Published in final edited form as:

Criminol Public Policy. 2012 August 1; 11(3): 525–534. doi:10.1111/j.1745-9133.2012.00828.x.

Building Evidence for Evidence-Based Policy Making

Kenneth A. Dodge and Adam D. Mandel

Duke University

With declining high-school graduation rates and comparatively high rates of adolescent violence and problem behavior in this country, we are in a moment of great need for effective federal and state policy to prevent juvenile delinquency. Leading intellectuals in the field, including Ron Haskins (Haskins and Baron, 2011), Jon Baron (Baron and Haskins, 2011), and Steve Barnett, have recently called for adoption of a technocracy: They have asked policy makers to use the *science* of prevention to guide policy making and funding. Haskins and Baron (2011) wrote persuasive essays arguing that if policy and funding decisions were made based on evidence of what works, then we would experience better population-level outcomes in education, crime, and child well-being; furthermore, we would save costs and solve the deficit crisis.

Faith in technocracy has won the day (mostly) in health care: It is virtually impossible to enter a hospital with a disease and not have both patients and physicians call up data on its prevalence, course, and treatment. Insurance providers make reimbursement decisions based (mostly) on evidence. We can point to improvements in the population-level health of our citizenry that result. One might quibble with the validity of the empirical evidence, but we cannot quibble that as public policy we have accepted the technocratic philosophy that empirical evidence *should* rule the day in medicine. The same can be said about energy, the environment, and the economy. But in matters affecting children, we are a long way from a technocratic culture. Jon Baron (2007) summed up the contrast well:

In medicine ... rigorous evaluations are common, and often drive policy and practice. By contrast, in education and most other areas of social policy, such studies are relatively rare. In these areas, policy and practice tend to be driven more by advocacy studies and anecdote than by rigorous evidence, costing billions of dollars yet often failing to produce meaningful improvements in educational achievement, employment and earnings, or rates of substance abuse and criminal behavior. (p. 32)

Call to Disseminate Evidence-Based Programs

Greenwood and Welsh (2012, this issue) lead with the fact that evidenced-based intervention programs (EBPs) such as Multisystemic Therapy (MST), Functional Family Therapy (FFT), and Multidimensional Treatment Foster Care (MTFC) have been developed in academic settings based on developmental science and have been shown through small-sample, well-controlled, randomized trials to alter the trajectory of a child's development, and they assert that the next step is to persuade state policy makers to align their funding to saturate the population with these programs. Their call is consistent with the Institute of Medicine's (IOM's; Mrazek and Haggerty, 1994) recommendation that prevention science and policy should follow a course that (a) begins with epidemiology and basic science; (b) moves to

replicated small-sample, randomized controlled trials in pristine conditions, called efficacy studies; (c) expands to field demonstrations, called effectiveness trials; and (d) culminates in population-wide dissemination.

Greenwood and Welsh (2012) review the efforts of seven states to disseminate EBPs, and they herald especially the efforts and progress of Connecticut and Pennsylvania. Their case stories communicate much wisdom that will be of use to state-level stakeholders. They propose that the benchmark of “success” for a state should become the number of these proprietary brand-name programs per million residents, which we call the penetration rate. Unfortunately, EBP penetration rates are often low.¹ Even if the penetration rate improves substantially, however, we are not satisfied that this outcome should suffice. What evidence supports the contention that increasing the EBP penetration rate will bring improved population-level impact on youth outcomes?

Application of the IOM Model to Behavioral Interventions

Evidence suggests that the impact of EBPs on population rates of disorder might be less than what policy makers are led to expect. We cannot think of a single demonstration in which the scaling up of an evidence-based social program for youths has led to a demonstrated change in the population rate of an important child outcome. Why has the IOM plan not yet succeeded? We suggest that two complementary factors operate, and together they suggest a new and different approach.

Program Implementation in Community Settings

When university-inspired social programs are disseminated in communities, not only do they yield low penetration rates (in terms of the percentage of eligible individuals who utilize services), but they also tend to degrade as a result of lower per-case funding levels, lower credentialing of interventionists, lower supervision, and lower fidelity of implementation. Welsh, Sullivan, and Olds (2010) called this effect the “scale-up penalty” and estimated it at 50%. Some changes in implementation are not merely degradation but are planned adaptations to accommodate a new population or context. The impact of these planned variations is not always positive, and the general impact of disseminated programs (called “clinic therapy” by Weisz and Jensen, 1999) on child outcomes tends to be lower than that reported in the original randomized trials (called “lab therapy” by Weisz and Jensen, 1999). We suggest that this slippage is not only a result of degradation but also of mismatching of the program with the population and context. MST, FFT, and MTFC were all developed with “volunteer” families that had some degree of motivation, and they were implemented with small numbers of families in a context in which the marginal demand on community resources was low. These interventions depend on cooperation from school teachers, ancillary after-school programs, and professional community service providers to maximize impact on the individual family. When the number of families involved is low, as in most randomized trials, there is little net strain on community resources, and the intervention families might have a comparative advantage. When a program is implemented at scale, however, the strain could become greater than the community’s ability to respond. The result could be a decreased net impact on child outcomes. All of these factors may contribute to a community’s sense that the program is not working.

¹In juvenile justice, only 5% of Californian youths who should receive an EBP actually receive one (Hennigan et al., 2007), and the rate is surely lower in some other states.

Program Framing in Community Settings

A second factor in scaling up programs is how they are framed. To the university researcher, the framing and goal of the intervention program might well have been to test a theory about how problem behaviors develop rather than to change a community. To community members, this framing, oriented toward community, family, and individual “deficiency,” is disconcerting, to say the least. The framing of a problem and its solution determines the willingness of community members to participate, alters stereotypes, and shapes outcomes (Dodge, 2008). A program that has been developed by university researchers in a distant context with a different population might not be well received when thrust on a new community, and it could result in less compliance, fewer resources, and negative expectations.

Utility of the IOM Model for Behavioral Interventions

We suggest that the IOM model, while fitting for pharmacotherapies, is not well suited to psychosocial programs for youths whose problem behaviors are contextually bound. Instead, we propose starting not at the basic-science end of the continuum but rather at the community-impact end. Program developers should imagine the community’s problem and possible population impact, and then they should work backward to envision how to achieve that impact. By doing so, program developers will take into account the overall community resource constraints and the framing of the problem and its solution, and they will be able to integrate those circumstances with the developmental science that will still undergird a program’s rationale. This new approach does not imply ignoring basic science, of course, but rather, it suggests integrating that science with an understanding of community context. The difference is between laboratory science and engineering, where the engineer takes the actual circumstances as a given in designing a solution. Program development and implementation might take longer, but it will come with greater stakeholder participation and fewer problems in future disseminations. The “transportation penalty” of disseminating a program from one community to another might well be less than the “scale-up penalty” that plagues current EBPs.

Whether a program is developed as wholly new within a community context or is adapted from an existing program, this discussion suggests the need for continued measurement of child outcomes and evaluation of impact, *even, or especially*, during dissemination. We fear that an exclusive emphasis on penetration rates could lead to apparently successful efforts that genuinely have little impact on public health. So we call for an effort to build evidence regarding the impact of strategies for implementing EBPs.

Policy Implications for Intervention Psychologists

An approach to program development that originates in the community has implications for program developers, researchers, and evaluators. First, we suggest that social behavioral interventionists should take control of their policy agenda. Second, in defining and promoting this agenda, interventionists should deploy the same scientific methods used in designing and evaluating EBPs. To these ends, the following section outlines recommendations to supplement Greenwood and Welsh’s (2012) proposals.

Generate Consensus on What “Evidence-Based” Means

Although psychologists have progressed in evaluating the evidence base for treatments, various competing standards have emerged. Intervention evaluators differ in terms of (a) the type and quantity of evidence they require to designate a program “evidence-based” and (b) the type and meaning of “evidence-based” labels they assign (Chambless and Ollendick, 2001). To write a persuasive core story about EBPs, the intervention community must

generate consensus and endow the label “evidence based” with reliable and valid meaning. A failure in this regard would place politicians in a position similar to consumers when shopping for “natural” foods; they would be forced to study a program’s jargonistic packaging to understand how and to what degree it is “evidenced based.”

One problem emerges when considering a disseminated program’s fidelity to an original model. Adaptations are often needed and may be inevitable, but less clear is whether they are meaningful. To clarify how programs may be adjusted during implementation, interventionists should consider emphasizing evidenced-based principles of change (EBPC) instead of, or as the foundation of, specific EBPs (Rosen and Davison, 2003). To illustrate the value of EBPCs, consider how physicians treat heart disease by managing a set of risk factors. Physicians, knowing that high blood pressure is associated with cardiac disease, use tools that reduce blood pressure (e.g., medication, exercise, weight loss, and low-sodium diets). The precise combination of methods employed is less important than reaching the theoretically sensible and empirically validated proximal and distal goals—decreasing distal heart disease by lowering proximal blood pressure. Moreover, in prescribing blood pressure medication, a physician does not prescribe a fixed dosage previously found to be effective in a randomized clinical trial; rather, best practice is to titrate the dosage until the specific patient’s blood pressure is lowered to a range associated with reduced risk. With this model in mind, it is time to extend Lipsey’s (2009) meta-analytic work to examine the precise mechanisms of change within treatment modalities. For example, change in which beliefs predicts effectiveness in cognitive-behavioral therapy for adolescents?

A shift in focus to EBPCs may provide three benefits. First, it will direct interventionists’ attention toward the basic science of change and away from proprietary programs that constrain access to treatment (Rosen and Davison, 2003). We have concerns that well-intended advocates for children might be channeling public funds toward proprietary corporations in a way that limits rather than improves the public’s options in the long run. The developers of MST, FFT, and MTFC have the loftiest of goals, no doubt, but public policy needs to be open to equally, or even more, effective options. Second, EBPCs will ultimately lead to greater effectiveness than high-fidelity EBPs because EBPCs allow for greater contextual specificity and sensitivity to individual differences. Third, EBPCs may promote greater cost effectiveness by allowing interventionists to streamline existing EBPs. Moreover, despite the research highlighting the importance of treatment fidelity, researchers may find that chasing higher levels of fidelity does not decrease scale-up penalties enough to justify increased implementation costs.

Examining Institutional Structures and Political Mechanisms of Change

Greenwood and Welsh (2012) examine how the decentralized administration of social services at the county level may serve as a barrier to change. Taking the issue a step further, behavioral interventionists should team with other social scientists to examine whether and how EBPs can effect population-level change when selected, implemented, and evaluated at the community and county level. For example, meta-analyses might be conducted to consider whether states with centralized control over the administration of social services observe different scale-up penalties than states with decentralized control. Regression discontinuity designs could be used to examine (a) how state-level legislation (or the establishment of centers to promote evidenced-based policies) affects the rate at which EBPs are adopted at the county and community level and (b) the impact that higher EBP penetration rates have on the rate of change for population-level child outcomes.

Translational Advocacy

Intervention scholars should replicate the efforts of developmental scholars and team with communications experts to formulate strategies to promote and evaluate evidenced-based practices. The first step in the communications process is to identify the “core story” interventionists want to share with their audience (Shonkoff and Bales, 2011). Second, interventionists need to identify frames that communicate such a story accurately and in a manner that promotes action (Dodge, 2008). Work at both stages should be informed by empirical research (see Shonkoff and Bales, 2011).

The advocacy goal could be the enactment of state-level legislation that is consistent with, but even more ambitious than, that outlined by Greenwood and Welsh (2012).² Two related mandates may be instrumental in effecting change. First, when selecting from competing programs addressing overlapping social problems, publicly funded providers should be obligated to fund an EBP first, if one is available. Second, whenever a publicly funded provider funds a non-EBP, it should be obligated to provide rigorous evaluations to its state governing body (thus promoting science and “policy learning”) (Weissert and Scheller, 2008). Although these legislative proposals require substantial clarification (e.g., when to treat programs as addressing the same or distinct social problems and whether to permit jurisdictions to fund low-cost non-EBPs when competing EBPs are unaffordable), requiring publicly funded social service providers to preference programs proven to work is on its face a relatively uncontroversial proposition.

Establish Best Practices for Economic Analyses

Interventionists should partner with economists to establish a set of best practices to employ when conducting economic analyses of intervention impacts. The first question to ask is what analytical method is most appropriate (see Adler, 2010). The second and related question is what variables should be included in economic analyses, as the selection of variables substantively defines intervention results and, therefore, the core story that is told. At least three factors should be considered.

1. Well-being—The type of cost–benefit variables included in economic analyses can alter results dramatically. More narrowly, there is a long-standing debate about the role of (see Bok, 2010; Diener, 2009) and methods used to calculate (e.g., Adler, 2010; Adler and Posner, 2008; Klose, 1999; Smith and Sach, 2010) nonmonetary factors such as well-being in health policy analyses (economists’ effort to monetize well-being notwithstanding). For example, the Pew Center for the States (Weiss, 2011) released a brief citing the cost of child abuse as \$30,000 per child abused when based on tangible costs alone and \$200,000 when intangible costs are included (U.S., price years not reported).

2. Scope of analysis—The scope of cost–benefit variables raises important theoretical questions to consider during intervention design and evaluation. The first issue of scope deals with the unit of analysis. When conducting economic analyses, should one constrain results by an “intent-to-treat” model, considering only those individuals directly and intentionally affected by the intervention, or should value be analyzed according to the treatment’s impact on the population (e.g., at a community level)? This question is particularly important when there is a distributive component to an intervention, as in the Moving to Opportunity trial (MTO), where certain families received lottery-granted vouchers enabling them to move from low-income housing developments to private

²Some have argued that legislative intervention should occur at the federal level (Greer and Jacobson, 2010); however, the historic inability of Washington to generate consensus on health-care policy suggests that state level action may be the only politically feasible path in the near term (Greer, 2010).

residences in less economically depressed communities (Kling, Liebman, and Katz, with the Joint Center for Poverty Research, 2001). Moreover, if valuing the MTO at the community level, should both the community of origin and the community of destination be considered? The second issue of scope relates to temporal constraints. For example, one trial of the Nurse Family Partnership yielded fewer pregnancies in nurse-visited mothers than among controls (Olds et al., 2009), a result that undoubtedly ripples across generations. Should one attempt to model the impact of such long-lasting results? If so, how?

3. Specificity of calculation—One can perform an economic analysis of an intervention's value based on how it was implemented in a single trial or how it might be implemented prospectively in other cultural, geographical, temporal, and political contexts. For example, an effort to quantify the value of decreased visits to emergency rooms prompts the question of whether to report net savings based on local hospitals' cost of services, the average cost of such services in the county, state, or nation, or some combination thereof. Moreover, one can perform economic analyses targeting a specific outcome or calculate costs and benefits holistically. For example, Greenwood, Model, Rydell, and Chiesa (1996) compared the relative cost effectiveness of implementing (a) four delinquency prevention programs (home visits/day care, parent training, graduation incentives, and delinquent supervision) and (b) a "three-strikes" law in California, as a function of (i) nominal costs, (ii) total crimes prevented, and (iii) costs-per-crime prevented, without regard to outcomes outside the realm of law enforcement, such as changes in expected lifetime earnings and social service utilization. Finally, by revisiting an issue raised earlier in the context of scope, an evaluator might select one or more perspectives when valuing a program, including taxpayers, victims, offenders, and implementing agencies, as costs and benefits are not uniformly distributed (Welsh and Farrington, 2000).

Additional considerations

Ultimately, best practices should be designed to make the core story valuable for the intended audience (service consumers, service providers, voters, and policy makers). A key feature of such value is consistency across studies. Without consistency, policymakers cannot use economic analyses to compare and prioritize competing programs and funding priorities.

Whenever possible, interventionists should employ experimental designs, particularly population-level experimental designs, that allow evaluators to use administrative and public records to measure effectiveness (Dodge, 2011). Doing so may increase external validity and make analyses more interpretable by, and salient to, policy makers. Finally, interventionists should partner with economists early in the design stage to identify ex-ante valuation strategies (Welsh and Farrington, 2000) and grant reviewers should consider the merits of such strategies in applications.

Conclusion

We applaud the effort by Greenwood and Welsh (2012) to tell the stories of how evidence-based programs are being brought to scale by various states. In doing so, however, we hope that the most important message is not lost. The goal is *not* to proliferate specific proprietary programs but to improve the well-being of a community's population. Thus, we advocate a rigorous evidence-based approach to evaluating the dissemination of evidence-based programs and the impact thereof, lest we find ourselves, a decade from now, lamenting misplaced faith in a technocratic agenda.

Acknowledgments

K.A. Dodge acknowledges the support of NIDA Grant K05DA15226.

Biographies

Kenneth A. Dodge is the William McDougall Professor of Public Policy at Duke University, where he directs the Center for Child and Family Policy, which is devoted to addressing contemporary issues in children's lives. As a clinical psychologist, Dodge receives support from a Senior Scientist Award from the National Institute on Drug Abuse and leads grant-funded projects in violence prevention.

Adam D. Mandel is an attorney and clinical-psychologist-in-training at Duke University's Department of Psychology and Neuroscience. Mandel is mentored by and works under Dodge at the Center for Child and Family Policy.

References

- Adler, Matthew D. Contingent valuation studies and health policy. *Health Economics, Policy and Law*. 2010; 5:123–131.
- Adler, Matthew D.; Posner, Eric. Happiness research and cost-benefit analysis. *The Journal of Legal Studies*. 2008; 37:S253–S292.
- Baron, Jon. Making policy work: The lesson from medicine. *Education Week*. 2007; 26:32–33.
- Baron, Jon; Haskins, Ron. Congress Should Use Cost-Effectiveness to Guide Social Spending. Washington, DC: Brookings Institution; 2011.
- Bok, Derek C. *The Politics of Happiness: What Government Can Learn from the New Research on Well-Being*. Princeton, NJ: Princeton University Press; 2010.
- Chambless, Dianne L.; Ollendick, Thomas H. Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*. 2001; 52:685–716.
- Diener, editor. *Well-Being for Public Policy*. Oxford, U.K.: Oxford University Press; 2009.
- Dodge, Kenneth A. Framing public policy and prevention of chronic violence in american youths. *American Psychologist*. 2008; 63:573–590. [PubMed: 18855489]
- Dodge, Kenneth A. Context matters in child and family policy. *Child Development*. 2011; 82:433–442. [PubMed: 21291450]
- Greenwood, Peter W.; Model, Karyn; Rydell, C. Peter; Chiesa, James. The economic benefits of diverting children from crime. *Challenge*. 1996; 39:42–44.
- Greenwood, Peter W.; Welsh, Brandon C. Promoting evidence-based practice in delinquency prevention at the state level: Principles, progress, and policy directions. *Criminology & Public Policy*. 2012; 11:493–513.
- Greer, Scott L. How does decentralisation affect the welfare state? Territorial politics and the welfare state in the UK and US. *Journal of Social Policy*. 2010; 39:181–201.
- Greer, Scott L.; Jacobson, Peter D. Health care reform and federalism. *Journal of Health Politics, Policy and Law*. 2010; 35:203–226.
- Haskins, Ron; Baron, Jon. *Building the Connection Between Policy and Evidence: The Obama Evidence-Based Initiatives*. London, U.K.: National Endowment for Science, Technology and the Arts; 2011.
- Hennigan, Karen; Kolnick, Kathy; Poplawski, John; Andrews, Angela; Ball, Nicole; Cheng, Connie, et al. *Survey of Interventions and Programs: A Continuum of Graduated Responses for Juvenile Justice in California*. Los Angeles: California Juvenile Justice Data Project; 2007.
- Kling, Jeffrey R.; Liebman, Jeffrey B.; Katz, Lawrence F. with Joint Center for Poverty Research. *Bullets Don't Got No Name: Consequences of Fear in the Ghetto*; Joint Center for Poverty Research Working Paper 225; 2001. from http://www.economics.harvard.edu/faculty/katz/files/bullets_jcpr.pdf.

- Klose, Thomas. The contingent valuation method in health care. *Health Policy*. 1999; 47:97–123. [PubMed: 10538292]
- Lipsey, Mark W. The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims & Offenders*. 2009; 4:124–147.
- Mrazek, Patricia J.; Haggerty, Robert J. *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Washington, DC: National Academies Press; 1994.
- Olds, David L.; Eckenrode, John; Henderson, Charles R., Jr; Kitzman, Harriet; Cole, Robert E.; Luckey, Dennis W., et al. Preventing child abuse and neglect with home visiting by nurses. In: Coleman, Doriane Lambelet; Dodge, Kenneth A., editors. *Preventing Child Maltreatment: Community Approaches*. New York: Guilford Press; 2009.
- Rosen, Gerald M.; Davison, Gerald C. Psychology should list empirically supported principles of change (ESPS) and not credential trademarked therapies or other treatment packages. *Behavior Modification*. 2003; 27:300–312. [PubMed: 12841586]
- Shonkoff, Jack P.; Bales, Susan Nall. Science does not speak for itself: Translating child development research for the public and its policymakers. *Child Development*. 2011; 82:17–32. [PubMed: 21291426]
- Smith, Richard D.; Sach, Tracey H. Contingent valuation: What needs to be done? *Health Economics, Policy and Law*. 2010; 5:91–111.
- Weiss, Elaine. *Paying Later: The High Costs of Failing to Invest in Young Children*. Washington, DC: The Pew Center for the States; 2011. Retrieved from pewcenteronthestates.org/report_detail.aspx?id=328408.
- Weissert, Carol S.; Scheller, Daniel. Learning from the states? Federalism and national health policy. *Public Administration Review*. 2008; 68:S162–S174.
- Weisz, John R.; Jensen, Peter S. Efficacy and effectiveness of child and adolescent psychotherapy and pharmacotherapy. *Mental Health Services Research*. 1999; 1:125–157. [PubMed: 11258738]
- Welsh, Brandon C.; Farrington, David P. Monetary costs and benefits of crime prevention programs. *Crime and Justice*. 2000; 27:305–361.
- Welsh, Brandon C.; Sullivan, Christopher J.; Olds, David L. When early crime prevention goes to scale: A new look at the evidence. *Prevention Science*. 2010; 11:115–125. [PubMed: 19936922]