



Program
Assessment



Budget
Development



Implementation
Oversight



Outcome
Monitoring



Targeted
Evaluation

Targeted Evaluations Can Help Policymakers Set Priorities

A policymaker's guide to building evaluation capacity

This brief is one in a series about the five key components of evidence-based policymaking as identified in "Evidence-Based Policymaking: A Guide for Effective Government," a 2014 report by the Pew-MacArthur Results First Initiative. The other components are program assessment, budget development, implementation oversight, and outcome monitoring.

Overview

State and local leaders allocate millions of dollars each year to programs designed to deliver services to their constituents, but often the results go unmeasured. Without data on the effect, if any, that programs are having on participants and communities, policymakers are unable to discern which ones work, which don't, and how best to distribute limited public resources.

Governments have a range of evaluation tools to gauge the effectiveness and efficiency of public programs. One type they can employ is called an impact evaluation, which is a targeted study of how a particular program or intervention affects specific outcomes. These evaluations allow governments to isolate the effect of a program or initiative on a group, similar to the way that medical researchers test the effectiveness of a treatment or drug on a population. When conducted correctly, impact evaluations can help policymakers decide when to scale up programs that work, when to scale back or eliminate those that don't work, how to reallocate resources to better use, and when to improve those that show potential.

But states and counties face real challenges in conducting impact evaluations, and typically don't have the resources, time, or ability to evaluate every program. Recent innovations in techniques, increased access to data, and expanded partnerships between governments and researchers have made it easier for governments to rigorously assess the effectiveness of their programs, and they can remain judicious with their limited capacity and target evaluations carefully.

This brief provides options for government leaders to build their capacity to conduct these targeted studies by:

1. Hiring or training staff to facilitate impact evaluations.
2. Building partnerships with external research entities to leverage expertise.
3. Making better use of existing administrative data systems to reduce costs.
4. Aligning data policies and funding to support impact evaluation.

The brief also includes examples of how state and local governments have used impact evaluations to assess their programmatic investments, and details when such an assessment is worthwhile.

What Is an Impact Evaluation and How Is It Different From Outcome Monitoring?

Impact evaluations use a randomized controlled trial (RCT) design or quasi-experimental design (QED) to rigorously assess effectiveness. Both RCTs and QEDs use treatment (program participants) and comparison (nonparticipants) groups to evaluate outcomes against what would have occurred without the program. When done well, these evaluations provide policymakers with evidence of a program's effectiveness, helping inform their programmatic, policy, and funding decisions.

Impact evaluations provide different information and answer different questions than outcome monitoring systems that many state and local governments operate. Those systems track performance data that can be used to illustrate trends in the program's outcomes and compare that performance to prior years or other benchmarks, but this information cannot show what is driving those results.

For example, consider a state that implements a summer program to help students who are reading below grade level. Performance data show that participants scored significantly higher on English language arts exams their senior year of high school than students who were similarly struggling to read at grade level but did not participate in the program. Is the summer reading program the reason the test scores went up? More importantly for state policymakers, is this a program the state should expand in order to improve student learning? Outcome monitoring can't answer those questions. But impact evaluations can isolate the effects of the program and control for other factors that could influence student test scores, thus revealing more details about possible linkages.

How impact evaluation can support more effective government programs

Government leaders have a lot to gain from impact evaluations. They can use information from these studies, alongside performance data, to decide when to:

Scale up what works. Policymakers in Chicago relied on evidence from impact evaluations when deciding whether to scale up an intervention aimed at reducing violent crime. They turned to the University of Chicago Urban Labs,¹ which helps governments and nonprofit organizations test programs across public safety, education, health, poverty, energy, and environment areas, to conduct multiple impact evaluations of the city's Becoming a Man² program, which targets at-risk youth. Results from a 2016 impact evaluation showed that people who participated in the program had up to 50 percent fewer arrests for violent crime and increased their on-time graduation rates by up to 19 percent compared with a similar group of individuals who did not participate.³ Because of the program's success, the city expanded the program later that year to serve an additional 1,300 youths.⁴

Improve programs that show promise. The New York City Mayor's Office for Economic Opportunity (NYC Opportunity) regularly conducts impact evaluations to test local anti-poverty programs and uses the findings to make needed improvements. For instance, the office's Justice Corps⁵ initiative has undergone multiple modifications since its inception in 2008, driven by findings from internal performance monitoring and evaluations, including an RCT study.⁶ The findings showed positive impacts on participants' employment rates and wages but no effect on educational attainment or recidivism. This led to several programmatic adjustments, such as an expansion of youth development training for staff and implementation of a new risk-needs assessment and case management toolkit, which further evaluation showed improved outcomes.⁷

Scale back or replace programs that don't work. The Iowa Department of Corrections decided to scale back a community-based domestic violence intervention after impact evaluations in other states did not show the desired results. The department redirected resources to an alternative intervention⁸ and contracted with a local university to test its effect on reducing the number of persons who reoffend. The impact evaluation demonstrated positive effects, including a significantly lower recidivism rate than another domestic violence program operating in the state.⁹ This evaluation showed Iowa policymakers that the funds redirected from the old program to the new one (\$4.85 million dedicated since fiscal year 2011) were not wasted.¹⁰

Is an Impact Evaluation Always the Right Tool?

Before state and local governments spend limited funds and staff time on an impact study, it is important that they understand the questions being asked—and that they match the questions with the appropriate type of evaluation.

The table below shows four types of evaluations commonly used to answer policymakers' questions, the basic questions each study best addresses, and how each can help policymakers make informed choices. For instance, an outcome study could show policymakers whether participants in a college access program reported increased understanding of college readiness resources available to them after completing the program. An impact study, on the other hand, would show policymakers whether the program, as opposed to other factors, contributed to the increased understanding of college readiness resources.

Evaluation type	What the evaluation can answer	What the evaluation does for policymakers
Process	<ul style="list-style-type: none"> Has the program been implemented as intended? What types of implementation issues have emerged and how can they be addressed? What new ideas are emerging that can be tried out and tested? 	<ul style="list-style-type: none"> Demonstrates whether programs are being delivered with fidelity to the program model
Outcome	<ul style="list-style-type: none"> Did participants report desired changes after completing the program? For whom did the program work best? 	<ul style="list-style-type: none"> Shows whether desired intermediate outcomes are being achieved
Impact	<ul style="list-style-type: none"> Is the program effective at achieving desired outcomes? How did program participants do in achieving desired outcomes compared with nonparticipants? 	<ul style="list-style-type: none"> Provides rigorous evidence on a program's impact
Cost-benefit Analysis	<ul style="list-style-type: none"> Do the program's benefits outweigh its costs? Is the program more cost-beneficial than other programs addressing the same problem? 	<ul style="list-style-type: none"> Compares the expected return on investment of multiple programs that address the same problem

Sources: Centers for Disease Control and Prevention, "Types of Evaluation," accessed July 28, 2017, <https://www.cdc.gov/std/Program/pupestd/Types%20of%20Evaluation.pdf>; The University of Minnesota, "Different Types of Evaluation," accessed July 28, 2017, <https://cyfar.org/different-types-evaluation#Formative>; Permanency Innovations Initiative, "Family Search and Engagement (FSE) Program Manual" (2016), https://www.acf.hhs.gov/sites/default/files/cb/pii_fse_program_manual.pdf; Kathryn E. Newcomer, Harry P. Hatry and Joseph S. Wholey, eds., *Handbook of Practical Program Evaluation*, 4th ed., (New Jersey: John Wiley & Sons, Inc., 2015), 27–29; Michael Q. Patton, *Utilization-Focused Evaluation*, 4th ed., (Thousand Oaks, CA: Sage Publications, 2008)

How to build impact evaluation capacity

To leverage the benefits of impact evaluations, governments need to build their capacity—expertise, data, and funding—to conduct them effectively. Building this capacity can be challenging, even to officials who understand the importance of these studies. But government officials can choose an approach that best fits their available resources. These include: developing internal staff; building partnerships with external research entities; making better use of administrative data systems that enable researchers to use existing data to conduct impact evaluations; and aligning data policies and funding to support evaluations.

Hire or train staff to facilitate impact evaluations

In order to effectively and more frequently complete impact evaluations, governments can either hire research staff with the requisite skills to conduct these studies, or hire new or train existing staff to manage impact evaluation contracts. Governments can develop staff capacity in individual agencies and offices, as well as in centralized offices such as the office of a mayor, county commissioner, governor, or legislature.

Some states have hired staff with the requisite skills to conduct impact studies. For example, the Washington Department of Social and Health Services' Research and Data Analysis (RDA) office has about 70 full-time employees who perform a range of analytical and data management tasks, including impact evaluations. About 70 percent of these staff are funded through time-limited grants and projects, and 30 percent are supported by legislative appropriation. RDA has evaluated innovative pilot programs as well as long-standing projects that had not previously been studied. These evaluations have significantly affected policy and program decisions in the state, such as a 2009 study of a chronic care management practice that led to a scaling up of two of the state's health projects.¹¹

While Results First researchers found few offices in the 50 states, other than RDA, that regularly conduct impact evaluations, they did identify several offices that conduct other types of evaluations. For example, some legislative audit and research offices, as well as some agency research divisions, perform outcome evaluations that aim to measure the results of state programs or policies but are unable to control for external factors likely to influence those outcomes. Policymakers could work with these offices to identify opportunities to conduct impact evaluations, particularly on high dollar programs, or programs being considered for expansion.

Even where governments contract out impact studies to universities or external evaluation firms—a viable option for jurisdictions with limited resources—Results First researchers found that maintaining a small level of staff knowledgeable in evaluation has substantial benefits. Such staff can, for instance, help select external evaluators, manage evaluation contracts, collaborate on choosing study designs, and assist with data access. These staff may also have a deeper understanding of the data and issues relevant to a program than external evaluators do, and can help facilitate communication and knowledge transfer between external evaluators and program staff.¹² Michael Martinez-Schiferl, research and evaluation supervisor at the Colorado Department of Human Services, noted, “Program knowledge is very important to the research aspect. Having internal research staff embedded with program staff promotes collaboration on research and provides opportunities for research staff to develop some program expertise, as there are so many nuances about the program that they couldn't understand from an outside perspective.”¹³

NYC Opportunity contracts out its evaluations while maintaining staff to oversee the work. The staff members manage the contracts of the independent evaluation firms, oversee impact studies of the anti-poverty programs and key mayoral initiatives, and work in partnership with city agencies to use the evaluation results to inform decisions to expand, improve, or discontinue programs.¹⁴ NYC Opportunity has a dedicated fund from the city

to support its staff and external evaluations, but also seeks funding from federal grants and philanthropy to supplement this work.¹⁵

In the past 10 years, the organization has launched more than 70 programs, most of which have undergone an evaluation and some of which have become a national model for success.¹⁶ One example is the City University of New York's Accelerated Study in Associate Programs, which provides extensive financial, academic, and personal support to working adults who are pursuing an associate's degree. The program's first impact evaluation, done in partnership with a research organization, showed promising early results on academic outcomes, including lower drop-out rates and higher total credits accumulated, and a subsequent study demonstrated increased graduation rates.¹⁷ Following these successful evaluations, the program is expanding from 4,000 students in fiscal 2014 to 25,000 students in fiscal 2019, and has been replicated in other parts of the country.¹⁸

Build partnerships with external research entities to leverage expertise

Governments can help fill evaluation capacity gaps by creating long-term partnerships with external research entities, such as local universities, to perform an evaluation in its entirety or provide technical assistance or training.

For example, Brian Clapier, former associate commissioner at the City of New York's Administration for Children's Services (ACS), attributed some of his office's most successful evaluations to partnerships with universities.¹⁹ "Based on my experience the research-to-policy gap is a real challenge. One key strategy is to bring in the research partners (often from universities), and have these partners conduct the evaluations. Once the evaluation has occurred, strategically placed public agency staff can bridge the findings of the evaluation to program staff responsible for the policy development."

This approach proved successful when, in 2012, ACS contracted with Chapin Hall at the University of Chicago to implement and test a pilot child welfare program to promote healthy families and child well-being. The study found that two of the program's interventions—KEEP (Keeping Foster and Kin Parents Supported and Trained) and Parenting Through Change—were 11 percent more likely than a comparison group to achieve permanency.²⁰ Based on the results, ACS continued funding both programs.

Results First researchers identified several government-university partnerships—some that perform impact evaluations for programs spanning a range of policy areas and others that evaluate the same program over a period of time. For instance, state and county agencies in Wisconsin frequently partner with the University of Wisconsin Population Health Institute to perform evaluations on a range of human services programs, including behavioral health, child welfare, juvenile justice, and health.²¹ The New York State Office of Children and Family Services, on the other hand, has partnered with the Center for Human Services Research at the University of Albany since 1995 to perform multiple evaluations—including impact studies—of one child welfare program, Healthy Families New York.²²

Several jurisdictions have joined forces with policy labs—typically housed in universities—to help them design and test the effectiveness of government programs. For example, the Colorado Evaluation and Action Lab is a new government-research partnership that will help state officials to evaluate public programs and policies.²³

Government leaders also partner with individual researchers, rather than a research organization, whose interests and skills align with a particular policy or program they want to evaluate. South Carolina's Department of Health and Human Services, with the help of the Abdul Latif Jameel Poverty Action Lab (J-PAL), a research center at the Massachusetts Institute of Technology, paired up with health-focused researchers at Northwestern University to develop a randomized evaluation of assignment to Medicaid managed care plans.²⁴

Policymakers should pair these research partners with trained government staff in the offices that operate (or oversee providers who operate) the programs being evaluated.

Make better use of existing administrative data systems to reduce impact evaluation costs

Most governments maintain reporting systems that collect administrative data, such as criminal arrest or education records, which could be used to help conduct impact evaluations at a lower cost than collecting this information from scratch.²⁵ For example, New Mexico's Department of Corrections conducted a quasi-experimental design (QED) study of a substance use disorder program using administrative data from three state correctional offices. With a small evaluation budget, the department was able to successfully answer policymakers' questions about whether a program affected recidivism or substance use disorder relapse rates.²⁶

Some states, such as Washington and South Carolina, have built sophisticated data warehouses that link data across multiple agencies and can be used for performing evaluations. For example, the South Carolina Revenue and Fiscal Affairs Office (RFA) operates an integrated data warehouse that receives copies of agency databases used for program administration and research.²⁷ While the originating agencies maintain control over the use of the data, the RFA provides guidance on sharing and usage agreements to help researchers access the information to evaluate programs.²⁸ A new Pay for Success project focused on expanding an evidence-based family support and coaching program in the state will use data from the warehouse to evaluate the program's impact and calculate its return on investment.²⁹

Washington state's RDA also uses administrative data for most of its impact evaluations, which enables the agency to conduct more frequent evaluations on a wide range of programs. "We can knock out a quasi-experimental evaluation, assuming there's no new data to collect, in a relatively short time and at a fraction of the cost of doing an external evaluation," said health economist David Mancuso.³⁰

Both federal and private entities are creating funding opportunities to support state and local governments in using administrative data to support low-cost RCTs. (See Appendix A for a list of funding sources.)

Align data policies and funding to support evaluation

Two key obstacles to conducting impact evaluations are accessing the data necessary for the study and finding the resources to fund it. Yet state and county officials have found creative ways to mitigate these seemingly formidable challenges.

To generate new evidence on what works, researchers need access to government data, but service providers and government agencies may be hesitant to share data due to privacy issues or concerns over how the data might be used. Government leaders can alleviate these sensitivities and facilitate information access by developing sharing and usage agreements that outline the purpose of the data sharing, how it will be shared with the public, and privacy protections.³¹

For example, the Colorado Department of Education established a data-sharing and usage agreement with Mathematica Policy Research to allow Mathematica to access the department's administrative data to conduct an impact evaluation of a new charter school program's effects on education achievement.³² The agreement outlines the types of data to be shared, as well as the responsibility of the requestor to use the data only for the purposes outlined in the agreement (in other words, the impact study), to secure and later destroy the shared data, and to share analyses with the department prior to publication. Because of the shared data, Mathematica could perform an impact evaluation that showed state officials that taxpayer investments in the program had positive impacts on reading and math skills at the elementary, middle school, and high school grade levels.³³

States and localities can finance impact evaluations by setting aside internal funding for the studies, allowing governments to select programs in most need of an evaluation rather than being subject to the priorities of external funders. Results First researchers identified several ways governments are setting aside funding for rigorous evaluations, including through legislation, grants, and budget allocations.

For example, California passed a law in 2014 that sets aside funds to award contracts to recipients who agree to partner with an independent evaluator to assess the effectiveness of programs funded through the contracts.³⁴ Three counties received \$1.25 million to \$2 million in 2016 to implement and evaluate selected social services programs.³⁵ Though the law does not require recipients to conduct an impact evaluation, Alameda County is performing an RCT of a life coaching and mentoring services program aimed at reducing recidivism and increasing employment.³⁶

Some state and local governments dedicate funds to support staff who oversee or conduct impact evaluations. Washington state's RDA receives approximately 30 percent of its funding from a legislative appropriation that includes support for research staff who manage evaluations.³⁷ RDA supplements this funding with federal grants, including Medicaid and a Substance Abuse and Mental Health Block Grant.

Even when state and local governments build impact evaluation staff or set aside funds to support these studies, additional federal and private funding can help fill remaining capacity and funding gaps.

The federal government has provided several competitive and formula grant opportunities. For example, the Institute of Education Sciences released a request for applications in 2017 for low-cost RCTs or QEDs of education interventions.³⁸ The U.S. Department of Health and Human Services has also provided grants that included funding to evaluate child welfare and teen pregnancy interventions.³⁹ While these opportunities provide substantial support for impact evaluations, they should not be a substitute for using existing government resources to support this work; many are one-time grants that limit support to one study, and some target programs or policies that might not be an area of need in a particular jurisdiction.

Even jurisdictions that have never completed an impact evaluation have opportunities to start building this capacity through external sources. For instance, in 2016 J-PAL launched the State and Local Innovation Initiative to help jurisdictions perform randomized studies of social programs,⁴⁰ with eight jurisdictions chosen to participate in the first two rounds.⁴¹ In addition to funding, each will receive technical support and custom trainings to expand the internal capacity to create, use, and share rigorous evidence. (See Appendix A for more information on funding opportunities.)

How to select programs for an impact evaluation

While state and local governments have demonstrated the value of impact studies to assess programmatic investments, it is not practical (or even necessary) to rigorously evaluate every program. Decision-makers can identify and prioritize which programs to study by considering four key questions:

1. **Does the program have an evidence base?** To identify programs that could benefit from an impact evaluation, governments can inventory the programs they fund and determine which ones have evidence supporting their effectiveness. National research clearinghouses—which review and aggregate impact evaluations in order to rate programs by their level of evidence of effectiveness—can help determine if local programs have existing evidence. Governments can prioritize evaluations for programs that do not have strong evidence of their effectiveness.
2. **Has the program been properly implemented?** Poorly implemented programs are less likely to achieve the outcomes that leaders and residents expect, which would impair the results of an impact assessment.

3. **Are the right data available for an impact study?** To conduct an impact study, evaluators need access to the right kinds of data. If the data are owned by other parties (e.g., another agency or program provider) or do not exist, governments should consider the feasibility of getting the data, which could entail developing data-sharing agreements or spending additional funds and time to collect new data.
4. **Does the program serve a significant number of people and/or is it a large budget item?** Programs that have a higher number of clients and/or are costlier typically have a larger impact on a government's budget than those that are less prescribed or less costly, and may be more attractive options for an impact evaluation.

Decision-makers may find that some of their untested programs are not good candidates for an impact evaluation. In that case, they can take other steps to ensure these programs are generating positive results, such as tracking outcomes of participants and reviewing implementation to identify any issues with operation and delivery. Decision-makers can review these programs again at a later time to determine if they have become evaluation-ready.

Conclusion

Policymakers care about funding what works, and impact evaluations are an important tool that can be used to inform better, data-driven decisions. Impact evaluations provide critical information on program effectiveness, which policymakers can consider when making decisions about when to scale up, scale back, or make adjustments to a particular program or initiative.

By building their jurisdiction's capacity to evaluate untested programs, policymakers can hold themselves accountable to the public, and ensure that the state's public dollars are directed to those programs that yield positive results. While challenges still exist for governments seeking to regularly evaluate their programs, new technology and opportunities to leverage impact evaluation expertise through partnerships or grants have made it more feasible than ever for state and local governments to conduct rigorous evaluations of local programs. By carefully prioritizing which programs are ripe for impact evaluations, governments can make the most of their resources and fill in gaps about which programs are working and which are not.

Appendix A: Impact evaluation resources

General impact evaluation resources:

- Actionable Intelligence for Social Policy's Technical Assistance and Training Program, <https://www.aisp.upenn.edu/resources/training-technical-assistance/>
- American Evaluation Association, Find an Evaluator, <http://www.eval.org/p/cm/ld/fid=108>
- Kathryn E. Newcomer, Harry P. Hatry and Joseph S. Wholey, eds., *Handbook of Practical Program Evaluation*, 4th ed., (New Jersey: John Wiley & Sons, Inc., 2015)
- Laura and John Arnold Foundation, "Key Items to Get Right When Conducting Randomized Controlled Trials of Social Programs" (2016), <http://www.arnoldfoundation.org/wp-content/uploads/Key-Items-to-Get-Right-When-Conducting-Randomized-Controlled-Trials-of-Social-Programs.pdf>
- Office of Juvenile Justice and Delinquency Prevention, "Evaluability Assessment: Examining the Readiness of a Program for Evaluation" (2003), <http://www.jrsa.org/pubs/juv-justice/evaluability-assessment.pdf>
- Paul J. Gertler et al., *Impact Evaluation in Practice*, Second Edition (2016), <https://openknowledge.worldbank.org/handle/10986/25030>

- Social Innovation Fund, “Introducing the Impact Evaluability Assessment Tool” (2014), https://www.nationalservice.gov/sites/default/files/resource/SIF_Impact_Evaluability_Assessment_Tool_Final_Draft_for_Distribution.pdf

Impact evaluation funding resources:

- Abdul Latif Jameel Poverty Action Lab, State and Local Innovation Initiative, <https://www.povertyactionlab.org/stateandlocal/apply>
- Federal grants, <https://www.grants.gov>
- Harvard Kennedy School Government Performance Lab, <https://govlab.hks.harvard.edu/apply>
- Institute of Education Sciences, <https://ies.ed.gov/funding/>
- Laura and John Arnold Foundation, <http://www.arnoldfoundation.org/request-for-proposals/>

Sample of national research clearinghouses*:

- California Evidence-Based Clearinghouse for Child Welfare, <http://www.cebc4cw.org>
- Penn State’s Clearinghouse for Military Family Readiness, <https://militaryfamilies.psu.edu/programs-review/>
- Robert Wood Johnson Foundation and University of Wisconsin Population Health Institute’s What Works for Health, <http://www.countyhealthrankings.org/roadmaps/what-works-for-health>
- U.S. Department of Education’s What Works Clearinghouse, <https://ies.ed.gov/ncee/wwc/>
- National Cancer Institute’s Research-Tested Intervention Programs, <https://rtips.cancer.gov/rtips/index.do>
- U.S. Department of Justice’s CrimeSolutions.gov, <https://www.crimesolutions.gov/>

**This is not an exhaustive list; readers can find additional clearinghouses in:*

- The Bridgespan Group and Results for America, “The What Works Marketplace, Helping Leaders Use Evidence to Make Smarter Choices” (2015), Appendix 3, <http://results4america.org/wp-content/uploads/2015/04/WhatWorksMarketplace-vF-1.pdf>
- Corporation for National and Community Service, “Clearinghouses and Evidence Reviews for Social Benefit Programs” (2016), <https://www.nationalservice.gov/sites/default/files/documents/Clearinghouses%20and%20Evidence%20Reviews.pdf>

Endnotes

- 1 The University of Chicago Urban Labs, “Crime Lab,” accessed July 27, 2017, <https://urbanlabs.uchicago.edu/labs/crime>.
- 2 Youth Guidance, “BAM—Becoming a Man,” accessed July 27, 2017, <https://www.youth-guidance.org/bam/#>.
- 3 Sara B. Heller et al., “Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago,” *The Quarterly Journal of Economics* 132, no. 1 (2017): 1–54, <https://doi.org/10.1093/qje/qjw033>.
- 4 City of Chicago, “Mayor Emanuel Launches First Phase of Mentoring Initiative by Immediately Expanding Becoming a Man Mentoring Program to Reach More Than 4,000 Male Youths This School Year,” news release, Oct. 3, 2016, https://www.cityofchicago.org/city/en/depts/mayor/press_room/press_releases/2016/september/Expanding-Becoming-A-Man-Mentoring-Program.html.
- 5 NYC Justice Corps, accessed July 27, 2017, <http://www.nycjusticecorps.org/>.
- 6 Erin L. Bauer et al., “Evaluation of the New York City Justice Corps” (2014), <http://www1.nyc.gov/assets/opportunity/pdf/Westat-Justice-Corps-Evaluation.pdf>.
- 7 David Berman (director of programs and evaluation, the New York City Mayor’s Office for Economic Opportunity), interview with Pew-MacArthur Results First Initiative, May 19, 2016.

- 8 The Pew-MacArthur Results First Initiative, "Iowa's Cutting-Edge Approach to Corrections: A Progress Report on Putting Results First to Use" (2013), http://www.pewtrusts.org/-/media/legacy/uploadedfiles/pcs_assets/2013/rfbriefresultsfirstiowaprogressreportfinal.pdf.
- 9 Amie Zaring, Sarah Bannon, and Meg Berta, "Evaluation of Acceptance and Commitment Therapy for Domestic Violence Offenders," *Psychology of Violence* (2017): <http://dx.doi.org/10.1037/vio0000097>.
- 10 Pew-MacArthur Results First Initiative internal data tracker on funding shifts and allocations to evidence-based programs.
- 11 David Mancuso (health economist, Washington State Department of Social and Health Services), interview with Pew-MacArthur Results First Initiative, Feb. 13, 2017.
- 12 Melissa Conley-Tyler, "A Fundamental Choice: Internal or External Evaluation?," *Evaluation Journal of Australasia* 4, nos. 1 & 2 (2005): 3-11, https://www.aes.asn.au/images/stories/files/Publications/Vol4No1-2/fundamental_choice.pdf; James Bell Associates, "Evaluation Brief: Building Evaluation Capacity in Human Service Organizations" (2013), http://www.jbassoc.com/ReportsPublications/building%20internal%20eval%20capacity%20final%2011_21_13.pdf; Administration for Children and Families, *The Program Manager's Guide to Evaluation*, 2nd ed., (Washington, DC: U.S. Department of Health and Human Services, n.d.), https://www.acf.hhs.gov/sites/default/files/opro/program_managers_guide_to_eval2010.pdf.
- 13 Michael Martinez-Schiferl (research and evaluation supervisor, Office of Performance and Strategic Outcomes, Colorado Department of Human Services), interview with Pew-MacArthur Results First Initiative, Aug. 2, 2017.
- 14 Berman interview.
- 15 Ibid.
- 16 Ibid.
- 17 Susan Scrivener, Michael J. Weiss, and Colleen Sommo, "What Can a Multifaceted Program Do for Community College Students?" (2012), https://www.mdrc.org/sites/default/files/full_625.pdf; Susan Scrivener et al., "Doubling Graduation Rates—Three-Year Effects of CUNY's Accelerated Study in Associate Programs (ASAP) for Developmental Education Students," (2015), <https://www.mdrc.org/publication/doubling-graduation-rates>.
- 18 Berman interview.
- 19 Brian Clapier (former associate commissioner for research and analysis, New York City Administration for Children's Services), interview with Pew-MacArthur Results First Initiative, Aug. 1, 2016.
- 20 New York City Administration for Children's Services, "ACS Launches Major New Foster Care Initiative, Strong Families NYC," news release, July 22, 2015, <https://www1.nyc.gov/site/acs/about/press-releases-2015/foster-care-initiative.page>.
- 21 University of Wisconsin Population Health Institute, <https://uwphi.pophealth.wisc.edu>.
- 22 University at Albany, State University of New York, Center for Human Services Research, "Healthy Families New York," accessed Jan. 29, 2018, <http://www.albany.edu/chsr/hfny.shtml>.
- 23 University of Denver, "Colorado Evaluation and Action Lab," news release, June 19, 2017, <http://news.du.edu/colorado-evaluation-and-action-lab/>.
- 24 Jason Bauman and Todd Hall, email to Pew-MacArthur Results First Initiative, Sept. 11, 2017; Abdul Latif Jameel Poverty Action Lab, "Lessons From the J-PAL State and Local Innovation Initiative, Case Study: South Carolina," accessed Feb. 12, 2018, https://www.povertyactionlab.org/sites/default/files/documents/stateandlocal-southcarolina_0.pdf.
- 25 National Conference of State Legislatures, "Child Welfare Information Systems" (2015), <http://www.ncsl.org/research/human-services/child-welfare-information-systems.aspx>; Coalition for Evidence-Based Policy, "Rigorous Program Evaluations on a Budget: How Low-Cost Randomized Controlled Trials are Possible in Many Areas of Social Policy" (2012), <http://coalition4evidence.org/wp-content/uploads/2012/03/Rigorous-Program-Evaluations-on-a-Budget-March-2012.pdf>; Laura Feeney et al., "Using Administrative Data for Randomized Evaluations" (2015), <https://www.povertyactionlab.org/sites/default/files/resources/2017.02.07-Admin-Data-Guide.pdf>.
- 26 New Mexico Sentencing Commission, "Report in Brief: Bernalillo County Metropolitan Court DWI-Drug Court Intent-to-Treat Outcome Study Stage 2" (2010), <https://nmssc.unm.edu/reports/2010/report-in-brief-bernalillo-county-metropolitan-court-dwi-drug-court-intent-to-treat-outcome-study-stage-2.pdf>.
- 27 Sarah Crawford, "Data Quality Report for the South Carolina Department of Probation, Parole, and Pardon Services (SCDPPPS) Offender Services Report," (n.d.), http://www.scdps.gov/ohsjp/stats/SpecialReports/Data_Quality_Report.pdf.
- 28 Chris Finney (South Carolina Revenue and Fiscal Affairs Office), interview with Pew-MacArthur Results First Initiative, Aug. 31, 2016.
- 29 South Carolina Department of Health and Human Services, "Pay for Success Contract Among South Carolina Department of Health and Human Services and Nurse-Family Partnership and The Children's Trust Fund of South Carolina" (2016), https://www.scdhhs.gov/sites/default/files/2016_0321_AMENDED%20NFP%20PFS%20Contract_vFinal%20Executed.pdf.

- 30 Mancuso interview.
- 31 Permanency Innovations Initiative Evaluation Team, "Using Child Welfare Administrative Data in the Permanency Innovations Initiative Evaluation, OPRE Report 2016-47" (2016), https://www.acf.hhs.gov/sites/default/files/opre/using_child_welfare_administrative_data_in_pii_compliant.pdf; Erika M. Kitzmiller, "IDS Case Study: The Circle of Love: South Carolina's Integrated Data System" (2014), http://www.aisp.upenn.edu/wp-content/uploads/2015/08/SouthCarolina_CaseStudy.pdf; William T. Grant Foundation, "Research Practice Partnerships: Developing Data Sharing Agreements," accessed July 31, 2017, <http://rpp.wtgrantfoundation.org/developing-data-sharing-agreements>.
- 32 Colorado Department of Education, "Research and Data Sharing Agreement: Between the Colorado Department of Education and Mathematica Policy Research, Inc." (2014), <https://www.cde.state.co.us/cdereval/mathematicadsa>.
- 33 Mathematica Policy Research, Inc., "Understanding the Effect of KIPP as it Scales: Volume I, Impacts on Achievement and Other Outcomes" (2015), <http://www.kipp.org/results/independent-reports/#mathematica-2015-report>.
- 34 California Government Code § 97010.
- 35 California Board of State and Community Corrections, "BSCC Awards Grants That Will Pay for Successes," last modified April 14, 2016, <http://www.bscc.ca.gov/news.php?id=87>.
- 36 Harvard Kennedy School Government Performance Lab, "Our Projects," accessed July 31, 2017, <https://govlab.hks.harvard.edu/our-projects>; California Board of State and Community Corrections, "BSCC Awards."
- 37 Mancuso interview.
- 38 Institute of Education Sciences, "Low-Cost, Short-Duration Evaluation of Education Interventions," accessed Jan. 29, 2018, <https://ies.ed.gov/ncer/projects/program.asp?ProgID=90>.
- 39 James Bell Associates, "Summary of the Title IV-E Child Welfare Waiver Demonstrations" (2013), http://www.acf.hhs.gov/sites/default/files/cb/waiver_summary_final_april2013.pdf; U.S. Department of Health and Human Services Office of Adolescent Health, "About the Teen Pregnancy Prevention (TPP) Program," last modified Feb. 13, 2017, http://www.hhs.gov/ash/oah/oah-initiatives/tpp_program/about/#.
- 40 Abdul Latif Jameel Poverty Action Lab, "State and Local Innovation Initiative," accessed Jan. 29, 2018, <https://www.povertyactionlab.org/stateandlocal>.
- 41 Abdul Latif Jameel Poverty Action Lab, "State and Local Innovation Initiative - Partner Jurisdictions," accessed Jan. 29, 2018, <https://www.povertyactionlab.org/stateandlocal/evaluations>.



MacArthur
Foundation

Contact: Catherine An, officer
Email: can@pewtrusts.org
Phone: 202-540-2088

Pew-MacArthur Results First Initiative, a project of The Pew Charitable Trusts and the John D. and Catherine T. MacArthur Foundation, works with states to implement an innovative evidence-based policymaking approach that helps them identify and invest in policies and programs that are proved to work.