more success.¹⁴

Exemplary Programs

The three programs highlighted during the opening workshop panel served as exemplars for hundreds of interventions that NSF has supported for decades to broaden participation in STEM.¹⁵ Their longevity alone sets them apart (positive change takes time). But all three programs—the 20+ years of Increasing the Participation and Advancement of Women in Academic Science and Engineering Careers (ADVANCE, represented by original cohort member University of Michigan), the 12-year-old National Center for Women in Technology (NCWIT), and the quarter-century of the Louis Stokes Alliances for Minority Participation (LSAMP)—are instructive for their focus on *institutional* change.

¹¹ NSF has two principal criteria for evaluating proposals for funding: intellectual merit and broader impacts (discussed below). The process is called “merit review.”

¹² *NSF’s Grant Proposal Guide* (https://www.nsf.gov/pubs/policydocs/pappguide/nsf15001/gpg_2.jsp#IIC2d) defines broader impacts as including, but not limited to: full participation of women, persons with disabilities, and underrepresented minorities in science, technology, engineering, and mathematics (STEM); improved STEM education and educator development at any level; increased public scientific literacy and public engagement with science and technology; improved well-being of individuals in society; development of a diverse, globally competitive STEM workforce; increased partnerships between academia, industry, and others; improved national security; increased economic competitiveness of the United States; and enhanced infrastructure for research and education. *Perspectives on Broader Impacts* (https://www.nsf.gov/od/oia/publications/Broader_Impacts.pdf) contains recent examples of broader impacts.

¹³ This is where most reports would present illustrative numbers, *by group*, on enrollments, degrees, faculty, and STEM workforce participation. Such documentation of under-participation is not our purpose. Rather, the workshop took the absence of diversity throughout the S&E enterprise as a condition, and aimed for systemic solutions to change both the composition of participants and the behavior of institutions.

¹⁴ What follows is a recounting based on notes from plenary and breakout sessions plus those shared by CEOSE members serving as breakout group note-takers.

¹⁵ In FY2015, the 28 programs in the broadening participation portfolio totaled \$600M (for details, see NSF 2016). This represents less than 10% of NSF’s total \$7.5B budget.

The University of Michigan (UM) Increasing the Participation and Advancement of Women in Academic Science and Engineering Careers (**ADVANCE**) program, as presented at the workshop, was born in the wake of the Supreme Court rulings on undergraduate and graduate admissions. UM's briefs submitted in support of affirmative action made the academic case for diverse learning environments. They also underscored how the student pathway leads to the faculty outcome: a lack of student diversity (despite demographic trends) will ensure that faculty remain overwhelmingly white and male. ADVANCE was designed to identify evidence-based practices to increase the participation of women on faculties of STEM departments. ADVANCE became a national model (see the portal at <http://www.portal.advance.vt.edu>).¹⁶ UM, one of the premier ADVANCE programs, demonstrated the power of ownership and accountability at the top (the Presidents, Provosts, Deans, etc.). With such leadership, faculty became increasingly open to changing the institutional climate. Taking stock, sharing key indicators, and offering a timeline for institutionalizing promising practices created the energy needed at all levels (within departments and programs) for change.

The second exemplary program presented at the workshop highlighted the National Center for Women in Technology (**NCWIT**), a “change-leader network” of 900 organizational members working to increase the participation of women and girls—across race, ethnicity, class, age, sexual orientation, and disability status—in computing. NCWIT epitomizes “thinking nationally, acting locally.” With 160 different resources on its website (<https://www.ncwit.org>), it is a one-stop shop for finding information to help develop a theory of change. NCWIT's success stems from working across sectors on a common agenda that “convenes-equips-unites” the computing ecosystem from K-12 into the workforce to increase the meaningful participation of women and girls. Perhaps its greatest strength is granting open access to its age/grade-specific resources with the motto “don't monetize change.” As a public-private partnership, NCWIT's nonprofit arm functions as a board of directors that can make/accept donations and protect its intellectual property without restricting its use.

In the case of the Louis Stokes Alliances for Minority Participation (**LSAMP**), NSF focused on undergraduates, seeking to keep students from underrepresented groups in STEM on a science path and their enrolled institutions accountable via cooperative agreements and a centralized database maintained by a contractor organization. (See, for example, the digital library at the Louis Stokes Midwest Center of Excellence, <http://ismce.org/>.) From the outset, LSAMP was intent on monitoring inputs and outputs that could be linked quantitatively to the program. Activities introduced later as LSAMP developed addressed two key transition points—from two- to four-year colleges (Bridge to the Baccalaureate) and into graduate study (Bridge to the Doctorate) (<https://www.nsf.gov/pubs/2015/nsf15594/nsf15594.htm>). Both activities were created in re-

¹⁶ A hallmark of ADVANCE is the use of diversity and inclusion benchmarks. Such metrics utilize institutional data augmented by periodic climate surveys to characterize movements beyond measurement by human resource offices, individual academic units, or categories of personnel. Monitoring change keeps organizational culture uppermost in everyone's mind and provides feedback on the effectiveness of promising practices (Worthington, Stanley, and Lewis, 2014).

sponse to findings that focused academic and research opportunities help keep undergraduate and graduate students (especially underrepresented minorities) on the path to earning degrees.¹⁷

Emerging from discussion (in the small groups as well as by the entire workshop assembly) of these three exemplary efforts to broaden participation in STEM was how data and evaluation are instrumental for guiding the evolution of programs. Each program served as an agent of change within institutions or systems whose goals and priorities were not necessarily designed to broaden participation. Nevertheless, each example demonstrates the difference between *knowing* (as derived from data mined in research and evaluation) and *bringing about change* (in various organizational contexts).¹⁸ Each also shows the interplay between policy and practice, with the latter, as one workshop participant noted, “often bubbling up to policy rather than policy imposing procedures that inhibit ownership and change behaviors.” To take one example, NCWIT work led to the development of new organizational guidelines and strategies for overcoming the bias that leads to treatment of girls and women in educational settings (e.g., science camps, classrooms, internship settings) as less capable and serious than boys and men in identical situations.¹⁹

Two other lessons surfaced through these panel presentations and previewed measurement concerns. First, although institutional-level change is an overarching goal, there are more micro-units of change, beginning with student participation and student learning.²⁰ Second, broadening participation measures, interestingly, should not be based on numbers alone. Within an “analytic horizon,” as one workshop member put it, students’ “stories matter.” These stories of overcoming institutional barriers to pursue STEM education and careers should be calibrated against time and data expectations. What is “sensed” by students, faculty, and those responsible for “process” evaluation might precede outcome metrics as an indicator of change. And “spillover effects” may elude measurement altogether.²¹

Metrics and Measurement

The panel on metrics and measurement at the workshop served as a bridge to conceptualizing an accountability system with broadening participation at its core. Data should enable us to make decisions and take actions. Data should also be regularly collected to measure progress, and such efforts should be made to identify what is facilitating and preventing success. Metrics

¹⁷ Focused academic and research opportunities and supports include summer bridge experiences (like the one LSAMP provides to students in the summer prior to matriculation into postsecondary institutions), research internships, mentorship by faculty, and financial aid.

¹⁸ As more than one workshop participant observed, translation is needed between knowledge and practice. Deeply-held stereotypes are not relinquished just because research confirms their existence. This is the foundation on which the decade-old organization, Understanding Interventions That Broaden Participation in Science Careers (www.understanding-interventions.org), is built.

¹⁹ Training to reduce “stereotype threat” (<http://reducingstereotypethreat.org/definition.html>), for instance, has been shown to be effective with all groups and is now used proactively.

²⁰ A longing for longitudinal measures that trace student progression on an educational path is constantly expressed by researchers and evaluators, but privacy concerns about “student unit records” continue to override this preference (Kreighbaum, 2016b).

²¹ Diving even deeper, workshop participants noted the difficulty in measuring “spillover effects,” the influence of programs on practice (from the grass-roots to institution-wide) that no one expected.

should guide implementation that is both flexible and effective. The panelists' presentations included crucial points for designing a measurement roadmap to a better STEM future, as summarized below.

The panel stressed that the S&E enterprise improves by including those who have not traditionally been included. Supporting the demographic diversity that STEM has historically excluded, unwittingly or not, thus connects broadening participation to "good science." This is a formidable challenge given what funding agencies currently honor as intellectual merit and what universities reward. Which measures, then, should be *used or devised* that illuminate progress toward a more responsible, responsive, and inclusive system?

To identify appropriate metrics, the first point made by the panelists was how to define broadening participation. The key to any such definition is inclusion of those not traditionally participating for reasons unrelated to ability or interest in STEM. This places the burden squarely on those responsible for appraising and nurturing talent, not the student. Can the discharge of this responsibility be evaluated more explicitly as part of the proposal review process?

As previously described, NSF has two criteria for determining which proposals should be funded: intellectual merit and broader impacts. Many workshop participants insisted that NSF's criterion of intellectual merit should *include* broadening participation.²² The ramifications of this are far-reaching because what passes for broader impacts is subordinated, in the conduct of reviews, to excellent science. As noted above, if broadening participation remains just one of many forms of broader impacts, it will not necessarily even be used as a funding criterion. *Rewarding efforts that elevate and integrate broadening participation with intellectual merit would have far greater impact.*

The measurement panelists also illustrated to the workshop participants several tensions inherent in what we measure, how we measure and report metrics, and the contradictions between field/research-driven metrics on the one hand and evaluation-imposed requirements on the other. As one of the panelists emphasized, we should "measure what we value," not "value what we measure." But this is easier said than done. The STEM higher education community, as was pointed out, is generally not as knowledgeable as might be assumed of the literature on student learning. Needless to say, student learning is a core component of any evaluation of the impacts of broadening participation.²³

²² Though not explicitly cited at the workshop, conceptualizing broadening participation as part of intellectual merit avoids the decades-old false dichotomy that trades off "excellence" for "equity" (National Academies, 2012: ch. 11). The concept of inclusive excellence is helpful for illuminating the interconnection between broadening participation and intellectual merit. For discussions of the concept, which was first developed by the Association of American Colleges & Universities, see <https://www.aacu.org/making-excellence-inclusive> and https://www.aacu.org/sites/default/files/files/mei/williams_et_al.pdf.

²³ Some suggested that the research community appears to favor evaluation criteria loaded against individuals and institutions that have demonstrated broadening participation as a project outcome. Furthermore, research on broadening participation is neither rewarded nor widely disseminated. Repositories of studies are known to too few faculty and provide, as one participant put it, "no owner's manual." STEM tends to function, despite the acronym, as a set of balkanized disciplines and subdisciplines—a construct with little operationally to connect its constituent parts. Each inadvertently limits its attention to a narrow sphere of journals and professional peers.

The panel stipulated that measurement is confounded by issues of language and communication. By virtue of its sponsorship role, NSF (arguably more than the other R&D agencies) can provide incentives to communicate across its programs (and their respective research communities), and opportunities to share metrics. But those metrics must become more inventive, as one panelist argued. Broadening participation, she suggested, is a kind of catalytic, S-curve, reaction. That is, instead of defining progress toward broadening participation in linear terms, “a paradigm shift may be needed,” with investments somewhere between those spurring typical “incremental” (grantee-driven) change and more ambitious “large-scale social change” (cross-sector driven). Many of the workshop participants found this insightful.

In sum, the metrics and measurement panel offered three key ideas: collect measures that drive change and improvement that could be coordinated in a strategic roadmap; exchange knowledge across STEM disciplinary and organizational boundaries; and adapt successful practices for use locally. Through the ensuing discussions, the workshop cited specific measurement needs within an accountability framework for broadening participation. Prominent among them are:

- Identifying the root causes of disparities
- Determining what works under what conditions
- The necessity to intervene at many levels
- The importance of longitudinal data in measuring change
- How to measure “meaningful” participation
- Respecting the unevenness of change over time and among units
- The importance of capturing changes in culture, language, and norms
- Identifying which metrics can be common to programs, populations, and other analytical categories

Metrics, after all, are pointers that illuminate processes summarized by numbers. This is where stories lead and measures may lag, where context and nuance guide interpretation. Weighted measures that account for the type of institution and the student population, as well as the character of the intervention and experience, help us view outcomes on a grayscale rather than simply as success or failure. Indeed, learning from experiments that are often identified as “failures” should be at the heart of broadening participation measures. Good science and engineering is innovative, revealing of something new and often unanticipated. Investing in the prospect of such unanticipated discoveries includes inviting more into the pool and committing resources for doing so. The workshop was calling for a STEM-wide systemic attempt.

Accountability Systems

The accountability panel spoke both to STEM as it is and visions of what it could be. The building blocks of an accountability system were presented as precursors emanating from participants’ values, resulting in outcomes that first “do no harm,” and ultimately improving the participation of all individuals and institutions. Such building-blocks reflect a holistic approach. The one presented by a panelist that seemed to resonate most with workshop participants was “multi-

cultural validity”—the social or cultural factors that may influence a particular assessment or measure.

*As workshop participants discussed, the most important aspects of an accountability system are a set of clear goals, assumptions, definitions, metrics, and a strategy for change.*²⁴ The general sense was that the ultimate goal is to “democratize STEM” such that barriers to full participation by all groups are reduced and that there is meaningful participation by all. Barriers to full participation cited at the workshop included: persistent inequality across a variety of dimensions, false expectations around who can “do” science, lack of mentors and support networks, high teaching loads for STEM faculty, widespread lack of awareness of STEM programs, and bias—explicit and implicit.²⁵

To help address these barriers, a general consensus formed around the proposition that an accountability system must ensure dedication to pre-K through 20+ pathways. Such an approach was seen as involving systemic cultural change in federal agencies and in institutions of higher education. Echoing the discussion prompted by the measurement panel, the accountability panel advocated for a change in mindset by sponsors, institutions, and grantees alike. Workshop participants reiterated the importance of broadening participation as a part of intellectual merit, and a commensurate commitment to institutionalize successful practices and policies.

Several points useful for an effective accountability system for broadening participation emerged from this set of panel presentations and discussions:

- Clarity about goals, assumptions, context, and definitions is needed. (Questions like these need to be asked: What are the strategies for change? What is the big goal? Why do we want to do this and what do we think will change?)
- Accountability involves shared learning about what does and does not work.
- A broadening participation accountability framework should involve developing and implementing impactful efforts and ensuring institutional commitment as part of evaluation.
- Change is a process with a goal. There should be continuous learning and evolution towards realizing the goal.
- Accountability of institutions and individuals should be to principles, not just results.
- To be accountable, programs might find it useful to:

²⁴ Some members were opposed to the concept of theory of change. But all agreed on the need to state a high-level strategy or theory to develop an accountability system that implements the bold new initiative that CEOSE (2013) previously recommended. More specifically, CEOSE encourages developing an accountability system to change NSF, other agencies, and higher education, influencing pre-K through 20+ pathways to broaden participation.

²⁵ Six weeks after the workshop, a White House Interagency Policy Group (2016) issued a report on reducing bias in the STEM workforce. It features an inventory of federal agency actions and includes recommendations to federally-funded institutions of higher education relevant to accountability.

- target expectations for improvement and calibrate to institutional type for raising STEM degree rates;²⁶
 - provide options consistent with national goals and local adoption of common metrics for key outcomes;
 - contain stakeholder agreement on locally-calibrated metrics linked to rigorous evaluation²⁷;
 - issue report cards that allow for public comparison of performance by peer institutions; and
 - emphasize systems-level thinking, that is, thinking about how things work together for wide benefit.
- Outcomes should reflect the accountability goals that have been co-constructed by collaborating stakeholder organizations.

As the panel further discussed, broadening participation outcomes should be stated as agreed-upon common metrics that are publicly visible and assessable.²⁸ They should be tailored to the accountability system that an organization has crafted and adopted.²⁹ They should also include: scope of effort, financial commitment, a lead institution, and goals. Once these data are in hand, there is the question of how to use them constructively in an accountability system.³⁰

Any accountability system is caught, in the words of one panelist, between “doing the right things” and “doing them right.” Incentives for both and disincentives for misuse or disregard are mandatory. Individual careers, institutional fortunes, STEM equity, and societal well-being hang in the balance.³¹

²⁶ For details on this approach, see Closing the Achievement Gap in the California State University system, <https://www2.calstate.edu/impact-of-the-csu/diversity/inclusive-excellence/Pages/default.aspx>.

²⁷ This is detailed in the NIH Diversity Program Consortium, <http://diversityprogramconsortium.org>.

²⁸ According to one panelist working with NIH grantees, these “hallmarks” can be found, for example, in the NIH National Research Mentoring Network (NRMN), a nationwide consortium to enhance the training and career development of individuals from diverse backgrounds who are pursuing biomedical, behavioral, clinical, and social science research careers (<https://www.nigms.nih.gov/training/dpc/pages/nrmn.aspx>). After two years, analysis indicates that NRMN is making progress (Vishwanatha et al., 2016).

²⁹ Tools such as the Equity Scorecard (<https://cue.usc.edu/tools/the-equity-scorecard/>) and the Inclusive Excellence Toolkit (<http://www.du.edu/gsg/media/documents/InclusiveExcellenceToolkit-DUCME3-09.pdf>) might be useful in this process.

³⁰ The gathering of evidence can be abetted, as was noted by one panelist, by “access to curated databases” that reflect community validation and “can shorten the lag time between what is learned and what applies to a particular case.” (There is skepticism on whether existing databases, such as the Department of Education’s “What Works” Clearinghouse, would be informative for accountability. Some workshop participants endorsed more research-specialized resources and others favored the creation of new databases suited to the task.) Either way, dissemination is seen as central to accountability, reinforcing community consensus on broadening participation.

³¹ Examples of solutions to problems that inhibit participation in science and engineering abound: retention becomes an institutional responsibility, not a problem of student attrition (Change the Equation, 2016); minority-serving institutions are recognized for their contributions to STEM education and entry to the workforce (Gasman and Nguyen, 2014); and inequalities in opportunity decrease, facilitated by financial aid and transfer policies that make negotiating a STEM degree more seamless, to attract more students from underrepresented populations and support them

If approached as a set of “if-then inferences” (the opposite of “one size fits all”), accountability systems can be potent motivators. Without context or regard for particulars (such as history, baseline measurement, demographic group), negative “unintended consequences,” that is, actions contrary to the values driving the system, can flourish.

To create a system that nurtures and utilizes more of its population’s talents, the U.S. S&E enterprise and its various stakeholders must share ownership of its successes and failures. Workshop participants agreed that STEM needs to treat failures more productively by fostering “an environment that encourages learning from failures, as opposed to discouraging those who have failed.” Currently, failure is not seen as a value inherent in the funding of “high-risk, high-payoff” ventures. Yet NSF’s twin mission of research and education is just such a venture. Broadening participation thus raises the stakes about what promises will be achieved when funding a particular proposal.³²

In sum, the three panel presentations defined an agenda for reconstructing, or at least reimagining, an S&E enterprise that values broadening participation. To change the trajectory of knowledge and practice, higher education institutions and systems must also collaborate to strengthen, spread, and sustain efforts that impact underrepresented groups. They must address vexing problems that were mentioned or alluded to at the workshop, including precollege preparation, campus climate, debt disparities, data recognition and use, graduate student transitions, and faculty diversity. Institutions need to ask tough questions: How will a democratizing system look and function? How will we know a different accountability scheme is having intended consequences? And what, in the words of workshop participants, are “reasonable horizons, short- and long-term,” for reaching “tipping points” and measuring “different paces of change”?³³

through degree completion (Ekowo and Palmer, 2016). Such changes democratize science as a vital part of societal well-being.

³² It is worth noting that this is why NSF investments are distributed over a vast number of investigators, teams, disciplines, and technical puzzles. There is an implied acknowledgement that most projects will at best add only in modest increments to what is known and how to act on new knowledge. Promising the production of new scientists and engineers at one or more degree levels as an outcome of intellectual merit would, as workshop attendees observed, change the status of broadening participation as a funding criterion.

³³ Workshop participants were sent an annotated bibliography on “accountability in the public, nonprofit and private sectors” (Desai and Frehill, 2016, contained in **Appendix C**) dating to the Government Performance and Results Act of 1993. This bibliography could be viewed as a starter list of the numerous dimensions of accountability.

Framing an Accountability System for Broadening Participation

Following the panel presentations (which were elaborated by small group discussions and a reconvening for reporting out to all participants), the core components of any improved accountability system began to emerge. To form the bedrock of an accountability system that is accepted and rewarded, leaders and decision-makers need to examine *empirically* how their current policies and practices aid or deter broadening participation. Such a self-study would reveal an institution's current accountability framework that values certain achievements and discounts others. Thus, it behooves NSF and other R&D agencies to revisit their fundamental assumptions about who and what it takes to do good science.

The workshop identified four broadening participation needs that an accountability framework could accommodate merely by applying additional performance criteria to those currently employed by R&D agencies:³⁴

- Leverage resources and measure successful partnerships;
- Widen participation in programs or projects to include groups underrepresented in STEM;
- Broaden participation among panelists in the peer review process; and
- Exercise leadership by amending program announcements to emphasize PI creativity in broadening participation and instructing merit review panels on the value of impactful broadening participation activities.

Workshop participants stressed that more discussions about how to change “the mindset and the culture of PIs and institutions of higher education” are imperative. PIs, both at their home institutions and at agency program meetings, should have opportunities to discuss among themselves ways to broaden participation, particularly what is “success” and “risk” with STEM students while they are enrolled and after they graduate. Who better to lead such discussions than NSF and its staff with intimate knowledge of STEM communities (where grantees and expert reviewers interact)?

All along the STEM pathway are indicators that participation by underrepresented groups is stymied by the PreK-12 to workforce system *as it is*. Workshop participants coalesced around the notion that *implicit* theories of change drive the work of engineering and science. How could those theories be revised to value the multicultural validity of broadening participation? “What,” some asked, “must change for whom?” And what would broadened participation (an outcome) look like?

³⁴ NSF performance indicators are aggregated and presented in eight chapters of *Science and Engineering Indicators—2016* (<https://www.nsf.gov/statistics/2016/nsb20161/#/data>). Four of the chapters contain demographic and participation data. For example, indicators of elementary and secondary school mathematics and science education relate to learning, course-taking, teachers, instructional technology, and transitions to higher education. Other indicators can be found in the higher education, labor force, and academic R&D chapters.

Indicators of broadened participation (as reported out from the small group discussions) include:

- “a reduction in the minority-majority student achievement gap” at various educational levels;
- a “more measurably inclusive workforce”;
- an “expansion of resources to minority-serving institutions”;
- “acceleration from incremental to a more rapid influx of students with certain attributes” into the S&E enterprise; and
- evidence of “investment in bold initiatives that yield high payoffs in degrees, research programs,” and advances that reach both the marketplace and the public consciousness.

Any theory of change that envisions broadened participation must declare the actions that will lead to the attainment of such changes. Far be it for one workshop to prescribe a *single* theory of change to broaden participation in S&E. That would be presumptuous. Rather, the workshop urged all stakeholders to conceptualize their own theories, with milestones and measures that can gauge processes and impacts of action.³⁵

Many problems of under-participation in STEM were mentioned or alluded to at the workshop. These fleeting comments implicated organizations that must change behavior—their policies, programs, and/or practices—to increase the participation of the students they enroll, educate, and train for careers in STEM.³⁶ Identifying significant and persistent problems helps to frame and focus the design of any accountability system. To re-state, these significant institutional problems include precollege preparation, campus climate, debt disparities, data recognition and use, graduate student transitions, and faculty diversity. There is neither space here, nor is it the purpose of this document, to describe these specific problems or possible solutions in detail.³⁷ But they do indicate that broadening participation as a whole is a researchable problem. Working actively to solve these problems will determine the resolve of the S&E enterprise to achieve its core purposes and be accountable for them.

Given the variety of settings in which students are educated, research is conducted, and scientists and engineers are employed, the recommendations presented in this report prescribe what should be done, not specifically how for any given setting. This is the essence of policy analysis, suggesting a menu of options that implement actions in accordance with local circumstances and

³⁵ In the words of one workshop participant, such a theory-construction “can fill conceptual cavities” and raise what Durkheim (1893) termed *collective consciousness* with the details of on-the-ground operations. Once shared, the participant continued, such “maps” can at once serve as a unifying force while inviting scrutiny from policy bodies and sponsors alike. Others agreed that this may fuel experimentation and subsequent evaluation. The entire process helps to establish accountability at the organizational level and declares publicly a commitment to self-study and strategizing in the name of change.

³⁶ As one labor economist and his team (Carnevale et al., 2015) counsel, students must be able to respond to market opportunities and contribute their skills to the science-based workforce.

³⁷ Comprehensive analyses of these problems can be found in reports from the White House Office of Science and Technology Policy and the President’s Council of Advisors on Science and Technology (PCAST, 2012); the National Academies of Sciences, Engineering and Medicine (2011); and numerous professional society “blueprints” and “roadmaps” (e.g., APLU, 2016).

needs. Multiple stakeholders will favor different courses of action, which themselves may change over time.³⁸ Nevertheless, each stakeholder has a responsibility to act in behalf of broadening participation for the good of science and the integrity of the entire S&E enterprise.

Workshop Prescriptions: Principles of Accountability

Self-study can reveal what an organization should consider in developing an accountability system for broadening participation, especially the outcomes to be measured. Principles of accountability, as outlined below, should guide any self-study by (institutional or individual) members of the STEM community. The sense of the workshop was clear: all stakeholders—but particularly gatekeepers that select and amplify messages such as the American Council on Education, the Association of American Universities, and other higher education advocacy organizations; policy bodies such as CEOSE; and catalytic agencies such as NSF—should weigh the following principles of practice in crafting an accountability system. Near-unanimity crystallized around these principles. But it took many conversations during the two-day workshop to yield a set of prescriptions for implementation in a full-blown accountability system.

Table 1 displays principles of a sound accountability system and the role they would play in a reimagined S&E enterprise. The table, distilled from workshop discussions, is a kind of template, but *not a menu*. Each stakeholder would interpret *each* element as a challenge to its current practices. Each would also need to develop a measurement and reporting regimen that communicates to all organizations to which it is accountable.

³⁸ See the annotated bibliography focused on “evaluation, performance measurement and change management in the public, nonprofit and private sectors” (Phillips and Frehill, 2016, reproduced in **Appendix D**, which was sent after the October 2016 workshop as a resource for all participants).

Table 1. Principles of a Sound Accountability System for Broadening Participation

- **The system should take account of local conditions, context and history.** All institutions should not be held to the same standards. Rather, the accountability framework should take into account “multicultural validity”—social or cultural factors that may influence a particular assessment or measure. Additional actions to recognize local context include:
 - *Creating a flexible theory of change:* There should be no slavish devotion to a theory of change; it should be updated and revised like a strategic plan would be as interim measures of broadening participation are produced.
 - *Measuring at multiple levels:* Changes in broadening participation are not confined to a single organizational level. Baseline measures must be created at the levels of the individual, organizational unit, institution as a whole, collaboration with partners, and the broader community.
 - *Developing a timeline that is informed by conditions:* Because organizations and the programs that house them differ in size, scope of operations, resources, mission, etc., each should specify the time-frame for realizing change over the short and long term.
- **Successful efforts should have a research foundation and an emphasis on accountability from the beginning.** Programs and projects need to have the evaluation structured into the research, not added as an afterthought. Multi-dimensional metrics, both quantitative and qualitative, are needed to ensure understanding and meaning beyond metrics alone. Measurement needs to be made at regular intervals, which includes monitoring contextual factors. Additional actions to build accountability into research projects include:
 - *Articulating goals:* Organizations should not assume awareness by their constituents. And they should invite periodic critiques of their broadening participation goals given the progress made.
 - *Using data to illuminate different outcomes:* Whatever the organization is trying to improve in broadening participation must be declared as an outcome linked to experimentation, learning, and/or adaptation of current practices. Define the indicators of change through collaboration with partners, and the broader community.
 - *Compiling repositories for action:* Resources to guide organizations are essential. But they must be curated, sometimes consolidated, and staff should be designated to train and be knowledgeable about what resources exist, are applicable, and available for use.
- **A successful evaluation system will involve periodic feedback that can be used to modify practice.** Evaluation should guide the evolution of programs and projects. Evaluation benefits from participatory methods that engage the groups affected, that involve learning from failure, and that are flexible. Specific actions to collect feedback include:
 - *Designing feedback mechanisms:* Feedback from both organizational staff and constituents outside (including community members) should be ongoing to inform and engage partners.
- **Learning from and through implementation of programs is crucial to evaluation as well as to the evolution and replication of successful approaches to broadening participation.** Learning through implementation includes the following key steps:
 - *Identifying barriers:* Devise a means of pinpointing practices that diminish access to organizational resources or reduce the performance of some, as revealed by studies (e.g., implicit biases or climate surveys) and self-evaluation.
 - *Changing practice, then policy:* A bottom-up approach that takes the experiences and concerns of organizational participants seriously also spreads ownership to the leadership. At the appropriate time, new practices can be codified in policy.
- **Connecting organizations within and across sectors can heighten the impact of accountability systems.** These must span education-to-workforce boundaries to form structural pathways. Such programs require incentives for collaborating organizations that increase participation system-wide. Additional actions to connect organizations include:
 - *Implementing pathway programs**:* As the STEM community has learned about the connectedness of stages preK-to-workforce, programs that focus on students’ educational transitions to the next stage have been implemented. These are more systemic in their support and advice about options and careers, as demonstrated here by NCWIT in the high school-to-college transition, by LSAMP in undergraduate-to-graduate degree achievement, and by ADVANCE in ascending faculty ranks.

* Source: NSF Workshop discussions, Oct. 13-14, 2016

** One nationally-renowned example of a pathway program (not cited at the workshop) is the Fisk-Vanderbilt Master’s-to-PhD Bridge Program, established in 2004. It has facilitated the transition of students to PhD completion and research careers. One in two of those students has been a woman and four in five underrepresented minorities. Six different federal agencies have supported the program over its short life (Szrom, 2015).

The above accountability principles are designed to help stakeholders focus organizational resources on assessing their current strengths and weaknesses, while envisioning alternative courses of action and remaining mindful of time horizons in measuring and using data that inform programs and partners alike. Engagement of all stakeholders is crucial for system-wide impact.³⁹

Stewards of Accountability Working in Concert

For accountability systems to take shape and thrive, stakeholders must become *stewards* of such a system for broadening participation. As noted repeatedly at the October 2016 workshop, these stewards can be found in all sectors and types of organizations. Each has a particular role, but the key is working in concert. And as workshop participants articulated, *institutions of higher education, on the front line of STEM research and education, must lead the way*. In summary:

Institutions of higher education are leaders in their communities—the sites where broadening participation occurs.⁴⁰ They also function as models of inclusion for other institutional stakeholders, notably K-12 schools and employers of new graduates. Colleges and universities are *the* bridge between STEM education and the workforce.⁴¹ Institutions of higher education also command talent—from U.S. citizens, the foreign-born (on temporary visas), and immigrants—and they are counted on to add value to the students they train, certify, and shape to advance the pillars of America’s national and global interests: innovation, economy, and equality.⁴² Given their many roles, these public and private institutions that receive federal funding should lead the way in implementing and documenting an accountability system, developing in the process, as one workshop participant termed it, a “science of accountability.” Finally, there is a necessity for higher education to be held accountable on the institutional level.

The Federal government is a key enabler. The White House Office of Science and Technology (OSTP) and its coordinating bodies such as the President’s Council of Advisors on Science and

³⁹ As the NSF Director asserted in her inaugural speech in April 2014, “Not enough of our fellow citizens understand how relevant the research . . . conducted is to their daily lives In both formal and informal venues, we need to engage the public in order to help improve understanding of the value of basic research and why our projects are worthy of investment” (NSF, 2014).

⁴⁰ Over the past decade, the concept of “anchor institutions” has emerged as a new paradigm for understanding the role that place-based institutions, particularly institutions of higher education and medical centers, could play in building successful communities and local economies. This concept recognizes that higher education institutions are not only educators and knowledge producers, but also major employers, real estate developers, purchasers, incubators for business and technology, and centers for arts and culture (Task Force on Anchor Institutions, 2009).

⁴¹ As one university president states, “. . . the most important task for universities . . . is one that we are uniquely well-suited to perform: to help society at large—not only our own campus communities—better understand the painful and still-unresolved historical context with which the need for affirmative action exists. This context includes a public education system that remains nearly as segregated and unequal today as it was at the time of *Brown* [v. *Board of Education*] more than six decades ago” (Bollinger, 2016).

⁴² One dean of engineering, for example, has suggested that rather than ranking U.S. universities by prestige, faculty compensation, research expenditures, or class size, *U.S. News and World Report* should rate institutions on the diversity of their student body, or its inclusive excellence (May, 2016). For a compelling dataset, the 2016-17 Almanac of *The Chronicle of Higher Education* relates 16 African-American presidents and chancellors of institutions of higher education who have served at least six years in their present capacities to changes in student, faculty, and managers’ diversity (Chubin, 2016).

Technology (PCAST) and the National Science and Technology Council (NSTC) can speak to the executive R&D agencies with a forceful voice. This voice can be amplified by the committees of Congress with oversight for the education, science, and workforce missions within their respective purviews. NSF can particularly serve as the catalyst and lead among the R&D agencies (detailed below). The federal laboratories (Departments of Defense and Energy, and NASA) and their industry patrons form a kind of extension service to advance these missions by reaching students and K-12 STEM teachers where they live, then convening them annually for interactions with peers from around the U.S. Accountability, as has been the case for decades (since the Government Performance and Results Act of 1993, modernized in 2010), is prominent in the ethos of government. Broadening participation should be seen as a prized benefit within that ethos.⁴³

Corporations and small businesses alike are increasingly consumers of science- and engineering-based talent. It is in their self-interest to support the universities that supply research on and by their future workforce. It is in the nation's interest that corporations and small businesses not only subsidize the research and education of underrepresented groups who resemble their consumer base, but also hire them. This would reflect the global marketplace for talent, enhance the S&E enterprise in the process, and demonstrate that accountability measures matter.⁴⁴ Public-private partnerships, such as that formed by the Council on Competitiveness, their member companies, and NSF (BEST, 2004), are a model for collaborations of national scope and impact.

Nonprofit organizations, national and local in the form of philanthropic foundations, think tanks, and technical assistance arms, are the connective tissue and honest brokers that can knit individuals and institutions into new social forms to advise, promote, and help sustain systemic change in local communities. (State and local government agencies also serve these connective and brokering functions.) The National Academies of Science, Engineering, and Medicine have been most durable in working across sectors and organizational types to bring complementary expertise to bear on STEM-based problem-solving. Such interactions could renew all stakeholders' commitment to accountability and demonstrate that the S&E enterprise seeks talent in every sector and type of organization.

⁴³ Recall that gender- and race-conscious policies are legal remedies under federal statutes to past discrimination. The key to their use, however, requires carefully crafted justification tied to *institutional mission statements* (Burgoyne et al. 2010: 23-24). Virtually all such statements in higher education explicitly cite access, diversity, and/or inclusion as essential to the achievement of educational objectives.

⁴⁴ In 2014, ten hi-tech companies released workforce diversity reports for the first time revealing their overwhelming White, Asian, and male composition (Forrest, 2014). This move by corporate giants and early-stage startups toward transparency in the workplace has also motivated affiliations with nonprofit organizations such as the Anita Borg Institute (<http://anitaborg.org>), NCWIT, and beyond computing-centered interventions, the STEM Education Coalition (<http://www.stemedcoalition.org/wp-content/uploads/2010/05/One-pager-on-STEM-Ed-Coalition.pdf>).

NSF's Pivotal Role

The role of the National Science Foundation in broadening participation, as the above attests, is both undeniable and singular. The October 2016 workshop was held at NSF under a grant awarded to the PI and Co-PI, who also serve as leaders of a chartered agency committee (CEOSE). The participants were independent actors, but all looked to *NSF's role as primary catalyst for incentivizing principal investigators and institutions of higher education to move with urgency toward the goal of broadening participation*. Thus, several suggestions were targeted specifically by the workshop to NSF (and could be generalized to all funding agencies that value broadening participation efforts):

- 1. Hold universities more directly accountable for broadening participation.** The proposal process, for example, could require information on the institutional performance in broadening participation, similar to asking what facilities and resources are available. NSF, as well as other funding agencies, could highlight “broadening participation as a funding criterion in all program solicitations.” Higher education grantees should be held accountable for: rigorous evaluation, details on who participates in each project and why, closing disparities in participation, raising degree completion rates, forming partnerships with local K-12 schools and other community-based organizations, and clarity on the consequences of the project/program/center success for various publics.
- 2. Convene grantees in national forums to discuss how to change the mindset and culture of PIs and institutions of higher education to embrace inclusive STEM learning and research environments.** Projecting broadening participation as central to the future of science and engineering through national forums would influence individual campuses, which could host forums of their own. Discussions of “how to reconcile the conflict between short-term grants and the long time it takes to measure change” would be fruitful. A convening of grantees on how to raise the visibility of individually-funded projects to the level of institutional impact would also be worthwhile.
- 3. Reward institutions for positive past and present contributions to broadening participation.** In the review process, reviewers could be encouraged to discuss the strengths and weaknesses of institutional performance in broadening participation when evaluating the resources available to the PI. The institution could also aggregate data from grants that document PIs’ broadening participation efforts. Combining data from distinct NSF-funded projects, as well as those from different agency programs, can increase understanding of the scale and duration of effort as well as the number of students impacted. This “would be seen as an infusion of resources” from NSF (or other federal sponsors) that the institution would credit rather than take for granted.⁴⁵

⁴⁵ To cite a non-NSF example, the Health Resources and Services Administration (HRSA) has funded four grant programs aimed at increasing opportunities for individuals from minority and disadvantaged backgrounds to pursue training in health professions. Collectively, these university and medical school “diversity programs” were authorized under Title VII and Title VIII of the Public Health Service Act. They are about to be evaluated for their relative

4. **Provide multiple levels of training—in-house with NSF program officers, and in the field with PIs and program evaluators.** Broadening the participation of undergraduates, graduate students, and postdocs in STEM through mentoring and supervision varies across faculty members. NSF could offer training through a “technical assistance corps” or a kind of “virtual sponsored research office” for those who need to know how to plan, implement, and evaluate broadening participation initiatives and projects.⁴⁶ This is not a problem to be solved by new research. It is real-time information dissemination, utilizing NSF-generated resources, including studies on and tools for performance assessment (such as those in Appendices C and D of this report). Access to tools should be provided to all involved in the grant-making process, including principal investigators, program officers, and evaluators.⁴⁷
5. **While the NSF merit review system is sound, criteria could be strengthened as regards broadening participation.** The sense of the October 2016 workshop participants was that so long as “broader impacts” remains separate from “intellectual merit,” broadening participation will remain more an “add-on or afterthought” than a guiding funding principle. If intellectual merit is conceptualized to *include* broadening participation, the connection between inclusion and better science would be made clear.⁴⁸ Given the impact that diversity has on intellectual creativity and the production of better science and engineering, conceptualizing broadening participation as a component of intellectual merit makes particularly good sense.⁴⁹

But refining definitions or sharpening criteria means nothing if the agency’s participants do not execute. Execution relies on how consistently program announcements, reviewers’ judgments, and program officers’ recommendations are *connected*. It is the process that justifies better deci-

and aggregate contributions to HRSA’s workforce diversity efforts (NORC, personal communication to D.E. Chubin. Feb. 4, 2016).

⁴⁶ The “technical assistance corps” could comprise multidisciplinary teams of NSF staff, PIs, evaluators, and specialists in online content who are familiar with certain NSF directorates and programs, and provide outreach on request.

⁴⁷ Note that all front-line players are partially informed. Principal investigators are often not aware of broadening participation literature or existing databases. Rotating program officers are drawn from this same population and need such resources to do an enlightened and efficient job. Evaluators may be conversant with theories of change and logic models, but not current on agency priorities or procedures. A 2009 NSF-funded workshop sought to speak to these three key populations precisely about evaluating the impacts of broadening participation projects (Clewell and Fortenberry, 2009).

⁴⁸ This was expressed at the October 2016 workshop as “the difference between research on gravitational wave detection and the composition of the research team.” Participation harbors the potential for seeing, interpreting, and discovering differently. A most recent example is found in the book and feature film *Hidden Figures*, which depicts the heretofore untold role of African-American women mathematicians and engineers at NASA in the early success of the U.S. space program (Shetterly, 2016).

⁴⁹ Scott Page (2007) argues that diversity fuels innovation, which enriches the intellectual environment: “Innovation provides the seeds for economic growth, and for that innovation to happen depends as much on collective difference as on aggregate ability. If people think alike then no matter how smart they are they most likely will get stuck at the same locally optimal solutions. Finding new and better solutions, innovating, requires thinking differently. That’s why diversity powers innovation.” Lisa Burrell (2016), senior editor at *Harvard Business Review*, highlights “the decades’ worth of studies showing that a diverse workforce measurably improves decision making, problem solving, creativity, innovation, and flexibility.”

sions about what and who is grant-worthy. This is the kind of organizational change that the workshop proposed—and the kind of accountability system that NSF should implement, document, and encourage its grantees to embrace and make their own.

Recommendations and Conclusions

Any accountability system dedicated to advancing the S&E enterprise must reasonably reflect the values of its practitioners by elevating those mechanisms that identify and develop talent. Each sector has a stake in diversifying the human inputs through equal opportunity.⁵⁰ That is where policy incentives and competitive funding should attract and reward those (individuals, institutions, and coalitions) that vigorously support a people-centered accountability system.

Table 2 distills what the workshop (and the literature relevant to broadening participation) recommends. While not a checklist—that would be an insult to the complexity of the tasks at hand—this table operationalizes ten actions to *institute* an accountability system that strengthens the S&E enterprise and the nation’s workforce. Any organization dedicated to broadening participation in STEM could undertake these actions in concert with its stakeholders and partners.

Table 2. Actions to Institute an Organizational Accountability System for Broadening STEM Participation

1. Conduct a self-study that takes stock of your organization’s current broadening participation portfolio and climate.
2. Construct a timeline (near- and long-term) for achieving broadening participation outcomes articulated by your theory of change consistent with the institutional mission and strategic plan.
3. Identify data and measures that are required—either extant or to be created—to gauge progress organization-wide (and within operating units) toward your broadening participation outcomes.
4. Engage stakeholders to define a common agenda and recruit partners to work toward agreed-upon outcomes, disaggregated by demographic, educational, and careers stages as much as possible.
5. Communicate gains and setbacks with national as well as local stakeholders through a variety of media, sharing information to reach out for new partners and ideas.
6. Update and revise plans and practices as new knowledge and experience (e.g., program cost, workforce turnover) reshape your organization’s thinking about how to experiment, accelerate, and expand broadening participation outcomes.
7. Incorporate what has been learned from ongoing longitudinal assessments of your organization’s broadening participation programs.
8. Re-examine how the roles of government, institutions of higher education, the private sector, and nonprofits are expediting or inhibiting outcomes that your organization values and report on the contributions of each role-player to broadening participation.
9. Appraise the performance of your organization in taking steps toward increasing accountability and institutionalizing a democratized science and engineering system.
10. Be ready to begin again, as accountability for broadening participation is a recursive, iterative, and ongoing process.

* Source: distillation of workshop and literature discussed in this report

⁵⁰ Again, to enlarge on the workshop, there is a profound difference between opportunity and outcome that critics of both broadening participation and affirmative action policies often get wrong. Affirmative action policies—not just Title IX, but Titles VI and VII of the Civil Rights Act of 1964 and the Equal Protection Clause of the 14th Amendment (Malcom et al. 2004)—are a tool to be used to counteract gender and racial biases that keep women and persons of color out of higher education, including STEM, and thereby provide equal opportunity.

This report has offered a series of principles and conditions that could guide organizations to reconsider with whom to partner, what to measure, and how to influence broader engagement in the preparation and participation of *all* students in STEM—all in the name of accountability.

The report speaks to stakeholders in organizations that produce and consume STEM talent and proposes how to get “from here to there”: from the current state of accountability to a desired future (and ongoing) state of broadened participation. Furthermore, it has illustrated what should be done, who should do it, and how to do it well. That would include a focus on a theory of change with NSF as the catalyst for developing an accountability system that permeates other R&D agencies, and in turn, motivates accountability systems that influence individuals and institutions, particularly in higher education. The higher education sector must build broadening participation into its operations and programs if it ever is to institutionalize it. Finally, this report has urged collaborations among the various stakeholders, including institutions of higher education and the private and public sectors, to implement long-term plans (with suitable timetables and metrics) that will impact an increasingly diverse U.S. student population.

The members of the October 2016 workshop look, then, to universities united with sponsors, notably NSF and other federal R&D agencies, to galvanize their myriad partners and function as the principal change agents for broadening participation in the S&E enterprise. This would begin to fulfill the workshop’s call to democratize science and engineering. The workshop participants highlighted the need to continue the conversation on accountability for broadening participation, and to engage all stakeholders in developing and implementing an accountability system that serves all.

References

APLU, Coalition of Urban Serving Universities, and AAMC. 2016. *Increasing Diversity in the Biomedical Research Workforce—Actions for Improving Evidence*, <http://www.aplu.org/library/increasing-diversity-in-the-biomedical-research-workforce/file>

Bollinger, L.C. 2016. Affirmative Action Isn't Just a Legal Issue. It's Also a Historical One. The New York Times, June 24, http://www.nytimes.com/2016/06/25/opinion/affirmative-action-isnt-just-a-legal-issue-its-also-a-historical-one.html?_r=0

Burgoyne, R., Dawson, R.C., Shaw, T.M., Scheikman, R., Coleman, A., Winnick, S.Y., Buck J.R., Palmer, S.R., Keith, J.L., Chubin, D.E., Malcom, S.M., and Vaughn, J. 2010. *Navigating A Complex Legal Landscape to Foster Greater Faculty and Student Diversity in Higher Education: The Law Governing Effective Faculty and Student Body Diversity Programs in STEM and Related Disciplines . . . and Its Implications for Institutional Policy*. American Association for the Advancement of Science and the Association of American Universities, April 2009, Retrieved Apr. 6, 2015 at <http://www.aaas.org/sites/default/files/LawDiversityBook.pdf>

Burrell, L. 2016. We Just Can't Handle Diversity. *Harvard Business Review*, July-August, <https://hbr.org/2016/07/we-just-cant-handle-diversity>

Bush, V. 1945. Science The Endless Frontier. A Report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development, United States Government Printing Office, July, <https://nsf.gov/od/lpa/nsf50/vbush1945.htm#ch4>

Carnevale, A.P., Cheah, B. and Hanson, A.R. 2015. The Economic Value of College Majors. Georgetown University: Center for Education and the Workforce, <https://cew.georgetown.edu/wp-content/uploads/The-Economic-Value-of-College-Majors-Exec-Summ-Web.pdf>

CEOSE (Committee on Equal Opportunity in Science and Engineering), 2013. *2011-2012 Biennial Report to Congress*. Broadening Participation in America's STEM Workforce, https://www.nsf.gov/od/oia/activities/ceose/reports/Full_2011-2012_CEOSE_Report_to_Congress_Final_03-04-2014.pdf

CEOSE (Committee on Equal Opportunity in Science and Engineering), 2015. *2013-2014 Biennial Report to Congress*. Broadening Participation in STEM, https://www.nsf.gov/od/oia/activities/ceose/documents/2013-2014%20CEOSE%20Biennial%20Report%20to%20Congress_Final%20Version_09-08-2015.pdf

Change the Equation, 2016. New Data: Inequities in K-12 Science, Oct. 27, <http://www.changetheequation.org/blog/new-data-inequities-k-12-science>

- Chubin, D.E. 2016. Giving the Gift of Data on Diversity, Driving Force, Sept. 7, <https://www.aaas.org/blog/driving-force/giving-gift-data-diversity>
- Clewell, B.C. and Fortenberry, N. (eds.) 2009. *Framework for Evaluating Impacts of Broadening Participation Projects*. Report from a National Science Foundation Workshop, June 30, https://www.nsf.gov/od/broadeningparticipation/framework-evaluating-impacts-broadening-participation-projects_1101.pdf
- Córdova, F. 2016. Dear Colleague Letter: NSF INCLUDES (Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science), Feb. 22, DCL NSF 16-048, <https://www.nsf.gov/pubs/2016/nsf16048/nsf16048.jsp>
- Durkheim, E. 1893 (1997). *The Division of Labor in Society*. Trans. W.D. Halls. New York: Free Press.
- Ekowo, M. and Palmer, I. 2016. *The Promise and Peril of Predictive Analytics in Higher Education—A Landscape Analysis*, New America, October, https://na-production.s3.amazonaws.com/documents/Promise-and-Peril_4.pdf
- Forrest, C. 2014. Diversity Stats: 10 Tech Companies that Have Come Clean, TechRepublic, Aug. 28, <http://www.techrepublic.com/article/diversity-stats-10-tech-companies-that-have-come-clean/>
- Garcia, J. and Serrata, W. 2016. Meeting the Challenge of Demographic Change, *The Chronicle of Higher Education*, Dec. 6, http://www.chronicle.com/article/Meeting-the-Challenge-of/238582?cid=trend_right_wc
- Gasman, M. and Nguyen, T-H. 2014. Historically Black College and Universities (HBCUs) Leading Our Nation's Efforts to Improve the Science, Technology, Engineering, and Math (STEM) Pipeline. *Texas Education Review*, 2: 75-89, http://txedrev.org/wp-content/uploads/2014/06/Gasman-Nguyen_HBCUs-and-STEM.pdf
- Harkavy, I., Cantor, N., and Burnett, M. 2015. *Realizing STEM Equity and Diversity through Higher Education-Community Engagement*, January, https://www.nettercenter.upenn.edu/sites/netter_internal/files/Realizing_STEM_Equity_Through_Higher_Education_Community_Engagement_Final_Report_2015.pdf.
- Harkavy, I. and Martin-Vega, L. 2016. Assessing Performance and Developing an Accountability System for Broadening Participation, NSF Workshop Grant, No. OIA-1639188, May.
- Kreighbaum, A. 2016b. ACLU Declares Opposition to Student Unit Records, Inside Higher Ed, Nov. 16, <https://www.insidehighered.com/quicktakes/2016/11/16/aclu-declares-opposition-student-unit-records>

Malcom, S.M., Chubin D.E., and Jesse, J.K. 2004. *Standing Our Ground: A Guidebook for STEM Educators in the Post-Michigan Era*. American Association for the Advancement of Science and National Action Council for Minorities in Engineering.

The National Academies. 2011. *Expanding underrepresented minority participation: America's science and technology talent at the crossroads*. Washington, DC: <http://www.nap.edu/read/12984/chapter/1>

National Academies of Sciences, Engineering, and Medicine. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press, <https://www.nap.edu/read/13165/chapter/16>

NSB (National Science Board). 2015. *Revisiting the STEM Workforce*, A Companion to Science and Engineering Indicators 2014, Arlington, VA: National Science Foundation (NSB-2015-10), <https://www.nsf.gov/nsb/publications/2015/nsb201510.pdf>

NSF (National Science Foundation). Merit Review, https://www.nsf.gov/bfa/dias/policy/merit_review/

NSF (National Science Foundation), 2014. *Perspectives on Broader Impacts*, April. https://www.nsf.gov/od/oia/publications/Broader_Impacts.pdf

NSF (National Science Foundation), 2016. *FY 2017 Budget Request to Congress*. NSF Programs to Broaden Participation, Summary Table 11-12.

Page, S. E., 2007. Diversity Powers Innovation. Center for American Progress. January 26, <http://www.americanprogress.org/issues/economy/news/2007/01/26/2523/diversity-powers-innovation/>.

PCAST (President's Council of Advisors on Science and Technology). 2012. *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: Executive Office of the President, https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast_engage_to_excel_release_powerpoint__final.pdf

Shetterly, M.L. 2016. *Hidden Figures: The American Dream and the Untold Story of the Black Women Mathematicians Who Helped Win the Space Race*. William Morrow/Harper Collins Publishers.

Spencer, D.B. and Dawes, S., 2009. *Report of the Advisory Committee for GPRA Performance Assessment FY 2009*, Arlington, VA: National Science Foundation (NSF 09-068).

Szrom, M. 2015. Master's-to-PhD Program Paves Smooth Path for Underrepresented Students, Aug. 18, <http://www.insightintodiversity.com/masters-to-phd-program-paves-smooth-path-for-underrepresented-students/>

Task Force on Anchor Institutions, 2009. Anchor Institutions as Partners in Building Successful Communities and Local Economies. In *Retooling HUD for a Catalytic Federal Government: A Report to Secretary Shaun Donovan*, by Paul C. Brophy and Rachel D. Godsil, 147-169. Philadelphia, PA: Penn Institute for Urban Research.

Vishwanatha et al. 2016. NIH's Mentoring Makes Progress. *Science*, 353 (Nov. 18): 840-841.

White House Interagency Policy Group. 2016. *Reducing the Impact of Bias in the STEM Workforce: Strengthening Excellence and Innovation*, Nov. 29. Office of Science and Technology Policy, and Office of Personnel Management, https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp-opm_bias_mitigation_report_20161129_final.pdf

Williams, J.C. 2016. Study Underscores Bias Obstacles for Women, Minorities in Engineering, *Diverse Issues in Higher Education*, Oct. 31, http://diverseeducation.com/article/88736/?utm_campaign=DIV161...20DAILY%20NEWSLETTER%20NOV1&utm_medium=email&utm_source=Eloqua

Worthington, R.L., Stanley, C.A., and Lewis, W.T., Sr. 2014. "National Association of Diversity Officers in Higher Education Standards of Professional Practice for Chief Diversity Officers." *Journal of Diversity in Higher Education* 7: 227-234.

Appendices

Appendix A: Workshop Participants

- Mary Atwater, Professor of Mathematics and African American Studies, College of Education, Dept. of Mathematics and Science Education, University of Georgia
- Garikai Campbell, Provost, Senior Vice President, Morehouse College
- Patricia Campbell, President, Campbell-Kibler Associates, Inc.
- Nancy Cantor, Chancellor, Rutgers University – Newark
- Liza Cariaga-Lo, Associate Provost for Academic Development and Diversity, Brown University
- Joanna Chae, Director of Moelis Access Science, Netter Center for Community Partnerships, University of Pennsylvania
- Daryl Chubin, Technical Writer, Independent Consultant
- Clemencia Cosentino, Senior Researcher, Mathematica Policy Research*
- Anand Desai, Section Head, Evaluation and Assessment Capability, Office of Integrative Activities, National Science Foundation*
- Peter Eden, President, Landmark College
- Sarah Eichhorn, Associate Vice Provost for Teaching and Learning, University of California-Irvine
- Jaqueline C. Falkenheim, Senior Science Resource Analyst, National Center for Science and Engineering Statistics, National Science Foundation
- Jose Fuentes, Professor of Meteorology, Pennsylvania State University
- John Gawalt, Division Director, National Center for Science and Engineering Statistics, National Science Foundation
- Leslie Goodyear, Principal Research Scientist, Education Development Center
- Melvin Hall, Professor of Educational Psychology, Northern Arizona University*
- Ira Harkavy, Associate Vice President and Founding Director, Barbara and Edward Netter Center for Community Partnerships, University of Pennsylvania*
- Stafford Hood, Sheila M. Miller Professor of Education, Professor, Curriculum & Instruction and Educational Psychology, Director, Center for Culturally Responsive Evaluation and Assessment, University of Illinois at Urbana Champaign
- Rodney Hopson, Professor, Division of Educational Psychology, Research Methods, and Education Policy, College of Education and Human Development; Senior Research Fellow, Center for Education Policy and Evaluation, George Mason University

- Sylvia Hurtado, Professor of Higher Education and Head of Higher Education and Organizational Change Division, Graduate School of Education and Information Studies, University of California Los Angeles*
- Charles Isbell, Professor and Associate Dean, Georgia Institute of Technology
- Bobby Jeanpierre, Associate Professor, Department of Educational and Human Sciences, University of Central Florida
- James H. Johnson, Jr., Director, National Center for Environmental Research, Office of Research and Development. U. S. Environmental Protection Agency
- Robert Jones, Chancellor, University of Illinois at Urbana-Champaign
- Alicia Knoedler, Associate Vice President for Research Director, Center for Research Program Development & Enrichment, University of Oklahoma
- Richard Ladner, Professor in Computer Science & Engineering; Adjunct Professor in the Department of Electrical Engineering and in the Department of Linguistics, University of Washington
- Frances Lawrenz, Associate Vice President for Research and Professor of Educational Psychology, University of Minnesota*
- Michele Lezama, Executive Director, National Consortium for Graduate Degrees for Minorities in Engineering
- Kelly Mack, Vice President for Undergraduate STEM Education; and Executive Director, Project Kaleidoscope, Office of Undergraduate Science Education, Association of American Colleges & Universities
- Daniela Marghitu, Faculty Coordinator and Director, Education and Assistive Technology Lab, Samuel Ginn College of Engineering, Auburn University
- Louis Martin-Vega, Professor & Dean, College of Engineering, North Carolina State University*
- Travis Mayo, Evaluation Officer, U.S. Agency for International Development
- William McHenry, Executive Director of the Mississippi e- Center and Professor of Organic Chemistry, Jackson State University*
- Tom McKlin, Senior Evaluation Consultant, SageFox Consulting Group
- Robert Megginson, Arthur F. Thurman Professor, Department of Mathematics, University of Michigan
- Loretta Moore, Vice President for Research and Federal Relations, Professor of Computer Science, Jackson State University*
- Julie Park, Assistant Professor, Counseling, Higher Education, and Special Education, University of Maryland
- Cynthia Phillips, Evaluator, Office of Integrative Activities, National Science Foundation*

- Claudia Rankins, Program Director, Human Resource Development, National Science Foundation
- Emilda Rivers, Deputy Division Director, National Center for Science and Engineering Statistics, National Science Foundation
- Havidán Rodríguez, Provost and Executive Vice President for Academic Affairs, University of Texas Rio Grande Valley
- Melody Russell, Associate Professor, Science Education, Department of Curriculum and Teaching, College of Education, Auburn University
- Lucy Sanders, CEO and Co-Founder, National Center for Women & Information Technology (NCWIT) and Executive-in-Residence for the ATLAS Institute at the University of Colorado Boulder*
- Abigail Stewart, Sandra Schwartz Tangri Distinguished University Professor of Psychology and Women's Studies, Director, U-M ADVANCE Program, University of Michigan*
- Bonnie Swan, Director of Program Evaluation & Educational Research, University of Central Florida
- Antoinette Torres, Former Associate Vice Provost, Drexel University
- Lydia Villa-Kamaroff, Board Member, Cytonome/ST, LLC, ATCC Massachusetts Life Science Center
- Norman Webb, Senior Research Scientist Emeritus for the Wisconsin Education Research, University of Wisconsin–Madison
- Nai-Chang Yeh, Professor of Physics, Fletcher-Jones Foundation Co-Director, Kavli Nanoscience Institute, California Institute of Technology

* *Panelists and Panel Chairs*

Appendix B: Workshop Agenda

Workshop on Assessing Performance and Developing an Accountability System for Broadening Participation

October 13-14, 2016

Stafford I, Room 375
National Science Foundation
Arlington, Virginia 22230

Meeting Agenda

1st Day, October 13, 2016

12:00 PM-1:00PM

Working Lunch – Introductions & Opening Remarks/Framing of the Workshop

Dr. Ira Harkavy, Associate Vice President and Founding Director, Barbara and Edward Netter Center for Community Partnerships, University of Pennsylvania (CEOSE Chair and Principal Investigator)

1:00PM-2:30PM

Panel presentation on exemplary programs, their implementation, and scalability

Panel Chair: **Dr. Loretta Moore**, Vice President for Research and Federal Relations and Professor of Computer Science, Jackson State University (CEOSE Member)

Panelists:

- **Dr. Abigail Stewart**, Sandra Schwartz Tangri Distinguished University Professor of Psychology and Women's Studies, Director of U-M ADVANCE Program, University of Michigan
- **Dr. Lucy Sanders**, CEO and Co-Founder of National Center for Women & Information Technology (NCWIT) and Executive-in-Residence for the ATLAS Institute, University of Colorado Boulder
- **Dr. William McHenry**, Executive Director of the Mississippi e-Center and Professor of Organic Chemistry, Jackson State University

2:30PM-2:45PM

BREAK

2:45PM-3:30PM

Small group discussion

3:30PM-4:30PM

Report out and large group discussion on lessons learned from the panel presentation on exemplary programs

4:30PM-6:00PM

Panel presentation on metrics and measurement

Panel Chair: **Dr. Louis Martin-Vega**, Dean of Engineering and Professor of Industrial & Systems Engineering, North Carolina State University (CEOSE Vice-Chair and Co-Principal Investigator)

Panelists:

- **Dr. Clemencia Cosentino**, Senior Researcher, Mathematica Policy Research
- **Dr. Cynthia Phillips**, Evaluator, Office of Integrative Activities, National Science Foundation
- **Dr. Frances Lawrenz**, Associate Vice President for Research and Professor of Educational Psychology, University of Minnesota

6:00 PM-6:15PM

BREAK

6:15 PM-7:00PM

Small group discussion

7:00 PM-7:15PM

Travel to Holiday Inn for Working Dinner

Location: Ballston Room at the Holiday Inn

4610 N. Fairfax Dr. Arlington, VA 22203

7:15 PM-8:45PM

Working Dinner - Report out and large group discussion on lessons learned from panel presentation on metrics and measurement

2nd Day – October 14, 2016

8:00 AM-8:30AM

Working Breakfast – Summary of Day 1

8:30 AM-10:00AM

Panel presentation on developing and implementing an accountability system for broadening participation

Panel Chair: **Dr. Ira Harkavy**, Associate Vice President and Director of the Netter Center for Community Partnerships, University of Pennsylvania (CEOSE Chair and Principal Investigator)

Panelists:

- **Dr. Anand Desai**, Section Head, Evaluation and Assessment Capability, Office of Integrative Activities, National Science Foundation
- **Dr. Melvin Hall**, Professor of Educational Psychology, Northern Arizona University
- **Dr. Sylvia Hurtado**, Professor of Education and Head of Higher Education and Organizational Change Division, Graduate School of Education and Information Studies, University of California Los Angeles

10:00 AM-10:15AM

BREAK

10:15 AM-11:00AM

Small group discussion

11:00 AM-12:30PM

Working Lunch – report out & large group discussion on lessons learned from panel presentation on developing and implementing an accountability system for broadening participation

12:30 PM-1:30PM

Next steps

Appendix C: Annotated Bibliography: Accountability in the Public, Nonprofit and Private Sectors

Annotated Bibliography

Accountability in the Public, Nonprofit and Private Sectors

Compiled by
Anand Desai and Lisa M. Frehill
Evaluation and Capability Section
Office of Integrative Activities
National Science Foundation

Workshop on Accountability

National Science Foundation
October 13-14, 2016

This document serves as a beginning draft of a list of documents and articles on accountability in the public, nonprofit and private sectors.

Additionally, information on key documents and current efforts associated with accountability in the federal government are included at the end of this document.

Ballesteros, Athena, Smita Nakhooda, Jacob Werksman, and Kaija Hurlburt. (2010). *Power, Responsibility and Accountability: Re-Thinking the Legitimacy of Institutions for Climate Finance. Executive Summary*. Washington, DC: World Resources Institute. [Online at <http://www.wri.org>. (11 October 2016)]

This report from the *World Resources Institute* takes a financial perspective in exploring the responsibility and accountability of international institutions addressing climate change.

Behn, Robert D. (2001). *Rethinking Democratic Accountability*, Washington, DC: Brookings Institution Press.

Draws on ideas from the literature on public management and political theory to develop a model of accountability that includes “compacts of collective, mutual responsibility.” This model of democratic accountability attempts to empower public employees to exercise discretion while the government can still be held accountable to the public.

Bovens, Mark. (2007) *Analysing and Assessing Accountability: A Conceptual Framework* *European Law Journal*, 13(4): 447–468.

Framing accountability as a relationship, this paper explores that relationship along multiple dimensions.

Bovens, Mark, Robert E. Goodin, and Thomas Schillemans, eds. (2014). *The Oxford Handbook of Public Accountability*. Oxford University Press, Oxford.

The **Oxford Handbook of Public Accountability** is one of the key volumes in this literature, with review essays covering a variety of research on the many dimensions of accountability. Its editors, Bovens, Goodin, and Schillemans, are widely-published. The following chapters were reviewed and informed the EAC presentation to CEOSE, each providing insights into a key dimension of accountability:

- Klijn, Erik Hans, and Joop Fm Koppenjan. "Networks and Accountability."
- Warren, Mark E. "Accountability and Democracy."
- Halachmi, Arie. "Accountability Overloads."
- Greiling, Dorothea. "Accountability and Trust."
- Olsen, Johan P. "Accountability and Ambiguity."
- Mashaw, Jerry L. "Accountability and Time."
- Meijer, Albert. "Transparency."
- Patil, Shefali V., Ferdinand M. Vieider, and Philip E. Tetlock. "Process versus outcome accountability."
- Smith, Steven Rathgeb, "Accountability and the nonprofit sector."

Bovens, Mark, Thomas Schillemans, and Paul 't Hart. (2008). "Does public accountability work? An assessment tool," *Public Administration* 86(1): 225-242.

Drawing on three—democratic, constitutional and learning—perspectives, the authors develop a multi-criteria assessment tool and propose "boards of oversight of agencies" as a vehicle for assuring accountability.

Brandsma, Gijs Jan and Thomas Schillemans. (2012). "The Accountability Cube: Measuring Accountability," *Journal of Public Administration Research and Theory* 23: 953-975.

The article presents a three-dimensional scheme for understanding accountability along three axes: information (ranging from little to much); discussion (ranging from non-intensive to intensive); and consequences (ranging from few to many) is proposed. The authors emphasize the relational nature of accountability, describing multiple accountabilities that operate at different levels, depending on the reference axis.

Busuioc, E. Madalina, and Martin Lodge. (2015). "The reputational basis of public accountability," *Governance* 29(2): 247-263.

Accountability is discussed in terms of one's reputation as perceived by different audiences. Accountability is described as conveying the impression of competently performing one's accountability roles to generate reputational benefits.

Desai, Jay P. (2009) "The Power of Public Accountability: Executive Summary." Mumbai, India: Universal Consulting India. [Online at http://online.wsj.com/public/resources/documents/mum_accountability.pdf, accessed 23 September 2016)]

Provides a review of the literature for a study on public accountability conducted in India shortly after the Mumbai terrorist attacks of November 26, 2008. Focusing on the country as the unit of analysis, comparisons are made among a number of factors using multinational data sources.

Dowdle, M.W. (2006) Public Accountability: Conceptual, historical and epistemic mappings, in *Public Accountability, Designs, Dilemmas and Experiences*, ed. M. W. Dowdle. Cambridge: Cambridge University Press, pp: 1-32.

As the title suggests, it gives view of accountability and how various "crises" in accountability have been addressed.

Dubnick, M.J. (2002) Seeking salvation for accountability. Online at <http://andromeda.rutgers.edu/~dubnick/papers/salvation.htm>

Arguing that accountability has "fallen from conceptual grace," Dubnick offers suggestions for its salvation.

Fisher, Elizabeth. (2004). "The European Union in the Age of Accountability," *Oxford Journal of Legal Studies* 24(3): 495-515.

Accountability is context dependent because there is no general consensus about the standards for accountable behavior, and because the standards differ from role to role, time to time, place to place and from speaker to speaker.

International Forum on Business Ethical Conduct (IFBEC) (2014). *Accountability Report*. [Online at <http://ifbec.info/wp-content/uploads/2013/07/2014-Public-Accountability-Report.pdf>, accessed on 23 September 2016)].

Report on IFBEC member compliance with the principles and best practices that align with those principles aimed at reducing corruption.

Lerner, J. S. and P. E. Tetlock, (1999) 'Accounting for the Effects of Accountability', *Psychological Bulletin* Vol. 125(2): 255-275

A psychological perspective on accountability and its consequences for judgments and choices.

Lewis, Jenny M. (2014). "Individual and Institutional Accountability: The Case of Research Assessment," *Australian Journal of Public Administration* 73(4): 408-416.

While most of the other articles that have been included in this bibliography deal with project or service accountability, Lewis uses the example of three universities in three different countries to illustrate variations in the measurement of research performance.

The article links these structural features to academics' perceptions of the research assessment system.

Mashaw, Jerry L. (2006). "Accountability in Institutional Design: Some Thoughts on the Grammar of Governance," in Michael W. Dowdle (ed.) *Public Accountability, Designs, Dilemmas and Experiences*. Cambridge: Cambridge University Press, pp. 115–56.

Offers an approach to thinking about accountability by seeking answers to six linked questions to build what Mashaw refers to as accountability regimes.

Newcomer, Kathryn E. (2015). "From Outputs to Outcomes," pp. 125-156 in Guy, Mary E. and Marilyn M. Rubin (eds.). *Public Administration Evolving: From Foundations to the Future*. New York, NY: Routledge, pp.125-156.

This article links performance measurement and program evaluation to accountability. It provides a history of accountability efforts, in the US, starting with the Chief Financial Officers Act of 1990, which required agencies to have strategic plans that specified performance goals and reporting of outcomes.

Romzek, B.S. and M. J. Dubnick, (1998) Accountability, in J. M. Shafritz (ed.), *International Encyclopaedia of Public Policy and Administration*, Vol. 1: A–C Westview Press.

Broad overview

Schillemans, Thomas. (2014). "Calibrating Public Sector Accountability: Translating experimental findings to real world public accountability." [Online at <http://accountablegovernance.sites.uu.nl>. (accessed on 16 September 2016)].

Drawing on examples from the European Union, this paper provides lessons for governance from studies of the effects of accountability on decision making.

Scott, Colin. (2000). "Accountability in the Regulatory State," *Journal of Law and Society* 27(1): 38-60.

Discusses the need to balance discretion with control, which becomes difficult in complex administrative state where the distance between those empowered to exercise discretion and those who are held responsible for exercising control.

Shkabatur, Jennifer. (2012). "Transparency With(out) Accountability: Open Government in the United States," *Yale Law & Policy Review* 31(1): 79-140.

Shkabatur provides background on efforts to increase transparency as a means of increased accountability of federal government agencies. She provides details about why more data do not, necessarily result in stronger government accountability.

Smyth, Stewart. (2007). "Public accountability: a critical approach," *Journal of finance and management in public services* 6(2): 27-45.

Smyth re-conceptualizes public accountability as "a dynamic social relationship through which civil society seeks to control and challenge the state" by tracing the past 30 years' history of social housing in England to documents how "collective self organization of

sections of civil society not only challenges but at times succeeds in changing the policies and actions of the local state organizations.”

Stapenhurst R. and M. O’Brien (no date) Accountability in governance. World Bank. Online at <http://siteresources.worldbank.org/PUBLICSECTORANDGOVERNANCE/Resources/AccountabilityGovernance.pdf>

Tenbenschel, Tim, Judith Dwyer, and Josée Lavoie. (2014). "How Not to Kill the Golden Goose: Reconceptualizing accountability environments of third-sector organizations," *Public Management Review* 16(7): 925-944.

The focus is on the relationships between three sectors: nonprofit (third sector) organizations, government, and members of ethnic or geographic communities. Tensions among the interests associated with these three sectors can be resolved through shared goals, longer timeframes, and on-going negotiations.

Tetlock, Phillip E. (1998). “Losing Our Religion: On the Precariousness of Precise Normative Standards in Complex Accountability Systems,” pp. 121- 144 in Kramer, Roderick M., and Margaret A. Neale. (eds.) *Power and influence in organizations*. Thousand Oaks, CA: Sage.

Tetlock focuses on the ways in which people are “intuitive politicians,” who carefully consider how others will judge them based on their actions, so that individuals assess actions in terms of justifiability.

Willems, Tom. (2014). “Democratic accountability in public-private partnerships: The curious case of Flemish school infrastructure,” *Public Administration* 92(2): 340-358.

This paper challenges the view in the literature that accountability in public-private partnerships is poor and offers suggestions for improving accountability.

Other Materials – Federal Context for Accountability

Commission on Evidence-Based Policy

In March 2016, the bi-partisan Evidence-Based Policy Commission Act (P.L. 114-140) was signed into law. The Act established a 15-member commission with an 18-month charge “to develop a strategy for increasing the availability and use of data in order to build evidence about government programs, while protecting privacy and confidentiality.” Through the course of the Commission’s work, members will study how data, research, and evaluation are currently used to build evidence, and how to strengthen the government’s evidence-building efforts.” Kathryn Abraham (University of Maryland) and Ron Haskins (Brookings) co-chair the commission with includes members from industry, academia, and government.

Five white papers have been developed in support of the Commission’s work, as follows:

- Overview of Federal Evidence-Building Efforts
- Barriers to Using Administrative Data for Evidence-Building

- Comprehensive Data Inventory
- Using Administrative and Survey Data to Build Evidence
- Privacy and Confidentiality in the Use of Administrative and Survey Data.

Online at https://www.whitehouse.gov/omb/management/commission_evidence.

The Commission is the most recent in a long line of work that has sought to reduce waste, fraud, and abuse, while increasing efficiency and effectiveness of federally funded programs. Other milestones are described, below.

- **Government Performance and Results Act (GPRA, 1993)**

- First time federal agencies were mandated to be results-oriented, in earlier efforts, accountability was focused on process.
- Requirements:
 - Develop long-term strategic plans
 - Develop annual performance plans showing expected progress towards strategic goals

- **GPRA Modernization Act of 2010 (PL 111-352)**

This act addresses the concern that OMB and Congress agencies were not making effective use of the information on results generated under GPRA for program management.

- **OMB Program Assessment Rating Tool (2002)**

- Four-part questionnaire, each with different scoring weights:
 - Program purpose & design (20%)
 - Strategic planning (10%)
 - Program management (20%)
 - Program results (50%)
- Ratings:
 - Effective (Rating: 85-100)
 - Moderately effective (Rating: 70-84)
 - Adequate (Rating: 50-69)
 - Ineffective (Rating: 0-49)
 - Results not demonstrated – for programs without performance measures agreed-upon with OMB
- Inform budget decisions and identify actions to improve results
- Agencies accountable for implementing PART follow-up actions

Series of OMB Memos

Following two memoranda from the President (“Open Government Memorandum” (74 Fed. Reg. 4685) and “Freedom of Information Act (FOIA) Memorandum” (74 Fed. Reg. 4683), both signed 21 January 2009) to enhance transparency, OMB has issued a series of memoranda to provide guidance to federal agencies related to evaluation and evidence. By connecting such evi-

dence to the budget, some of the issues raised in the literature, (e.g., Shkabatur, 2012) associated with transparency and accountability have been addressed.

OMB memos include:

- 29 July 2010: M-10-32 “Evaluating Programs for Efficacy and Cost-Efficiency”
- 17 August 2011: M-11-31 “Delivering an Efficient, Effective, and Accountable Government”
- 18 May 2012: M-12-14 “Use of Evidence and Evaluation in the 2014 Budget”
- 14 February 2014: M-14-06 “Guidance for Using Administrative Data for Statistical Purposes”

Appendix D: Annotated Bibliography: Evaluation, Performance Measurement and Change Management in the Public, Nonprofit and Private Sectors

Annotated Bibliography

Evaluation, Performance Measurement and Change Management in the Public, Nonprofit and Private Sectors

Compiled by

Cynthia Phillips, PhD & Lisa M. Frehill, PhD
Evaluation and Capability Section
Office of Integrative Activities
National Science Foundation

Workshop on Accountability
National Science Foundation
October 13-14, 2016

This annotated list of documents is a start on building a bibliography on evaluation, performance measurement and change management in the public, nonprofit and private sectors.

Ashkenas, Ron. (January 2015). “We still don’t know the difference between change and transformation.” *Harvard Business Review* [Online at <https://hbr.org/2015/01/we-still-dont-know-the-difference-between-change-and-transformation> (accessed on 12 October 2016)].

Change management has become a common catch-phrase in business, yet the incremental “tweaks” to organizations are not often well-delineated from the more fundamental conceptualization of transformation. In this article, Ashkenas builds upon classic work by Kotter (2007) to emphasize the need to be more precise in the use of terminology and the expected results from efforts to bring about change in organizations. [Reference: Kotter, John P. (January 2007). “Leading change: Why Transformation efforts fail.” *Harvard Business Review* [Online at <https://hbr.org/2007/01/leading-change-why-transformation-efforts-fail> (accessed 12 October 2016)].

Christensen, Clayton M., Michael E. Raynor, and Rory McDonald. (December 2015). What is Disruptive Innovation? *Harvard Business Review*. (Online at <https://hbr.org/2015/12/what-is-disruptive-innovation> (accessed 12 October 2016)].

A disruptive innovation creates a new market and value network, eventually displacing established market leaders and alliances. The term was defined and phenomenon analyzed by Clayton M. Christensen beginning in 1995. More recent sources also include "significant societal impact" as an aspect of disruptive innovation. An innovation that is disruptive allows a whole new population of consumers at the bottom of a market access to a product or service that was historically only accessible to very few consumers.

Eckel, Peter D. and Adrianna Kezar. (2011). *Taking the Reins: Institutional Transformation in Higher Education*. Lanham, MD: Rowman & Littlefield.

Based on an American Council on Education project supported by the W.K. Kellogg Foundation, this book consolidates information Eckel and Kezar originally developed in a series of reports to Kellogg in the early 2000's that evaluated the five-year Project on Leadership and Institutional Transformation. The book focuses on six of the institutions that had the most success at the end of the five-year project. The authors develop a Mobile Model of Change to show how multiple change strategies work to bring about lasting transformation.

Marschke, Robyn, Sandra Laursen, Joyce McCarl Nielsen, and Patricia Rankin. (2007). "Demographic inertia revisited: An immodest proposal to achieve equitable gender representation among faculty in higher education." *Journal of Higher Education* 78(1): 1-26.

The authors use institutional data on faculty representation by sex at a large research university with a National Science Foundation Advance: Institutional Transformation program to perform simulations of faculty demographic change. The framework was developed in 2002 by Hargens and Long (2002), who found that faculty gender composition change was very slow, even when the pool of new PhD graduates had equivalent numbers of women and men. Marschke and her colleagues deploy a differential equations model of demographic inertia to simulate various potential administrative interventions to show that faculty gender composition can be changed more rapidly. [Reference: Hargens, Lowell L. and J. Scott Long. (2002). "Demographic inertia and women's representation among faculty in higher education. *Journal of Higher Education* 73(4): 494-517.]

Newcomer, Kathryn E. (2015). "From Outputs to Outcomes" pp. 125-156 in Guy, Mary E. and Marilyn M. Rubin (eds.). *Public Administration Evolving: From Foundations to the Future*. New York, NY: Routledge.

Newcomer provides a timeline of federal performance measurement, including a discussion of the Government Performance and Results Act (GPRA) 1993 and the GPRA Modernization of 2010. She describes how performance measurement has been compartmentalized in administrative units situated away from units involved with evaluation of other agency programs and services. Newcomer asserts that performance measurement and evaluation have recently been connected via a series of presidential memos and subsequent OMB memos that requiring both activities to develop an evidence base about "what works."

Styhre, Alexander. (2002). “Non-linear change in organizations: Organization change management informed by complexity theory.” *Leadership and Organization Development Journal* 23(6): 343-351.

In contrast to typical linear modeling of organizational change Styhre uses a non-linear model to study the change process in a Swedish telecommunication company, TelCo. By integrating complexity theory and non-linear approaches to organizational change, Styhre argues that disruptive, fluid processes of change may be better understood.

Twersky, Fay, Jodi Nelson, and Amy Ratcliffe. (2012). *A Guide to Actionable Measurement*. Bill and Melinda Gates Foundation Author: Seattle, WA. [Online <https://docs.gatesfoundation.org/Documents/guide-to-actionable-measurement.pdf> (accessed 12 October 2016)].

This document provides a basis for a common understanding of results measurement at the enterprise level. The guide is intended to inform organizational choices about strategic measurement and the setting of performance goals. “Actionable measurement is defined as, “measurement that has the potential to be acted upon, or is designed with action in mind.”

Federal Sources:

(March 2014). “Chapter 7: Evaluation as a tool for improving federal programs” Economic Report of the President” pp. 269-298 in Economic Report of the President. Washington, DC: U.S. Government Printing Office.

This chapter amplifies the President’s emphasis on use of evaluation within the federal government to document impact of programs and improve the delivery of federal services. The chapter describes the distinction between impact evaluations, process evaluations, and performance measurement. All of these methods are described as ways that program managers can “assess how programs operate and how well they work.”

General Accountability Office (GAO). (May 2011). Performance Measurement and Evaluation: Definitions and Relationships. GAO-11-646SP.

A simple GAO brochure that defines the relationship between performance measurement and evaluation within the federal government, providing definitions of four types of evaluation: process (implementation) evaluation; outcome evaluation; impact evaluation, and cost-benefit and cost-effectiveness analyses.

_____. (January 2012). Designing Evaluations: 2012 Revision. GAO-12-208G.

A document providing guidance to federal agencies on measurement and evaluation, which covers the four types described in the 2011 GAO brochure.