Ranking States in Their Use of Evidence-Based Programs for Juvenile Offenders: A 20-Year Progress Report

Peter W. Greenwood[†]

In 2004 Frank Zimring invited me to write a book about progress in juvenile rehabilitation programs as part of a series he was editing for the MacArthur Foundation's Network on Adolescent Development and Juvenile Justice.[1] That book completely changed the course of my career and made me a dedicated advocate for evidence-based programs (EBPs) in juvenile justice, social services, and mental health.

Widespread interest in preventive approaches to juvenile delinquency date back to the founding of the Juvenile Court, in the early nineteenth century. Since that time state and local governments have experimented with a wide variety of approaches running the gamut from long term secure institutional confinement to group homes, boot camps, preparatory academies and community supervision. Most of these "experiments" were never evaluated with sufficient rigor to determine

whether they worked well or not. Results from those that were evaluated with sufficient rigor were seldom replicated with any degree of fidelity to the original design and failed to demonstrate that any one particular approach was consistently better than any other. The accepted wisdom in the 1970s and 1980s was that spending money on any type of preventive or rehabilitative programs for juvenile or adult offenders was a waste. Residential programs became warehouses for delinquents, many of whom recidivated soon after they left. Probation officers focused on supervision rather than services. In a 1993 review commissioned by the National Academy of Sciences, the experts in the field determined there was insufficient evidence to identify any program that worked better than any others.[2]

At the same time this review was taking place, a number of psychology professors were securing long term grants, from the National Institute of Mental Health (NIMH), for the purposes of further developing their programs for at-risk youth, and conducting clinical trials to rigorously evaluate their outcomes.[3] Long term funding for these efforts was, and still is, critical to provide program developers time to design, implement, and then evaluate clinical trials that involved months of services to hundreds of youth.

In 1995, Professor Delbert Elliot and his team on the Blueprints Project, at the University of Colorado, secured funding to review these recent studies for the purpose of determining whether any of the new models consistently produced better outcomes than the usual services youth would have received. The review was conducted using strict criteria to determine whether the evaluation(s) for programs showing positive results were sufficiently rigorous. Those criteria were: 1) the evaluation must utilize a strong research design, preferably a randomized control trial (RCT); 2) the evaluation must provide evidence of significant negative

---

[2]   *See* NAT'L RES. COUNCIL, UNDERSTANDING AND PREVENTING VIOLENCE V.1, 289 (Albert J. Reiss Jr. & Jeffrey A. Roth, eds. 1993).

[3]   *See* J.F. Alexander & B.V. Parsons, *Short-term behavioral intervention with delinquent families: Impact on family process and recidivism*, 81 J. OF ABNORMAL PSYCHOL. 219-225 (1973); C. Barton, et al., *Generalizing treatment effects of Functional Family Therapy: Three replications*, 13 THE AM. J. OF FAM. THERAPY 16-26 (1985); C. M. Borduin, et al., *Multisystemic treatment of adolescent sexual offenders*, 35 INT'L J. OF OFFENDER THERAPY AND COMP. CRIMINOLOGY 105-114 (1990); P. Chamberlain, *Comparative evaluation of specialized foster care for seriously delinquent youths: A first step,* 2 COMMUNITY ALTERNATIVES: INT'L J. OF FAM. CARE 21-36 (1990); P. Chamberlain & J. Reid, *Comparison of two community alternatives to incarceration for chronic juvenile offenders*, 5 J. OF CONSULTING AND CLINICAL PSYCHOL. 857-863 (1998).

effects on delinquent or criminal behavior; 3) the must be evaluations of the program in multiples sites demonstrating that it can be replicated in different settings; and 4) outcome data must show sustained program effects after youth are no longer in the program.

Blueprints published their first list of model programs for reducing delinquency and violence among juveniles in 1996.[4] There were 10 program models on that initial list that met the Blueprints selection criteria. Over time the selection criteria for Blueprints have been tightened and expanded to distinguish three levels of certainty regarding program effectiveness: Promising, Model, and Model Plus. All three levels require programs to meet three minimum requirements. First, the program description must clearly identify the outcomes the program is designed to change, the specific risk and/or protective factors targeted to produce this change, the population for which it is intended, and how the components of the intervention work to produce this change.

Secondly, the preponderance of evidence from the high-quality evaluations must indicate significant positive change in the intended outcomes that can be attributed to the program, and there is no evidence of harmful effects.

The third requirement requires that the program be currently available for dissemination and have the necessary organizational capability, manuals, training, technical assistance and other support required for implementation with fidelity in communities and public service systems. Some of the early Model programs identified by Blueprints could not meet this last requirement and were later removed from the list.

Promising programs require a minimum of one high quality randomized control trial or two high quality quasi-experimental evaluations. Model programs require a minimum of (a) two high quality randomized control trials or (b) one high quality randomized control trial plus one high quality quasi-experimental evaluation and evidence that the positive intervention impact is sustained for a minimum of 12 months after the program intervention ends. Model Plus programs require that in at least one high quality study demonstrating desired outcomes, authorship, data collection, and analysis has been conducted by a researcher by a researcher who has no financial interest in the program and is neither a current or past member of the program developer's

---

[4]   *See* Delbert Elliot, *BLUEPRINTS FOR VIOLENCE PREVENTION SERIES*, BOULDER, CO: CENTER FOR THE STUDY AND PREVENTION OF VIOLENCE, 1996.

research team.[5]

Blueprints began with a focus on youth programs to prevent violence, delinquency, and drug use, but its scope has recently expanded. It now recommends programs to improve mental and physical health, self-regulation, and educational outcomes. At the current time there are more than 70 programs ranked on the Blueprints website, 15 of which are rated as "Model" or "Model Plus." Some of these programs are designed to reach very young at-risk children or those already in school. Only three of the Model programs are appropriate for use with actual juvenile offenders. These three have been rated as Model programs since the start of the Blueprints project. They are: Functional Family Therapy (FFT); Multi-systemic Therapy (MST); and Treatment Foster Care Oregon (TFCO), which was originally called Multi-Dimensional Treatment Foster Care (MTFC) when it was first evaluated.

For all three of these models the focus is on family dynamics and behavior. FFT therapists work with troubled families (delinquents and their parents) in their homes. Over the course of approximately 14 sessions they first engage and then motivate the family to adopt a problem-solving approach to each behavioral issue (truancy, staying out all night, chores, etc.). Family members complete a series of questionnaires that are used to monitor how well the therapy is going and how the family feels about it.

MST works with more difficult parental situations where the family needs assistance in dealing with the school, social services, and mental health systems. TFCO works with families that cannot provide an adequate home for their child because of their own behavioral, law enforcement or substance abuse issues. TFCO places the youth with specially trained foster care parents, while they work with the parents on their own issues and plans for reunification. TFCO is most often used as an alternative to group home placement.

The "treatment team" is the basic unit of operation, supervision, training and accountability. Each team, that consists of a lead therapist supervising 5-7 therapists, costs approximately $500,000 per year to support. Adopting one of these "brand name" programs is something like buying into a franchise. Instead of an up-front licensing fee, agencies adopting one of these programs has to pay the upfront costs of hiring and training the new therapists. For a modest annual fee, the program

---

[5]   *See* BLUEPRINTS FOR HEALTHY YOUTH DEVELOPMENT, http://www.blueprintsprograms.org (Nov. 1, 2018).

developer provides all of the software and printed material required to administer, track and evaluate all their cases, comparing their performance with other sites. The program developer usually provides access to a master therapist, who has several years' experience with the program, and who will usually participate in the weekly team meetings, usually through some internet conferencing service, where they review progress on every open case. Supervision within the team utilizes a coaching, rather an accountability, approach.

For more than a decade the Washington State Legislature has required the Washington State Institute for Public Policy (WSIPP) to review all the evaluations for any program that might possibly improve the state's effectiveness in dealing with delinquent and at-risk youth, along with programs for adult corrections, mental health, public schools and several more. For more than a decade WSIPP has been publishing reports showing that all three of these programs produce savings in future government expenses well in excess of what they cost to operate (WSIPP). These are programs that both reduce future criminality and save money for the taxpayers.

Given this state of affairs, one might expect that most states would be in the process of revising their programs and case disposition processes to increase the participation of youth in programs that have been proven effective. In fact, a few states have responded to this knowledge by taking explicit steps to facilitate the implementation of these proven programs, often as alternatives or replacements for their more traditional programming. Some of these states have set up special resource centers to provide technical assistance to local providers and to monitor their progress in implementing these programs. Some have established local "compacts" for sharing the expected savings in state prison costs with counties who cut their admission rates through the use of evidence-based programs. Others have established special funding streams to support the launch of new evidence-based programs. Yet, many others have not taken any but the most rudimentary steps toward embracing this new opportunity in the field of delinquency prevention.

Evidence Regarding State Progress in Adopting EBPs

In 2011 a national study examined the utilization of FFT, MST, and MTFC across states by comparing the number of licensed teams in

each state, adjusted for their population.[6] The number of licensed teams in each state was obtained from the program developers. Population data came from the U.S. Census. The availability of programs was measured in terms of the number of "therapist teams" available on a per capita basis. The research team also conducted on-site case studies in all of the leading states and several who were farther back in their progress.

Figure 1 from that study shows the total number of FFT, MST, and TFC teams per million population in each of the states that have at least one of these EBPs. It is easy to see that there is a very wide spread between the top five states (Connecticut, Hawaii, Louisiana, Maine, and New Mexico) and the others. There is also a big difference between those in the middle range of progress and those who have made very little progress.

In New Mexico, Louisiana, Maine, Hawaii and Connecticut, the availability of these programs averages more than 10 per million population. The availability of these programs is more than double that in the four states with the next highest availability (Colorado, North Carolina, Pennsylvania, and Rhode Island). Figure 1 also shows that MST is the most available of these three family-focused proven model programs.

---

[6]   *See* Peter W. Greenwood & Brandon C. Welsh, *Promoting Evidence-Based Practice in Delinquency Prevention at the State Level: Principles, Progress, and Policy Directions*, 11 CRIMINOLOGY & PUB. POL'Y 493-515 (2012).

Figure 1: States rank ordered by family therapy teams per million population in 2010.
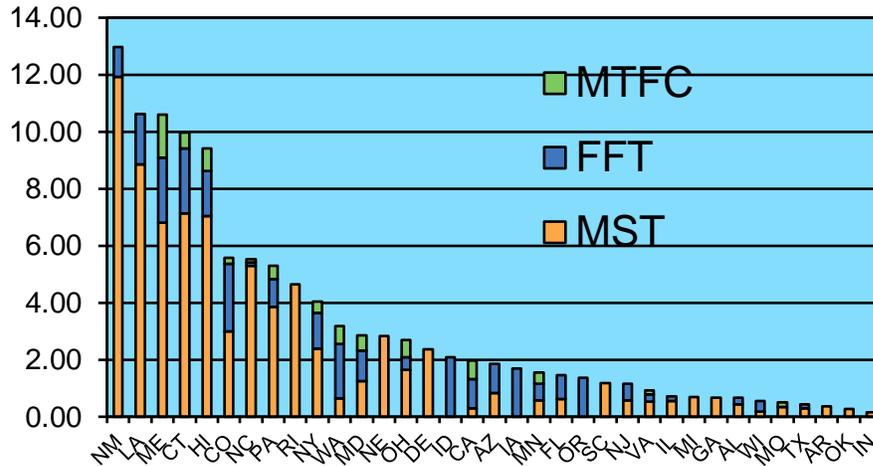
The case studies found that the top five states shared several common characteristics in their use of EBPs. They all benefited by providing a separate chain of command for juvenile and adult corrections. In all five leading states the administration of juvenile justice programs is separate from, and not subservient to, adult corrections.

They all adopted some strategy for developing local expertise. All the leading states identified at least one person to become fully informed about the available EBP options and made the time available for them to do this, including travel to operational sites and training in specific models.

Effective changes in juvenile justice programming efforts require the cooperation of many state and local agencies, including state departments of children and families, mental health, probation, law enforcement, and school systems. All the leading states created high-level stakeholder groups to oversee the process of rolling out EBPs. In Connecticut it was the Governor's Blue Ribbon Commission on Mental Health in 2000; in Maine it was the Juvenile Justice Advisory Group; in New Mexico it was the Behavioral Health Collaborative; in Louisiana it was the Juvenile Justice Implementation Commission; and in Hawaii it was the EBS Task Force and the local Community Councils.

All but one of the leading states picked one or two sites in which to test the program models they had selected as the best to suit their needs.

These pilot tests were closely monitored with progress reports and the final results shared widely among interested parties.

Most of the leading states created some type of "information resource center" that became the primary bridge between the science of EBP (e.g., review articles, assessment instruments, training consultants) and the practitioners. Staff from these centers would sit in on practitioner meetings to better understand their needs, and then develop analytical or informational tools to help address them. Practitioners would ask center staff for information about particular problems, or programs they may have heard about, and receive timely, unbiased answers.

All of our leading states started out supporting just one EBP, either MST or FFT. All of them gradually added programs to the list of what they supported, albeit slowly. The availability of funds to support the very important but non-revenue producing pre-implementation aspects of a new EBP can be a challenge. That challenge is reduced if the state can support some of those costs.

Since counties in most states, are far from uniform in size or demographics, it is seldom likely that a one-size policy reform will fit them all. In order to help local policy makers make intelligent choices about which EBPs to adopt, most of the leading states set up some process to provide them with training and technical assistance. Research has demonstrated that local communities get better outcomes if they receive proper training in how to assess their needs, and then select, and implement the best program for their needs.[7] It has been demonstrated that the spread of EBPs becomes much more rational and effective when states are able to serve local communities in this way.[8]

Three of the leading states were being sued by the federal Department of Justice over conditions in their juvenile institutions. In the other two there was a growing political consensus that many youth being sent to placement did not belong there. All five leading states were able to capitalize on this crisis of confidence by bringing appropriate stakeholders together and identifying capable individuals to take charge.

---

[7]   Issac Rhew, et al., *Effects of Exposure to the Communities That Care prevention system on youth problem behaviors in a community-randomized trial: Employing an inverse probability Weighting Approach*, 41(2) EVALUATION AND THE HEALTH PROF. 270-289 (2018).

[8]   Davis Hawkins, et. al., *Youth Problem behaviors 8 years after implementing the Communities That Care prevention system: A community-randomized trial*, 168(2) JAMA PEDIATRICS 122-129 (2014).

More Recent Evidence Regarding State Progress

Six years have now passed since that original study. Once again, we have replicated the 2011 study's methods, by contacting the purveyors of FFT, MST and TFCO to find out how many licensed teams currently existed in each of the states. Figure 2 shows the current rate of these programs availability per million population, compared to the rate in 2011.

Figure 2: Availability of Model Programs in July of 2017

State Use of EBPs

Leading State Use of Most Popular EBPs for Juvenile Justice
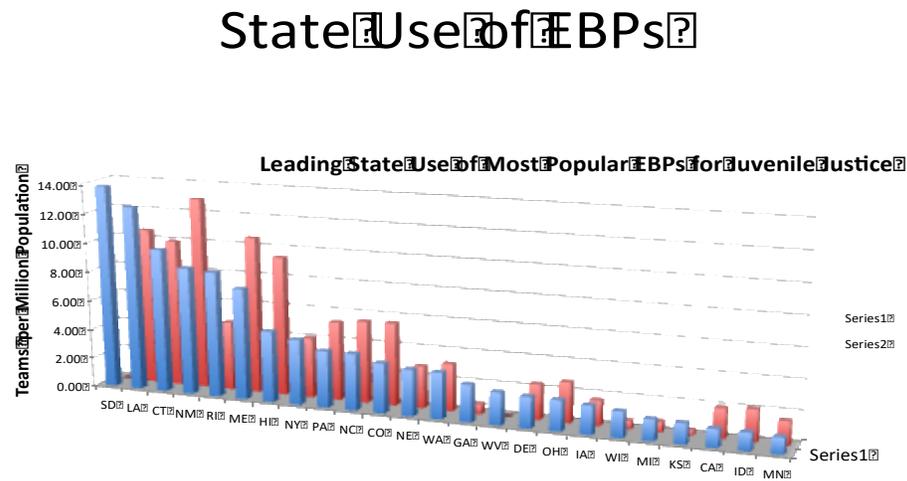
The blue front row in Figure 2, representing the rates in 2017 shows that there has been very little if any growth in the use of these program models, compared to the red back row representing 2011. The only new states in the top 5 are South Dakota and Rhode Island. In the last six years South Dakota went from the lowest ranking, with none of these programs, to first place.

In that same time period Rhode Island, almost doubled their use of Model programs. Many other states, including some of the most progressive, are making less use of the Model Programs than they were in 2011. Clearly, the dissemination and adoption of Blueprint Model Programs has hit something of a wall. This lack of progress may be due to a combination of factors including: the lack of a sustained advocacy efforts; restrictions on Medicaid funding that can be used to fund such

programs, but may not be able to cover all cost required to fully implement the program; and competition from programs claiming to be evidence-based, but which are not.

Other Methods of Measuring State Progress

There are a number of other methods that might be used to measure state progress. Entirely different methods of measuring state progress could involve measuring the actions states have taken to support EBP, or examining the specific programming afforded to a representative sample or the entire population of youth who were served.

The approach of measuring state actions was the approach taken in a recent study for the PEW-MacArthur Results First Initiative.[9] That study: 1) identified six distinct actions that states can use to incorporate research findings into their decisions, 2) assessed the prevalence and level of these actions within four human service policy areas across 50 states and the District of Columbia, and 3) categorized each state based on the final results.

The six critical actions the PEW-MacArthur team decided to look at include:

1. **Defining levels of evidence** can allow state leaders to distinguish proven programs from those that have not been evaluated. Thirty-nine states and the District of Columbia have defined at least one level of evidence, such as "evidence-based;" 23 of the 40 have created an advanced definition that distinguishes multiple levels of rigor, such as "evidence-based" and "promising." If appropriate criteria are not used to distinguish these programs, the results will be meaningless.

2. **Inventorying state programs** can help governments to manage available resources strategically. Forty-nine states and the District have produced an inventory of state-funded programs; 29 of the 50 have created an advanced inventory that classifies programs by evidence of effectiveness.

3. **Comparing program costs and benefits** would allow policymakers to weigh the costs of public programs against the outcomes and economic returns they deliver. Seventeen states

---

[9]   *See* PEW-MACARTHUR RESULTS FIRST INITIATIVE, HOW STATES ENGAGE IN EVIDENCE-BASED POLICYMAKING: A NATIONAL ASSESSMENT 4-5 (2012).

have conducted cost-benefit analyses; 16 of the 17 have created an advanced analysis that monetizes benefits to calculate return on investment. If the cost benefit studies are not based on rigorous evaluations that meet the Blueprints standards, then their results will not be accurate.

4. **Reporting outcomes and program effectiveness** can help policymakers identify which investments are generating positive results and use this information to better prioritize and direct funds. Forty-one states and the District reported or required key outcome data during the fiscal year 2013-17 budget cycles; 13 of the 42 have created advanced budget materials that include findings from program evaluations. Once again, without a rigorous evaluation design that can control for differences in youth risk levels, any attempts to compare program effectiveness will be meaningless.

5. **Targeting funding to evidence-based programs**, such as through a grant or contract, can help states implement and expand these proven approaches. Forty-nine states and the District of Columbia have such a funding mechanism; five of the 50 have created advanced mechanisms to dedicate at least 50 percent of program funds for a specific policy area toward these initiatives.

6. **Requiring action through state law**, which includes administrative codes, executive orders, and statutes, can help states sustain support for evidence-based policymaking. Thirty-three states and the District have developed a framework of laws to support one or more of the five advanced actions listed above in at least one policy area; 11 of the 34 states have created an advanced framework of laws to support two or more advanced actions.

The PEW-MacArthur study team collected data from individual states by reviewing documents and conducting interviews with key officials. The study team then scored each state in regard to their taking the six critical actions identified above. That study concluded that:

- Washington, Utah, Minnesota, Connecticut, and Oregon are *leading* in evidence-based policymaking by developing processes and tools that use evidence to inform policy and budget decisions across the areas examined.

- 11 states show *established* levels of evidence-based policymaking by pursuing more actions than most states but either not as frequently or in as advanced a manner as the leading states.

- 27 states and the District of Columbia demonstrate *modest* engagement in this work, pursuing actions less frequently and in less advanced ways.

- Seven states are *trailing*, taking very few evidence-based policymaking actions.

There is not much overlap in the list of leading states identified by this approach and the one developed by measuring program availability. As shown in Figure 3, only one of the states identified as leading in policy actions, scores in the top 10 when they are measured by actual use. Two of the top 5 states based on policy actions score in the lower half of states when it comes to use. There are several factors that could cause this divergence.

Figure 3: Ranking State Progress by Alternative Means

State Rank by Model Program Availability	State	State Rank by PEW Study
1	SD	7 Worst
2	LA	BELOW 26
3	CT	TOP 5
4	NM	Top 26
5	RI	BELOW 26
6	ME	BELOW 26
7	HI	BELOW 26
8	NY	Top 26
9	PA	Top 26
10	NC	BELOW 26
11	CO	Top 26
12	NEV	BELOW 26
13	WA	TOP 5
14	GEO	BELOW 26
18	IOWA	BELOW 26
19	WIS	BELOW 26
21	DEL	BELOW 26
24	MIN	TOP 5
27	AZ	BELOW 26
29	OR	TOP 5
32	FL	Top 26
47	TEN	Top 26
48	UT	TOP 5

The policy actions measured by PEW may be necessary but not sufficient to produce higher rates of Model Program use. They do not specifically include the pilot testing of model programs or the establishment of an information resource center, both of which seemed critical to ramping up use of Model Programs in "high use" states.

It could be that some of the states with lower current use of Model Programs have scored high on the "policy action" measure because they see the need to improve their performance. Only time will tell whether

the policy actions measured by Pew will lead to wider use of "Model Programs" in the states where they were taken.

Despite the many positive state policy actions identified by PEW it appears that the wider use of Blueprint type Model Programs in juvenile justice has encountered unforeseen obstacles. Even though cost-benefit studies show clear financial reasons why such programs should be adopted, most states have not seen fit do more than dabble in the world of evidence-based programs. Their rate of use in the U.S. has been flat over the past six years.

Alternative Types of Evidence

The identification, dissemination, and replication of effective model programs is not the only way of using evidence to improve local practice. An entirely different approach involves determining the common features of effective programs, and then helping local service providers to incorporate those features into their existing programs. In juvenile justice Professor Mark Lipsey and his team at Vanderbilt University most actively champion this latter approach. Using meta-analysis to comb through a database containing information on 548 controlled studies of interventions with juvenile offenders,[10] Lipsey and his team have identified the 4 most important characteristics that appear to determine program effectiveness. These are: 1) the generic program type (cognitive-behavioral therapy, family therapy, individual counseling, etc.); 2) the amount of service provided (duration and contact hours); 3) the quality of service delivery; and 4) the recidivism risk of the youth served. SPEP (Standardized Program Evaluation Protocol) is a scheme developed by Lipsey for helping program providers and public officials assess the degree to which these characteristics are present within any of their program, and predict the expected effects of the programs on recidivism. A number of states and local communities have used the SPEP to assess all of their programs, and provided feedback to providers on how they can raise their SPEP scores and hopefully the effectiveness of their program.

The advantage of the "Blueprint Model Replication" approach guarantees that the adopting agency will have all the training resources and guidance it will need to successfully adopt the model, and not cut any corners if they wish their program to be licensed by the developer. FFT, MST, and TFCO have each been implemented in hundreds of sites with

---

[10] *See* Mark W. Lipsey & James C. Howell, *Research-Based Guidelines for Juvenile Justice Programs*, 14 JUST. RES. & POLICY 17- 34 (2012); Mark W. Lipsey, *The Primary Factors that Characterize Effective Interventions with Juvenile Offenders: A Meta-Analytic Review*, 4 VICTIMS & OFFENDERS 124-147 (2009).

the developer's training and technical support. The disadvantage of this approach is that it requires a significant upfront investment in training and a minimum level of effort to support at least 2 therapists and a part time supervisor. The annual cost for a full team of 4-6 therapists and supervisor now runs close to $500,000. Another disadvantage of the Model approach is that policies and procedures are specified in such great detail that there is very little room for local adaptation. In their inventory of all therapeutic juvenile programs in 10 states, the SPEP team found only 7.3% were programs listed in the Blueprints, OJJDP Model Programs Guide, or CrimeSolutions.gov evidence-based program registries.

The advantage of the SPEP approach is that it is much more flexible and accommodating to local conditions and resources. Its primary disadvantages are its lack of technical assistance, susceptibility to short cuts, and the fact that it has not been adequately evaluated as a strategy for actually improving program outcomes.

The proponents of these two approaches do not exactly see eye to eye regarding their relative merits. The people behind Blueprints will argue that SPEP provides communities with an excuse for not adopting the more rigorously evaluated Models that they list on their website. The SPEP folks argue that evaluations of the Blueprint Model Programs usually find much better results when the Model and evaluation are both implemented by the program developer, suggesting that local communities should not expect to achieve the benefits reported in the literature.[11]

As a matter of fact, neither of these two evidence-based approaches seems to have gained much traction among state and local policy makers. Statewide program assessments with SPEP have only been attempted in 10 states, as shown in Table 4. In conducting those assessments, the SPEP discovered that only 7.3% of the programs in use were listed in the Blueprints, OJJDP Model Programs Guide, or CrimeSolutions.gov evidence-based program registries. The count of Blueprint Models by state conducted for this paper shows that half of the states have less than 1/8th of these programs, per million population, than the 5 leading states. One quarter have none at all.

Table 4 also shows how a number of states rank according to the three rating methods described above: number of family-based Blueprint Model Programs per million population; PEW/MacArthur Policy Assessment; and SPEP. Only one state, Connecticut, is highly ranked by all three methods. South Dakota, which ranks first in program availability,

---

[11] *See, e.g.*, Mark W. Lipsey, *Effective Use of the Large Body of Research on the Effectiveness of Programs for Juvenile Offenders and the Failure of the Model Programs Approach*, 17 CRIMINOLOGY & PUB. POLICY 189-198 (2018).

is ranked by Pew/MacArthur as among the 7 worst in regard to policy. Louisiana and Rhode Island, ranked among the 5 best in regard to Model Program availability are both ranked below 26th by Pew. The other 4 states ranked most highly for their evidence-based policy making in regard to juvenile justice (WA, Min, OR and UT) are ranked 10th, 24th, 29th and 48th by the availability of programs. None of these four have implemented SPEP. On the other hand, 7 of the 10 states that have implemented SPEP are also ranked in the top 21 according to program availability.

Figure 4: Ranking of States by Alternative Methods

States by Rank	Model Program Availability	PEW POLICY REFORMS	SPEP
1	SD	7 Worst	
2	LA	BELOW 26	
3	CT	TOP 5	YES
4	NM	Top 26	
5	RI	BELOW 26	
6	ME	BELOW 26	
7	HI	BELOW 26	
8	NY	Top 26	
9	PA	Top 26	YES
10	NC	BELOW 26	YES
11	CO	Top 26	
12	NEV	BELOW 26	
13	WA	TOP 5	
14	GEO	BELOW 26	YES
18	IOWA	BELOW 26	YES
19	WIS	BELOW 26	YES
21	DEL	BELOW 26	YES
24	MIN	TOP 5	
27	AZ	BELOW 26	YES
29	OR	TOP 5	
32	FL	Top 26	YES
47	TEN	Top 26	YES
48	UT	TOP 5	

So what gives? How can two different methods of measuring state level commitment to evidence-based programming lead to such different conclusions. There are at least 4 possibilities: <u>inaccuracies</u> in the data on which the scales are built; <u>irrelevance</u> of items measured by the scales; <u>omission</u> of critical items on some of the scales, and lack of correlation between progress in juvenile justice and the other policy areas reflected in the PEW scores.

Although it is possible that there has been a mistake in counting Blueprint programs across states, those numbers are readily available at any time from the model purveyors and can be easily checked. The same is true for SPEP. The data utilized by the PEW/MacArthur team to characterize state policies is not readily available and had to be collected by means of: 1) an exhaustive review of statutes, administrative codes, executive orders, and publicly available documents released between 2010 and 2015; and 2) an email survey of more than 200 state officials, including agency directors with control over the policy areas examined in this study, budget directors, and directors of commissions and entities that influence policy in these areas (such as sentencing commissions). As part of the email survey, respondents were asked to review researchers' initial findings for relevant policy areas and identify additional examples for inclusion. In other words, state officials were given an opportunity to review and improve the data on which they were being rated. This self-report approach is inherently more subject to bias or exaggeration than the objective approach used to score the use of Blueprint programs or SPEP.

Irrelevance

The Pew study was not just limited to juvenile justice but also investigated evidence-based policy making in three related area: behavioral health, child welfare and criminal justice. Since each state's final score combines results from all four of these policy areas, it is possible that the final score is not an accurate reflection of evidence-based policy making in juvenile justice. The highest score that can be achieved in any of the four policy areas covered by the Pew survey is 12, 2 points for each of 6 items. The 5 states that ranked in the top 5 overall had scores for their juvenile justice policies ranging from 11 for WA to 6 or 7 for the other four. The state with the lowest total score (25) to be ranked in the top 5 on policy was CT with a 7 for CJ policy and sixes for the other 3 areas. So 6 or better is a pretty good score. Only 3 of the 10 top states, as ranked by availability scored 6 or higher on the Pew JJ measure. In summary, the combined scores across 4 policy areas introduces considerable uncertainty regarding how well states are doing with EBP in

particular sectors, but is not entirely uncorrelated with their JJ scores. A low standing for the total is predictive of a low score on JJ.

Omission

One major aspect of evidence-based policymaking not included is: how the infrastructure and tools developed to inform decision-making translates into a tangible change in the programs and services offered by states. The study does not attempt to assess 1) the weight given to evidence in decision-making, or other tools developed to inform the process, 2) the number of evidence-based programs being implemented within the state, or 3) the quality of evidence-based programs being implemented, including efforts to manage them through fidelity monitoring.[12]

It could be argued that by restricting the model programs we counted to only listed by Blueprints, we are not giving states credit for adopting model programs from other EBP lists such as DOJ's Crime Solutions.gov or NIDA's NREP. However, an examination of those programs listed as Models by CrimeSolutions.gov but not Blueprints shows that either their positive effects were not demonstrated by sufficiently rigorous evaluations, or there is no organization or infrastructure in place to adequately replicate them, or both; the two requirements of Blueprints but not crimesolutions.gov. In other words, there are no real evidence-based substitutes for the three family-based Blueprint Model programs we used to rank states. They have no competitors.

In summary, ranking states according to the self-reported evidence-based policy steps used by PEW/MacArthur leads to very different conclusions than if they are ranked by their actual use of EBPs. Public officials and citizen advocates should be made aware of these differences.

The five states that lead in establishing EBPs support more than 8 teams per million population. Half of the states support less than one team per million population. Fourteen have none. The average rate of EBP support across the U.S. has declined from 2.1 teams to 1.8 teams per million population over the past seven years.

Using the availability of Blueprint EBPs as a measure of progress in advancing evidence-based practice, it appears that the movement toward more evidence-based programming for juveniles is currently stalled in the water, at least in the United States. A number of countries

---

[12]   *See* PEW-MACARTHUR RESULTS FIRST INITIATIVE, *supra* note 9, at 36.

in Europe and other parts of the world are making considerably more progress.