Behavioural Public Policy (2021), 5: 1, 90–102
© The Author(s) 2020. Published by Cambridge University Press doi:10.1017/bpp.2020.25

The realities of scaling within evidence-based policy

LAUREN H. SUPPLEE*
Child Trends, Bethesda, MD, USA
MAGGIE C. KANE
Child Trends, Bethesda, MD, USA

Abstract: The economic model for scaling described by Al-Ubaydli and colleagues offers recommendations to policymakers who make decisions about whether or not to implement evidence-based programs. The core economic model does not currently acknowledge the broader context of policy decision-making and therefore may fail to achieve its objectives. The model focuses primarily on the generation and use of available research in the decision on whether to scale a program. Research studying the use of evidence in policymaking points to a complex set of factors beyond just the strength of the evidence such as leadership, relationships, timing and financial resources that contribute to decisions to scale a program. Second, there is already a robust evidence-based policy movement at the federal, state and local levels. The economic model should leverage this movement rather than providing recommendations that might stall or redirect the movement. The economic model can push the field to strengthen the available evidence while providing recommendations on selecting models to scale within the currently available evidence. This commentary finishes with suggestions for moving forward.

Submitted 23 December 2019; accepted 31 March 2020

The economic model for scaling described in the Al-Ubaydli *et al.* paper (this issue) suggests a path toward studying the potential for evidence-based interventions to scale up to the broader population and offers recommendations for policymakers[1] who make decisions about whether or not to implement these programs. The core tenets of the economic model have the potential to

* Correspondence to: Child Trends, 7315 Wisconsin Ave #1200w, Bethesda, MD 20814, USA. E-mail: lauren.supplee@gmail.com

1 The term 'policymaker' here is used broadly to include federal, state and local legislative and executive branch staff.

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

move the field toward more effective design and testing of evidence-based programs; however, it falls short in acknowledging the broader context of policy decision-making using that evidence. The model's developers acknowledge that factors like political will fall outside of the current model. These factors, however, have implications for the feasibility of the recommendations in the model offered to policymakers.

This commentary will discuss two key challenges for the model. First, research studying the integration of evidence into policymaking points to a complex set of factors influencing implementation decisions, including leadership, maintaining relationships, windows where policies can advance, financial resources and availability and campaign promises (Kingdon, 1995; National Research Council, 2012; Oliver *et al.*, 2014; Cairney *et al.*, 2016; Bogenschneider, Corbett & Parrot, 2019a, 2019b; Nutley *et al.*, 2019). These factors complicate the model's recommendations, which rely primarily on making decisions about selecting and implementing evidence-based programs based on findings from research. In addition, the model implies that the policymaker will be the one weighing the strength of the evidence on a program or intervention. Policymakers, however, are not often researchers by training and therefore look to trusted internal or external partners to explain the evidence. Research on the use of research evidence points to the vital role of research–practice or –policy partnerships (RPPs) to integrate individuals trained in research methods with the policymakers (Tseng *et al.*, 2017).

Second, there is currently substantial political attention on evidence-based policies and programs at national and state levels. Funding that prioritized evidence-based programs began rolling out during the Bush administration, and since then, a range of evidence-based programs in fields from health to early childhood development have already been implemented at the local and state levels (Haskins & Baron, 2011). The model appropriately advocates for strengthening the available evidence base, but in the meantime, policymakers and practitioners must choose existing programs to fulfill these evidence-based mandates using the currently available research.

After highlighting these two key challenges, we offer suggestions for adapting the model's recommendations to fit the existing policy landscape.

What we know about evidence and the policy decision-making process

The economic model's proposals for policymakers implementing programs at scale include recommendations about post-study probability thresholds for adoption decisions, leveraging multi-site trials with an eye toward both internal and external validity and understanding an intervention's non-negotiables to know what must be implemented in order to maintain effects. While we

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

believe that these recommendations reflect best practices in research-based policymaking, we also know that the reality of the policy environment presents barriers to following these recommendations. Scientific decision-making is often described as a linear process, with decisions informed by evidence generated through hypothesis testing, but the points at which research influences policy decision-making can be harder to identify (Bogenschneider *et al.*, 2019a). Increasing the supply of quality research will not change the fact that policymakers are often weighing multiple factors when making decisions, including building or maintaining relationships, beliefs or confirmation bias about the topic at hand or the value of research and political or financial pressures (National Research Council, 2012; Oliver *et al.*, 2014; Cairney *et al.*, 2016; Boaz *et al.*, 2019; Bogenschneider *et al.*, 2019a). The four factors described below – financial drivers, the policymaking environment, policy windows and political promises – present challenges for following the recommendations outlined by the economic model for scaling.

Financial drivers

Budget constraints shape the implementation of policy and practice. In a large systematic review, Oliver and colleagues (2014) identified political and financial pressure as influencing whether policymakers supported evidence-based practices. Funding and policy mandates, for example, shape decisions about adopting evidence-based programs at the national and state agency levels (Nutley *et al.*, 2007; Aarons *et al.*, 2011). Budget constraints and policy mandates limit the resources available for generating the research, as well as the implementation of evidence-based projects and practices (Nutley *et al.*, 2007). In the face of budget constraints, policymakers may not be able to make decisions about programs based on the recommendations in the economic model. Effective programs, for example, may be too costly to implement with existing funds. In other cases, restrictions on the use of funds may limit policymakers' ability to allocate funding toward the replication studies that the model recommends.

Policymaking environment

In addition to funding constraints, the collaborative, coalition-building nature of the policymaking environment presents challenges for implementing the economic model's recommendations. To advance priorities, policymakers must develop and maintain strong relationships with others who will advocate for their proposals; they also need to make compromises to reach consensus (Bogenschneider *et al.*, 2019a). This is particularly true for interventions that address contentious policy topic areas, where politicization often sidetracks

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

evidence-based decision-making (Head, 2010). Research often gets used in the policymaking process to build support from others for policy priorities (Bogenschneider *et al.*, 2019b) in legislatures and in non-partisan policy settings. In the education sphere, for example, research by Penuel and colleagues (2018) found that school district leaders turned to research in part to foster buy-in for initiatives they were implementing. The model recommends that researchers publish and receive rewards for publishing null studies and policymakers use the body of evidence to make decisions about implementing based on the strength of the totality of the evidence. However, the existence of evidence does not ensure they will be used when deciding whether or not to implement a program.

Policy windows

While some policy is developed slowly, with coalitions building over time, other policy opportunities, or windows, open and close quickly (Haskins & Margolis, 2014). These windows often arise as a result of external factors, like changes in administration or external events that attract public and news media attention (Kingdon, 1995), and require policymakers to make decisions quickly regardless of the availability of evidence. In response to school shootings, for example, states and districts have implemented a wide range of different interventions, from arming teachers to increasing school security, to address the issue (Modan, 2019), despite a dearth of evidence about whether these policies are effective (Harper & Temkin, 2018). The economic model recommends waiting to implement programs until they have a sufficient body of research evidence pointing to their efficacy. In the real world, however, the timing of evidence may not coincide with the political opportunity to implement a program. Research timelines, which can extend across years or decades, are a key barrier for policymakers' use of research in decision-making (Nutley *et al.*, 2007; Oliver *et al.*, 2014). When the opportunity arises to implement a large-scale program that addresses a critical social need, such as in the wake of a tragedy like a school shooting, policymakers need to act. Policy windows do not stay open for long, and then can close due to a lack of action or a lack of alternatives to the status quo (Kingdon, 1995). The standard for evidence on interventions proposed by the economic model will require multiple replications that may take decades to build; in many policy areas, a strong research base to inform evidence-based policymaking does not yet exist (Head, 2010). While we agree that it would be ideal to have a robust body of evidence supporting the scale-up of interventions, there will likely be more situations in which emerging evidence will need to inform adaptations to existing large-scale programs rather than the implementation of new programs.

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

Political promises

A related challenge comes from the fact that policymakers face pressure to make short-term decisions that will fulfill promises made on the campaign trail or their party's political agenda (Head, 2010). Depending on their position, legislators in the USA hold office for two to six years before facing re-election, at which time they will need to demonstrate the progress they have made on issues they promised to address during their campaign. This pressure can affect the timeline for scaling programs as well as the types of programs that policymakers champion. In the context of scaling timelines, political pressure can cause politicians to scale-up programs rapidly rather than incrementally in order to fulfill promises. In Peru, for example, the government implemented a large-scale home visiting program in rural districts as part of a program for young children proposed by candidate Ollanta Humala during the country's 2011 presidential campaign cycle. The program was based on smaller-scale initiatives implemented in other countries, and while it reached tens of thousands of children, it faced challenges to quality caused by rapid expansion to meet political promises within Humala's five-year term (Araujo *et al.*, in press). Political pressure can also make policymakers hesitant to abandon programs they have championed in the face of evidence that the program is not as effective as originally assumed (Head, 2010). While policymakers do value research in their decision-making processes, they can also be hesitant to seek out evidence that contradicts their decisions (Bogenschneider *et al.*, 2019a). Even if researchers receive higher incentives to report null results in studies, as the economic model recommends, this will not guarantee that policymakers will seek out those results and walk back their efforts to push certain programs forward or modify existing programs.

Evidence-based policy movement

The economic model implies that policy should wait until the highest bar of evidence of multiple replicated studies that have representative populations and situations with economic analyses have been conducted prior to scaling up the program. It would take decades to meet this threshold, given that even some of the programs that are cited as having the highest evidence still do not meet this threshold. We are concerned that not only has the horse left the barn already as far as evidence-based policymaking is concerned, but pulling it back will be risky in its own sense. This may disempower all of the stakeholders who at this point see their role as being in the evidence-generation system and have embraced the importance of integrating evidence into their decision-making processes to a more passive role of waiting many years before research can help them.

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

The political environment has shifted over the last 15 years to prioritize generating evidence and funding programs that demonstrate some evidence of effectiveness. While increased interest in evidence-based practices has created opportunities to expand evidence-based interventions, it also creates challenges for waiting for the bar of evidence for programs recommended by the economic model. Beginning in the Bush Administration, there were some funding opportunities that emphasized the selection of evidence-based programs (i.e., Supporting Evidence-Based Home Visiting) (Haskins & Margolis, 2014). During the early years of the Obama Administration, this surge became a tsunami, with multiple federal social programs structured as tiered evidence programs, or programs where the majority of the funds were reserved for programs that had evidence of effectiveness (Haskins & Margolis, 2014). Programs such as the Department of Education's Investing in Innovation (i3; now called Education Innovation and Research) and Maternal Infant and Early Childhood Home Visiting (MIECHV) reserved the majority of grant funds for programs that had the highest-quality evidence, with smaller amounts of funds allocated to testing innovative ideas. Through a mix of Legislative and Executive Branch actions, evidence in policymaking became a central part of the conversation. The White House Office of Management and Budget began including a chapter in the guidance for federal agencies' budget requests that included an emphasis on evidence and building evidence (Office of Management and Budget, 2019). Congressional support for the central role of evidence has continued in bills such as the recent Family First Prevention Services Act (FFPSA) of 2017, which reserves the majority of the funds for child abuse-prevention programs that have undergone two impact studies in real-world settings.

States had their own similar evidence-based movement. Initially, the Washington State Institute for Public Policy led the movement by providing non-partisan assessments of every program that may receive state funding using the level of evidence available on that program and any economic analysis (Maki & Lieb, 2009). The legislature uses this information to inform decisions about investment. Pennsylvania made a smaller yet similar investment, requiring evidence-based youth substance abuse-prevention programs across the state (Pennsylvania Commission on Crime and Delinquency, 2019). Now another wave of actions in states is occurring, where most states are doing some investment in evidence-based policymaking (Pew-MacArthur Results First Initiative, 2017). States are defining levels of evidence in statutes to distinguish programs that have some evidence of effectiveness and directing the majority of funding to these programs. States are creating inventories of programs that meet these thresholds, using any available economic analysis

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

on the programs, and requiring programs to report outcome data to show progress (Pew-MacArthur Results First Initiative, 2017).

To support these evidence-based policy movements at the federal and state levels, there has been a proliferation of systematic reviews with the aim of assessing the quality of the available evidence and actively disseminating it to increase the accessibility of evidence (Supplee & Metz, 2015). These systematic reviews can provide powerful additional levers to forward many of the goals of the economic model. This is because these systematic reviews often influence the future creation of evidence by defining which research designs are the highest quality, which measures count as policy relevant, where federal or state funding is directed and the value of factors such as external validity in determining which models should be scaled (Avellar *et al.*, 2017).

We admit there have been challenges in implementing some of these tiered evidence programs. Many of the evidence-based home visiting programs funded under the MIECHV program, one of the flagship tiered evidence statutes, have not conducted replications or collected information on external validity (Supplee & McCann, 2019). The recent independent randomized evaluation of four of these models did find significant effects at scale, though these were reduced in magnitude (Michalopoulos *et al.*, 2019). In the hundreds of teen pregnancy-prevention impact study replications, few of the well-implemented studies found similar results (Feldman Farb & Margolis, 2016). More recently, the FFPSA reserves the majority of child welfare funding for models that have replications in real-world settings. There are few programs that will meet this threshold, and even fewer that have been tested with populations similar to those served by child welfare (McKlindon, 2019). These challenges in implementation led to a second bill, the Family First Transition and Support Act of 2019, which will slow down the implementation timeline to give states a longer on-ramp to the highest bar of evidence.

Given the challenges in scaling programs, we agree with the model that there is a large inherent risk in proceeding to scale up programs prior to generating this level of evidence, as we may invest millions of dollars in programs that may not actually achieve the promised outcomes at the magnitude everyone is hoping. There is an additional risk that policy and practice may therefore conclude that evidence-based policy as a concept does not work. Given the context we are in, proceeding with initiatives to scale while also expanding the evidence base for future scaling may not be mutually exclusive activities. The thresholds in the economic model are the long game, and we can begin to pursue this ideal while continuing to nurture the support for evidence at all levels of social service funding and implementation.

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

Recommendations

We believe that the recommendations in the economic model can be adopted within the existing evidence-based policy movement based on knowledge about the policy decision-making process. The evidence-based policy movement has already been embraced by policymakers, decreasing some of the challenges to the model posed by the decision-making process, and it provides a natural infrastructure and inherent incentives to leverage.

Expand research partnerships

While we agree that the economic model's recommendations reflect best practices in scaling implementation, the barriers listed above present challenges for fulfilling the model's recommendations. One approach to overcoming some of these barriers is using research partnerships to strengthen the relationships between researchers, policymakers and practitioners. RPPs are long-term, trusted partnerships between policymakers and researchers to identify research priorities, share and interpret findings and ensure the research created is useful to policy or practice. Research in the field of the use of evidence in policy points to the critical role of relationships between policymakers, practitioners and intermediary organizations in the integration of evidence in decision-making (Davies *et al.*, 2008). Intermediary organizations sit between the research community and policy and are seen as honest brokers of evidence, which is critical in the policymaker's use of evidence (Bogenschneider *et al.*, 2019a). They serve a critical function of translating and interpreting research for policymakers to advise on whether or not to move forward using particular evidence (Haskins & Margolis, 2014).

The utility and relevance of research can influence policymakers' decisions (Oliver *et al.*, 2014). Researchers must understand the policy decision-making process in order to create more useful research and to share it in non-partisan, unbiased ways (Oliver *et al.*, 2014). This not only points to the critical role of RPPs in the use of evidence, but RPPs will also help improve the quality of the evidence generated by ensuring that the questions posed for future studies are more aligned with the real-world constraints and need for information (Tseng *et al.*, 2017), which is encouraged by the economic model. RPPs can help researchers identify policy windows and the relationships needed to advance evidence-based practices when those windows open (Carter *et al.*, in press). Researchers can provide evidence to inform a particular policy window efficiently if they are already at the table as a trusted partner.

Raise the standard of evidence incrementally through systematic reviews

Many of the tiered evidence initiatives and state evidence movements point to independent systematic reviews to determine the quality of the impact evidence

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

and which programs meet these thresholds. This would allow the current framework of an 'on-ramp' for programs to build their evidence and increased confidence in the existing evidence in the higher tiers, while also pushing the field to address some of the critical gaps noted in the economic model. The statutes that support these initiatives often include an ongoing portfolio of research and evaluation. As more evidence is accumulated, the systematic reviews affiliated with these initiatives should increase the standards for evidence over time. Subsequent review standards can incorporate recommendations from the model, such as examining, reporting and accounting for post-study probabilities and increasing evidence of external validity. The broader open science movement can be leveraged to support the more careful examination of the impact evidence, including allowing the integration of multiple comparison adjustments or prioritizing confirmatory tests through the transparency of the pre-registration of studies. In addition, increased pre-registration will instill more transparency in all of the design decisions that happen throughout a study, allowing the field to provide comments at each stage and to better assess issues that influence evidence interpretation, such as whether research had been conducted but found non-significant findings and therefore was not published. If systematic reviews begin prioritizing (and eventually requiring) the studies they review to have pre-registration, many of the gaps in our current knowledge can begin to be addressed.

Prioritize real-world studies

Some of the current tiered evidence programs include language about more 'real-world' settings (e.g., FFPSA, Education Innovation and Research (EIR)), but do not go far enough to define what that means, nor do they include elements such as the representativeness of the intended program recipients, contexts and staff. This means regulations that cover these programs or their systematic reviews must prioritize or require statistics that compare the participants in the study to data relating to the larger intended populations and contexts. Systematic reviews can begin prioritizing studies that skip the efficacy trial stage and include more elements of a pragmatic trial, such as directly reporting and testing the program with the intended populations of participants, intended contexts and intended staff. This approach will both lead to study impacts that more closely approximate what may be expected at scale and generate economic evidence of costs in real-world contexts. Policymakers can use the concepts in the economic model to understand the properties of the population and situation that can help them to assess the relevance for a study to the context in which they are scaling a program.

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

Leverage new methods to improve existing models

We do not believe that we need to reserve all of these improvements in the generation of evidence for full-scale impact tests included in systematic reviews. We can leverage new methods to test and improve the models we do have now by conducting impact tests of active ingredients and asking questions about how to make models more efficient and effective. This work is being explored through empirical frameworks such as precision research methods (e.g., Home Visiting Applied Research Collaborative), which aim to examine what works best for whom in which contexts by exploring the relationship between core intervention components and outcomes in order to make evidence-based models more efficient and effective. The field currently has little to no empirical support for many of the factors in the economic model, such as which aspects of the model are non-negotiable to maintain impacts. However, there are research designs that can start to build this evidence and then push evidence-based programs to focus on those areas where we lack evidence, such as non-negotiables. In the phases of both designing new programs and refining existing programs, methods such as Multiphase Optimization Strategy (MOST) (Collins *et al.*, 2011; Broder-Fingert *et al.*, 2019) can add to the empirical evidence called for by the economic model. Programs already being implemented can use continuous quality-improvement methods, leveraging random assignment or testing enhancements to specific aspects of their programs to generate new knowledge. We can also leverage meta-analysis and other similar methods to unpack the empirical evidence on components of programs (e.g., Chorpita *et al.*, 2005; Lipsey, 2009). The results from these studies can further refine and strengthen evidence-based programs at scale in more real-world contexts.

Broaden the factors that researchers consider when designing studies for policymaking

We support the model's recommendation that researchers design studies with the parameters in mind that policymakers will need to make decisions on evidence. The research on the use of evidence in policy would suggest broadening the variables researchers need to consider to include not only properties of the population and situation, but also factors such as relationships, confirmation bias on the topic or utility of the research and political pressures. Some of these factors may be built into the study design, while others may come into play during dissemination of the findings. For all of the variables, conducting research within the context of a RPP will provide researchers with insights into the factors that influence policy decision-making and increase policymakers' trust in the research results. This may include researchers working

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

with advocates and other intermediaries that work closely with policymakers to support the dissemination and interpretation of the findings within an existing trusted relationship.

We believe tying funding decisions for federal and state funds to standards of evidence in partnership with their systematic reviews will incentivize the type of research called for in the model – more replicable and more relevant to real-world settings. We believe we can leverage the existing momentum to integrate the gaps identified in the economic world in order, over time, to improve the evidence available for it to be higher quality and more actionable in evidence-based policymaking.

References

Aarons, G. A., M. Hurlburt and S. M. Horwitz (2011), 'Advancing a Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors', *Administration and Policy in Mental Health*, **38**(1): 4–23. doi:https://doi.org/10.1007/s10488-010-0327-7

Araujo, M. C., M. Rubio-Codina and N. Schady (In press), 70 to 700 to 70,000: Lessons from the Jamaica Experiment. In J. List, D. Suskind and L. Supplee (eds), *The Scale-Up Effect: Why interventions lose impact at scale and what we can do about it*. Routledge.

Avellar, S. A., J. Thomas, R. Kleinman, E. Sama-Miller, S. E. Woodruff, R. Coughlin and T. R. Westbrook (2017), 'External Validity: The Next Step for Systematic Reviews?' *Evaluation Review*, **41**(4): 283- 325. https://doi.org/10.1177/0193841X16665199

Boaz, A., H. Davies, A. Fraser and S. Nutley (eds), (2019), *What Works Now? Evidence-Based Policy and Practice Revisited*. Bristol, UK: Policy Press.

Bogenschneider, K., T. J. Corbett and E. Parrot (2019a), 'Realizing the Promise of Research in Policymaking:', *Journal of Family Theory & Review*, **11**: 127–147. https://doi.org/10.1111/jftr.12310

Bogenschneider, K., E. Day and E. Parrot (2019b), 'Revisiting Theory on Research Use: Turning to Policymakers for Fresh Insights', *American Psychologist*, **74**(7): 778–793. https://doi.org/10.1037/amp0000460

Broder-Fingert, S., J. Kuhn, R.C. Sheldrick, *et al.* (2019), 'Using the Multiphase Optimization Strategy (MOST) framework to test intervention delivery strategies: a study protocol', *Trials* **20**: 728. doi:10.1186/s13063-019-3853-y

Cairney, P., K. Oliver and A. Wellstead (2016), 'To Bridge the Divide between Evidence and Policy: Reduce Ambiguity as Much as Uncertainty', *Public Administration Review*, **76**(3): 399–402. https://doi.org/10.1111/puar.12555

Carter, S., I. Dhaliwal, S. Friedlander and C. Walsh (In press), Forging Collaborations for Scale: Catalyzing Partnerships Among Policymakers, Practitioners, Researchers, Funders, and Evidence to Policy Organizations. In J. List, D. Suskind and L. Supplee (eds), *The Scale-up Effect in Early Childhood and Public Policy: Why interventions lose impact on scale and what we can do about it*. Routledge.

Chorpita, B.F., E.L. Daleiden & J.R. Weisz Ment Health Serv Res (2005), 7: 5. https://doi.org/10.1007/s11020-005-1962-6

Collins, L. M., T. B. Baker, R. J. Mermelstein, M. E. Piper, D. E. Jorenby, S. S. Smith, … M. C. Fiore (2011), 'The multiphase optimization strategy for engineering effective tobacco use interventions', *Annals of behavioral medicine: a publication of the Society of Behavioral Medicine*, **41**(2): 208–226. doi:10.1007/s12160-010-9253-x

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

Davies, H., S. Nutley and I. Walter (2008), 'Why 'knowledge transfer' is misconceived for applied social research', *Journal of Health Services Research & Policy*, **13**(3): 188–190. https://doi.org/10.1258/jhsrp.2008.008055

Feldman Farb, A and A. L. Margolis (2016), 'The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings', *American Journal of Public Health*, **106**: S9-S15. https://doi.org/10.2105/AJPH.2016.303367.

Harper, K. and D. Temkin (2018), New school safety bills could make schools less safe from kids. Retrieved from https://www.childtrends.org/new-school-safety-bills-could-make-schools-less-safe-for-kids

Haskins, R. and J. Baron (2011), *Building the Connection between Policy and Evidence: The Obama evidence-based initiatives*. NESTA. Retrieved from http://coalition4evidence.org/wp-content/uploads/2011/09/Haskins-Baron-paper-on-fed-evid-based-initiatives-2011.pdf

Haskins, R. and G. Margolis (2014), *Show Me The Evidence: Obama's Fight for Rigor and Results in Social Policy*. Washington, DC: Brookings Institution Press.

Head, B. W. (2010), 'Reconsidering evidence-based policy: Key issues and challenges', *Policy and Society*, **29**(2): 77–94. https://doi.org/10.1016/j.polsoc.2010.03.001

Kingdon, J. W. (1995), The Policy Window, and Joining the Streams. In *Agendas, Alternatives, and Public Policies* (Second ed., pp. 165–195). Addison-Wesley Educational Publishers.

Lipsey, M. W. (2009), 'The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview', *Victims & Offenders*, **4**(2): 124–147. https://doi.org/10.1080/15564880802612573

Maki, J. and R Lieb (2009), 'Washington State Institute for Public Policy: Origins and Governance, Olympia: Washington State Institute for Public Policy', *Document* No. 09-06-4101.

McKlindon, A (2019), Applying the research and evaluation provisions of the Family First Prevention Services Act. Retrieved from https://www.childtrends.org/wp-content/uploads/2019/10/FamilyFirst_ChildTrends_October2019.pdf

Michalopoulos, C., K. Faucetta, C. J. Hill, X. A. Portilla, L. Burrell, H. Lee, . . . V. Knox (2019), *Impacts on Family Outcomes of Evidence-Based Early Childhood Home Visiting: Results from the Mother and Infant Home Visiting Program Evaluation*. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Modan, N. (2019, August 27). *50 States of Ed Policy: States, districts tackle gun policies as lawmakers remain gridlocked*. Retrieved from https://www.educationdive.com/news/the-50-states-of-education-policy-tackling-school-safety-in-an-age-of-mass/561670/

National Research Council. (2012), *Using Science as Evidence in Public Policy*. Washington, DC: The National Academies Press. doi:https://doi.org/10.17226/13460

Nutley, S., A. Boaz, H. Davies and A. Fraser (2019), 'New development: What works now? Continuity and change in the use of evidence to improve public policy and service delivery', *Public Money & Management*, **39** (4): 310–316. DOI: 10.1080/09540962.2019.1598202

Nutley, S. M., I. Walter and H. T. Davies (2007), *Using evidence: How research can inform public services* (1 ed.). Bristol, UK: Bristol University Press, Policy Press. https://doi.org/10.2307/j.ctt9qgwt1

Office of Management and Budget (2019) Circular A 11, Section 290 Evaluation and Evidence Building Activities. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2018/06/s290.pdf

Oliver, K., S. Innvar, T. Lorenc, J. Woodman and J. Thomas (2014), 'A systematic review of barriers to and facilitators of the use of evidence by policymakers', *BMC Health Services Research*, **14**(2). Retrieved from http://www.biomedcentral.com/1472-6963/14/2

Pennsylvania Commission on Crime and Delinquency (2019), Cost-Benefit Analysis for PCCD's Evidence-based Initiatives – Full Report: Investing in Effective Programs to Improve Lives

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press

and Save Tax Payer Dollars. Retrieved from https://www.pccd.pa.gov/Juvenile-Justice/Documents/Cost-Benefit_Analysis_for_PCCDs_Evidence-based_Initiatives_FULL-Report_5-28-19.pdf

Penuel, W. R., C. C. Farrell, A. R. Allen, Y. Toyama, C. E. Coburn (2018), 'What research district leaders find useful', *Educational Policy*, **32**(4): 540–568. DOI:10.1177.0895904816673580

Pew-MacArthur Results First Initiative. (2017), *How States Engage in Evidence-Based Policymaking*. Retrieved from https://www.pewtrusts.org/en/research-and-analysis/reports/2017/01/how-states-engage-in-evidence-based-policymaking

Supplee, L. H. and T. McCann (2019), There's No Place Like Home: Improving Outcomes One Home Visit at a Time. Chapter in N. Hart and M. Yohannes (eds), *Evidence Works: Cases Where Evidence Meaningfully Informed Policy*. Washington, D.C.: Bipartisan Policy Center.

Supplee, L. H. and A. Metz (2015), 'Opportunities and challenges in evidence-based social policy', *Social Policy Report*, **28**(4): 1–16.

Tseng, V., J. Q. Easton and L. H. Supplee (2017), 'Research-Practice Partnerships: Building Two-Way Streets of Engagement', *Social Policy Report*, **30**(4): https://doi.org/10.1002/j.2379-3988.2017.tb00089.x

https://doi.org/10.1017/bpp.2020.25 Published online by Cambridge University Press