Principles of Evidence-Based Policymaking

September 2016

Evidence-Based Policymaking Collaborative

The potential for a golden era of evidence-based policymaking has never been greater. The increased availability of data, new investments in rigorous research, and a growing public focus on results have given policymakers expanded abilities to determine whether public investments achieve their intended outcomes and to make informed choices based on evaluations of social costs and benefits. We believe that policy and practice rooted in evidence can lead to better use of taxpayer dollars and better outcomes for children, adults, and families.

This brief describes four principles of evidence-based policymaking that policymakers, agency heads, and other public leaders can use to improve results in the public sector. These principles represent the consensus of researchers and practitioners from different backgrounds who are interested in improving how policy decisions are made and how programs are managed. Our democratic process sets goals for policies and programs, and evidence-based policymaking is an important tool to help achieve those goals.

These principles can serve as a foundation for a common policymaking framework. Although policymakers at different points on the ideological spectrum disagree on the proper size and scope of government services, all agree that the services government *does* provide should work well. An evidence-based policymaking framework may ultimately promote bipartisan collaboration in important policy areas.

What is evidence-based policymaking and why is it important?

Evidence-based policymaking has two goals: *to use what we already know* from program evaluation to make policy decisions and *to build more knowledge* to better inform future decisions. This approach prioritizes rigorous research findings, data, analytics, and evaluation of new innovations above anecdotes, ideology, marketing, and inertia around the status quo.

Evidence-based policymaking can take many forms: using research findings to inform new policies or improve effectiveness of existing programs, supporting data collection and analysis for research and management, developing policies that incentivize the use of evidence, and so on. It has most frequently been applied to social and human services programs, but a wide variety of government programs could benefit from building and using evidence.

The movement toward evidence-based policymaking has seen significant progress in recent years at the federal, state, and local levels. But these efforts are still nascent, and many government actors could strengthen their use of evidence-based approaches. Some public agencies lack the capacity, skills, or funding to effectively use and build evidence; others lack the commitment from agency leaders to thoroughly integrate evidence into decisionmaking. We are encouraged, however, by the growing number of agencies that, with the support of political leaders and career public servants, are building a results-focused culture of learning and continuous improvement.

In an era of intense partisanship and constrained public resources, evidence-based policymaking can help bridge the partisan political divide and support research-based debate about what outcomes we

want to achieve, for whom, and at what cost. It encourages transparency and accountability by clearly stating the goals of policies and programs and then independently evaluating their results to see if those goals were achieved. By centering on outcomes, an evidence-based framework focuses policymaking on effectiveness of social interventions and efficiency in use of resources, an approach that greatly increases the chances of bipartisan agreement.

Evidence-based policymaking also encourages a virtuous cycle of knowledge building. By evaluating policies and programs and using program data, we can learn how well programs are working. We can then use that information to improve programs or to terminate consistently ineffective programs and find better approaches. From there, the cycle of learning and improving continues.

Principles of evidence-based policymaking

These principles are guidelines that apply to policymaking at every level of government and reflect a consensus among researchers across the ideological spectrum. They create a set of core ideas that a range of stakeholders can champion to promote evidence-based policymaking. They also establish a foundation for the work of the Evidence-Based Policymaking Collaborative. In the pages that follow, we will define each principle and its component elements.

First Principles of Evidence-Based Policymaking

1. Build and compile rigorous evidence about what works, including costs and benefits.

2. Monitor program delivery and use impact evaluation to measure program effectiveness.

3. Use rigorous evidence to improve programs, scale what works, and redirect funds away from consistently ineffective programs.

4. Encourage innovation and test new approaches.

Principle One: Build and compile rigorous evidence about what works, including costs and benefits.

DEFINING KEY TERMS

Evidence	• The available body of facts and other information indicating whether a belief or proposition is true or valid,[1] in this case regarding the impacts of programs
Rigorous evidence about what works	• The evidence produced by rigorous evaluations, such as randomized controlled trials and well-designed quasi-experimental studies, that measure program impact • Impact is the net effect of a program relative to the status quo[2]
Build evidence	• To conduct research and evaluations to identify social programs and policies that produce desirable outcomes (e.g., educational skills attainment, improved health, economic mobility, and economic well-being) by employing a range of rigorous methodologies
Compile evidence	• To assemble evaluations—in "what works" clearinghouses, for example—to clearly describe programs, the methodology by which they were evaluated, and their effects
Costs and benefits	• To analyze how much a program or policy costs per outcome achieved (cost analysis) and, in some cases, also measure the monetary value of benefits produced by the program so that costs and benefits can be directly compared (cost-benefit analysis)[3]

WHY IS THIS PRINCIPLE IMPORTANT?

To make the best use of limited public resources and deliver better outcomes for taxpayers, it is necessary to *build evidence* about what works through research and evaluation. Many government-funded programs have not been rigorously evaluated for outcomes or effectiveness. The absence of rigorous evaluation does not necessarily mean that these programs are ineffective, but it does mean we are in the dark about their true impact.[4]

Subjecting programs and policies to *rigorous* evaluation is the best way to build evidence about what works. Programs that have undergone rigorous evaluation contribute to an evidence base and help us

---

[1] *Oxford Dictionaries Online*, s.v. "evidence," accessed August 26, 2016, http://www.oxforddictionaries.com/us/definition/american_english/evidence.
[2] Peter Tatian, "Performance Measurement to Evaluation" (Washington, DC: Urban Institute, 2016), http://urbn.is/2bY6Ch3.
[3] "Benefit-Cost Results," Washington State Institute for Public Policy, last updated June 2016, http://www.wsipp.wa.gov/BenefitCost.
[4] Justin Milner and Matthew Eldridge, "From Evidence to Outcomes: Using Evidence to Inform Pay for Success Project Design" (Washington, DC: Urban Institute, 2016), http://urbn.is/2bY8Ddh.

estimate the likely effects of scaling up programs. As with all areas of scientific inquiry, this evidence base is not static and is updated and expanded as new evidence is created.

Several other elements are critical to the evidence-building process. One is independence: evaluators must be impartial to ensure the validity and rigor of new evidence. Agency leaders can help evaluation staff identify the most relevant questions to examine, but the evaluation itself and the publication of results must be free of influence from program or political leadership that might have a vested interest in a particular outcome. Privacy must also be protected when using or sharing data, particularly administrative and statistical data containing personally identifiable information.

Moreover, it is also important to *compile evidence* in ways that can be used by policymakers and practitioners in the field to select and implement programs proven effective. Online "clearinghouses" of research help users distinguish between programs with stronger or weaker evidence. Examples include the US Department of Education's What Works Clearinghouse and the University of Colorado Boulder's Blueprints for Healthy Youth Development. The Pew-MacArthur Results First Initiative has compiled data from eight research clearinghouses into a single online resource, the Results First Clearinghouse Database, which summarizes evidence on a range of social policy interventions.

The building of an evidence base should focus not only on determining what works, but also on measuring the costs and benefits of interventions. Cost analysis measures the monetary costs of interventions per outcome achieved. It allows different interventions with the same goal to be directly compared in terms of cost efficiency. Another tool is cost-benefit analysis, which measures both the costs and benefits of programs to calculate a return on investment. Cost-benefit analysis can help policymakers compare the relative benefits of spending on programs with the same general goals (e.g., boosting school readiness) or programs with different goals (e.g., boosting school readiness versus reducing delinquency). Cost-benefit analysis models, such as the one developed by the Washington State Institute for Public Policy, can help governments do more with limited resources.

Principle Two: Monitor program delivery and use impact evaluation to measure program effectiveness.

DEFINING KEY TERMS

Monitor program delivery	• To clearly define the key components of a program model and track the inputs, activities, outputs, and outcomes of program service delivery through process evaluation and performance management
	• To check whether services are delivered as intended, in terms of quantity and quality, and whether programs are meeting their goals
Use impact evaluation to measure program effectiveness	• The systematic collection of information about a program to identify (or estimate) the specific contribution of that program to intended outcomes
	• That specific contribution, in the language of evaluators, is known as a program's *impact*[5]

WHY IS THIS PRINCIPLE IMPORTANT?

There are two key elements to this principle: monitoring a program as it is delivered and evaluating the success of the program. Both steps are important for building knowledge about what works and acting on that knowledge (as discussed in principle three).

Monitoring program delivery ensures basic accountability by making sure that programs operate as intended. Monitoring programs is also critical for identifying problems or opportunities for improvement in program delivery so that programs can continually improve. In cases where programs are replicating established models—such as a job training program in Minneapolis that aims to use the same approach as one in Saint Paul—program monitoring can help ensure fidelity to the model by tracking how well the new program is implementing the core characteristics or intentions of the original.[6]

Impact evaluation, on the other hand, aims to answer a specific question: how are things different because of this program relative to the status quo that would have existed without it? Rigorous impact evaluation provides policymakers and program managers with hard evidence about what a program actually achieves or which versions of a program are more effective. It also helps ensure that public investments go beyond simply paying for activities and places emphasis on better results and outcomes for society.

---

[5] Ibid.
[6] Pew-MacArthur Results First Initiative, "Implementation Oversight for Evidence-Based Programs: A policymaker's guide to effective program delivery" (Washington, DC: Pew-MacArthur Results First Initiative, 2016), http://www.pewtrusts.org/~/media/assets/2016/05/rf_programimplementationbrief.pdf.

One useful strategy to reduce cost barriers to rigorous evaluation is to conduct low-cost randomized controlled trials (RCTs). RCTs are widely recognized as the gold standard for impact evaluations, but they can be expensive because of the cost of data collection. Low-cost RCTs reduce or eliminate the need for data collection by using existing program data (e.g., test scores, arrest records, and earnings records) already being collected. These administrative data, if they are high quality, can be repurposed to conduct low-cost RCTs at a fraction of the traditional expense. They can also potentially increase evaluation quality by providing a more accurate or complete picture of program outcomes than survey results. For all these reasons, low-cost RCTs have the potential to make RCTs the norm for program evaluation.[7] A competition funded by the Laura and John Arnold Foundation and the Annie E. Casey Foundation has already supported low-cost RCTs in education, health care, workplace safety, and other issue areas.[8]

Principle Three: Use rigorous evidence to improve programs, scale what works, and redirect funds away from consistently ineffective programs

DEFINING KEY TERMS

Improve programs	• To adapt a program so that it produces better results, particularly around participant outcomes
Scale what works	• To take a program that has worked for one population and implement it in a larger population, a different population, or a population in a different context
Redirect funds	• To defund a program that has failed to achieve desirable outcomes in favor of a program that is more effective or is likely to be more effective

WHY IS THIS PRINCIPLE IMPORTANT?

Program evaluation is too often thought of as a "thumbs up or thumbs down" determination—does this program work or not? In reality, evaluation can also be an important tool for program *improvement*. For example, impact evaluation can test different approaches within the same program, such as when the National Evaluation of Welfare-to-Work Strategies compared the effects of two available strategies for

---

[7] Maya Shankar, "How Low-cost Randomized Controlled Trials Can Drive Effective Social Spending," *The White House Blog*, July 30, 2014, https://www.whitehouse.gov/blog/2014/07/30/how-low-cost-randomized-controlled-trials-can-drive-effective-social-spending.

[8] Laura and John Arnold Foundation, "Request for Proposals: Low-Cost Randomized Controlled Trials to Drive Effective Social Spending," (Houston, TX: Laura and John Arnold Foundation, 2015), http://www.arnoldfoundation.org/wp-content/uploads/Request-for-Proposals-Low-Cost-RCT-FINAL.pdf.

preparing welfare recipients for employment.[9] More broadly, findings from program evaluations can and should be used to refine program strategy and improve results.

Scaling programs is another way to apply impact evaluations and is an important way to broaden the reach of approaches backed by strong evidence. A program proven effective in one neighborhood, for example, might be scaled up across a city. If shown to be effective across the city, the program might be further scaled up across the state. Federal policymakers might also scale up evidence-based approaches across the nation. For example, the nationwide scaling of permanent supportive housing—a model with strong evidence—has helped drive a dramatic reduction in veterans experiencing homelessness.

Scaling programs is not easy, but it is an important tool in helping local and federal policymakers realize the potential of evidence-based policymaking. However, even as programs are being scaled up and implemented in new settings, it is important to include an evaluation component to continue building evidence. For example, the Teen Pregnancy Prevention Program and the Home Visiting Program at the US Department of Health and Human Services are helping to scale up evidence-based approaches to new jurisdictions while using program evaluation to learn from these expanded efforts. Lessons from these initiatives can then inform broader policy efforts in the future.

When the evidence base is strong enough to convincingly show that a program is ineffective, inefficient, or even harmful, policymakers should *redirect funds* away from that program. Policymakers might instead invest in programs with stronger supporting evidence or decide to test newer, innovative approaches. Just as it is important to do what works, it is also important to stop what is not working.

For the past few years, Results for America has published the *Federal Invest in What Works Index*. The *Index* tracks the progress of seven federal agencies in building the infrastructure necessary to use data, evidence, and evaluation in budget, policy, and management decisions. It rates agencies on the extent to which they use evidence in their largest grant programs, effectively apply data to their decisionmaking, and redirect funds away from ineffective programs. The *Index* can be a resource for any federal, state, or local agency that wants to monitor its own capacity to build and use evidence.

---

[9] Gayle Hamilton et al., *National Evaluation of Welfare-to-Work Strategies: How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs* (Washington, DC: US Department of Health and Human Services and US Department of Education, 2001), http://www.mdrc.org/sites/default/files/full_391.pdf.

Principle Four: Encourage innovation and test new approaches.

DEFINING KEY TERMS

Innovation	• Building on theory and existing research to develop new ways to address specific program or policy challenges
Test new approaches	• To determine with rigorous evaluation whether new approaches actually work and achieve desired outcomes

WHY IS THIS PRINCIPLE IMPORTANT?

Although a key goal of evidence-based policymaking is to focus resources on approaches backed by strong evidence, *only* using approaches backed by strong evidence prevents us from finding new and innovative ways to address the nation's challenges. That is why encouraging *innovation* should always be an important part of evidence-based policymaking. This is especially true in policy areas where the evidence base is thin and there is limited research to guide funding and programmatic decisions. Testing innovative approaches in these areas is an essential way to advance evidence-based policy.

For grant programs, a tiered-evidence design can be a useful strategy to spur innovation. Tiered-evidence grant programs include a dedicated funding tier for innovative and relatively untested approaches, often called the development or proof of concept tier. Funding in this tier supports both program activities and evaluation so that grantees can build evidence about their new approach. Grantees that have built more evidence and demonstrated program effectiveness can move up tiers and receive larger grants to support validation and scaling. The US Department of Education's Investing in Innovation Fund uses a tiered-evidence design to fund practices aimed at boosting student achievement and other education-related goals. The fund includes a development tier designed to establish an evidence base for untested programs. Between 2010 and 2015, it funded 105 development grants totaling over $350 million.[10]

Waiver demonstrations with evaluation requirements are another useful tool to advance both evidence-based policy and innovation. These demonstrations typically involve federal agencies allowing states or localities to waive certain federal laws or regulations to test new ways to serve program participants and advance community goals. In exchange, those jurisdictions must rigorously evaluate their approaches to determine if they work better than the status quo. This combination of innovation and learning can be a powerful tool for identifying new and more effective policies and programs.

Although testing program design choices is an explicit goal of waiver demonstrations, in other cases, unexpected local conditions can sometimes create opportunities for innovation. In these situations, promising variations to the program model should be analyzed. Rapid-cycle testing of the new program

---

[10] "Awards," Investing in Innovation Fund (i3). accessed August 26, 2016, http://www.ed.gov/programs/innovation/awards.html.

can use subgroups of program participants to test innovations to the model as they are implemented. Depending on evaluation results, the innovative program can be implemented with the full population or participants can be returned to the initial program.

Conclusion

Evidence-based policy is a critical resource as we strive to meet our nation's most important challenges while ensuring that public funds are used as effectively and efficiently as possible. It is a tool to help government learn what works. Of course, evidence and research will not be the only factor in policy decisions. Rather, the goal is to give evidence of program effectiveness a seat at the table when decisions are being made. The principles presented here are basic building blocks in that process. They can be applied at every level of government and at various points in the policymaking process, including the testing of new approaches, improvement of existing programs, identification of ineffective programs, and scaling of programs with strong evidence. Incorporating the principles of evidence-based policymaking into decisionmaking on a regular basis will improve the effectiveness of government programs and help solve the nation's social problems.