THE PEW CHARITABLE TRUSTS | MacArthur Foundation

Evidence-Based Policymaking

A guide for effective government

External Reviewers

Acknowledgments

The Pew Charitable Trusts

901 E St. NW
Washington, DC 20004

pewtrusts.org

The Pew Charitable Trusts is driven by the power of knowledge to solve today's most challenging problems. Pew applies a rigorous, analytical approach to improve public policy, inform the public, and invigorate civic life.

The John D. and Catherine T. MacArthur Foundation

140 S. Dearborn St. Chicago, IL 60603

macfound.org

The John D. and Catherine T. MacArthur Foundation supports creative people and effective institutions committed to building a more just, verdant, and peaceful world. In addition to selecting the MacArthur Fellows, the Foundation works to defend human rights, advance global conservation and security, make cities better places, and understand how technology is affecting children and society.

Contents

Overview

Governments make budget and policy choices each year that have long-term effects on both their fiscal futures and the outcomes they deliver for constituents. Recognition is growing that policymakers can achieve substantially better results by using rigorous evidence[1] to inform these decisions, enabling governments to select, fund, and operate public programs more strategically. Until now, however, no comprehensive road map has provided clear guidance on using this approach.

To fill this gap, the Pew-MacArthur Results First Initiative has developed a framework that governments can follow to build and support a system of evidence-based policymaking. Based on an extensive review of research and in-depth interviews with government officials, practitioners, and academic experts, the framework identifies steps that both the executive and legislative branches can take to drive the development, funding, implementation, and monitoring of policies and programs.

The framework has five key components, each with multiple steps that enable governments to make better choices through evidence-based policymaking: (1) program assessment, (2) budget development, (3) implementation oversight, (4) outcome monitoring, and (5) targeted evaluation.

1. **Program assessment**. Systematically review available evidence on the effectiveness of public programs.

    a. Develop an inventory of funded programs.
    b. Categorize programs by their evidence of effectiveness.
    c. Identify programs' potential return on investment.

2. **Budget development**. Incorporate evidence of program effectiveness into budget and policy decisions, giving funding priority to those that deliver a high return on investment of public funds.

    a. Integrate program performance information into the budget development process.
    b. Present information to policymakers in user-friendly formats that facilitate decision-making.
    c. Include relevant studies in budget hearings and committee meetings.
    d. Establish incentives for implementing evidence-based programs and practices.
    e. Build performance requirements into grants and contracts.

3. **Implementation oversight**. Ensure that programs are effectively delivered and are faithful to their intended design.

    a. Establish quality standards to govern program implementation.
    b. Build and maintain capacity for ongoing quality improvement and monitoring of fidelity to program design.
    c. Balance program fidelity requirements with local needs.
    d. Conduct data-driven reviews to improve program performance.

4. **Outcome monitoring**. Routinely measure and report outcome data to determine whether programs are achieving desired results.

    a. Develop meaningful outcome measures for programs, agencies, and the community.
    b. Conduct regular audits of systems for collecting and reporting performance data.
    c. Regularly report performance data to policymakers.

5. **Targeted evaluation**. Conduct rigorous evaluations of new and untested programs to ensure that they warrant continued funding.

    a. Leverage available resources to conduct evaluations.
    b. Target evaluations to high-priority programs.
    c. Make better use of administrative data—information typically collected for operational and compliance purposes—to enhance program evaluations.
    d. Require evaluations as a condition for continued funding for new initiatives.
    e. Develop a centralized repository for program evaluations.

This report discusses how and why evidence-based policymaking is a growing national trend and reviews the framework in detail to provide tips and strategies that policymakers can use to instill evidence in decision-making at all levels of government.

Why evidence-based policymaking?

Evidence-based policymaking uses the best available research and information on program results to guide decisions at all stages of the policy process and in each branch of government. It identifies what works, highlights gaps where evidence of program effectiveness is lacking, enables policymakers to use evidence in budget and policy decisions, and relies on systems to monitor implementation and measure key outcomes, using the information to continually improve program performance. By taking this approach, governments can:

* **Reduce wasteful spending.** By using evidence on program outcomes to inform budget choices, policymakers can identify and eliminate ineffective programs, freeing up dollars for other uses.

* **Expand innovative programs.** Requiring that new and untested programs undergo rigorous evaluation helps determine whether they work and identifies opportunities to target funding to innovative initiatives that deliver better outcomes to residents or reduce costs.

* **Strengthen accountability**. Collecting and reporting data on program operations and outcomes makes it easier to hold agencies, managers, and providers accountable for results.

A new era in responsible governance

Support is growing across the country for using evidence to inform policy and budget decisions and guide the implementation of programs, in good times as well as bad. Although the need to improve government performance has long been recognized, researchers from the Results First Initiative identified several factors that are driving renewed attention to this issue, including ongoing fiscal pressures, the increasing availability of data on program effectiveness, federal funding incentives, and state legislation that support—and in some cases require—the use of evidence-based programs and practices.

Previous attempts to address these challenges by linking program performance to budget allocations—for example, performance-based budgeting—have met with limited success because of insufficient analytical capacity or limited data, among other reasons.[2] Now, with better technology, easier access to data, and the ability to more accurately measure the performance and cost-effectiveness of government services, policymakers have an opportunity to put their jurisdictions on a sustained path of evidence-based decision-making.

Ongoing fiscal pressures

In recent years, many governments were forced to make major budget reductions due to revenue shortfalls that occurred during the Great Recession. Although some states have seen tax revenue rebound, others continue to confront tight budgets due to lagging revenue, increasing costs in areas such as Medicaid, and other pressures.[3] Many governments at both the state and local levels also face long-term fiscal challenges, such as meeting retirement benefit obligations for public employees.[4] This has increased demands by policymakers for better information on the outcomes that programs deliver for constituents and better tools to identify activities that fail to deliver desired results.

Increasing availability of evidence on what works

Over the past two decades, a growing body of research has evaluated the effectiveness of public programs. Multiple clearinghouses are compiling this information by reviewing and categorizing hundreds of research studies to identify effective and promising programs across a range of policy areas.[5] As a result, policymakers have access to more information about what works than ever before.[6] States and local governments can avoid duplication of effort and use this evidence to inform their policy and budget decisions.

Federal funding incentives

Increasingly, federal grant recipients, including states and localities, are required to target federal funds to evidence-based programs. Since 2009, for example, the U.S. departments of Education, Health and Human Services, and Labor have directed approximately $5.5 billion to seven initiatives that support proven programs.[7] Although they represent only a small percentage of total federal spending, these grants provide incentives for recipients to implement proven programs.[8] These include the Investing in Innovation (i3) Fund, which prioritizes education programs with strong evidence of effectiveness and evaluation of innovative programs; the Maternal and Infant Early Childhood Home Visiting program, which requires grantees to direct 75 percent of federal dollars to evidence-based programs and to evaluate the impact on key outcomes; and the Workforce Innovation Fund, which supports projects that use data to design new approaches to improving employment and training outcomes.[9]
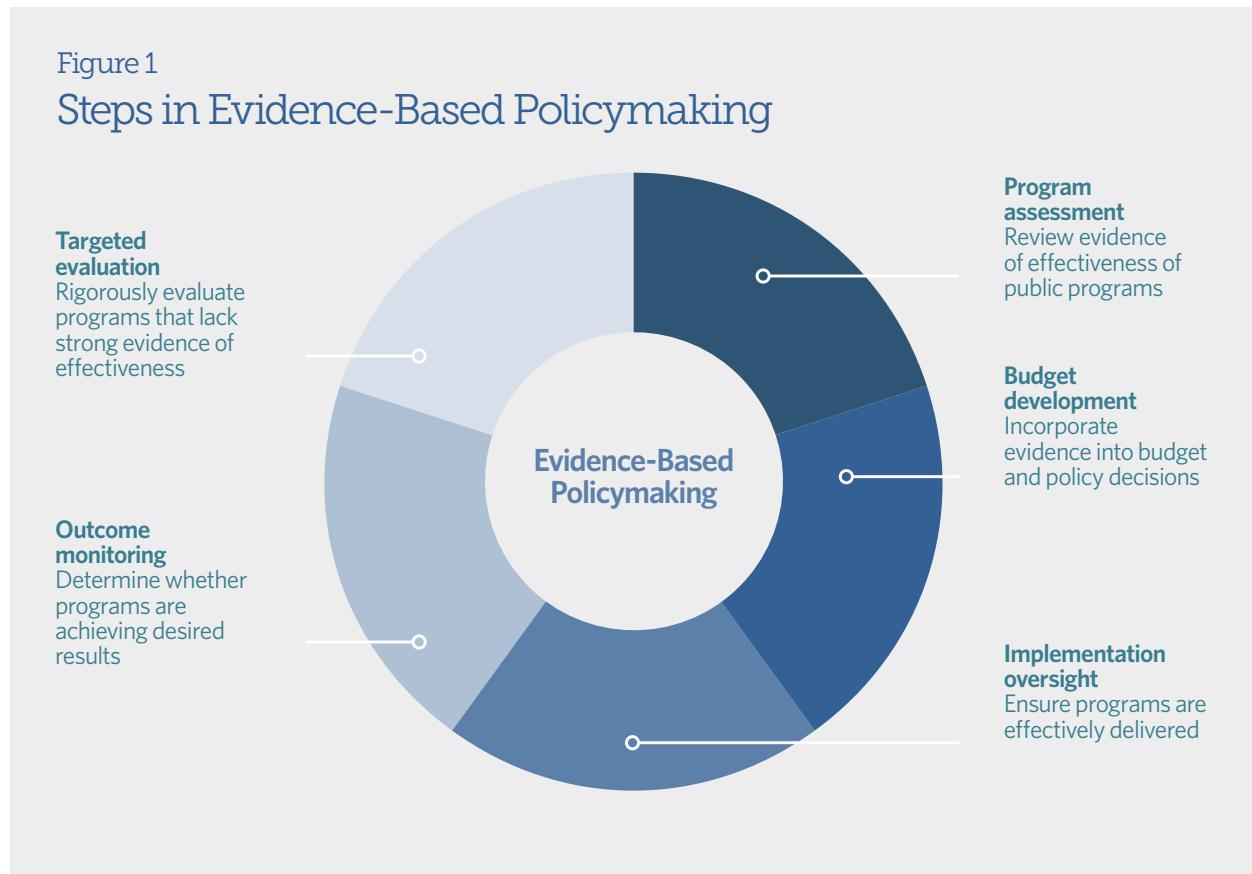
Growing interest from state leaders

State policymakers are using legislation as a vehicle to encourage investment in programs that have been proved effective. Results First researchers identified over 100 state laws across 42 states passed between 2004 and 2014 that support the use of evidence-based programs and practices.[10] These laws provide incentives for agencies to implement proven programs and help establish common standards with which to compare programs.

State leaders are also using cost-benefit analysis to inform their policy and spending decisions. A recent Results First study found that the number of states assessing the costs and benefits of programs and policy options increased 48 percent between 2008 and 2011, and 29 states reported using cost-benefit studies to inform policy or budget decisions.[11] In addition, since 2011, 16 states and four California counties have partnered with the Results First Initiative to apply a customized, innovative cost-benefit approach to policy and budget decision-making.

Key components of evidence-based policymaking

Results First researchers identified five key components that support a system of evidence-based policymaking (see Figure 1). In developing this report, our research found that while many states have put one or more of these in place, none has developed a comprehensive approach across all branches of government. For each of the components, our framework includes specific steps that help to ensure successful implementation. Governments may lack capacity to implement all of the elements at once, but they can still strengthen their use of evidence-based policymaking by focusing on particular features highlighted in this report.

Figure 1

Steps in Evidence-Based Policymaking

Targeted evaluation
Rigorously evaluate programs that lack strong evidence of effectiveness

Outcome monitoring
Determine whether programs are achieving desired results

Evidence-Based Policymaking

Program assessment
Review evidence of effectiveness of public programs

Budget development
Incorporate evidence into budget and policy decisions

Implementation oversight
Ensure programs are effectively delivered

© 2014 The Pew Charitable Trusts

Program assessment. Systematically review available evidence on the effectiveness of public programs

Government leaders should develop an inventory of the programs they currently operate and then assess the available evidence of effectiveness and return on investment for each one. This provides important baseline information that enables government leaders to identify which programs are working and achieving high returns on taxpayer dollars, which need further evaluation, and which are not delivering expected outcomes (see Appendix B: Potential roles in state government).

Develop an inventory of funded programs

Many state and local governments do not have a complete catalog of the programs they fund, which is a necessary starting point for determining which are effective and which are not. Government leaders can require agencies to conduct a census to identify all publicly operated and contracted programs and collect standard information about each, including their funding levels, services delivered, and populations served. To help facilitate this process, governments often find it beneficial to develop a common definition of "program" to provide consistency across agencies.

In 2014, Rhode Island's Office of Management and Budget worked with the state's departments of Corrections and Children, Youth, and Families and the judiciary to develop an inventory of 58 state-funded programs intended to reduce recidivism in adult and juvenile justice systems. In its initial report, published in March 2014, the office found that 33 percent of the programs inventoried were not evidence-based, and only two had been recently evaluated to determine whether they were implemented according to research-based standards. As a result of this process, the office recommended additional evaluations to ensure fidelity to these standards.[12]

Categorize programs by their evidence of effectiveness

Policymakers need clear information about the effectiveness of the programs they fund. By requiring agencies to categorize the programs they operate according to the rigor of their evidence of effectiveness, lawmakers and agency leaders can ensure they have access to the information they need to make this determination. A first step is to develop definitions for each category, based on the strength of evidence. For example, some states use "evidence-based programs," which may be defined as requiring multiple evaluations that use rigorous methods such as randomized controlled trials. A second is "promising programs," which may include those that have been evaluated and shown effective but through a less rigorous research design. State or local governments can use resources from national clearinghouses or other states in developing these definitions.

Embedding such standards of evidence in statute can increase the likelihood that they will be enforced consistently and endure political changes. In 2012, Washington passed legislation to increase the number of evidence-based children's mental health, child welfare, and juvenile justice services.[13] The law has three key requirements:

1. The Washington State Institute for Public Policy and the University of Washington Evidence-Based Practice Institute, in consultation with the Department of Social and Health Services, will publish definitions of "evidence-based," "research-based," and "promising practices." To be considered an evidence-based program, the law requires that the benefits produced outweigh its cost. In addition, the institute and the university will review existing national and international research to identify programs that meet the criteria based on these definitions.

2. The state's Department of Social and Health Services and the Health Care Authority will complete a baseline assessment of evidence- and research-based practices in child welfare, juvenile rehabilitation, and children's mental health services. This includes the extent to which currently funded programs meet the standards of evidence, the utilization of those services, and the amount of funding received by each program.

3. The Department of Social and Health Services and the Health Care Authority must report to the governor and Legislature on strategies, timelines, and costs for increasing the use of evidence- and research-based practices.

In 2014, Mississippi passed similar legislation mandating that its Legislative Budget Office and Joint Committee on Performance Evaluation and Expenditure Review, known as PEER, categorize programs in four state agencies as evidence-based, research-based, promising practices, or other programs and activities with no evidence of effectiveness.[14] The legislation includes definitions of each evidence level to guide the work of the budget office and PEER.

Leveraging National Research Clearinghouses

In recent years, several national research clearinghouses have been established that conduct systematic literature reviews to identify effective public programs across a range of policy areas, including adult criminal and juvenile justice, child welfare, mental health, pre-K to higher education, and substance abuse.[*] Although the clearinghouses use slightly different criteria for evaluating the strength of evidence, most have adopted a tiered structure that allows researchers and policymakers to easily determine the relative effectiveness of each program. For example, the What Works Clearinghouse, an initiative of the U.S. Department of Education's Institute of Education Sciences, uses a system of recognizable symbols to convey this information: two plusses mean a program has positive effects, while an oval means there is no evidence of an effect on outcomes.[†] The What Works Clearinghouse has rated the impact of approximately 130 education programs on 26 educational outcomes.

Policymakers and agency leaders can use these clearinghouses to compare the programs that their state or locality operates to those the clearinghouses have deemed to be effective. For example, a state might find that only a small percentage of its adult criminal justice programs had nationally recognized evidence of positive outcomes, which would raise questions about whether the remaining programs should continue to receive funding.[‡]

---

* There are several widely recognized national research clearinghouses, including the U.S. Department of Education's What Works Clearinghouse, the U.S. Department of Justice's CrimeSolutions.gov, Blueprints for Healthy Youth Development, the Substance Abuse and Mental Health Services Administration's National Registry of Evidence-Based Programs and Practices, the California Evidence-Based Clearinghouse for Child Welfare, What Works in Reentry, and the Coalition for Evidence-Based Policy.

† What Works Clearinghouse, U.S. Department of Education Institute of Education Sciences, accessed July 29, 2014, http://ies.ed.gov/ncee/wwc/findwhatworks.aspx.

‡ The Pew-MacArthur Results First Initiative recently created a central database that compiles information from eight research clearinghouses to enable policymakers and their staffs to readily identify effective, evidence-based programs in multiple policy areas, including adult criminal justice, juvenile justice, mental health, substance abuse, early education, K-12 education, and child welfare. For more information, please see: http://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2014/09/results-first-clearinghouse-database.

> Over the past two fiscal years, five states—Iowa, Massachusetts, New Mexico, New York, and Vermont—have used the Results First model to target $81 million in funding to more effective programs that the model shows will achieve higher returns.

Identify programs' potential return on investment

In addition to knowing whether programs have been rigorously evaluated, it is also important for government leaders to know if investing in them would generate enough benefits to justify their costs. Governments can use cost-benefit and cost-effectiveness analyses to answer this question. These studies calculate the dollar value of the outcomes that different programs achieve and weigh them against the costs. Conducting such analyses requires technical expertise and extensive fiscal and outcome data and may not be practicable for all programs. When feasible, however, this approach enables governments to rank programs by their potential return on investment, providing policymakers with critical information on which alternatives can achieve the greatest returns for constituents.

The Pew-MacArthur Results First Initiative is working with 16 states and four counties to implement cost-benefit analysis models that enable policymakers to use this approach in their budget and policy decisions. Results First uses a nationally recognized, peer-reviewed model and a three-step process:

1. Employ the best national research on program outcomes to identify what works, what doesn't, and how effective various alternatives are in achieving policy goals.

2. Apply jurisdiction-specific data to predict the impact each program would achieve.

3. Compare the costs of each program to its projected benefits and produce a report that ranks each alternative by the relative value it would generate for taxpayers.

Over the past two fiscal years, five states—Iowa, Massachusetts, New Mexico, New York, and Vermont—have used the Results First model to target $81 million in funding to more effective programs that the model shows will achieve higher returns.[15]

Budget development. Incorporate evidence of program effectiveness into budget and policy decisions, giving funding priority to those that deliver a high return on investment of public funds

For evidence-based policymaking to be successful, governments must systematically use evidence of program effectiveness to inform their processes for making budget and policy decisions. This requires regular communication between researchers, budget staff, and policymakers as well as the development of strong executive and legislative champions. Analytic results must be reported to policymakers in timely and accessible ways.

Integrate program performance information into the budget development process

Executive branch agencies should use performance information when developing their budgets to ensure funds are directed to programs that have strong evidence of effectiveness and away from those that are not delivering results. To accomplish this, agencies can develop output and outcome measures for all major programs and report those metrics in their budget requests. Agencies should develop numerical performance targets that can be used by policymakers to measure progress against key benchmarks and goals. For evidence-based programs, the targets should reflect outcomes predicted by research.

A well-functioning performance measurement system can help governments decide where to pull back on funding as well as where to provide greater support. Connecticut's Result-Based Accountability system has been operating for eight years and has become an important part of the state's appropriations process. When outcome measures showed that the state's $20 million annual investment in early reading programs was having no positive effect on reading skills, they were first denied funding and later analyzed in-depth to identify potential solutions. The study found that reading specialists, a central element of the initiative, lacked sufficient training to achieve expected results and that funding to support early reading efforts was often used for other purposes. Based on this, the state has turned to other approaches, such as adding reading-related graduation requirements for education degrees and implementing techniques based on a reading program in Norwalk that has had success. "Our reading scores are now creeping up instead of going down," said Representative Diana Urban, co-chair of the Connecticut General Assembly's Select Committee on Children.[16]

Present information to policymakers in user-friendly formats that facilitate decision-making

To increase the likelihood that policymakers will use evidence to inform critical budget decisions, complex information must be presented in ways they can readily understand and act on. For any program, policymakers need answers to at least three important questions:

- Is the program working?

- Do its benefits outweigh its costs?

- How does the program compare to alternative programs?

To provide this information, agencies can produce annual rankings that compare programs targeting similar outcomes based on effectiveness, cost, and benefits produced. When practicable, governments can use cost-benefit analyses to calculate a return on investment for each program, providing policymakers with data on how to best allocate resources to achieve each agency's goals.[17] At a minimum, policy staff should compare programs with common goals according to their documented impact on specific outcomes—for example comparing a set of programs that all have as their primary goal reducing child abuse and neglect.

When practicable, contracts and grants should include performance goals that encourage organizations to provide evidence-based programs and to implement those services as designed.

Several states, including Washington, Iowa, and New Mexico, have developed *Consumer Reports*-type analyses, which rank programs by their benefit-to-cost ratios.[18] In 2012, the Iowa Department of Corrections issued a report highlighting the costs and benefits of various criminal justice programs over a 10-year period.[19] The analysis showed that among prison-based programs, cognitive behavioral therapy programs were inexpensive to operate and highly effective in reducing recidivism, returning $37.70 in benefits for every dollar spent. In contrast, correctional educational programs, although also effective, returned only $2.91 in benefits per dollar invested.[20] As a result, the department is considering expanding its cognitive behavioral therapy programs and plans to reduce other, less effective activities proportionally.

Include relevant studies in budget hearings and committee meetings

Policymakers can use executive and legislative budget hearings and committee meetings as opportunities to discuss key findings from program evaluations, audits, cost-benefit analyses, and other research. Governments can establish procedures requiring research offices to provide relevant reports to budget and policy committees, which should, in turn, be encouraged to consider the findings in their deliberations.

The New Mexico Legislative Finance Committee regularly presents program evaluations, agency performance report cards, and cost-benefit analyses during budget hearings and committee meetings to support its budget and legislative recommendations. In 2013, for example, the committee presented a report in budget hearings showing that reducing recidivism by 10 percent using proven programs could save the state $8.3 million in prison costs and approximately $40 million in avoided costs to victims.[21] The findings, in addition to other analyses, helped inform decisions to allocate $7.7 million to effective criminal justice programs.

Establish incentives for implementing evidence-based programs and practices

Governments can use grant competitions to encourage adoption or expansion of evidence-based programs. Agencies can also partner with private philanthropies or businesses to scale up promising programs—those that demonstrate the potential to achieve a positive return on investment.

Wisconsin's Treatment Alternatives and Diversion grant program provides funding to counties to implement data-driven alternatives to prosecution and incarceration of criminal offenders with a history of substance abuse. A county is eligible for a grant if, among other criteria, the services provided are consistent with evidence-based practices. Between 2006 and 2013, these grants funded nine county diversion or drug court programs. A recent evaluation found that grant-funded projects averted 231,533 incarceration days for offenders, 57 percent of whom were not convicted of a new crime three years after being discharged from the program.[22]

Governments can also develop pay-for-success models and social impact bond agreements, both of which raise capital from private investors or philanthropic organizations to scale up programs that have the potential to achieve better outcomes and save the government money. Although these efforts are still in their infancy, several states, including Massachusetts and New York, are moving forward with plans to provide incentives for data-driven programming.

New York raised $13.5 million through its social impact bond to support the Center for Employment Opportunities, which provides evidence-based employment services to ex-offenders including job training, transitional employment, and job placement. Bank of America Merrill Lynch (BAML) and Social Finance Inc. raised funding from more than 40 individual and philanthropic investors, which included several BAML clients, as well as foundations, among them the Laura and John Arnold Foundation and the Robin Hood Foundation. The Rockefeller Foundation agreed to guarantee up to 10 percent of the investors' principal. An independent evaluator will determine whether the program is reaching its goals of reducing recidivism and increasing employment.[23] The state will repay investors only if the outcomes outlined in the bond agreement are achieved.

Build performance requirements into grants and contracts

When practicable, contracts and grants should include performance goals that encourage organizations to provide evidence-based programs and to implement those services as designed. To realize the benefits of performance-based contracts, program administrators should work closely with providers and program developers to create measures that accurately gauge performance, while striking a balance between the need for accountability and the importance of continuous quality improvement and increased capacity. These contracts need to be carefully crafted and monitored to protect against unintended consequences, such as creating incentives for providers to take only those clients most likely to succeed and to reject those considered high-risk.

In the early 2000s, the Connecticut Judicial Branch's Support Services Division, which oversees state-run juvenile justice programs, developed a Center for Best Practices to review research on evidence-based interventions and integrate effective strategies into current programs, most of which were contracted out.[24] The center determined that several programs were achieving poor outcomes, and the division began working with contractors to identify the aspects of service delivery that yielded desired outcomes and to incorporate those elements into their contracts. Through this process, the division developed a standard report card, which includes performance data and other quality assurance information, that is updated semiannually and is reported to the Legislature each year. Division staff members also meet quarterly with contractors to review performance data, identify areas for improvement, and determine technical assistance needs.[25]

When properly designed, performance-based contracts can help move agencies away from a fee-for-service model, which pays providers for the amount of services they deliver, toward a system that rewards results. For example, in Tennessee, under more traditional fee-for-service contracting methods, foster care providers that were most successful in finding permanent homes for children could suffer financially because the children no longer needed their services. In contrast, the state's pay-for-success program, which was introduced in 2009, provides contracts that pay more to agencies that achieve permanent placements for children. Over a five-year period, this helped reduce the time children spent in foster care by 235,000 days and saved $20 million, which has been reinvested to further improve services.[26]

Implementation oversight. Ensure that programs are effectively delivered and are faithful to their intended design

The quality of program implementation can dramatically affect outcomes: Even the most effectively designed interventions can produce poor results when poorly run. To ensure proper implementation, governments should establish strong monitoring systems that assess all funded programs, including those administered by nongovernmental entities. This monitoring should ensure that evidence-based programs are carried out with fidelity to their design and incorporate the elements that are critical to their effectiveness, and it should include processes that improve quality by using information gathered through monitoring to make adjustments that improve performance.

Too often, program support and oversight is one of the first areas cut when budgets are tight, resulting in inadequate implementation and poor outcomes. To sustain the positive results, policymakers should include funding for support and monitoring in the base budgets of programs. Then, if budgets are reduced, effective services can still be delivered to high-need clients, which is preferable to serving more people ineffectively by poorly implemented programs.

Establish quality standards to govern program implementation

Broad-based implementation standards can promote the consistent delivery of high-quality services by providing baseline requirements for monitoring and oversight. These criteria should also be included in agency contracts to help ensure that providers understand and comply with expectations. Evidence-based programs frequently have detailed implementation manuals that managers can use to set quality standards.

For example, state leaders tasked the Washington State Institute for Public Policy with developing standards to implement evidence-based juvenile justice programs after an evaluation found that sites where the programs were not implemented with fidelity had poor results.[27] The standards address four key elements of quality assurance—program oversight, provider development and evaluation, corrective action, and ongoing outcome evaluation—and include protocols for hiring, staff training and assessment, and management and oversight of service delivery. Providers are required to undergo an initial probationary period during which they receive training and feedback. Thereafter they are evaluated annually. The state regularly monitors program completion and recidivism rates for juveniles who receive certain services. The implementation standards are credited with helping the state achieve greater reductions in crime and juvenile arrest rates compared with the national average and a decrease of more than 50 percent in youth held in state institutions.[28]

Build and maintain capacity for ongoing quality improvement and monitoring of fidelity to program design

Governments can support effective implementation by offering—or partnering with organizations that offer—training, technical assistance, and other services to program providers. They can also offer infrastructure support, including computer systems that facilitate data collection and outcome reporting. Some nationally recognized evidence-based programs also provide training or technical assistance services to assist implementation.

The Evidence-based Prevention and Intervention Support Center, or EPISCenter, provides technical assistance to communities and service providers in Pennsylvania to support the implementation of evidence-based prevention and intervention programs.[29] Since 2008, the center has assisted in the establishment of nearly 300 evidence-based programs in more than 120 communities throughout the state.[30] The center is a collaborative partnership among the Pennsylvania Commission on Crime and Delinquency and Penn State University. It receives funding and support from the commission and from the Pennsylvania Department of Public Welfare. Experts from the center provide technical assistance to local staff on implementation, evaluation, and sustainability and help develop the infrastructure to monitor the program for fidelity to its original design. Over time, providers build internal capacity for these operations and many continue to report data to the EPISCenter even after their initial funding has ended. These efforts have been highly beneficial.

Balance program fidelity requirements with local needs

Many evidence-based programs have identified the key service elements that are critical to achieving desired outcomes but they also note that some services may need to be modified for local conditions. Administrators monitoring programs should ensure that key elements are implemented with fidelity while allowing other features

to be adapted to meet community and cultural differences. Administrators, program developers, and service providers should work together to ensure that program adaptations do not negatively affect outcomes.

In 2009, the Oregon Legislature passed a bill to utilize the nationally recognized "Wraparound" system of care for emotionally disturbed and mentally ill children, with statewide programs in place by 2015.[31] A fundamental part of Oregon Wraparound is fidelity monitoring, overseen by the Oregon Department of Human Services. The National Wraparound Initiative has provided assessment tools to ensure that programs remain faithful to its 10 basic principles. However, administrators may adapt other services to local conditions and needs, which can vary across the state. "The goal is to meet communities where they are so that this is sustainable. Whatever you're building needs to be part of the community you're working with. You maintain the fidelity of the model, but ensure that it's tailored to the community," says William Baney, director of the Systems of Care Institute, at Portland State University's Center for Improvement of Child & Family Services, which provides training and systems support to Oregon Wraparound.[32]

> ❝ The goal is to meet communities where they are so that this is sustainable. Whatever you're building needs to be part of the community you're working with. You maintain the fidelity of the model, but ensure that it's tailored to the community."
>
> **William Baney,** *director of the Systems of Care Institute at Portland State University's Center for Improvement of Child & Family Services*

Conduct data-driven reviews to improve program performance

Regularly scheduled data-driven performance management meetings enable agency and state leaders to discuss performance data, develop or refine performance objectives, identify areas for improvement, promote innovative strategies, foster coordination, and hold managers accountable for results.[33] Agencies should hold similar meetings with their staffs and service providers to pinpoint opportunities for improvement and address performance barriers.

This approach was developed by the New York City Police Department and popularized by the city of Baltimore through CitiStat. The CitiStat model allowed Baltimore leaders to focus on performance goals, improve service delivery, and generate $350 million in savings over a seven-year period, enabling it to reinvest $54 million in new programming for children.[34]

Using a similar approach, Maryland StateStat measures statewide performance and tracks key indicators from biweekly agency data, which are analyzed for trends to inform strategies for improvement. Regular meetings are held with the governor, agency heads, and StateStat staff to clarify goals, refine approaches for achieving outcomes, and track performance.[35] This use of data has engendered a culture of organizational learning in which program managers and agency leaders discuss challenges and solve problems.

Aligning Existing Services With Key Elements of Evidence-Based Programs

Governments can often improve the outcomes from programs that are not evidence-based by aligning their key characteristics with those that are. For example, a locally developed program for juvenile offenders may be able to improve its results by incorporating features of programs that research shows are highly effective in reducing recidivism.

The Standardized Program Evaluation Protocol, or SPEP, developed by the Peabody Research Institute at Vanderbilt University, provides a standardized measure to determine how closely a particular program conforms to the most effective practices, according to scientific research, in juvenile justice.[*] The tool assesses programs in four primary areas that research has identified as critical to effectiveness, including the primary service provided, the quantity of service, the quality of delivery, and the risk level of the juveniles served. The tool is currently being implemented in three jurisdictions—Milwaukee County, Wisconsin, and in Iowa and Delaware. They are part of the federal Office of Juvenile Justice and Delinquency Prevention's Juvenile Justice Reform and Reinvestment Initiative, established to support improvements to current service delivery models.[†] The information gathered through the tool is used by states and localities to improve existing juvenile justice services and align them with evidence-based practices without having to redesign entire service systems. Arizona and North Carolina have also used the SPEP tool to assess the effectiveness of their juvenile justice programs, and initial data show that larger reductions in recidivism correlated with higher SPEP ratings.

"The SPEP tool allows states to look at programs that may not be name brand, but to determine whether they have the common elements that research suggests works," says Mark Lipsey, Ph.D., director of the Peabody Research Institute.[‡] "From a practical standpoint, in some policy areas there are relatively few evidence-based programs; they can be expensive and require significant training to get providers up to speed. We see our approach as complementary with model programs which are also part of our scheme, but it allows states to look at a broader set of programs."

---

[*] Peabody Research Institute, "Standardized Program Evaluation Protocol," accessed July 29, 2014, https://my.vanderbilt.edu/spep. The protocol was developed by Mark Lipsey, Ph.D., of the Peabody Research Institute, Vanderbilt University.

[†] Shay Bilchik and Kristen Kracke, "How Do You Scale Evidence-Based Programs: A look at OJJDP's Juvenile Justice Reform and Reinvestment Initiative," *Cost-Benefit Knowledge Bank on Criminal Justice*, (Dec. 4, 2013), accessed July 29, 2014, http://cbkb.org/2013/12/how-do-you-scale-evidence-based-programs-a-look-at-ojjdps-juvenile-justice-reform-and-reinvestment-initiative.

[‡] Pew-MacArthur Results First Initiative interview with Mark Lipsey, director, Peabody Research Institute, Vanderbilt University, Jan. 8, 2014.

Outcome monitoring. Routinely measure and report outcome data to determine whether programs are achieving desired results

Many governments have made significant investments to build and implement performance reporting systems, but these too often focus on outputs, such as the number of programs provided or clients served, rather than results, such as reduced recidivism or increased graduation rates, and are of limited use to policymakers. Governments should make sure that performance measurement systems collect and report essential outcome data for all major programs.

Develop meaningful outcome measures for programs, agencies, and the community

Performance monitoring systems should provide output and outcome data that meet the information needs of various stakeholders, including program administrators, policymakers, and constituents. For example:

- Administrators can monitor operations by using data on program outputs, such as the number of families served, the percentage of families achieving program milestones, and the caseloads of field staff.

- Agency leaders can use intermediate outcome data to assess progress toward key goals, such as reducing the percentage of participating mothers who deliver low-birth-weight babies.

- Policymakers and constituents can use measures that gauge long-term trends, such as the percentage of children graduating from high school, to determine whether public programs are achieving their overall objectives.

For example, Virginia Performs is an interactive, publicly available database that collects and reports performance data on a wide range of government functions at multiple levels—including program, agency, department, and cross-cutting strategic government priority—and for diverse audiences such as program administrators, agency leadership, policymakers, and the public. As part of Virginia's strategic planning process, state agencies identify performance measures, which are then tracked through the Virginia Performs system.[36] These data are one set of inputs used to generate the annual Virginia Report, a balanced accountability scorecard created by the bipartisan Council on Virginia's Future, which is headed by the governor.[37] Where data are available, Virginia's performance is compared with the national average, the top performing state in the nation, and three similar states. The data allow users to consider high-level strategic goals and a wide range of performance indicators at the department, agency, and program levels.

When determining what measures to track, governments can consult resources available from several national organizations. For example, in 2012, the Federal Department of Housing and Urban Development launched the Healthy Communities Transformation Initiative to provide governments with the tools to assess the "physical, social, and economic roots of community health." The initiative's first deliverable, a collection of 28 key indicators that governments can use to track outcomes across 10 policy domains, was created following review of existing models and is now being tested in select jurisdictions. Many of the indicators can be derived from publicly available data and customized by state, municipality, or neighborhood.[38]

Agencies can also visit the national clearinghouses to identify the outcomes predicted for various programs by rigorous research and use those findings to set performance targets for funded programs. Governments can require programs that lack strong evidence of effectiveness to develop theories of change or logic models that specify their expected results and can then use this information to establish outcome measures and performance targets for those programs.

Performance measures also should periodically be examined to ensure they still serve as reliable indicators of success. For contracted services, governments should ensure that providers collect and report common outcome metrics so that officials can compare performance and aggregate the overall program effects.

Conduct regular audits of systems for collecting and reporting performance data

Effective performance measurement systems should be user-friendly and provide data that meet the needs of multiple stakeholders. Even the best-designed system, however, will be of little value if the reported data are inaccurate or misleading. Governments should provide training to agency staff and contracted providers on how to collect, analyze, and report performance data, and develop processes for regularly verifying that these data are accurate.

Performance measurement systems can easily fall into disuse without strong leadership supporting them or adequate training for providers and agency personnel. In 2012, Louisiana's auditors confronted this issue during a review of the state's performance budgeting system, once considered a model program.[39] The audit noted that many statutory processes were no longer being followed and that reported information was not being used to inform budget decisions. The findings emphasized the need to increase awareness of the system, improve how performance data were presented to policymakers, and ensure reliability. The report also noted the importance of

training, for both legislative and agency staff, on using the system. "State agencies have all of this data but do not necessarily have the tools or the skill set to analyze the data and use it for performance management purposes," says Karen LeBlanc, performance audit manager at the Louisiana Legislative Auditor's office.

Regularly report performance data to policymakers

Performance data can be a valuable tool for managing, overseeing, and assessing the value of programs, but it is critical to provide the information to policymakers on a regular basis, in easy-to-digest formats that highlight key findings, and readily translate to budget and policy decision-making. Several state and local governments have developed report card systems that focus on agency or program performance on key outcomes. Report card data are often reported through public websites and may be presented to policymakers through regular hearings and meetings. Data dashboards, interactive business tools that display a set of performance indicators, can also be beneficial in tracking and focusing on high-level outcomes in real time.

In Michigan, for example, a frequently updated performance dashboard provides past and current data on a variety of indicators relevant to the administration's key policy objectives, including economic strength, health and education, quality of life, and public safety.[40] Policymakers and the public can quickly see which programs are succeeding or struggling based on simple graphics such as a green "thumbs up" for progress and a red "thumbs down" for a lack of achievement. For example, in spring 2014, third-grade reading test scores were slowly continuing to trend upward. The dashboard featured this information using a graph showing proficiency increasing from 63.5 percent in fiscal year 2011 to 70 percent three years later. On the other hand, the dashboard provided a warning signal that the self-reported percentage of students being bullied rose from 22.7 percent in 2011 to 25 percent in 2013.[41]

Targeted evaluation. Conduct rigorous evaluations of new and untested programs to ensure that they warrant continued funding

Programs with little or no evidence of effectiveness carry a higher risk of yielding poor outcomes. Governments should therefore direct evaluation resources to programs that lack rigorous outcome data, receive significant funding, or pose other risks in order to ensure they are delivering desired results and that further support is warranted.

Governments should also allocate funding for evaluation to limit the risk that investments are made in programs that do not work or that are less effective over time. Rather than assuming that programs can find money within existing budgets, governments should dedicate resources for this purpose once existing evaluation capacity and expertise have been maximized.

Leverage available resources to conduct evaluations

Almost all states have offices that conduct program evaluations and performance audits, and these can provide unbiased information to help policymakers assess program effectiveness. Governments should develop an inventory of their resources and dedicate at least a portion of them to conducting rigorous outcome evaluations.

For example, legislative audit and research offices can be a critical resource in conducting independent program evaluations, but historically much of their work has focused on assessing compliance and management issues rather than outcomes. Legislators can work with these offices as they set their research agendas to identify opportunities to dedicate a larger portion of their resources to determining whether programs are achieving desired results.

Target evaluations to high-priority programs

No government has the capacity to regularly evaluate all of its funded programs, so it is important to set priorities. Governments can develop a list of programs to be evaluated, weighing factors such as the program's purpose, existing evidence of effectiveness, spending level, potential for cost savings, and risk of poor outcomes.

Make better use of administrative data—information typically collected for operational and compliance purposes—to enhance program evaluations

Over the past decade, researchers have made significant advances in using existing data sources to conduct rigorous program evaluations, for example, linking education, child welfare, and juvenile and criminal justice records to determine child outcomes.[42] Because much of this information is already collected for other administrative purposes, the costs are much lower than more traditional program evaluations.

For example, Hawaii's Opportunity Probation with Enforcement program, a supervision program for offenders at high risk for probation violation, was evaluated in a randomized controlled trial using existing administrative data sources. The state's existing probation case-management system included records on supervision activities, drug test results, offenses, and other probationer interactions with the criminal justice system, and the Criminal Justice Information System provided comprehensive criminal record data. By linking these data sources, the evaluation was able to determine that the program was effective in reducing recidivism. Participants were 55 percent less likely to be rearrested and 53 percent less likely to have their probation revoked compared with high-risk offenders who did not participate in the program.[43]

Require evaluations as a condition for continued funding for new initiatives

Governments frequently operate small-scale programs as a way to test innovations before fully implementing them. When designing these programs, governments should specify the desired results to help managers and evaluators focus on specific objectives, and before financial support is renewed, outcome studies should be required to determine whether tested programs are effective.

In New York City, the Center for Economic Opportunity requires rigorous evaluations of all pilot programs to determine whether they were effective in achieving one or more of three primary goals: reducing poverty, encouraging savings, or empowering low-income workers to advance their careers. Center staff oversee monitoring and evaluation activities, working in partnership with city agencies and external research organizations. The center uses the results to help determine whether to expand or discontinue each program.[44]

Develop a centralized repository for program evaluations

As noted earlier, several national research clearinghouses are reviewing studies to identify what works in public programming across policy areas. Governments can support these efforts by designating a central entity to house the studies they conduct and requiring all agencies to submit copies of outcome evaluations and performance audits. This agency or unit should screen the reports, identify significant outcome findings, and incorporate the information into a comprehensive list of local programs. Governments can also report these studies to the national research clearinghouses to help expand the available knowledge base and help governments across the country more effectively direct funding to programs that have demonstrated strong results for residents.

Getty Images/ TC Lin

Conclusion

Government leaders are increasingly using rigorous evidence to identify policies and programs that work and are cost-effective. To date, however, policymakers had no comprehensive road map to guide them in this endeavor. The framework presented in this report identifies the steps that all levels and branches of government can take to build and support a system of evidence-based policymaking for strategically selecting, funding, operating, monitoring, and evaluating public programs that deliver the best returns on taxpayer investments.

Appendix A: Methodology

Developing a framework for a system of evidence-based policymaking required a two-step approach. Results First staff began by reviewing extensive academic research on systems that support evidence-based policymaking. Second, we conducted 46 interviews with academics, practitioners, and government experts to discuss their research and experiences with this approach and used the information to identify activities governments should undertake to establish and sustain a system of evidence-based policymaking. An external panel of experts in this area reviewed our findings and provided valuable input on the key components and the overall report.

Appendix B: Potential roles in state government

Key steps	Governor's office	Agency or program leadership	Legislature
Program assessment: Systematically review available evidence on the effectiveness of public programs.			
Develop an inventory of funded programs	Issue an executive order requiring agencies to develop program inventories	Lead the agency through the inventory process	Enact legislation requiring a program inventory
Categorize programs by their evidence of effectiveness	Create a workgroup to lead the development of research standards	Lead the agency through the categorization process	Enact legislation establishing criteria for the levels of research rigor
Identify programs' potential return on investment	Require agency budget requests to include cost-benefit information when practicable	Conduct or contract out cost-benefit analyses	Direct legislative research staff to conduct cost-benefit analyses
Budget development: Incorporate evidence of program effectiveness into budget and policy decisions, giving funding priority to those that deliver a high return on investment of public funds.			
Integrate program performance information into the budget development process	Create standard operating procedures and formats for agencies to report performance data in budget requests	Develop performance measures and benchmarks	Direct research and agency staff to develop performance measures and benchmarks
Present information to policymakers in user-friendly formats that facilitate decision-making	Require executive branch agencies to develop *Consumer Reports*-type summaries	Develop *Consumer Reports*-type summaries	Support the development and use of *Consumer Reports*-type summaries
Include relevant studies in budget hearings and committee meetings	Include relevant studies in budget hearings and committee meetings; direct agencies to support budget requests with evidence of outcomes	Simplify evidence-based requests to include clear, concise, and verifiable information about program results	Require agencies to regularly report on program outcomes and evaluations and to use a standard format for reports
Establish incentives for implementing evidence-based programs and practices	Set aside funding for competitive grants	Administer grant competitions	Set aside funding in budgets for grant competitions
Build performance requirements into grants and contracts	Require that performance measures be incorporated in contracts where practicable	Work with contracted providers to develop common outcomes for reports and provide training	

Key steps	Governor's office	Agency or program leadership	Legislature
Implementation oversight: Ensure that programs are effectively delivered and are faithful to their intended design.			
Establish quality standards to govern program implementation	Direct agencies to develop statewide standards	Meet with contract providers to gather input on standards	Incorporate standards into statute
Build and maintain capacity for ongoing quality improvement and monitoring of fidelity to program design	Emphasize the importance of building internal capacity to faithfully implement evidence-based programs	Provide training and technical support to contract providers and local governments charged with implementing programs	Provide resources for training and technical support
Balance program fidelity requirements with local needs		Hold regular meetings to review implementation practices and gather feedback from providers	
Conduct data-driven reviews to improve program performance	Develop a structure for program review meetings, define the roles of the participants, and provide leadership and support	Analyze data and provide it for the meetings	
Outcome monitoring: Routinely measure and report outcome data to determine whether programs are achieving desired results.			
Develop meaningful outcome measures for programs, agencies, and the community	Provide leadership, emphasizing the importance of measuring outcomes; create workgroups to guide the process	Establish consistent processes to review measures and goals	Provide input on performance measurement process to increase its usefulness to decision-makers
Conduct regular audits of systems for collecting and reporting performance data	Create work groups to guide the process	Establish systems for collection and validation of data; administer the process	Direct legislative staff to participate in work groups
Regularly report performance data to policymakers	Develop a standardized format for reporting outcome data (e.g., report cards, data dashboards)	Collect data and develop reports (e.g., report cards, data dashboards)	Request performance data at relevant committee hearings

Key steps	Governor's office	Agency or program leadership	Legislature
Targeted evaluation: Conduct rigorous evaluations of new and untested programs to ensure that they warrant continued funding.			
Leverage available to conduct evaluations	Create a work group to study resources available in the executive branch	Provide information to the work group	Create a work group to study resources available in the legislative branch
Target evaluations to high-priority programs	Develop criteria for prioritizing programs and provide funding for them in program budgets	Create a prioritized list within each agency	Prioritize program evaluation for legislative, fiscal, or research offices
Make better use of administrative data—information typically collected for operational and compliance purposes—to enhance program evaluations	Facilitate data-sharing among agencies	Identify university or other partners with experience in using administrative data for evaluations	Facilitate data-sharing among agencies
Require evaluations as a condition for continued funding of new initiatives	Review evaluations in the budget development process	Administer and monitor test projects	Enact legislation or include language in appropriations act requiring the evaluation of test projects, and review evaluations in the budget review process
Develop a centralized repository for program evaluations	Work jointly with the legislature to identify and staff the central repository	Contribute data and analysis to repository	Work jointly with the executive branch to identify and staff the central repository

Note: The roles of each branch of government described in this table may differ by state depending on laws and budget rules.

Endnotes
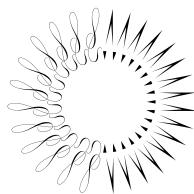
1   Although there are several classification systems that rank the strength of evidence on program outcomes, our framework uses four categories that consider the rigor of research methods used and the amount of evidence available:

• Evidence-based programs and practices have been evaluated multiple times and found to be effective using rigorous methods such as randomized controlled trials, statistically controlled evaluations, or a single large multisite randomized or statistically controlled evaluation. Typically, these programs have specified a set of procedures that allow for successful replication.

• Research-based programs or practices have been tested using rigorous methods (usually a single randomized control study or multiple studies that use strong comparison group designs) but do not meet the evidence-based standard. These programs typically have specified a set of procedures that allow for successful replication.

• Promising programs and practices have been tested using less rigorous research designs that do not meet the research-based standard. These programs and practices typically have a well-constructed logic model or theory of change.

• Non-evidence-based programs and practices lack sufficient evidence to meet the promising standard.

Three of the four categories—evidence-based, research-based, and promising—are based on standards initially developed in Washington state. See University of Washington, Washington State Institute for Public Policy, "Inventory of Evidence-Based, Research-Based, and Promising Practices for Prevention and Intervention Services for Children and Juveniles in the Child Welfare, Juvenile Justice, and Mental Health Systems" (Sept. 30, 2012), http://www.wsipp.wa.gov/ReportFile/1332/Wsipp_Inventory-of-Evidence-Based-Research-Based-and-Promising-Practices_Full-Report.pdf.

2   Yilin Hou et al., "States Performance-Based Budgeting in Boom and Bust Years: An Analytic Framework and Survey of the States," *Public Administration Review* (May/June 2011), 71(3): 370-388.

3   The Pew Charitable Trusts, "State Tax Revenue Grows, but a Full Recovery Eludes 26 States" (May 19, 2014), http://www.pewtrusts.org/en/research-and-analysis/analysis/2014/05/19/state-tax-revenue-grows-but-a-full-recovery-eludes-26-states.

4   The Pew Charitable Trusts, "The Fiscal Health of State Pension Plans Funding Gap Continues to Grow," (April, 8, 2014), http://www.pewtrusts.org/en/research-and-analysis/analysis/2014/04/08/the-fiscal-health-of-state-pension-plans-funding-gap-continues-to-grow.

5   Examples of clearinghouses include: Blueprints for Healthy Youth Development, *University of Colorado Boulder Institute of Behavioral Science Center for the Study and Prevention of Violence*, http://www.blueprintsprograms.com; Campbell Systematic Reviews, the Campbell Collaboration; "What Works in Social Policy?" *Coalition for Evidence-Based Policy*, http://evidencebasedprograms.org; Cochrane Reviews, *the Cochrane Collaboration*, http://www.cochrane.org/cochrane-reviews; Crimesolutions.gov, Office of Justice Programs, *U.S. Department of Justice*, https://www.crimesolutions.gov/Programs.aspx; Evidence-Based Home Visiting Service Delivery Models, Health Resources and Services Administration, *U.S. Department of Health and Human Services*, http://mchb.hrsa.gov/programs/homevisiting/models.html; SAMHSA's National Registry of Evidence-Based Programs and Practices, Substance Abuse and Mental Health Services Administration, *U.S. Department of Health and Human Resources*, http://www.nrepp.samhsa.gov; Teen Pregnancy Prevention Resource Center, Office of Adolescent Health, *U.S. Department of Health and Human Services*, http://www.hhs.gov/ash/oah/oah-initiatives/teen_pregnancy/db/tpp-searchable.html; and What Works Clearinghouse, *Institute of Education Sciences, U.S. Department of Education*, http://ies.ed.gov/ncee/wwc/findwhatworks.aspx.

6   Ron Haskins and Jon Baron, "Building the Connection Between Policy and Evidence: The Obama Evidence-Based Initiatives," NESTA (September 2011), http://coalition4evidence.org/wp-content/uploads/2011/09/Haskins-Baron-paper-on-fed-evid-based-initiatives-2011.pdf.

7   This number is based on total funding for the seven federal grant programs listed in "A Guide to Evidence and Innovation," Interagency Working Group on Youth Programs, accessed July 21, 2014, http://evidence-innovation.findyouthinfo.gov/investingEvidence.

8   The U.S. Office of Management and Budget, "Next Steps in the Evidence and Innovation Agenda" (letter dated July 26, 2013), http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-17.pdf.

9   U.S. Department of Health and Human Services, Health Resources and Services Administration, "The Maternal, Infant, and Early Childhood Home Visiting Program," http://mchb.hrsa.gov/programs/homevisiting/; Health Resources and Services Administration, "Affordable Care Act Maternal, Infant, and Early Childhood Home Visiting Program: (2011) 10, http://www.hrsa.gov/grants/manage/homevisiting/sir02082011.pdf; and U.S. Department of of Labor, Employment and Training Administration, "The Workforce Innovation Fund," last modified March 26, 2014, http://www.doleta.gov/workforce_innovation.

10   Results First researchers used LexisNexis to identify all legislation passed between 2004 and 2014 that included the following phrases: "evidence-based policymaking" (or "policy making"), "evidence-based decision making," "evidence-based program(s)," "evidence-based

practice," "research-based decision making," "research-based practice," "research-based program(s)." Some laws were omitted from the final list because they did not focus specifically on evidence-based policymaking. Researchers also reviewed laws that did not meet these criteria but were identified in previous work as supporting evidence-based policymaking.

11 The Pew-MacArthur Results First Initiative, *States' Use of Cost-Benefit Analysis: Improving Results for Taxpayers* (July 2013), http://www.pewtrusts.org/~/media/legacy/uploadedfiles/pcs_assets/2013/PewResultsFirst50statereportpdf.pdf.

12 Rhode Island Office of Management and Budget, "Results First—Adult & Juvenile Justice Program Inventory" (March 2014), http://www.omb.ri.gov/documents/performance/results-first/Results%20First%20Program%20Inventory%20March%202014.pdf.

13 H.B. 2536, 2012 Regular Session, Washington State Legislature, http://lawfilesext.leg.wa.gov/biennium/2011-12/Pdf/Bills/House%20Passed%20Legislature/2536-S2.PL.pdf.

14 The four agencies are the state's departments of Corrections, Health, Education, and Transportation; H.B. 677, 2014 Regular Session, Mississippi Legislature, http://billstatus.ls.state.ms.us/documents/2014/html/HB/0600-0699/HB0677SG.htm.

15 The Pew-MacArthur Results First Initiative regularly tracks budget allocations that were informed by each state's Results First model analyses. Results First regularly publishes the total allocations of Results First partner states, which go through a strict verification process. Total allocations have increased since our most recent publication, "Achieving Success With the Pew-MacArthur Results First Initiative: A State Progress Report 2011-13" (February 2014), http://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2014/02/05/achieving-success-with-the-pewmacarthur-results-first-initiative, from $38 million to $81 million. This figure and any additional 2014 allocations will be published in early 2015.

16 Pew-MacArthur Results First Initiative interview with District 43 Representative Diana Urban, Connecticut General Assembly, April 7, 2014.

17 Although cost-benefit analysis is one important tool that can help policymakers determine the best way to allocate scarce resources, the cost-benefit ratio produced should not be the only data used in making policy decisions. For certain programs, such as those concerning residents' health and safety, policymakers may opt to maintain a program whose benefits are only slightly greater than or are less than the costs.

18 Iowa Department of Corrections, "Return on Investment: Evidence-Based Options to Improve Outcomes" (May 2012), http://www.doc.state.ia.us/Research/DOC_HandoutROI_OffenderPrograms.pdf; New Mexico Legislative Finance Committee, "Evidence-Based Programs to Reduce Recidivism and Improve Public Safety in Adult Corrections" (July 2013), http://www.nmlegis.gov/lcs/handouts/CCJ%20072213%20Item%201%20LFC%20Results%20First%20Brief.pdf; Washington State Institute of Public Policy, "Prison, Police, and Programs: Evidence-Based Options That Reduce Crime and Save Money" (November 2013), http://www.wsipp.wa.gov/ReportFile/1396/Wsipp_Prison-Police-and-Programs-Evidence-Based-Options-that-Reduce-Crime-and-Save-Money_Full-Report.pdf.

19 Iowa Department of Corrections, "Return on Investment. Evidence-Based Options to Improve Outcomes" (May 2012), http://www.doc.state.ia.us/Research/DOC_HandoutROI_OffenderPrograms.pdf.

20 Ibid., 5.

21 The Pew-MacArthur Results First Initiative: "New Mexico's Evidence-based Approach to Better Governance" (August 2014), http://www.pewtrusts.org/~/media/Assets/2014/08/NM_Results_First_Brief_web.pdf.

22 Kit R. Van Stelle, Janae Goodrich, and Stephanie Kroll, "Treatment Alternatives and Diversion (TAD) Program: Participant Outcome Evaluation and Cost Benefit Report (2007-2013)" (July 2014), http://uwphi.pophealth.wisc.edu/about/staff/van-stelle-kit/tad-2014-outcomes-report.pdf.

23 Center for Employment Opportunities, Social Finance US, and the State of New York, "Investing in What Works: 'Pay for Success' in New York State—Increasing Employment and Improving Public Safety Detailed Project Summary," accessed Oct. 5, 2014, http://www.budget.ny.gov/contract/ICPFS/PFSProjectSummary_0314.pdf.

24 Connecticut Judicial Branch, Court Support Services Division, "Connecticut's Probation Risk Reduction Program: A Blueprint for Evaluation and Increased Effectiveness of Community Based Services" (spring 2002), 6, http://www.jud.state.ct.us/external/news/SpringSanctions.pdf.

25 Pew-MacArthur Results First Initiative interview with Cynthia Theran, assistant director of programs and services, Connecticut Court Support Services Division (Jan. 16, 2014).

26 The Center for Foster Care and Adoption Data, "Performance-Based Contracting in Tennessee's Foster Care System," https://fcda.chapinhall.org/wp-content/uploads/2012/10/2012_TN-PBC-case-hx2.pdf.

27 Washington State Institute of Public Policy, "Recommended Quality Control Standards: Washington State Research-Based Juvenile Offender Programs" (December 2003), http://www.wsipp.wa.gov/ReportFile/849.

28 Washington State Institute for Public Policy, "Outcome Evaluation of Washington State's Research-Based Programs for Juvenile Offenders" (January 2004), http://www.wsipp.wa.gov/ReportFile/852/Wsipp_Outcome-Evaluation-of-Washington-States-Research-Based-Programs-for-Juvenile-Offenders_Full-Report.pdf.

29 The EPISCenter is a project of the Prevention Research Center at Pennsylvania State University, jointly funded by the Pennsylvania Commission on Crime and Delinquency (PCCD) and the Pennsylvania Department of Public Welfare as part of PCCD's Resource Center for Evidence-Based Practice.

30 Pew-MacArthur Results First Initiative interview with Brian Bumbarger, director, Pennsylvania State University EPISCenter, and assistant director for knowledge translation and dissemination, Prevention Research Center (Oct. 29, 2013).

31 H.B. 2144, 2009 Regular Session, Oregon State Legislature, https://olis.leg.state.or.us/LIZ/2009R1/Measures/Text/HB2144/Enrolled.

32 Pew-MacArthur Results First Initiative interview with William Baney, director, Systems of Care Institute, Portland State University, March 17, 2014.

33 The data-driven performance management meeting approach was initially developed by the New York City Police Department in the 1990s. Commonly referred to as CompStat or PerformanceStat, this format has been adopted by other state and local governments, notably Baltimore's CitiStat and Maryland's StateStat. See *A Guide to Data-Driven Performance Reviews*, http://www.urban.org/UploadedPDF/1001559-Data-Driven-Performance-Reviews.pdf.

34 Robert Behn, What All Mayors Would Like to Know About Baltimore's CitiStat Performance Strategy" (2007), http://www.businessofgovernment.org/sites/default/files/Behn_BOG.pdf.

35 "Maryland StateStat," accessed July 21, 2014, http://www.statestat.maryland.gov.

36 "Virginia Performs," accessed July 21, 2014, http://vaperforms.virginia.gov.

37 Other members of the Council on Virginia's Future include the lieutenant governor, as well as eight members of the Virginia General Assembly, including the speaker of the House, the president of the Senate, chairs from both House and Senate finance committees, the majority and minority leaders from each house, and cabinet and citizen appointments made by legislators and the governor.

38 Healthy Housing Solutions Inc., "HUD Healthy Communities Transformation Initiative (HCTI): National Advisory Panel Review of the Healthy Communities Index 'Straw Man'" (May 3, 2013), http://www.housingnm.org/files/Library/hud_healthy_communities_index_may2013.pdf.

39 Pew-MacArthur Results First Initiative interview with Nicole Edmonson, director of performance audit services, Louisiana Legislative Auditor and Karen LeBlanc, performance audit manager, Louisiana Legislative Auditor, Dec. 16, 2013.

40 "Michigan Dashboard," accessed July 21, 2014, http://www.michigan.gov/midashboard.

41 "Michigan Education Dashboard," accessed June 27, 2014, http://www.michigan.gov/midashboard/0,4624,7-256-58084---,00.html.

42 See, for example, *Actionable Intelligence for Social Policy*, http://www.aisp.upenn.edu/about-us/, and Coalition for Evidence-Based Policy, "Demonstrating How Low-Cost Randomized Controlled Trials Can Drive Effective Social Spending: Project Overview and Update" (Dec. 13, 2013), http://coalition4evidence.org/low-cost-rct-competition.

43 Angela Hawken and Mark Kleiman, "Managing Drug Involved Probationers With Swift and Certain Sanctions: Evaluating Hawaii's HOPE" (December 2009), https://www.ncjrs.gov/pdffiles1/nij/grants/229023.pdf.

44 NYC Center for Economic Opportunity, "Evaluation Reports," accessed May 19, 2014, http://www.nyc.gov/html/ceo/html/data/reports.shtml.

PEW

THE

CHARITABLE TRUSTS

pewtrusts.org

MacArthur Foundation

macfound.org