

## APPROVAL SHEET

Title of Dissertation: Results First? Exploring Evidence Utilization by State Policymakers

Name of Candidate: John A. Olszewski, Jr.  
Doctor of Philosophy, 2017

Dissertation and Abstract Approved: Roy Meyers  
Dr. Roy Meyers  
Professor  
Political Science

Date Approved: November 27, 2017

## ABSTRACT

Title of Document: RESULTS FIRST? EXPLORING EVIDENCE  
UTILIZATION BY STATE POLICYMAKERS.

John Anthony Olszewski, Jr., Ph.D., 2017

Directed By: Professor Roy Meyers, Department of Political  
Science

This dissertation explores issues impacting the use of evidence during the policymaking process at the state legislative level. Specifically, the role of Results First, a cost-benefit analysis program of the Pew and McArthur Foundations designed to identify optimal policy outcomes, is examined for its impact on the evidence utilization behaviors of state policymakers. Despite steady calls from the public policy discipline for a better integration of evidence into the policymaking process, research efforts that directly measure (and therefore understand) the extent to which state lawmakers engage with evidence has been limited. Thus, there is a need to expand our understanding of how these policymakers engage with evidence, including the circumstances that inhibit or facilitate evidence utilization. The research begins to address these knowledge gaps, providing an update and expansion of scope to Hird's research on non-partisan research organizations. Primary data collection includes both surveys and interviews with state legislators, as well as interviews with legislative staff. Third-party data included in the analysis are news articles and legislative records. Mixed analytical methods are deployed to review the data, including quantitative analysis of survey responses, content analysis of interview data and a case study on justice reinvestment. The research finds that Results First has

mixed impacts on influencing the use of evidence in the policymaking process, marginally encouraging evidence utilization by state policymakers. It affirms the tensions between evidence and other legislative priorities, such as addressing the demands of constituents, the personal beliefs and value systems of legislators and navigating personal relationships with legislative colleagues. Specifically, the research finds that evidence-informed policymaking is possible when the following conditions exist: a policymaker believes evidence should be used; the policymaker does not feel constrained by constituent desires or legislative leadership; the policymaker's desire to use evidence outweighs the desire to meet the needs of other interests (such as colleagues and lobbyists); evidence is available and the policymaker has the time to review it as well as the ability to access and validate it. Evidence can save money and lives when effectively deployed in the state legislative context, but there remains considerable opportunity for growth in the evidence utilization space. The policy significance in the research findings relates to both producers and consumers of evidence at the state legislative level. Producers of evidence such as Results First can fuel growth of evidence use by opening their analytical processes to outside review and by more actively engaging lawmakers -- cultivating trust and credibility through stronger relationships. Producers of evidence should also change their language from evidence-based policymaking to evidence-informed policymaking to more accurately depict how evidence is used in the policy process. Consumers of evidence must recognize the value evidence can play in the policymaking process and make or continue to make significant investments into independent and bipartisan/nonpartisan research organizations.

RESULTS FIRST? EXPLORING EVIDENCE UTILIZATION BY STATE  
POLICYMAKERS.

By

John A. Olszewski, Jr.

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, Baltimore County, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2017

© Copyright by  
John A. Olszewski, Jr.  
2017



## Acknowledgements

This dissertation began as and remains a labor of love but ended, as I imagine is the case with most dissertations, as a test of will. Making it to completion would have been impossible without the support of countless individuals. Most important, my wife Marisa has been an incredible source of encouragement and help throughout the entire process. I am grateful for the ways in which my parents, John Sr. and Sherry, have consistently supported my educational pursuits – and the ways in which my entire extended family has pitched in to make this dissertation possible. Finally, I would like to acknowledge the extraordinary assistance provided by my entire dissertation committee. One could not ask for a better committee chair than Dr. Roy Meyers, who consistently provided feedback and advice, with no shortage of additional material to consider as I conducted my research. The other two designated readers, Dr. Mileah Kromer and Dr. Laura Hussey, were likewise especially helpful in offering constructive feedback and support throughout the writing process. I am sincerely grateful to all of you.

# Table of Contents

Acknowledgements.....	ii
Table of Contents .....	iii
List of Tables .....	vii
List of Figures .....	viii
Chapter 1: The promise of evidence-based policymaking.....	1
1.1 What exactly is the deal with evidence and government? .....	1
1.2 Outline of the dissertation .....	3
1.3 “Evidence” Revisited .....	5
1.4 Does Pew’s Results First deliver? The research question .....	8
1.5 Other issues impacting evidence utilization .....	11
1.6 What is Results First? .....	12
1.7 Alternative approaches: Virginia’s JLARC as an example .....	15
1.8 The leading model, or the only model of evidence-based policy? .....	17
1.9 What’s missing? Filling a gap in knowledge – why does it matter? .....	18
1.10 Study structure: methodology and research design .....	21
1.11 Details on case study selection rationale .....	24
1.12 Data analysis methods.....	30
Chapter 2: Where have we been and where are we going: a literature review .....	34
2.1 Evidence and policymaking: a steady drumbeat starting in the 1960’s.....	34
2.2 Evidence defined and its role in decision-making .....	42
2.3 Performance management and tensions with evidence based policymaking ...	46
2.3.1 Using evidence in policy.....	50
2.3.2 Communicating information about evidence.....	53
2.4 Doubts and shortcomings preventing evidence-based policymaking .....	55
2.5 A nine-layered view of policy worlds.....	59
2.6 Variations of evidence production for state-level policymaking: trends over time .....	64
2.6.1 Washington State Institute of Public Policy (WSIPP).....	64
2.6.2 South Carolina Government Research and Service Technical Assistance	66
2.7 From big ideas to specific insights about evidence and policy: the literature ..	68
2.8 Trusteeship among political representatives: a delegate or a trustee? .....	71
2.9 How elected officials make decisions .....	74
2.9.1 What goes into legislative decision making? A critical question .....	74
2.9.2 Single minded seekers of election?.....	75
2.9.3 Willing to give up a seat: Not single minded, but a bounded rationality?..	76
2.9.4 State legislative behaviors.....	81
2.10 Chapter conclusion from the literature to new data .....	85
Chapter 3: Evidence utilization among state legislators: the survey data.....	86
3.1 Survey instrument development – preliminary interviews .....	86
3.2 Survey methodology .....	87
3.3 Independent variables: demographic data and beyond .....	94
3.4 Comparing population demographics to sample demographics .....	97



3.5 Survey results.....	99
3.5.1 Question set one: factors that influence decision making.....	99
3.5.2 Question set two: evidence, personal values, and constituent opinion ....	103
3.5.3 Question set three: finding evidence and its influence by type .....	107
3.5.4 Question set four: personal values, constituent opinion, and policy area	114
3.5.5 Question Set five: evidence and legislative work .....	116
3.5.6 Question set six: Results First.....	120
3.6 Dependent variable: evidence utilization index .....	123
3.6.1 Index correlation .....	126
3.6.2 Alternative constructs of the evidence utilization index.....	126
3.7 Crosstabs of interest between evidence index and select independent variables	
.....	128
3.8 Conclusion: from surveys to interviews .....	132
Chapter 4: Evidence utilization among legislators: in their own words .....	135
4.1 Interview methodology .....	136
4.2 High evidence utilizer: in his own words .....	140
4.2.1 Massachusetts House member, score of 16 .....	140
4.3 Low evidence utilizers: in their own words.....	143
4.3.1 West Virginia House member, score of 1 .....	144
4.3.2 New York House member, score of 2.....	145
4.4 Unifying themes.....	147
4.4.1 Constituents matter most.....	148
4.4.2 Personal beliefs not far behind constituent feedback.....	152
4.4.3 Institutional leadership may matter more than elected officials admit ....	154
4.5 Evidence: what is it, how it's used, good versus bad evidence .....	158
4.5.1 Evidence matters, but there are varying views of what "evidence" means	
.....	159
4.5.2 Trying to identify "good" evidence that can be trusted .....	161
4.5.3 Social science research: "It's a huge black box." .....	163
4.5.4 Evidence rarely overruling other considerations .....	165
4.6 Colleagues do not use evidence, "but I do" .....	168
4.6.1 Use of colleagues for answers.....	169
4.6.2 Lobbyists as a proxy for evidence – "You start to rely on lobbyists" .....	170
4.7 Results First unknown to legislators – even in states where it exists .....	172
4.8 Results First: in their own words .....	175
4.8.1 Variation among the states.....	175
4.8.2 Ability for the model to "travel" .....	177
4.8.3 Selection criteria / selection bias / implementation standards .....	178
4.8.4 Time to value .....	179
4.8.5 Leading as enlightenment or through power channels .....	179
4.9 Embarking on change – how to push evidence.....	183
4.9.1 The lack of time .....	183
4.9.2 Making evidence more available .....	184
4.9.3 Executive summaries over complicated stories .....	185
4.9.4 Credible messengers .....	186
4.9.5 The power of relationships.....	188

4.10 Transition to chapter five – lessons learned, next steps, etc. ....	191
Chapter 5: Prison Break and justice reinvestment: a survey of legislative activity showing success with limitations.....	194
5.1 Justice reinvestment defined .....	196
5.2 Justice reinvestment overview: states discussed in this study .....	202
5.2.1 Pennsylvania: action taken.....	204
5.2.2 West Virginia: justice reinvestment action taken .....	207
5.2.3 Virginia: no recent activity – but plenty in the past.....	209
5.2.4 New York: action not taken .....	210
5.2.5 Rhode Island: attempts at progress stymied.....	213
5.2.6 Massachusetts: progress also thwarted .....	215
5.3 Maryland: “A step in the right direction” with Pew .....	218
5.4 Conclusion: a focused intervention in Pew’s wheelhouse with limited results .....	221
Chapter 6: lessons learned and cultivating an evidence-informed culture .....	226
6.1 Does Results First increase the use of evidence?.....	228
6.2 Other related questions .....	232
6.2.1 Can Results First help overcome institutional, policy, political, and communication challenges identified by research in the past? .....	233
6.2.2 Does Results First make legislators look more favorably upon evidence than in prior attempts to push evidence?.....	233
6.3 Evidence informed policymaking .....	234
6.4 Recommendations for Results First.....	238
6.5 Recommendations for increasing evidence-informed policymaking .....	241
6.6 Obstacles and limitations of this research.....	243
6.7 Final concluding remarks.....	245
Appendices.....	246
Appendix A: sample Results First report (Colorado) .....	246
Appendix B: updated benefit-cost rankings.....	247
Appendix C: NCSL Legislative professionalism definitions.....	249
Appendix D: Narrative interview script for Maryland pre-test, given on various dates .....	252
Appendix E: Email recruitment scripts/prompts used in the data collection process	253
If you wish to no longer receive emails from me regarding this project, please reply to this email address with “unsubscribe” in the title or body of your response.....	258
Appendix F: phone script for recorded call to legislators .....	259
Appendix F: final version of online survey .....	260
Appendix G: Olszewski – Hird comparison of means.....	269
Appendix H: Breakdown of responses on Results First questions .....	270
Appendix I: Sample interview questions mapped to research questions / hypotheses .....	272
Appendix J: Demographics of interview subjects by state .....	274
Appendix K: States and local governments in the Results First program .....	276
Bibliography .....	277

*This Table of Contents is automatically generated by MS Word, linked to the Heading formats used within the Chapter text.*

## List of Tables

- Table 1: A distribution of case study states, page 26
- Table 2: State population, median income, and percent of population 25+ with Bachelor degree, page 29
- Table 3: Partisan dynamics by state, page 29
- Table 4: Professionalism indices, page 29
- Table 5: A nine-layer view of governing, page 59
- Table 6: Survey contact calendar, page 90
- Table 7: Survey response rates, page 91
- Table 8: Survey demographics, individual level characteristics, page 95
- Table 9: Survey demographics, electoral characteristics, page 96
- Table 10: Survey demographics versus population demographics, page 98
- Table 11: Frequency distribution; factors that influence decisions – bail & gun control, page 100
- Table 12: Test of significance between bail and gun control scores on factors influencing legislative decision making, page 102
- Table 13: Frequency distribution; Legislators relying on evidence compared to select influences, page 104
- Table 14: Sources regularly consulted by legislators for legislation they are considering, 107
- Table 15: Comparison of Olszewski to Hird (2005, p. 139) means, page 109
- Table 16: Influence of evidence by type, page 111
- Table 17: Reliance on evidence over personal or constituent opinion by policy area, page 115
- Table 18: Evidence produced by organizations with formal ties to the legislature, page 116

## List of Figures

- Figure 1: Relative use of applied cost-benefit use among states, page 25
- Figure 2: RCT use in health vs. other policy areas over the decades, page 37
- Figure 3: Varying likelihoods of utilizing evidence in policy-making decisions, page 60
- Figure 4: From big ideas to specific insights regarding the role of evidence in policy-making, page 69
- Figure 5: Predicted outcomes for trustee vs. delegate with think vs. know dynamics; an evidence-informed delegate vs. trustee dichotomy, page 71
- Figure 6: Factors that can influence decision-making for elected officials, with evidence highlighted, page 74
- Figure 7: Ability of evidence and other factors to influence decision-making in the ACF, page 79
- Figure 8: Decline of swing districts over the last 30 years, page 150
- Figure 9: States in the Results First program, page 175
- Figure 10: An overview of JRI reforms between 2007 and 2016, page 202
- Figure 11: PA interactive dashboard for justice reinvestment performance, page 202 measurement, page 204
- Figure 12: Evidence-informed policy flowchart, page 228

# Chapter 1: The promise of evidence-based policymaking

## 1.1 What exactly is the deal with evidence and government?

When faced with answering important questions about the lives we live, many people hope their leaders in government stake out positions in the best interest of the people they represent. To meet these best interests, one could argue value congruence is as necessary as having political solutions be grounded in the best evidence available. Whether the questions are what level of school funding is appropriate or the best policies to reduce incarceration rates, the average citizen might reasonably expect that their elected leaders turn to available empirical evidence to inform their decision-making.

But is this reasonable expectation of evidence use borne out in practice? And what exactly does the term “evidence” mean? In conducting their work, our leaders – especially elected leaders – face any number of factors that influence their decision-making process. Whether it is their own opinion, their perceptions of constituent feedback, desires of the executive branch or legislative leadership, or other considerations, evidence is unlikely to stand as a sole (or even prominent) determinant of a legislative outcome.

These tensions were very real to me while serving as a member of the Maryland House of Delegates from June 2006 until January of 2014. Having worn both state legislators and public policy student hats, I have an appreciation of the tensions existing in the policy-making process and I also have a real passion for policy improvement, increasing the availability of quality analysis and promoting the use of strong evidence.

Competing legislative priorities was one major challenge. Limited knowledge of quality analysis and how to leverage that analysis was another. Complicating matters in this space are the ways in which the term “evidence” means different things to different people. To some, evidence means any type of information that affects opinion. To others, especially those in the academic community, evidence requires a higher standard – one that incorporates validity, reliability and that has been thoroughly vetted. I share in this latter perspective on what the term evidence should mean. Based on this view, all evidence is a type of information but not all information meets the standard to be considered evidence. The terminology matters: many state policymakers, included those I once served with in the Maryland legislature and others with whom I interacted for this project, classify even the most basic types of information (including citizen testimony rooted in nothing more than individual opinion) a form of “evidence.”

Conflating these terms can have serious implications on policy and needs to be further examined. The intention of this research is to investigate ways in which the more rigorous, higher standard of information (what I define as “true” evidence) is utilized among state policymakers. Doing so requires an understanding not only how and when the academic definition of evidence is utilized, but also how lawmakers use (or misuse, or misunderstand) the word in a way that discourages decision-making practices that are connected to the most scientifically rigorous forms of data available.

To get at these questions, I examine two approaches currently being deployed by the Pew Charitable Trusts to promote evidence greater utilization among policy-makers. I focus primarily on a Pew program called Results First, which is designed to help standardize and operationalize the use of benefit-cost analysis in governmental decision-

making. Additionally, I spend time examining what I label as Pew’s “red zone” program, which also deploys a benefit-cost approach but that operates differently from Results First, including focusing only on justice reinvestment.

This analysis was robust. It included 100 informed consent surveys and nearly 30 interviews with state legislators and relevant staff members. There are no major surprises, but I do find that Pew – both generally and through Results First – seems to be incrementally improving the policy process by making evidence a more prominent part of the policy agenda. Improvement is by no means perfection: not only does more work remain within state legislatures seeking to make better-informed decisions, the Pew evidence programs themselves can and must enact additional reforms to fully realize their potential.

## **1.2 Outline of the dissertation**

**This first chapter** provides a big picture overview of the intersection of evidence and policymaking, especially at the state legislative level. It provides a rough sketch about why it is important, and how new research will be undertaken to better understand the issues at hand. The balance of the dissertation is dedicated to going deeper into the topic.

**Chapter two** runs through some of the history on this subject and addresses the question of “how did we get here?” While there has been a proliferation of interest and an expansion of the gathering of the “best” kinds of evidence in recent years, there exists a long history of academia attempting to positively influence policymaking through their research, data points and general insights. Surveying the existing literature, the void



between the creation of evidence and its use in the policymaking process is better identified – especially as it relates to policymakers themselves.

Consequently, new exploratory research from this dissertation begins to be presented in **chapter three**. In this section, new survey research data is presented from state legislators in six states. The chapter provides some initial feedback and insights regarding connections as it relates to evidence utilization, including the identification of higher and lower users of evidence among respondents.

Case study data / interview data is explored and presented in **chapter four**, with more in-depth insights directly from individual legislators in states representing the full range in status assigned to them by the Pew Charitable trusts on the extent to which they have developed a mature benefit-cost analysis process to inform policymaking.

**Chapter five** builds on the qualitative data shared by presenting several mini case studies on state-level justice reinvestment efforts for the jurisdictions included in this research project. The area of justice reinvestment was selected as the focus for this chapter because of the ways in which the topic has permeated legislatures across the country and because of the ways in which it aligns very closely with the “wheelhouse” of Results First. In doing so, I explore the differences in approach that Pew has utilized – comparing their “red zone” approach in the justice reinvestment policy space to the more deliberate and consistent Results First model.

Finally, **chapter six** ends the research with some conclusions, suggestions for future research, and other insights on the topic of evidence utilization at the state legislative level.

### 1.3 “Evidence” Revisited

Empirical evidence originates in observation and experiences in addition to experimentation (Merriam-Webster 2015), and can include a wide array of inputs for government leaders to consider: program performance data, budget numbers, program audits, and a wide range of policy research. Information presented to legislators also comes from a diverse set of sources: partisan staff, non-partisan staff, not-for-profits, institutions of higher education, think tanks, lobbyists, and other interest groups. Finally, each state has its own set of institutional norms and routines for presenting information to decision-makers. Even if evidence were the only factor legislators considered – which it is not – there is considerable variation from among the types of information officials can choose.

Academic and political literature often argues the need for finding ways to have policymakers use more evidence. It likewise calls for a more robust discussion about the barriers that preclude the effective use of evidence and ways in which these barriers can be taken down. Hardly any of these calls for enhanced evidence use, however, have been anchored in conversations with legislators, asking if, why and how evidence is utilized in their work. Such conversations are the only way to get beyond arguing for the merits of removing barriers to evidence utilization in policy-making and moving to the question of *how* that might be done. Oliver et al. (2013) for example, conduct a thorough review of over 100 studies to identify barriers and facilitator of evidence use – but these reviews focused primarily on the perceptions of researchers themselves and the bulk of the studies were from the health field.

On September 7, 2017, the Commission on Evidenced-Based Policymaking issued its report envisioning a future where rigorous evidence is created efficiently and used to construct policy. Whether these calls to leverage the advances in technological and statistical methodologies will be followed-through by legislators remains to be seen. One measure of the likelihood of success could be the extent to which lawmakers currently rely on evidence in their work. However, with limited direct feedback from legislators, and with so many inputs on legislative decision-making, it is difficult to know to what extent our legislators currently engage in utilization of different forms of policy-relevant information and analysis, making the direct engagement of this research a meaningful contribution. This dissertation engages with state legislators in select states to better understand the individual, institutional, and state-level characteristics associated with evidence utilization. In addition to that general understanding, through an investigation of evidence utilization practices in states with and without the Results First program, it explores ways in which Results First, a more sophisticated impact evaluation mechanism tied to state budgeting practices, impacts state policymakers' views and behaviors on evidence utilization. The states identified for investigation offer substantial variation in several key categories that can inform in what ways, if at all, programs such as Results First matter.

The Pew Center on the States (2013) describes Results First as “a cutting-edge cost-benefit analysis model that identifies policy options that provide the best outcomes for citizens while improving states' fiscal health” (p. 1). It is Pew's effort to replicate the work of Washington State Institute for Public Policy (WSIPP) in states across the country. WSIPP is a non-partisan legislative research unit that undertakes robust benefit-

cost modeling on various policy options, ranging from crime to housing (Aos 2012). This replication has had considerable success: Results First is now in 24 states and in 32 individual jurisdictions (counties and cities) across the country (Dube 2017). This successful diffusion across adopting states has created a timely opportunity to investigate how the model is helping to inform legislative views on evidence utilization in the policymaking process.

Just a couple of years ago, such an exploration would have made little sense, with Results First's cost-benefit evaluation model constrained to its now-model state of Washington and a handful of states just beginning their own programs. The move toward integrating a sophisticated level of policy analysis into the policymaking process has been a long one. Despite progress over many years, questions remain about if and how it might ultimately be integrated into the deliberative democratic process. This research provides insight into the efficacy of these efforts. Considering the ways in which Results First touts their ability to meaningfully inform the legislative process is especially important within the context of a policy world seemingly on the edge of an evidence-driven approach.

Indeed, rave reviews have been showered upon the WSIPP. These accolades have come from organizations such as the Annie E. Casey Foundation and the National Conference of State Legislators (NCSL) and promoted the Pew and MacArthur Foundations to use the state as its model in formulating their Results First program for lawmakers in the rest of the country. WSIPP is cited as a "leader in the use of cost-benefit analysis" by the GovInnovator Podcast (Feldman 2013) and has been featured in publications such as *Governing* magazine for their efforts (see Urahn 2012). It is quickly

becoming *the* “go to” program for evidence based policy, having now been implemented (or planned to be implemented) in two dozen states.

Despite these accolades, and especially in light of the program’s growth, critical questions remain: Are there reasons to believe Results First can overcome the institutional, political, and communication challenges policy research has encountered in the past? Does the program make legislators look more favorably upon evidence in their policy making decisions than prior attempts? Does the existence of Results First drive lawmakers to utilize evidence more frequently as a matter of practice, and if so, in what kinds of ways? Do legislators understand what goes into the process of policy recommendations, or can they differentiate between types of evidence? Can the successes enjoyed in Washington be replicated in other states? In similar fashion, how will lawmakers across the country relate to new evidence-centered approaches to policy making? Is the evidence produced by Results First considered superior to other information produced or does it serve as one more data point to be considered during a legislative process among partisans akin to Lindblom’s (1980, p. 32) “competition of ideas?” In other words, to what extent are state-level policy makers even taking the evidence produced into account in their decision-making process?

#### **1.4 Does Pew’s Results First deliver? The research question**

An important question driving this work deals with the degree to which Results First, a program specifically designed to provide policymakers with empirical evidence, encourages them to consider it more meaningfully in their legislative deliberations and informs their decision-making. The research does not go so far as to answer whether a program can change the attitudes and behaviors of legislators, but it initiates that

discussion. The primary intent of this research is to serve as an investigation into the use of evidence by state legislators. Because the existing body of literature in this specific area is limited and the Results First program is a relatively new approach to policymaking, participant states were selected to reflect a range of current evidence utilization levels and include both Results First adopters and non-adopters.

Entering this work, the following hypothesis is set forward regarding the work of the Results First program: Results First, an evidence based policy program at the state legislative level, marginally encourages those policymakers predisposed toward the use of evidence to further incorporate evidence in their legislative decision-making, particularly in areas of policy that *do not* relate to core personal or constituent doctrine. This is a slight variation on the null hypothesis, which would suggest that the existence of Results First bears no impact on the extent to which state policymakers incorporate evidence into their decision making at the state legislative level. It should be noted that, while the emphasis of the research is on the Results First program, Pew has a longer history with another approach to pushing evidence into state level policymaking – described in greater detail in chapter five. Chapter five focuses on what I describe as Pew’s “red zone” approach to policymaking, based on Kirp’s (2007) analysis of the organization in *The Sandbox Investment*.<sup>1</sup> Chapters three and four, on the other hand, focus on the Results First program, finding marginal impacts. I find that state legislators largely behave as expected, based on the theoretical arguments I set forward in chapter two.

---

<sup>1</sup> While there are differing approaches between Pew’s Results First and “red zone” efforts, I find mixed policy success with both efforts.

Expanding upon this hypothesis, the literature (covered more extensively in chapter two) shows that there are circumstances in which lawmakers might be more concerned with evidence than in others. Similarly, research suggests that there is variation in the priorities that lawmakers give to keeping their elected positions – some are focused more on re-election while others prioritize the passage of “good” legislation. My research confirms the tensions between these and other conflicting priorities, accomplished partially by the creation of an index that measures evidence utilization that helps to begin putting a range of evidence preference in the context of these myriad factors. Equally important, this effort begins to connect not just relationships between the Results First program and evidence utilization, but a whole range of evidence utilization – for example, chapter three includes data suggesting that political party is not an indicator of evidence utilization, although ideology might be.

Finally, this research has a hypothesis (based on the limited, prior research and anecdotal experience) that the term *evidence-informed policy* is the more accurate descriptor of the role evidence plays in the policymaking process. While evidence based might be the idealized version towards which scholars and proponents work, individual preferences, institutional challenges, and political realities have been and will always be a challenging constellation of issues acting as a counterweight to complete evidence based policy approaches. More realistically, however, are the insights offered in the concluding chapter about if, when, how, and for whom evidence can play a role – and to use this knowledge to pursue a more evidence informed policy environment. As is so often the case, unpacking the answer to the questions surrounding evidence use is not immediately clear, nor evident – instead leading to more questions.

## **1.5 Other issues impacting evidence utilization**

Many of the conclusions I reach are tentative. This reflects an appreciation for the extent to which there are many factors that impact the policy calculus of legislators. Teasing these factors out is a difficult process. Exploring the Results First connection to evidence utilization must include a discussion with and about the individuals themselves and the institutions in which they work. Thus, a secondary exploration of this research examines how individual, institutional, and state-level factors might facilitate a policymaking environment that relies more upon evidence. Understanding these factors is critical to unpacking how a program such as Results First can play within a state.

In addition to exploring the impact of Results First and other institutional factors affecting evidence utilization, I also explore what individual characteristics might also affect the use, or extent of use, of evidence in policymaking. We know from the literature that any number of individual and institutional factors can influence policymaking outcomes, but do they have a relationship to how evidence is viewed, valued, and used?

It is the goal of this dissertation to take a deeper dive into how these programs interact with individual policymakers. While this research cannot (and did not) exhaustively answer these questions, it offers additional and timely analysis into the realm of evidence utilization by policy-makers. Before getting further into this discussion, it is necessary to begin with some background information on the primary variable of focus: Results First.



### 1.6 What is Results First?

As Pew (2017) explains the program, Results First seeks to “implement cutting-edge cost-benefit analysis tools to help identify options that provide the best outcomes for citizens while improving the states’ fiscal health.” That idea seems to be catching on, with 23 states signing up to partner with the Results First initiative since 2011 (Pew 2017). These partnerships signal a meaningful commitment to begin, continue, or grow rigorous empirical studies and cost-benefit analysis principles in states’ policymaking processes.

The Results First approach takes the WSIPP cost-benefit matrix and adjusts inputs based on state-specific data, calculating a potential return on investment of varying policy options, risk level associated with those options, and creating a dashboard for policy makers to consider in eliminating, changing, and creating programs (Pew 2017). The Results First program specifically seeks state partners in the issue areas of criminal justice, education, child welfare, substance abuse, mental health, healthcare, public assistance, housing and teen pregnancy prevention programs (Pew 2017). Results First engages interested states and facilitates a review of a state’s analytical and data capabilities and partners with the policy work groups established in each location tasked to oversee implementation of the analysis work (Pew 2017). The partnership costs collaborating states nothing other than the staff time necessary to collect and analyze data, but the Pew and MacArthur team does require a meaningful commitment from legislative leaders to pursue the new approach.

According to the program, the analysis model has enabled legislators to make policy decisions on evidence, overcome partisan gridlock, and invest in cost-effective programs over politically popular considerations (Pew 2017). While the Results First analysis process (including their analytical models) are considered proprietary material and closely guarded, we do have some guidance using the experience of the WSIPP model, upon which Results First was founded. In their evaluations, WSIPP takes the following steps in conducting evidence-based reviews: 1) all available studies on a chosen subject are selected for review; 2) research designs meeting certain criteria, including appropriate treatment and comparison groups are forwarded to the next stage; 3) from available studies, meta-analysis is conducted to obtain an average effect size to determine if outcomes are, on average, realized (Lee et al. 2012).

Once analysis has been completed, reports are generated for each subject area investigated within a state. These reports typically come as both an exhaustive review of the analysis conducted along with an executive report. The reports vary by state, but all of the executive summaries include a *Consumer Reports*-style comparison of benefits and costs for each individual program studied. For a sense of what these reports might look like, an example from a Colorado juvenile justice system review is included in Appendix A of this chapter. These analyses all provide information on what makes economic sense (projecting costs and describing the economic benefits of a project or policy) along with tests of robustness via Monte Carlo simulations (Lee et al. 2012).

The Results First program has many positive qualities to offer to state legislative leaders. Primary arguments in favor of the program are its delivery of evidence that meets higher-quality review standards and has limited (if any) partisan influences and leanings,

and its service as a mechanism to overcome previous policy concerns. For the six states that have fully implemented their Results First model, the program claims that \$38 million has already been shifted, cut, or allocated in new ways and there is an expected return of \$38:\$1 over the next decade (Pew 2014, p. 2). Whether these returns will come to fruition and if these programs will change the way lawmakers perceive evidence in policy making remain open questions.

On the other hand, there are reasonable questions to be asked about how effectively the analysis mechanism fully compensates for variation among states. Moreover, the selection criteria for the Results First program suggest that there are states that might not fit for any number of reasons. Among these are those that are both unable and unwilling to enter a partnership. This begs the question – is there an inherent selection bias when the program overlooks those with low capacity and/or implementation ability?

Users of the program have plenty of other questions that they can and should be asking, too. Among them are: How do the assumptions on the WSIPP model “travel” into other states when implemented through the lens of Results First? At what level of confidence can we make assumptions? Does the Results First model presuppose a certain implementation standard, both in terms of quality and consistency? How long should policymakers reasonably have to wait after changes in the law to see results in outcomes? Finally, how can an abundance of quality information on a subject make decisions be based, in whole or in part, upon that evidence? These questions are addressed in the proceeding chapters, though not all are able to be answered.

Thus, it is not suggested that Results First is without limitations or findings from it should be embraced whole cloth; quite the contrary, the best legislators will consider possible limitations and do their best to bolster their understanding of respective problems by exploring what else is and can be known about the problems facing our states and our country. The debate can (and should) be sparked about the quality of Results First and similar programs. This is especially true as Pew and MacArthur's efforts continue to spread, with five states using the Results First model over fiscal years 2013 and 2014 to either re-direct or create \$81 million of new government spending (Pew, November 2014). That debate needs to happen, if it is changing, even marginally, the way in which elected officials approach their role of shaping policy outcomes.

### 1.7 Alternative approaches: Virginia's JLARC as an example

With widespread buy-in to the Results First model that relies heavily upon conducting meta-analysis and benefit-cost calculations on vast clearinghouses<sup>2</sup> of information, there exists an important question regarding alternative approaches. Should there be a widespread extension of influence over the policymaking process led by the Results First charge, or are there alternative approaches that engage lawmakers with quality information and data and still lead to positive policy outcomes?

---

<sup>2</sup> The Pew Clearinghouse contains information from eight national clearinghouses, having "reconciled the difference systems and vocabularies" and providing "the data in a clear, accessible format." The clearinghouses included are: Blueprints for Healthy Youth Development; California Evidence-Based Clearinghouse for Child Welfare; Coalition for Evidence-Based Policy; US Department of Justice's CrimeSolutions.gov; US Substance Abuse and Mental Health Services Administration National Registry of Evidence-based Programs and Practices; Promising Practices Network; US Department of Education What Works Clearinghouse; What Works in Re-entry Clearinghouse (Pew June 2015).

The 1960s saw the beginning of the legislative professionalization movement. Most reforms to legislative staffing occurred in the 1960s and 1970s, with additional incremental changes occurring since the 1980s (Squire and Hamm 2005). In Virginia, one change during the legislative professionalization movement was the creation of the Joint Legislative Audit and Review Commission (JLARC). JLARC was authorized in statute by the Virginia General Assembly in 1973, with subsequent responsibilities and authority added through the Code of Virginia and Appropriation Acts (JLARC 2017).

Virginia's JLARC is one such alternative approach to WSIPP, where it acts as the oversight mechanism of the state's General Assembly. Like WSIPP, JLARC has a full-time staff and a budget supported by the state legislature to conduct independent oversight. members of the House and Senate serve either on the JLARC Commission or the board governance of the WSIPP. JLARC's objectives are program/agency savings, improved efficiency and effectiveness, an informed legislature, and compliance with legislative intent (JLARC 2014).

There are, however, still clear differences between JLARC and the Results First model. While Results First seeks to put research first and layer state-specific data upon it to forge recommendations for policy actions, the JLARC (2014) research process focuses more on agency-specific interviews, data collection, review of financial records – and utilizes literature, other research, and secondary data as partial inputs – a basic but potentially critical distinction. Despite these differences, the collaboration between the Pew Charitable Trusts and the MacArthur Foundation find that Virginia, like Washington (the model state for Results First) are among the few states that are “leading the way”

when it comes to implementing benefit-cost analysis within their jurisdictions (Pew 2013).

### 1.8 The leading model, or the only model of evidence-based policy?

Across the options available to states when it comes to engaging with benefit – cost analysis, it appears that Results First is both a leading model and the only nationally available model. While some states are engaging in benefit – cost evaluation activity, they are doing so through and with their own direction. As discussed above, Virginia is perhaps the most obvious example of a state’s independent efforts, as the state produced an impressive 13 benefit – cost evaluations to help inform policy between 2008 and 2011 (Pew July 2013).

Virginia’s JLARC provides useful information and evidence like what is provided by a “sophisticated” Results First review. Even within states that do not benefit from a robust in-house evidence operation such as JLARC or Results first, they do often have access to good data and there is rarely a shortage of researchers willing to share their work for utilization in legislative affairs. Nor is it necessarily the case that only legislators working with formal evidence-based programs (such as Results First) view evidence as useful in their work. Figuring out the intersection of these competing methods of evidence production and understanding the extent to which evidence-production programs matter in the state legislative process, is a gap in knowledge toward which this work modestly contributes.

Still, as a leading model, it is logical to want to understand exactly what role Results First is playing in connecting policymakers to evidence. Given the variation in

states that have engaged with the program and an equal variation of states that have engaged in evidence utilization regardless of whether they have a relationship with Results First or not, this research adds to the existing knowledge around evidence utilization. The importance of embarking on that research is described next.

### *1.9 What's missing? Filling a gap in knowledge – why does it matter?*

While the general topic of evidence use in policymaking has been a subject of interest for a long time in the social sciences, research on this connection at the state legislative level is needed, as comparatively little work has been conducted in this area, leaving a research gap. It is also true that, despite best efforts, a sizable divide remains for scholars hoping to have the evidence they create to influence public policy (Hird 2009). This fact was brought further to light when policy practitioners from across the world assembled in Ile-Ife, Nigeria for a conference on Evidence-Informed Policy Making. At the conference, participants believed the following: evidence played a small role in policy decisions, research was needed to better understand the ability of policymakers to use research, and that links that could better connect evidence to policy remained largely missing (Newman et al. 2013). While this dissertation addresses just a small portion of these needs, it certainly helps make progress on two critical questions. First, it generally adds to the limited research about evidence utilization habits of state legislators. Second, it helps provide a more complete understanding of how evidence intersects with other individual, institutional, and state level-factors that might influence legislative decision-making – specifically investigating whether the most prominent evidence-based program in the country, Results First, has any connection to evidence use by state legislators. By

Results First's own admission, there have been a few assessments of the program that have been done for the Pew and McArthur Foundations, but none of this data has been released and there have been no public evaluations of the program (Dube 2017).

While there has been considerable work identifying the myriad factors facing legislators in their decision-making (including evidence), there is a much smaller set of literature available examining actual use (Jennings and Hall 2011). As to the second question about the impact of Results First and other evidence-creation programs, there have been no comprehensive efforts to measure the impact that these have had on the end user (state elected policymakers), either in term of their perceptions or actual usage of the information. Considering the rate at which the Results First program is becoming a national product influencing state legislative outcomes, this work represents an important first step in that direction.

Indeed, if one believes empirical evidence should be used to drive our policy-making decisions, it is critical that we (a) understand the extent to which state legislators believe that the use of empirical evidence is important in their decision-making, (b) understand the extent to which they use evidence in their decision making, and (c) identify those things which both encourage and discourage its use. By exploring these areas of concern, this research proposes to fill a gap in knowledge about if, how, and why good empirical evidence is (or is not) integrated into important public policy decision-making processes at the state level for elected policy-makers.

The quality of "evidence" is also a critical question in this investigation. Those studying the policymaking process rightfully continue to scrutinize the quality of the mechanisms producing evidence in the years ahead, even as members of the public (and



even those involved in the policy making process) fail to differentiate between the methods of evidence production. These conversations are important, and even more so if programs such as Results First are associated with attitudes and behaviors in the policymaking world. As has been discussed, there is a wide range of available empirical evidence that might be provided to legislators, notwithstanding the fact that some believe higher standards (such as RCTs and meta-analysis) should be more of the baseline standard for what is utilized in the decision-making process. Further complicating matters, there is also a wide range of what individuals and lawmakers consider evidence.

Recognizing the differences in the quality of these approaches, for purposes of this study, and with the caveats noted elsewhere, the assumption used along with this research is that, if it works as promoted, the Results First program can be a preferable method of incorporating evidence into policymaking than alternatives<sup>3</sup>. As this research helps confirm, such a standard is often an unrealistic expectation for legislators both for practical purposes, including the time constraints under which lawmakers work, as well as the ways in which it is difficult for non-academics to understand the contours of differences between sophisticated research methods and basic data reporting.

Sometimes the methods deployed, such as randomized controlled trials, “may have limited value in some cases because they may take years to complete, and may not provide a sound basis to apply a particular intervention on a larger scale or to groups with different characteristics than those tested originally” (Hoagland et al. 2017, p. 5).

However, to the extent that Results First and programs like it are being utilized for

---

<sup>3</sup> Results First is the only known organization that engages in a comprehensive evidence generation process that includes rigorous methodologies, a benefit-cost approach, and the evaluation of several policy alternatives. Individual parts of this work *is* conducted by various state policy organizations from time to time, but I found no organization that is a complete match.

legislative decision-making, the more sophisticated mechanisms of evidence production will be given more prominence. There are many mechanisms across state legislatures that are producing outputs considerably more thoughtful and thorough than those released even a decade ago. To the extent that the more rigorous approaches championed by Results First are changing lawmakers' worldviews on using evidence to inform their decision-making, a more in-depth exploration of that program is a logical place to start in understanding the impact it (and perhaps others) is/are having on the views and practices of elected leaders. I turn next to how that exploration is structured for this research.

#### 1.10 Study structure: methodology and research design

For this study, analysis was completed in six locations, with the states selected (New York, Virginia, Massachusetts, Pennsylvania, Rhode Island, and West Virginia) providing sufficient diversity in two key categories: (1) acting as a Results First state and (2) the extent to which they are rated in connecting benefit-cost evaluations to their policymaking processes. These states also provide interesting variance in terms of population, ideology, and technical capacity.

As Results First is a relatively new initiative, there are few policy places in which its analytical methods have been deployed – even in Results First states. To that end, the research focused on issues of criminal justice, the area in which Results First has the most experience, in which the case study states have engaged with Results First, and in which all six states have legislative oversight. Choosing a topic that enjoys consistency across states mirrors the case study efforts of Moynihan (2008), who engaged in a similar process when he explored performance management issues. Given the limited research

conducted on evidence utilization by state policymakers, as well as the similarities between performance management and evidence-based policymaking, I believe that Moynihan's research design offers a solid framework from which to begin my own work. Like Moynihan, states in this research were selected in such a way as to represent the full range of performance standards (high, middle, and low).

This research deploys a mixed-methods approach leveraging both qualitative and quantitative methods. Such an approach includes interviews of key informants in the selected states as well as a survey instrumentation delivered to the entire population of lawmakers within each state. The case for combining research methods, particularly when both quantitative and qualitative methods are present, is strong (Gable 1994).

Such a design offered a variation of six systems, with a potential N of varying numbers within each state: New York (150 House and 63 Senate members for a total possible N of 213), Virginia (100 House and 40 Senate members for a total possible N of 140), Massachusetts (160 House and 40 Senate members for a total possible N of 200), Pennsylvania (203 House and 50 Senate members for a total possible N of 253), Rhode Island (75 House and 38 Senate members for a total possible N of 113) and West Virginia (100 House and 34 Senate members for a total possible N of 134). Across all the six states, an overall N of 1,053 was possible, though a relatively low response rate was expected and realized<sup>4</sup>.

The dissertation also doubles the amount of states employed for case study comparison by Moynihan. While he used three, this research attempted to interact with six. This was done to allow for in-depth exploration of Results First and non-Results First

---

<sup>4</sup> Details on response rates and demographics of the survey population are outlined in more detail in chapter three.

states that also represent all three levels (high, medium, and low, as used by Moynihan) of applying benefit-cost principles to policymaking. This connection is a critical distinction, as making this connection is precisely what Results First claims to do in the states with which it works. Several individual, institutional, and state level factors were examined as a part of the research. The key independent variable of interest, an institutional factor, was the existence of Results First (as of June 2014).

Select insights from the state of Maryland are also included in later chapters – namely, qualitative data is deployed in chapters four and five. Maryland was excluded from the quantitative analysis of chapter three, as I was concerned that my proximity to the lawmaking process could skew the results from legislators – that they would be even more likely to provide socially desirable answers. However, legislative and staff informants from Maryland were intentionally selected among those whom I believed, based on personal experience, would provide honest answers about their experience with evidence in the legislative process. These interactions not only helped to shape the development of the survey instrument utilized in the six other states, but were used to validate insights from other lawmakers and staff in chapter four and add an additional case study to the justice reinvestment review conducted in chapter five.

Literature suggests that many are factors worthy of examining for impact on evidence utilization. These factors, discussed in greater detail during chapter two, are analyzed to the extent available data permitted such examination<sup>5</sup>. At the individual level, these include respondent gender, ideology, party affiliation, number of years in office,

---

<sup>5</sup> Some of the variables described were unable to be evaluated within the context of this research. The most common reason preventing such analysis was unavailable data (for example, there was no consistent way to get a measure on number of legislative staff) or an insufficient population so as to have sufficient statistical power (such as there being only one presiding officer / committee chair within the survey sample) to run analysis.

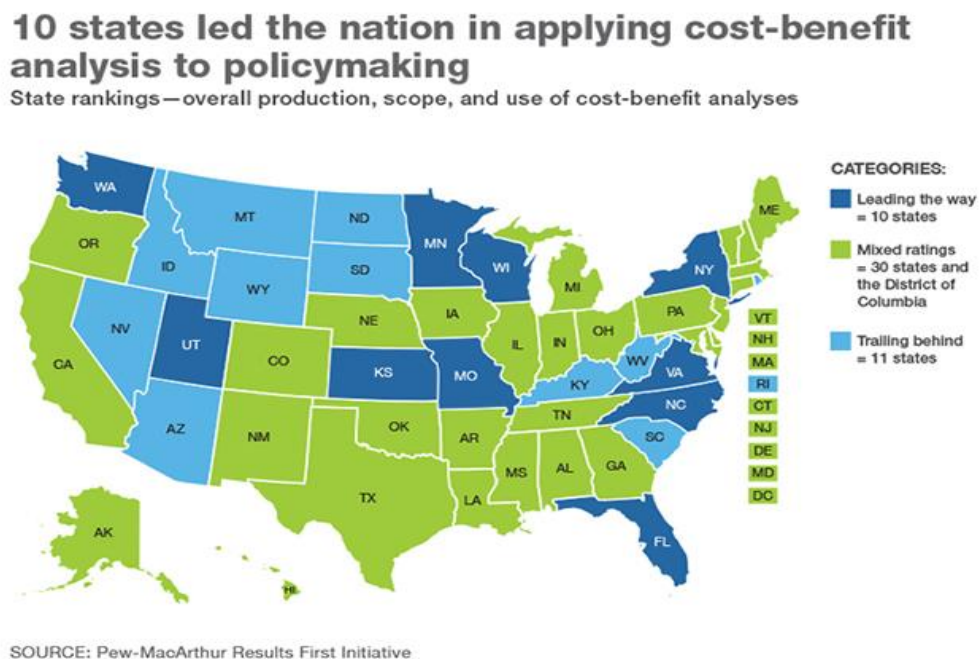
whether a presiding officer or committee chair, legislative productivity, educational attainment, issue salience (whether or not a member serves on the committee of jurisdiction for criminal justice issues) and whether the respondent will seek election in the next cycle for any office. At the institutional level, in addition to the key variable of the Results First program, the extent to which one serves in a divided government, number of staff, legislative professionalism (as determined by the National Conference of State Legislatures) and the extent to which the state connects benefit-cost analysis with their state policymaking (a ranking from the Pew Center on the States). State level factors of education level and per capita income are also considered, as are variables such as if a member is within the majority party in their own chamber. The dependent variable against which the Results First and other independent variables will be measured is the ranking on an evidence utilization scale, which is described in more detail in chapter three. To the extent other factors appear to impact evidence utilization at the state legislative level (based on survey and/or interview data), they are discussed at greater length.

### 1.11 Details on case study selection rationale

Each of the survey and interview states have been selected not only for their relative relationships to Results First, but also to how they are identified when it comes to applying a cost-benefit approach to policymaking, predicated upon the use of good evidence (Pew July 2013). The Pew Trust ratings, which were categorized as “leading the way” (10 total states), “mixed ratings” (30 states and the District of Columbia), or “trailing behind” (11 states), were developed based on three criteria. Those criteria were

the number of benefit-cost analyses produced, if the studies assessed multiple program alternatives, and if the results of those studies were incorporated into budgeting and policy decisions. Researchers made their determinations of category through a systematic searching of benefit-cost assessments by state, as well as interviews with legislative and program staff, executive officials, writers of the reports, and other agency personnel (Pew July 2013). In other words, this is Pew’s own determination of the extent to which states are engaging in the type of work being promoted by its own Results First program – regardless of if that work was being done because of the program.

*Figure 1: Relative use of applied cost-benefit use among states*



This research has strong variation between Results First versus non-Results First states as well as where states are along the cost-benefit analysis continuum. For example, New York and Virginia are both considered to be “leading the way” when it comes to the type of benefit-cost valuations Results First supports, although Virginia is not a “Results

First” state. New York also was the only state other than Washington (where the Results First model was created) to be leading in all three of the criteria assessed (Pew July 2013). While Virginia has not used the Results First program in any area (including criminal justice), they are among a select handful of states that require specialized criminal justice fiscal notes for any proposed legislation that either increases or decreases the length of criminal sentences (Leachman et al. 2012). Massachusetts is another Results First state, but considered to have “mixed results” in their benefit-cost efforts, while Rhode Island is considered behind in its use, even though it too is a Results First state. Similarly, Pennsylvania is considered to have mixed results even though not a Results First state, and West Virginia is a non-Results First state that scores low on the application of benefit-cost evaluation to policy. In doing so, this research explores all levels of the type of benefit-cost evaluation Results First promotes in both adopter and non-adopter states. Such an approach helps to “maximize theoretical replication” and to “avoid the tendency of best-practice research to generalize from high performers” (Moynihan 2008, p. 22)<sup>6</sup>. While not intentional, it also happened to be that the states utilized for the study were conveniently located, creating a regional control.

*Table 1: A distribution of case study states*

	Adopted Results First	Did Not Adopt Results First
Leading the way applying benefit-cost evaluation to policymaking	New York	Virginia
Mixed results applying benefit-cost evaluation to policymaking	Massachusetts	Pennsylvania
Trailing behind applying	Rhode Island	West Virginia <sup>7</sup>

<sup>6</sup> The Pew Foundation issued a new version of this map measuring the extent to which states deploy evidence-based policymaking. Details on this new view of the states, released in early 2017, can be found in Appendix B.

<sup>7</sup> West Virginia announced that it would begin the process of implementing some Results First practices in July 2014. For purposes of this study, data was collected prior to any actual use of

benefit-cost evaluation to policymaking		
---	--	--

Another slight variation from the Moynihan model of research was a state's "deployment mechanism." While there are any number of similarities between evidence-based policymaking and performance management, they are not the same (these distinctions are discussed at greater length in the literature review). Moynihan's work used three states to compare their experiences with performance management practices – something that, even if only to minimal levels, was occurring in *every* state. Results First, however, is in some states and not in others – and there are varying levels of the benefit-cost valuations connected to policymaking occurring in both adopter and non-adopter states. Thus, getting the most robust comparison about if and how Results First is correlated to higher levels of evidence utilization in policymaking was believed to be best accomplished by considering the experiences of adopter and non-adopter states at all levels.

In seeking to model this research after Moynihan, Seawright and Gerring (2008) offer useful ways to think about case selection and sharing methods to achieve purposeful modes of sampling. As they report, case selection in case study research shares the same research objectives as random sampling, creating a representative sample but also producing useful variation in the areas of theoretical interest (Seawright and Gerring 2008). In the case of this research, attempts were made to select states that represent the full range of benefit-cost capacity as tied to legislative work, which is the model and approach for Results First – consistent with Seawright and Gerring's diverse method.

---

Results First products in the decision-making of policy, which afforded the opportunity for state to remain as a "non-Results First state" for research purposes. Data presented in chapter four regarding the stalled implementation efforts of Results First in this state following the 2014 elections confirmed this decision.



Specifically, this case selection addresses two key diversity “values.” This first value is the use of Results First, and the other variation is a commitment to linking benefit-cost valuations to policymaking. Due to the robust selection of these six cases, valuable exploratory work can be conducted to better understand the full variation of these two characteristics relative to each other. Finally, the diverse method of case selection is one of the best mechanisms to be able to provide some level of representativeness of the full variation in the population (of all the states), with the important caveat that the case selection might not mirror the distribution of that variation (Seawright and Gerring 2008).

A second case selection strategy has as its primary objective the achievement of maximum variance along relevant dimensions. We refer to this as a diverse case method. It requires the selection of a set of cases—at minimum, two—which are intended to represent the full range of values characterizing X, Y, or some particular X/Y relationship. The investigation is understood to be exploratory (hypothesis seeking) when the researcher focuses on X or Y and confirmatory (hypothesis testing) when he or she focuses on a particular X/Y relationship. Where the individual variable of interest is categorical (on/off, red/black/blue, Jewish/Protestant/ Catholic), the identification of diversity is readily apparent. The investigator simply chooses one case from each category. For a continuous variable, the researcher usually chooses both extreme values (high and low), and perhaps the mean or median as well. The researcher may also look for natural break points in the distribution that seem to correspond to categorical differences among cases. (Seawright and Gerring 2008).

Given the limited literature on evidence based policy, especially among legislators, this research is first and foremost exploratory – investigating how lawmakers view and interact with Results First (or variable “X” in the above passage). There are, however, also two elements of hypothesis testing, leveraging a mixture of quantitative and qualitative data points to understand evidence utilization (variable “Y” in the above passage) among state policymakers, especially in relationship to Results First (again, “X”) and any number of other independent variables.

As Results First is categorical (exists in a state / does not exist in a state), the diversity among the states is apparent. However, this dissertation also mixed in all three levels of benefit-cost analysis relative to Results First – including not only the high and low extreme values mentioned above, but the middle value as well – and doing so in both Results First and non-Results First states.

Beyond the existence of Results First, other types of variation exist within the state in the areas of population, educational attainment, income, partisan dynamics, and various factors influencing the level of professionalism within state legislatures. These variables help demonstrate the ways in which the states selected can be representative of the larger body of legislatures across the country. There is no perfect combination of cases that would allow for absolute external validity. Quite the contrary, it might be that state legislative processes are more diverse than most other enterprises within states. However, this distribution of states does engage the most critical variations – a mix of states engaged and not engaged with the Results First program, as well as all the varying degrees to which those states deploy benefit-cost work to their legislative activities. In doing so, there is also considerable variation in other factors that could influence evidence utilization (including of the Results First kind) and opportunities to argue that the insights from these cases might be able to be applied in other locations. Information about some of these differences is highlighted in the tables below.

*Table 2: State population, median income, and percent of population 25+ with Bachelor degree*

	Population	Median Income	Bachelor Degrees
New York	19,651,127	\$58,003	33.2%
Virginia	8,326,289	\$63,907	35.2%
Massachusetts	6,745,408	\$66,866	39.4%
Pennsylvania	12,787,209	\$52,548	27.5%
Rhode Island	1,055,173	\$56,361	31.3%
West Virginia	1,850,326	\$41,043	18.3%

Data source: US Census (2015)

*Table 3: Partisan dynamics by state*

	Governor Party	House Partisan Balance	Senate Partisan Balance
New York	Democrat	105 D / 44 R / 1 I	33 R / 25 D / 5 IDC
Virginia	Democrat	67 R / 32 D / 1 I	21 R / 19 D
Massachusetts	Republican	123 D / 35 R / 2 vacancies	34 D / 6 R
Pennsylvania	Democrat	119 R / 83 D / 1 vacancy	30 R / 20 D
Rhode Island	Democrat	63 D / 11 R / 1 I	32 D / 5 R / 1 I
West Virginia	Democrat	64 R / 36 D	18 R / 16 D

Data source: National Conference of State Legislatures (2015)

*Table 4: Professionalism indices*

	Permanent Staff ('09)	Squire Professionalism Index ('03)	Type of Legislature ('14) <sup>8</sup>
New York	2,676	0.480	Full-time, well paid, large staff
Virginia	391	0.131	Hybrid
Massachusetts	903	0.385	Full-time lite
Pennsylvania	2,918	0.339	Full-time, well paid, large staff
Rhode Island	284	0.133	Part-time lite
West Virginia	219	0.125	Part-time lite

Data source: National Conference of State Legislatures (2009 and 2014) and Squire (2007)

In 2009, the range of permanent staff for legislatures was 32 in North Dakota up to the high of 2,676 in New York. The 2003 Squire scores measuring legislative professionalism ranged from 0.064 in South Dakota to 0.626 in California.

### 1.12 Data analysis methods

Structuring the research and collecting data is important, but so too are the methods used to analyze that data. In this project, the use of scaled data was an important

---

<sup>8</sup> Refer to appendix C for more information on what these terms mean, as identified by NCSL.

feature of the survey instrument and subsequent data analysis. Its use allowed the researcher to compare interview responses regarding the use of evidence against the priority ranking of situations where evidence is weighed against other priorities in the policymaking process. More specifically, the survey portion included considerable use of slider questions, which creates a continuous scale upon which respondents can place themselves rather than feeling forced into a category or into checking a box, which leads to better data (Thevenau 2015). The survey analysis includes analysis on the many variables of interest that are included at the individual, institutional, and state level. To test the extent to which high salience and issues of greater partisan divides lessen the likelihood evidence is utilized (as suggested from the literature), the online survey attempted to include opportunities for lawmakers to give responses on examples of both a highly-charged and salient topic (the state's gun laws) as well as an issue that could be somewhat less controversial and partisan (changes to a state's bail requirements). This approach neared but did not fully adopt a paradigm from the cognitive sciences, providing a small embedded experiment into the survey work (Fienberg and Tanur 1989) – an approach that has been successfully executed in similar research conducted with city managers (Matkin and Frederickson 2009). A more complete embedded experiment would have presented the issues as a “split ballot” between randomly assigned legislators. Due to the anticipated risk of a low response rate, this approach was not utilized. However, it could be leveraged in future research that has fewer questions and/or a larger sample size.

While a more complete experimental design was considered, it was ultimately decided that the information would most accurately be drawn through survey research methodologies. Quasi-clinical experimental designs have the benefit of eliciting the most

honest responses from participants (in that they are responding to situations that more indirectly measure their reliance upon evidence, rather than relying upon only direct questions about their use of evidence). This type of approach, however, also relies upon the writer to interpret *why* elected officials responded in certain ways, or would still require interview or survey follow-up to ask for direct questions about the decision-making process.

Similar research work exploring the use of evidence (Jennings and Hall 2011; Talbot and Talbot 2014) has been conducted via survey methodologies. While the available work focused on agency heads and this research emphasizes elected officials, the information sought from those being asked remains essentially the same. As such, without any clear examples of a quasi-experimental design that can effectively solicit and interpret data on evidence use, survey responses were selected to inform this research.

Surveys clearly limit the ability to know that elected officials are using evidence in their work (as discussed above, the data would be limited to what these policy makers *report* as their use), but some validation of the responses from the survey can be accomplished through the work of the case studies, where in-depth interviews with those associated with the policy making process (but who are not elected officials themselves) can provide opportunities to clarify how evidence is used in a given state.

Once the data is collected, it was analyzed using both qualitative and quantitative methodologies. Case study interviews were recorded and coded by hand; survey data was placed into SPSS where analysis was conducted on the responses. Both quantitative and qualitative data is presented individually and then considered collectively in the conclusion of this work.

Before turning to any of the new data, however, I first review the relevant literature to focus and ground the conversation on the topic of evidence.

## Chapter 2: Where have we been and where are we going: a literature review

Before jumping into the data from this study, it is important to understand the journey public policy scholars and government administrators alike have taken in relationship to the role of evidence. In conducting such a review, one finds that there is consistent – and growing – support for the greater utilization of evidence in government over a span of nearly 70 years. One also finds, however, that despite strong support for evidence utilization in policy making, crucial questions remain as to the actual application of available evidence.

Indeed, as is so often the case in public policy, the ideal does not match with the reality. This chapter seeks to examine the ways in which the use of evidence in the public policy sphere has changed over time, helping to answer the question of “where are we?” and “how did we get here?” Such a review sets the stage for the introduction of the new research with state legislators regarding their evidence utilization experiences and perspectives.

### **2.1 Evidence and policymaking: a steady drumbeat starting in the 1960’s**

While the recent spike in activity related to evidence-based policymaking could give the impression that evidence based policymaking is a new call for reform, the truth is that it represents a movement that has been over half a century in the making. The 1960’s marked the beginning of what has become a steady drumbeat calling for more “practical knowledge that can help policy makers make informed decisions” (Haskins et al. 2009, p. 1). Rivlin (1971), for example, released her seminal writing that called for

both a “widespread implementation of social experiment and acceptability by the federal government” in such a way to measure effectiveness and drive up accountability (Brookings 2015). In more recent years, the social science and management communities have continued the push toward the use of evidence to inform policy decisions.

Such a push for social experimentation has remained relatively consistent, although the available mechanisms for their conduct continue to become more refined and technical along the way. Examples of such evidence-based efforts abound domestically. They are also becoming increasingly sophisticated abroad. In the United States, numerous think-tanks, non-profit organizations, institutions of higher learning, and government agencies are coming together to create mechanisms to better inform policy practitioners using empirical data. One such example is the development of the Institute of Education Sciences (IES) in the United States Department of Education. The mission of the IES is to “provide rigorous and relevant evidence on which to ground education practice and policy and share this information broadly” (“About IES” 2014). Quality research is pointed to as a key component of bridging the political gaps which are inherent in a political system – bringing, for example, the “Witch Doctors” of social science together and delivering on comprehensive welfare reform efforts (see Baum 1991)<sup>9</sup>. Meanwhile, in the United Kingdom, an even more elaborate and robust system – the “Evidence Information System” (EIS) – is in the making and intends to inform policy efforts. As its sponsor, British newspaper *The Guardian* writes, the EIS is focused on making sound evidence the cornerstone of policymaking, when appropriate, by

---

<sup>9</sup> Baum specifically points to the important role of research to bring the “Witch Doctors” of social science to agreement on the welfare reform legislation Family Support Act of 1988, a feat he claims contrasted with the 1970’s. At the time, he opined that “whether this marks a new phase in the connection between social policy and research remains uncertain” (Baum 1991, p. 603).



“widening access to research evidence and injecting it at the ground floor of policy making” (Chamber and Lawrence 2014).

There has been a proliferation of information available to policymakers in recent years as not only methods and measurements have continued to improve, but also as interest continues to grow. This information enters the debate from many directions. Research involving observation and qualitative methods remains, as do more traditional research methods such as matching, differences in differences, and regression discontinuity designs (RDDs). All the while, a considerable rise in the use of randomized control trials (RCTs) – a method scholars have increasingly advocated for the use of in policy-making settings – has taken place. RCTs are one type of program evaluation method that randomly assigns subjects to one or more treatment groups and others to a control group, followed by a comparison in change between the two groups (CRS 2006). The Congressional Research Service (CRS) reports that “there is wide consensus that, under certain conditions, well-designed and implemented RCTs provide the most valid estimate of an intervention’s average impact for a large sample of subjects, as measured on an outcome of interest” (2006, p. 13). Furthermore, coupling RCTs with the practices of systematic review and research synthesis, where multiple studies are combined to determine program effectiveness, has only helped facilitate the belief that a new day of evidence-based policy may be on the horizon (Pawson 2006).

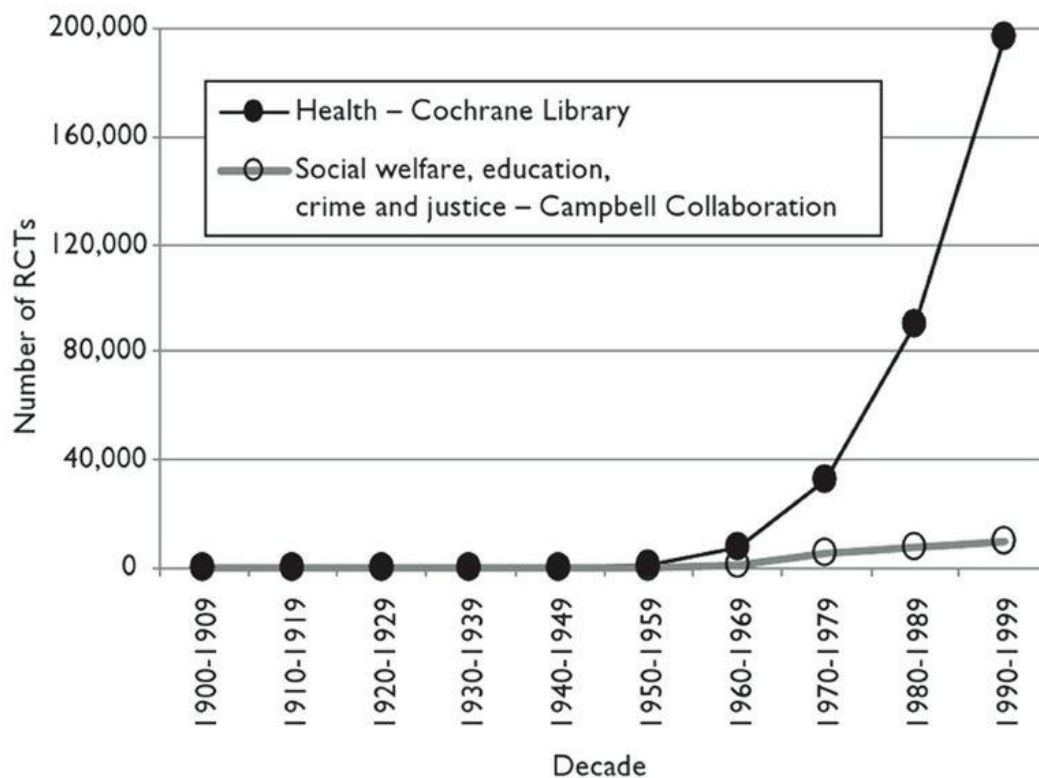
As important as the new push for experimentation might be, more “traditional” forms of evidence will continue to be produced and will continue to play important roles in the policymaking process. In fact, considerable evidence production continues to be conducted and held in high regard to inform the policymaking process that remains

“empirical” but short of the “gold standard” associated with randomized controlled trials. Maryland’s Department of Legislative Services (DLS), as a state government example, has a robust nonpartisan staff of over 350 people that provide “legal, fiscal, committee, research, reference, auditing, administrative, and technological support to the members of the legislature and its committees” (DLS 2015). About half of these are roles that relate to evidence production. According to Warren Deschanaux, Director of Policy Analysis at the Maryland Department of Legislative Services, there are about 100 staff members assigned to legislative audits. As it relates to the state’s legislative session, there are 23 professionals assigned to budget analysis and another 25 (including managers) who complete fiscal notes on legislature. In addition to these 148 staff, there are another eight who are there for support, for a total of 156 staff members dedicated toward some sort of evidence production, which Deschanaux (2017) defined as “producing information to be relied upon in decision-making through written documents.” On *every* piece of legislation considered by the state’s legislature, DLS creates a fiscal and policy note that summarizes the bill, provides a five-year projected fiscal impact on state and local governments, reviews existing law, and offers an expected economic impact on small business (DLS 2015). While not grounded in experimentation of any sort, the data and observations offered in the reporting is an important part of Maryland’s legislative process – and an important contrast to traditional policy evaluation efforts, which are largely retrospective.

Traditional forms of evidence, such as Maryland’s DLS, are important inputs to the policy process, as the number of RCTs in the social welfare, education, crime and justice categories – all areas subjected to considerable public policy intervention – remain considerably smaller compared to their health counterparts. As seen in the figure below,

there are approximately 10 health RCT studies produced for every 1 in *all* of the other fields. As seen below, there has been substantial growth in the number of RCTs performed in social welfare, education, crime and justice, but that it has also not experienced the type of exponential attention that health industries have offered to the approach. Perhaps for this reason it has been observed that public health is “the only area where evidence-based policy and practice has flourished” (Jennings and Hall 2011, p. 248).

*Figure 2: RCT use in health vs. other policy areas over the decades*



Source: Belluz and Hoffman (2015)

One might expect that “evidence” will continue to be viewed as and pulled from testing and practical experience alike in the years ahead. However, it also appears true that the calls of Rivlin and those following her are becoming more of a reality: widespread experimentation coupled with accountability measurements are the norm now

more than ever. Moreover, it may be that improvements to research design (such as the increased use of RCTs) coupled with study replication and systematic reviews will have scholars arguing that we can know (with relative certitude) in many places (though, perhaps not all) which intervention and policies are likely to lead to better outcomes. Others might take the call to its next logical step: *if we know the policies most likely to lead to the outcomes we desire, policy decisions should then be based upon that knowledge.*

But based upon that knowledge to what extent? It comes as little surprise that these reform efforts calling for more rigorous methodical approaches appear to be yielding some fruit. Policymakers have increasingly accepted the need for quality evidence in their decision-making. The Obama administration's 2009 budget blueprint, for example, included the initiation of a multi-billion-dollar federal support for a model of home visiting that received positive results in several rigorous social science evaluations (Haskins et al. 2009). Similarly, the President's head of the Office of Management and Budget (OMB) Peter Orszag penned a blog in which he argued for the need for government investments to be smarter and more results-oriented. As he writes, "Rigorous ways to evaluate whether programs are working exist. But too often such evaluations don't happen ... This has to change, and I am trying to put much more emphasis on evidence-based policy decisions here at OMB"<sup>10</sup>.

Even more recently, respective US House and Senate Budget Chairs Paul Ryan and Patty Murray introduced and passed legislation establishing a Commission seeking to promote evidence-based policymaking at the federal level. The 15-member Commission

---

<sup>10</sup> Dick Darman wrote similar pieces to those penned by Orszag while Darman was serving in the same OMB Director role under President George W. Bush.

on Evidence-Based Policymaking was appointed by President Obama and leadership of Congress, beginning its work in the summer of 2016, delivering their report in September of 2017 (Hoagland et al. 2017). The Commission was tasked with studying how to best expand the use of data in evaluating the effectiveness of federal programs and tax expenditures as well as determining if there should be a federal clearinghouse for program and survey data (US Committee on Budget 2014). In trumpeting the legislation, Representative Ryan talks about the desire to “change the focus in Washington” from one on effort to results (US Committee on Budget 2014). This legislation comes on the heels of Orszag and Ron Haskins’ (who happens to be a co-Chair of the Commission) publication of *Show Me the Evidence*, which chronicles the efforts of the Obama administration to incorporate evidence into the grant selection process, as well as *Moneyball for Government*, where a host of United States Senators, political operatives of both parties, and bureaucratic budget leaders suggest that the time has come for policy makers to pay more attention to the available data and evidence in their decision-making.

Orszag, Ryan, Murray and others in government join the chorus of scholars who have convincingly argued that advances in research design coupled with replication efforts allow us to know in many instances (though, again, perhaps not all) policies most likely to lead to the best outcomes; or, in the alternative, at least to know which policies are *more likely* to accomplish desired outcomes versus those which we can reasonably expect to fail to accomplish desired outcomes. As the IES project within the US Department of Education explains about its “What Works Clearinghouse,” the goal of the government agency is to “provide educators with the information they need to **make evidence-based decisions**” (emphasis original, Institute of Education Sciences 2014).

At the state legislative level, similar efforts are underway. The most prominent example of such effort is the collaborative effort between the Pew Charitable Trusts and the MacArthur Foundation's Results First program described in Chapter One. However, despite encouraging developments, critical questions remain about the usage of evidence among policy makers. Are there individual, institutional, or state-level factors that facilitate a greater use of evidence? Are there legislative issues amenable to evidence being utilized? How heavily do policy-makers weigh evidence relative to other factors? And finally, are there evidence-production mechanisms (namely, Results First) that are associated with policy makers being more favorable towards its usage?

These questions are important, as failing to connect the research to action leaves the empirical work to sit as an academic exercise alone. The designers of the EIS system, for example, speak directly to this dilemma in the following statement: "If we want to move evidence up the pecking order we need to make it easier for policymakers to access and understand research evidence" (Chamber and Lawrence 2014). What remains unsaid is that policymakers need to also *want* to use the available evidence.

That might be the biggest challenge in the years ahead. Even the staunchest American supporters of the use of evidence in policymaking concede that considerable room for more utilization exists. As the Coalition for Evidence-Based Policy (2014) shares, "In most areas of social policy ... government programs often are implemented with little regard to evidence, costing billions of dollars yet failing to address critical needs of our society" (paragraph 1). The What Works Clearinghouse is a good example of this dilemma, with many of those involved in education not knowing that the information is available to them, or knowing what to do with the evidence should they

access it (Kolata 2013). Despite good arguments regarding effectiveness, the push for evidence-driven decision-making remains largely elusive. To wit: Bridgeland and Orszag (2013) estimated that only one percent of spending on government programs is backed by *any sort of evidence*<sup>11</sup> and a staggering 75 percent of government-funded social programs were found to have *no effect at all* for the people they were presumably serving when evaluations were conducted (Haskins 2014).

## **2.2 Evidence defined and its role in decision-making**

It is important to understand the variation that also exists in the type of evidence and information available to policymakers. What exactly one means by evidence makes a big difference in a study of its effect on the policy making world. Part of the value of this research involves an opportunity to better understand what state lawmakers consider evidence and how they engage with it.

One of the well-documented histories of the rise of the use of evidence – and the champion of its use – comes from Gurrón and Rolston (2013), whose detailed experiences in the world of welfare reform efforts and research lead them to the conclusion that well-designed RCTs can have a significant effect on producing the type of information that can most easily be assimilated into policy work. As they write, the “[random assignment] technique could be used to address most of the policy options in a wide range of conditions, and furthermore, that the distinctive quality of the evidence was

---

<sup>11</sup> One could argue this calculation is unreasonably conservative, as they use a very ambitious standard for what constitutes evidence.

often recognized and valued” (Gurron and Rolston 2013, p. 19). Based on their work in the welfare world, they believe that similar reforms are possible in other policy areas.

Most scholars would agree with that point: there exists some variety of evidence that may be more sophisticated than others, with well-executed randomized controlled trials (RCTs) standing out as the obvious example. To some, this has led to a new minimum standard for evidence use – and these people are likely to believe that the use of a published study is no longer sufficient and that only published, peer reviewed studies with stronger methodologies should be the standard for which evidence is to be used in helping to drive policy decisions, even if only in small measure. In many respects, this is the argument being advanced by proponents of the Results First program. Their most recent report chronicles the framework they have designed that “identifies the steps that all levels and branches of government can take to build and support a system of evidence-based policymaking for strategically selecting, funding, operating, monitoring, and evaluating public programs that deliver the best returns on taxpayer investments” (Pew November 2014, p. 18). If there were ever an abundance of confidence in one’s evidence production program over other mechanisms, this is it.

Still, there remains room for a diversity of research methodologies utilized to inform policymaking. Even WSIPP, the program upon which Results First was designed, includes quasi-experimental designs with approved statistical techniques as well as studies utilizing regression discontinuity and instrumental variables (Lee et al. 2012). Moreover, it remains true that practical realities usually mean that “policymakers frequently have little choice but to consider and use a mix of different types of information” and from varying sources (Heinrich 2007, p. 273).



In addition to practical considerations, Heckman and Smith (1995) challenge the conventional wisdom that randomized experiments should always be considered the preferred methodology. In their view, a gap remains between the theoretical capabilities of random assignment evaluations and their results, that promising non-experimental evaluation developments are often overlooked, and that randomized experiments can jeopardize the long-range accumulation of knowledge in exchange for “simple black box evaluations” (Heckman and Smith 1995, p. 108).

For the foreseeable future, it appears evidence other than the most “rigorous” and “scientific” will remain an important part of the policymaking process – even in states with Results First and similar programs. That fact alone offers sufficient reason to take a broad perspective on evidence when engaging with state policymakers, especially considering the extent to which most would understand the intricacies of these two parallel movements and the data collection methodologies associated with them. While the most comprehensive evidence-based policy approaches are those that seek to involve a systematic review of all available research (not just relying upon one or a few studies) and subjected to a rigorous screening process such as those conducted by the Campbell Collaboration<sup>12</sup> (Petrosino et al. 2001) whenever possible, this research process leaves open the possibility that any form of empirical evidence is useful to inform and improve legislative outcomes. The justification is straightforward: even the most sophisticated approaches to evidence are unlikely to resolve all doubt or to “oversell

---

<sup>12</sup> Other examples of organizations seeking to apply a more rigorous academic standard to create evidence that can then be utilized within the policy-making process include the Coalition for Evidence Based-Policy, Alliance for Useful Evidence, and the What Works Clearinghouse.

scientific evidence” (Petrosino et al. 2001, p. 29), and the full range of evidence production will need to exist to most wholly inform policy leaders.

This approach is consistent with the idea that there remains considerable uncertainty about what evidence based policymaking means. As Heinrich (2007) reminds us, “we are still far from a consensus – intellectually or politically – regarding what should count as evidence, how it should be produced and validated, and how it should be used to influence policymaking” (2007, p. 259). If Heinrich is right, policy making will continue to have evidence incorporated that goes far and beyond distinguishing between quasi-experimental designs and RCTs (Jennings and Hall 2015). Given the wide range of approaches to this topic, we say that evidence-based policymaking is defined as making well-informed decisions about policies, programs and projects by putting the best available information at the heart of policy development and implementation (Davies 1999). In fact, I as outline through the course of the research, a more appropriate term for these discussions is evidence-informed policymaking – and not evidence-based policymaking.

As a practical matter, this would incorporate a wide range of data, studies, and evaluations (including, but not limited to fiscal notes, budget numbers, and program evaluations) prepared by academics and/or professionals in a relevant field of policy analysis as evidence to be used. When appropriate in the data collection process, these definitions will be the lens through which policymakers are asked about their interaction with evidence during their work.

### **2.3 Performance management and tensions with evidence based policymaking**

As evidence based policymaking has struggled to insert itself into the governance process, it has been forced to contend not only with its own challengers and skeptics, but also with the possibility of being crowded out by parallel movements. These other movements exemplify the challenges inherent in having a wide range of evidence types that are available to be incorporated into the policy making process.

There may be no better example to prove the point than the ways in which evidence based policymaking has found itself competing for attention with the practice of performance management. While evidence based policymaking is just now making its way into the federal decision-making consciousness, performance management has been active for over a decade. Meaningful use of performance management data began in 1993 with President Clinton signing the Government Performance Results Act (GPRA) with a promise to hold federal agencies accountable by measuring their progress (Heinrich 2007). For those in the evidence-based policy crowd, there is ample reason to remain frustrated that GRPA remains the primary source of law cementing “evidence” (in this case, performance data) in the federal administrative state. That could have something to do with the state and local level government approach now being executed by the Results First program, even as there is a continued push at the federal level to rethink and retool in such a way as to make room for evidence other than performance data to drive decision-making.

Since, and in addition to, GRPA there are now many public agencies at the federal, state, and local levels that use performance data to demonstrate results in the programs for which they are responsible. Often, these reports are to external

constituencies, including legislative oversight bodies. The performance measures help to enhance the accountability of public agencies in several ways. For example, Behn (2003) describes a multitude of ways in which performance systems can improve agencies: through evaluation, control, budgeting, motivation, celebration, learning, and improvement. Likewise, Hamilton (2005) offers an array of ways in which performance measures enhance accountability to the extent in which they include meaningful goals, measures, targets, and consequences. Performance measures have become popular because of the ways in which they sharpen public accountability by providing justification to enact the type of improvements identified by Behn. Regardless of to whom a manager uses the data to be responsive (including the organization itself, the legal system, auditors, or political figures), performance measures have been shown to establish a clear process of accountability (Bovens 2005).

Ideally, the performance data provided will function best by turning performance data inward and applying learning toward ways in which an organization can best adapt in the future (Moynihan 2005). For performance data to be useful, it needs to be linked to outcomes of interest, used as a mechanism to force managers to review data on a regular basis and to require performance improvement actions when necessary – such has been the case with the monitoring system established in Baltimore City known as CitiStat (Behn 2006).

Yet performance measurement systems are not perfect. As Moynihan (2008) reminds us, one part of the problem is that even when performance information is made available, it is not always widely used and does not always lead to better decision-making – with individual backgrounds and institutional roles making a big difference. For

example, even though most states now develop and use performance information in their executive branch budget development process, key performance indicators remain primarily used by executive branch decision-makers, not state policymakers (Melders and Willoughby 2004).

In many respects, the challenges facing performance management are shared by evidence based policymaking. Moynihan (2008) reminds us that “performance management is not (and cannot possibly become) the solution to governmental problems that its advocates promise ... Elected officials and senior bureaucrats would have to change the way they make decisions ... [and] these changes require a different system of governance, not just management changes” (p. 23). This may also represent a shared opportunity. Judy Zelio (2008), for example, recognizes that different performance information needs of state agencies and legislatures, and proposes five actions to bridge that gap, key among them creating performance information that “emphasizes policy results rather than administrative measures” (p. 13).

Indeed, performance reporting and evidence based policymaking have a shared goal of making government more effective and in using more rigorous information to accomplish that change. To that end, “in ideal circumstances, governments would use a full range of information in decision/policy making in a logical flow” in such a way that includes the “linking performance monitoring of ongoing processes and results [performance reporting] to the scientific evaluation of impacts and cost-effectiveness [or evidence based policy]” (Hienrich 2007, p. 256). However, we know that differences remain and keep these systems from working together as envisioned under ideal circumstances. One key reason is time. While performance measurements are more “real

time,” with usually no less than annual (and often quarterly) reporting, the time needed for conducting exhaustive research frequently does not align with performance management timelines. Thus, what is counted as “evidence” under performance management and who decides if it should count (a program’s manager versus an academic) differ. In other words, the “need to know” of performance management can put the knowing above the methods – at stark contrast with the more deliberative approach of evidence based policy. Consequently, “research has shown that in the absence of high-quality, readily available information on outcomes and under high-stakes pressure to demonstrate performance improvements, data collected for these purposes are more likely to be misused or manipulated in performance analysis” (Heinrich 2007, p. 273).

The irony in the performance management – evidence-based policy conflict is that there are also similarities between the two and ways in which they can work together towards common ends. In many respects, it could be a question of how state leadership chooses to engage with the information. A high crime rate in a performance management system should prompt reform, certainly. But what kind of reform effort: Any change that simply lowers the number in the years ahead? A careful evaluation of whether one new policy will generate better numbers through an elaborate program evaluation or benefit – cost analysis? Or an in-depth comparison of many program options that might include political options that are more effective in combatting crime though not as palatable to the electorate? Such a dichotomous choice tracks with the difference between single and double-loop learning identified by Donald Moynihan. As his research showed, there are more examples of states that have used results-based reforms to accomplish narrow

process improvements (what he calls single-loop learning) than those that have taken a broader look at policy choices and effectiveness (what he calls double-loop learning) – even though the double-loop approach is “more critical for long-term organizational success” (Moynihan 2006, p. 203).

Rather than be at odds, the idealized version of performance management (Moynihan’s double looped learning) tracks closely with the approach advocated for by organizations such as Results First. In fact, such a comprehensive approach to policymaking could be argued to represent both the most rational and natural argument for the benefit-cost valuations Results First seeks to implement in state legislatures. It remains to be seen how state lawmakers understand this interaction (if at all) between performance management and evidence based policymaking.

#### 2.3.1 Using evidence in policy

While performance management has grown in both research and in practical governmental use, the evidence based policy literature remains young and developing (Jennings and Hall 2011). Even the staunchest of supporters of evidence recognize that it has limitations, noting that even the highest quality evidence (well executed experimental findings) plays a limited role in final policy outcomes. Gurrón and Rouston (2013) share that they knew their efforts, regardless of their sophistication, would remain a “minor element at best in a highly value-laden and political debate” (p. 426). Moreover, as the amount of evidence designed to influence policy decisions continues to mount, it remains questionable if those making the decisions know how that evidence is produced – or care, for that matter. That is in line with the work of Lupia and McCubbins (1994), who investigated the principal – agent relationship between legislators and non-elected experts

and found that picking the right agents and/or learning enough about the issues is sufficient to protect legislators from the pitfalls of delegation. In the realm of evidence-driven policy, policy makers are well served to understand both the underlying issues and methodologies of analysis. In that regard, “when agents believe that legislators can learn about the consequences of their actions, agents who desire change will be induced to take actions that are beneficial for legislators” (Lupia and McCubbins 1994, p. 374).

Such an understanding is in short supply, with scant attention paid to the issue in the current literature. The Oliver et al. (2013) systematic review of 145 studies found that “little is known about the role of research in policy making.” That review also confirmed the facilitators and inhibitors of the use of evidence in policy making found in prior reviews. Unsurprisingly, the study found that facilitators of evidence use include relationships (between evidence producers and users), contact and collaboration, availability and access to research, and research findings which are reliable (Oliver et al. 2013). Conversely, a shortage of relevant and reliable research, as well as poor access to research findings are most often cited as barriers for policy makers’ use of evidence in their work (Oliver et al. 2013). Jennings and Hall (2011) found considerable variation in state agency use of source type (some draw heavily from scientific and professional sources, while others rely heavily on political sources) and the respective weight applied to these various sources. Such variation in the types of evidence used surely has an impact on the types of decisions that are reached within these state agencies. Talbot and Talbot (2014) found similar variation in the United Kingdom, with research showing that while most senior civil servants engage with academic production, there is a preference for prepackaged results and general expertise is given more value than specific research.



Moreover, they found a “significant minority does not engage at all with academics,” a concerning finding especially in a place like the UK where the push for evidence has been so strong (Talbot and Talbot 2014, p. 4).

Jennings and Hall (2011) issued a “study [that] opens the door to understanding state agency information consumption,” but was also a report in which they conceded “we have much yet to learn” (p. 265). If the door is only being opened regarding evidence utilization by state level bureaucrats, the same can be said of evidence utilization by state elected officials. Larsen’s (1980) review pointed to the lack of studies focused on knowledge utilization in state and local governments, and “relatively little has changed since then” (Hird 2005, p. 57). Where limited research exists, it is often focused on a singular state and has found that officials cared less about empirical findings of research as compared to program operations (Greenberg, Mandell and Onstott 2000) and that personal values and the views of constituents ranked as the most leaned upon feedback source (Gray and Lowery 2000).

Much like Jennings and Hall opened the door on evidence utilization by state agencies, this research opens the door on state policymakers’ utilization characteristics and behaviors. This is an important task, as legislators remain responsible for oversight and funding of state agencies. If state lawmakers actively reject evidence, can one expect state agencies to embrace it? On the other hand, if state lawmakers are found to embrace evidence – even in limited capacities – those insights can have important impacts not just on state policymaking, but likewise on the agencies that respond to legislative directives.

### 2.3.2 Communicating information about evidence

Evidence quality is one issue, and availability is another. Utilization is the goal, however, and the research suggests that the use of available evidence is predicated on the ability to have it easily understood and shared among those expected to use it. Whiteman (1996) found that in Congress, while having access to an impressive base of comprehensive information to inform decision-making, very few individual members were conducting exhaustive information expeditions.

Wolanin (1971) offers several reasons why Congress fails to obtain information it needs to make more fully informed decisions – or why they fail to use the relevant information when do have it. These reasons include: 1) questions being unanswerable – at least not without treating “citizens like bacteria in a Petri dish”; 2) internal dynamics within the Congress including disjointed timelines, the hierarchical structure, and adversarial systems in place 3) having multiple sources of information, each with their own perspectives and interests; and 4) a limited competency in meeting the overall needs of the legislative branch (Wolanian 1971, p. 383). Recent advances might address some of these concerns. For example, refined and sophisticated research designs allow for more insights on more issues than ever before and some policy solution systems are designed to produce recommendations without an interest in which solution is chosen. Congress and state legislatures, however, remain partisan institutions with hierarchical systems in place.

The question of interest and bias in information production is examined at length by Weiss (1991), who supports a more honest approach in the creation of evidence. With the rare exception of highly controversial issues, she sees a wide band through which

knowledge can help influence and shape policy, and that researchers can and should produce this knowledge from their own base of values. Coupled with extensive checks on data and analysis and/or a candor about values and assumptions of the researcher, she believes that research can be nestled in a productive political debate (Weiss 1991). Still, many scholars have doubted the ability of policy analysis to leave a larger mark upon legislators and the formulation of policy than it has in the past (Jones 1976; Dreyfus 1977; Weinberg 1979; Furubo 1994; Lampinen 1992). Kingdon (1989) and Downs (1967) argued that decision-makers only seek out information searches in those instances where problems arise, while Webber (1984) feels that it remains “unlikely that many decision makers will make widespread use of policy research if left totally to their own inclinations” (p, 117). Similarly, even with the proliferation of professional support staff and through other structural arrangements, Congressional staff continues to use analysis primarily as a tool to score political advantage – to the point where staff “were clear that analysis doesn’t have a chance of setting broad direction for public policy” (Weiss 1989, p. 428).

More directly connected to the exploration of this research, Hird’s (2005) exhaustive review of state-level nonpartisan policy organizations concludes that “political influence over institutional form is profound and generally consistent with theoretical expectations” (p. 124). These findings help to explain the reactive nature of legislative organizations producing evidence for legislative consumption. Even at the state legislative level, Hird reminds readers that scholars and policymakers have different challenges and motivations, that legislatures face severe time constraints, utilize the research presented to them in different ways, and finally that personal connections have

largely remained absent between the scholars creating the evidence and the lawmakers who are the intended audience (Hird 2009).

In the context of these challenges, the goal for research production remains utilization by the end user. Communication remains one key factor in facilitating this use. In Congress, as is the case in state legislatures, communication is a key component of decision-making (Whiteman 1996). For Weiss (1998), who seeks not necessarily more, but rather more effective knowledge utilization (p. 30), this means both engaging the potential users of the information and finding as many routes to communicate findings as possible. Given the proliferation of professional staffing and the push to establish evidence-based programs, it has become important to measure the extent to which these new routes of communication can make an impactful difference in the process of policymaking. This dissertation is one step in measuring the extent to which traditional obstacles, partisan considerations, and limited attention to research and evidence remain obstacles in state policymaking environments.

## **2.4 Doubts and shortcomings preventing evidence-based policymaking**

With an abundance of strong evidence available to direct or at least inform policymaking, one might reasonably inquire as to what barriers impede progress. The reluctance to more wholeheartedly adopt evidence-based policymaking standards has been associated with several causes. Despite successes in placing evidence into the policy conversation, there remains doubt regarding if the connection between data and decision-making is, can, or will be completed. Black (2001) encourages us to “proceed with care” on evidence based policy, noting several reasons about why research evidence is limited on its impact around service policies. These include goals other than clinical

effectiveness, the dismissal of evidence produced as irrelevant, a lack of consensus about the research, competing evidence, an environment not conducive to policy change, and a relative inability of those charged with sharing knowledge being able to effectively do so (Black 2001). Similarly, Weiss et al. (2008) note shortcomings of research and researchers, shortcomings in policy and practitioners, and shortcomings in the links between the two.

On the research side, it has been argued that “the more we have come to know, the more aware we are of how tentative, limited, and sometimes erroneous the bases of our information and evidence are” (Heinrich 2007, p. 274). Another factor that could limit the policy impact of research: who is conducting the research – is there a think tank with an ideological slant paying for (and potentially influencing) the outcome? Yet another obstacle could be in the research methodology itself: how valid are the measures used, how generalizable are the findings, and how/can the results be replicated? These are all legitimate questions that need to be asked.

It is perhaps this dose of reality that has lead those like Heinrich (2007) to advocate less for “high stakes” systems of policy that bind decisions to data and more for informational clearinghouse systems. Such systems could help inform the decision-making of government leaders regarding program selection and performance improvement but would not mandate them. Others argue that we should continue to lean on the best available evidence today even as leaders push for more and continued rigorous analysis (Nussle and Orszag, eds. 2014, p. 27). Failing to recognize that there is some quality evidence available and/or failing to utilize that information in political decision-making, even if only modestly, inevitably leads to baseless experimentation and

wasted resources in proportions of which the average citizen would be outraged to learn. Yet in many respects that is precisely what we do as a society. Considering the vast amount of money we pour into “societal goals we generally support, we often know next to nothing about whether what we’re doing is working, or how” (Nussle and Orszag, eds. 2014, p. 46).

Competing forms of information being supplied to lead governance decisions also complicates the ability for quality evidence to drive all policymaking. Legislators are inundated with innumerable types of information. These include program performance data, budget information, fiscal and policy notes, verbal testimony, written testimony, academic studies, program evaluations, and benefit-cost reviews. These pieces of information also come from myriad sources: constituents, agencies, academics, interest groups, lobbyists, think tanks, the media, pollsters, and other organizations. Sorting through the various types of information, and trying to decipher if the value is different based on by whom the information is created, makes it possible that the best available evidence is not making it to the forefront of policymaking decisions.

Then there is, of course, the thorny and much more complicated issue of politics itself. Government officials must not only contend with the realities of attempting to govern with policies that will improve the welfare of the people they represent, but must do so within an environment where each decision carries a real or potential political consequence. Such an intersection can sharply curtail the ability to put evidence at the forefront of policy decisions. Put bluntly,

Anyone who has watched policymakers in action knows that they will rarely allow evidence on program effectiveness to be the sole or even major factor driving the policy process. Politicians focus on costs, the needs and desires of their constituents, the position of their party leaders, public opinion, their own

political philosophy, pressure from lobbyists, the position favored by people and groups that finance their campaigns, and a host of other factors in making decisions about how to vote on program proposals (Haskins and Brown 2011). In fact, political considerations (real or perceived) are where the largest obstacles

to a full implementation of evidenced-based policymaking reside. Even if concerns regarding who conducts the research, how it is conducted, and the results are communicated in a clear and convincing manner to policymakers, other influences on legislative decision-making will continue to thwart a full and complete embrace of policy tied wholly or mostly to evidence. Members of the Republican Right (such as the Tea Party) are unwilling to move beyond anti-government rhetoric and invest *more* in places that can make a difference in long-term outcomes, especially if it requires additional taxes to fund such programs. Liberal Democrats are unwilling to admit drawbacks to the Head Start program – namely, the finding that participants lose most of the advantages of the program by the time they reach the third grade (Maxwell 2012).

Deborah Stone agrees that it is impossible to divorce the policy analysis process from politics. She argues that even the creation of facts and evidence in the research process allows our politics to become part of the research process, meaning that “there can be no neutral facts” (Stone 2002, p. 310). The numerous people within the organizations producing our evidence and data, Stone argues, make any number of choices including who gets to testify at legislative hearings, what types of data to collect, how to categorize data, and how strongly to go after missing information (2002). She even questions evidence developed using the gold standard methodology of RCTs. Using the medical experience with randomized controlled trials, Stone finds flaws in the gold standard because of the ways in which it has produced “disappointingly mushy results,” the fact that there will always be a human influence and skewing in experiments that

study humans, and physicians (presumably like elected officials) are reluctant to accept RCT findings because of the ways in which it takes away from their “power, autonomy, and indeed clinical effectiveness” (2002, p. 312-313).

Stone’s arguments are worthy of consideration. Evidence-based policymaking, for example, is very different than evidence-*informed* policymaking. The former implies a whole-hearted embrace of making decisions based on what the evidence suggests are best outcomes. For anyone involved or familiar with creating evidence (especially through the most rigorous methods) it might seem disheartening to learn that the world of politics would intentionally set aside evidence on how to better educate people, improve a criminal justice system, and even save lives. To be sure, however, *this happens*. Still, there are other times when the evidence may be *the* critical factor in decision-making. We need to better understand to whom, when, and why evidence plays a role in these decisions – and when it does not. As outlined in the introduction, this research proposes to further investigate these questions as it relates to state legislative action.

## **2.5 A nine-layered view of policy worlds**

There are any number of ways to think about the places where governing and policy decisions take place. America’s federalist system has created at least three (and in some places, more) layers of government across the country. The governmental structures then adopted at each level also have varying types of representatives at work: elected, appointed, and civil servants (Ricucci and Naff 2008). This combination of federalism and the way government has been structured together form what can be considered a nine-layered view of governing.



*Table 5: A nine-layer view of governing*

<b>Federal Government</b>	<b>State Government</b>	<b>Local Government</b>
Federal Elected Officials	State Elected Officials	Local Elected Officials
Federal Political Appointees	State Political Appointees	Local Political Appointees
Federal Civil Servants	State Civil Servants	Local Civil Servants

Such layering is not meant to suggest that each layer of government type is equally capable of delivering effective management outcomes. For example, agencies led by politically-appointed leaders earned systematically lower Program Assessment Rating Tool (PART)<sup>13</sup> scores than those drawn from the civil service (Lewis 2007). It is intended to suggest however, that our governing reality means that there are at least nine different places one could investigate if, how, and when evidence-based policy making is being carried out. One of the clearest lines of separation among government representatives in the nine-layered view is whether officials are directly elected by, and thus most directly accountable to the desires of, a constituency. This is a key theme in our American democracy – and, for that matter, all democratic institutions: an accountability mechanism which takes public opinion into account for decision-making (Eulau et al. 1959; Key 1961).

For elected officials, evidence-informed decision-making may be the hardest to achieve of all, considering the ways in which they have more and other pressing inputs to their decision-making model. Namely, elected officials worry about being just that: being elected and re-elected (Mayhew 1971). Political appointees may want to see the benefactors that helped get them into position returned to office, but neither they nor civil

---

<sup>13</sup> OMB used PART to evaluate the effectiveness of programs using a set of performance-related criteria, including program design, strategic planning, program management and results to inform the budget process and drive program improvements (US Department of State 2008).

servants need to fear a direct backlash at the ballot box for making an unpopular decision – regardless of if that decision is empirically valid. This allows us to create a continuum of sorts to understand the likelihood of each type of political actor utilizing evidence in their decision-making, relative to the other types of actors.

*Figure 3: Varying likelihoods of utilizing evidence in policy making decisions*

<b>Elected Officials:</b> Most likely to need/want to include considerations other than evidence in policy decisions.
<b>Political Appointees:</b> Moderately likely to need/want to include considerations other than evidence in policy decisions.
<b>Civil Servants:</b> Least likely to need/want to include considerations other than evidence in policy decisions.

As suggested in Figure 3 above, there is ample reason to believe that, of all the layers of policy making, those associated with an elected office would be the least likely to be drawn toward evidence-based policymaking. Whether this is because an elected official acts as a single-minded seeker of re-election, believes that normative values (their own or those of their constituents) have an important place in the policy discourse, dislikes the loss of control over the legislative process somewhat inherent in deferring to empirical evidence, or some combination of these and/or other issues, there are ample reasons to believe that if evidence can be used more among elected officials, it can be used more in any of the other layers of policy. Within the elected legislative branch of government, a venue exists where the intersection of politics and evidence (and the competing interests thereof) are more acute than in any other area of governance. All this makes exploring ways to encourage elected leaders to more fully utilize evidence in policy making an excellent place to conduct research.

For this research, the scope is further narrowed by focusing on three key issues. To that end, this study investigates the layer of policymaking that contains a) ample variation, b) has a treatment that is intended to increase the use of evidence, and c) is

most difficult to cultivate an environment conducive to evidence-based policymaking. The first requirement, variation, is met best at both the state and local government level. While a study of the federal government could be informative and interesting, there would be no other groups against which to compare subjects. Justice Brandeis popularized the notion of states as laboratories of democracy in the *New State Ice Co. v. Liebmann* decision (1932), and today states can serve as the great testing ground for efforts to inject evidence into policy making. The second item, a relevant treatment, has been established at the state level through the efforts of the Pew Center on the States and the MacArthur Foundation. Collectively, these groups are working to provide state lawmakers with access to a systematic way to conduct benefit-cost analysis on issues upon which legislative reform actions can be taken through their Results First program (Pew 2013). As previously discussed, Results First is certainly not the only approach to incorporating evidence into policy making, but it is clearly a preeminent intervention geared toward a specific audience related to improving the use of evidence in the policy world. Moreover, the program has a robust quality that allows it to span across many states of varying political ideology and background. This best allows for comparisons across the states. Finally, exploring if and how Results First impacts state lawmakers satisfies the final requirement, exploration of the policy layer most difficult to introduce evidence, elected office.

Finding breakthroughs at the state legislative level might help in the other layers of the policy world, which with even fewer limitations, continue to face obstacles instituting evidence as a guiding policy principle. As discussed in the first chapter, for

example, the federal government under the Obama administration was beginning to move earnestly toward making the connection between evidence and policy, as evidenced by their proposed approach for funding of a home visiting program (Hastings et al. 2009). What this means in a Trump administration is even less clear. The resignation of officials such as Elizabeth Southerland, alleging policies of “myth over truth,” suggest that the new administration may be approaching science and evidence in very different ways (Green 2017).

However, even a presidential administration focused on evidence-based policy yielded to the myriad forces influencing the political process – President Obama’s team ultimately did not require programs to have the strongest evidence of effectiveness to continue receiving some level of federal support (Haskins et al. 2009)<sup>14</sup>. Even with a chief executive wholly endorsing an evidence-based approach in one area of policy, the legislature was found to back off this direction. As the Obama administration discovered<sup>15</sup>, and as is argued here, political considerations that elected lawmakers (and to some extent, all government officials) must confront are daunting and often discourage the use of evidence in policy making. Jewell and Bero’s (2008) work helps to confirm this, suggesting that contrasted to the “comparatively insulated and technical policy world of administrative officials, the work of legislative officials is shaped by forces that present different and formidable obstacles to the regular use of scientific research” (p. 184). Moreover, even amongst elected legislative leaders, there remain opportunities for vast differences in individual and institutional approaches to the use of evidence –

---

<sup>14</sup> While President Obama initially signaled his intention to use evidence as the measure by which programs would receive funding, political pressure caused the administration to change course.

<sup>15</sup> The Obama administration found it “challenging to move beyond production of performance data to its use” (Joyce 2011, p. 356).

dependent upon the political realities within a given policy environment. The experiences of state legislatures around the country helps to confirm this, and indicate that evidence-based and evidence-informed policymaking continues to be sporadic in nature.

## **2.6 Variations of evidence production for state-level policymaking: trends over time**

Looking at just two states helps to put the level of variation in the quality of evidence produced for the purposes of policy making in perspective. Consider the differences between Washington State – the model state for Results First – and South Carolina. On the one hand, Washington has developed a robust and highly touted evidence based policy shop with recommendations based on careful reviews and multiple studies; on the other, South Carolina has a legislative policy support mechanism that largely produces individual studies in response to legislative inquiry without much of a review process.

### **2.6.1 Washington State Institute of Public Policy (WSIPP)**

As discussed in chapter one, this institute is perhaps the longest-running example of a state legislative body that has undertaken efforts to incorporate evidence into the decision-making associated with the creation of policy. Established in 1983 by the Washington legislature, WSIPP is a nonpartisan, general purpose legislative research unit that produces information based on projects assigned through legislative bills (Aos 2012). In recent years, WSIPP has undertaken empirical explorations of the costs and benefits of varying policy options in a host of areas. These areas include crime, early education, child abuse and neglect, substance abuse, mental health, developmental disabilities, teen births, employment, public assistance, public health, and housing (Aos 2012). Also

noteworthy are all the policy areas in which the specific approach of WSIPP is *not* utilized, including transportation, environmental and tax structures.

The work of WSIPP is carried out by an ongoing staff that is supported by contractual employees when work is at its peak. WSIPP takes a three-step research approach regarding their analysis of investments from the legislature. Step one involves a full review of all available, rigorous evaluations of real-world approaches to addressing key outcomes; step two monetizes the benefits, costs, and risk associated with each approach; and step three develops an analysis of how a “portfolio” approach to the options might impact statewide outcomes (Aos 2012).

Upon the completion of this three-part analysis, WSIPP produces a report outlining the relative benefits, costs, and risk associated with every project. In Washington State, this resembles a straightforward and easy to read report fashioned after the highly popular *Consumer Reports* list (Aos 2012)<sup>16</sup>. In a concise fashion, the report explains which programs work, which do not, which programs offer the best (likely) return on taxpayer investment, and what remains unknown given current levels of available, rigorous research. For those supporting the use of empirical evidence in policymaking, WSIPP stands out as a clear trailblazer in the use of tying evidence to public policy. For example, the Washington legislature used an evidence-based review conducted by WSIPP to fund a portfolio of evidence-based criminal justice programs in adult and juvenile corrections and prevention in 2007. Consequently, the state averted prison sentences above and beyond what was expected (using national and regional data to establish a baseline), leading to better outcomes for those staying out of prison, 2000

---

<sup>16</sup> As discussed in chapter one, this approach is replicated in the Results First program.

fewer prison beds being required, and state taxpayers saving approximately \$250 million annually (Aos 2012).

#### 2.6.2 South Carolina Government Research and Service Technical Assistance

In stark contrast to the elaborate work of WSIPP, South Carolina's Legislative Council could be considered on the opposite end of the spectrum when it comes to rigorous evaluation and feedback to inform policy decision-making. As explained on the website of the South Carolina Legislature (2013), "Legislative Council's purpose is to provide research, reference, and bill drafting services to the General Assembly." However, in comparison to the rigorous empirical work embarked upon by the WSIPP, content analysis of the organization's website suggests that the South Carolina Legislative Council's research remains straightforward in nature, opting to merely provide single studies to individual legislators rather than the more robust research offered in Washington. The South Carolina approach does not appear to have a mechanism that evaluates the rigor of methodology of the research provided, nor does it contemplate consolidating results of many available studies in the type of meta-analysis methods employed in Washington. Consequently, policymakers in South Carolina are less likely to be exposed to the full spectrum of views on a given policy topic, potentially leading to decisions being made with incomplete (or even incorrect) information.

While the Legislative Council appears to supply limited critical information to policymakers in South Carolina as they weigh policy options, legislators in the Palmetto State do at least have the shell of an organization that *could* have the capacity to employ more rigorous and far-reaching research to support the legislative process at their disposal. The University of South Carolina's Governmental Research and Service

Technical Assistance program boasts of offering training, technical assistance, research, and support for the South Carolina City and County Management Association (USC 2013). This work, however, is limited in nature, often narrowed to no more than providing a forum through which local governments can provide best practices in addressing a variety of public policy issues. Moreover, the website for the Government Research and Service Technical Assistance (2013) program indicates that research work and technical assistance is only undertaken on a contractual basis, and a review of the link on research produced only four total topic areas. The approach adopted (and support available to legislators) in South Carolina is sufficiently lacking if rigorous empirical research is thought to be a critical component of the policymaking process. Individual policymakers might still engage evidence in their work, but a comprehensive system of producing such evidence is not readily available within the state.

Such a content analysis is consistent with other findings. For example, when it comes to the creation and use of benefit – cost analysis in its policymaking decisions, South Carolina is decidedly behind other states. Between 2008 and 2011, the state produced only three reports, and even those reports did not consider multiple or alternative programs nor did they drive legislative or executive branch action (Pew July 2013).

Across the remainder of the country, there is considerable variation in methods of production, professionalization, and quality of evidence produced by policy shops for state legislatures. As Hird (2005) points out in his thorough review of all non-partisan research organizations among state legislatures, there are both “large, sophisticated, non-partisan agencies” in several states as well as examples where “offices are understaffed



or, in several instances, nonexistent” (p. 69). Perhaps because of this variation, and sensing an opportunity to inject more evidence-based decision-making into the policy-making process, the Pew Center on the states stepped forward to increase and streamline the application of evidence and benefit – cost evaluation work across the country with Results First. This program’s existence and experience serves as confirmation to the degree to which great discrepancy remains between state-level elected leaders on the issue of evidence and policymaking.

But do these varying evidence production institutions make a difference? Are there connections between what is happening with the Results First program and meaningful evidence utilization? After all, even a 2011 WSIPP-created benefit-cost analysis was prevented from setting policy reforms in Washington because of “entrenched sentencing laws and justice policy values that cannot be factored into a cost-benefit ratio” (Pew July 2013, p. 13). Are there other individual, institutional, or state characteristics that impact evidence utilization by state legislators? The literature provides some insights, but also suggests that more work remains to be done to better understand this issue.

## **2.7 From big ideas to specific insights about evidence and policy: the literature**

In recent years, the debate in academia regarding the importance of reliable evidence to inform policymaking has largely shifted from *if* evidence can be utilized to *how* it can be most effectively incorporated into existing policy frameworks. Indeed, the concerns regarding evidence now most often center on how seriously those in the policy-making field have taken the availability of this evidence. In her Association for Public Policy and Management Presidential Address, for example, Rebecca Maynard (2006)

calls the growing emphasis on policy “music to the ears” of her members but concedes that “we are far from a world in which evidence is routinely and smartly produced and integrated into decision making” (p. 249). She proceeds to outline things to keep in mind when embarking on such an endeavor, namely that lenses for framing questions and interpreting evidence can change; sometimes multiple rounds of research with varying methods are necessary; and that context matters in framing questions, designing research, and interpreting findings (Maynard 2006).

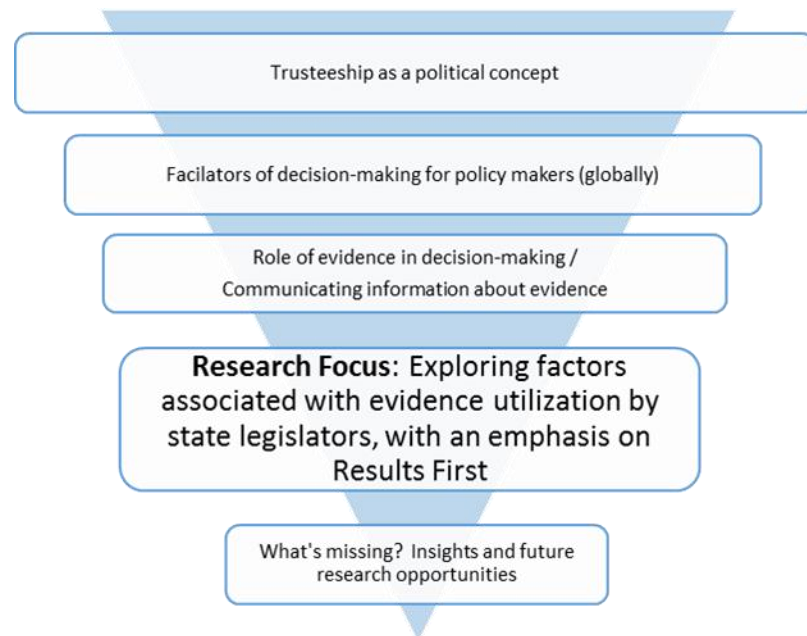
It seems reasonable that Maynard’s recommendations (and similar recommendations from others) are built upon conversations with those charged to implement policy. However, the available literature largely omits from the conversation direct feedback from those making various policy calls regarding if, how, and why evidence makes its way into their work. This is surprising, as these important questions regarding policy and evidence integration are best accomplished by taking the time to speak directly with policy leaders about how they utilize evidence in their decision-making and how they reconcile its importance with the many other factors that drive their work. On the other hand, there are others who believe “there is no one version of the truth that will convince or satisfy the various interested parties” (Gow and Wilson 2014, p. 129) which reflects the multiple and politically-motivated norms of practitioners of public administration (Riccucci 2010).

Given this uncertainty, it makes sense to determine if there are specific ways (such as the Results First program) that align legislators with empirical evidence as they conduct their policy making? Furthermore, it merits exploration if are there issues at play

beyond those programmatic in nature – personal, state-level, or other institutional characteristics – that more accurately predict when evidence is employed?

This means gathering information on a delivery mechanism need not just be about the program itself, but also consider the role of evidence in decision making generally, thinking about how factors other than evidence influence the decision-making process for policy makers, and even exploring the idea of what trusteeship<sup>17</sup> means for elected representatives. A clearer way to understand the connections of these ideas – and how they impact how evidence is viewed and used – is to think about them as a funneling down from big picture ideas (trusteeship) into specifically programmatic outcomes (in this case, Results First). The visual representation of this type of mapping might look like an inverted triangle, as seen in Figure 4 below.

*Figure 4: From big ideas to specific insights regarding the role of evidence in policy making*



<sup>17</sup> In this context, trusteeship refers to the public trust conferred upon elected officials to make decisions on the public's behalf. How officials carry out this trust while in office varies on a continuum between reflecting the views of one's constituency and making decisions thought to be in their best interests.

## **2.8 Trusteeship among political representatives: a delegate or a trustee?**

Questions about how legislators should act in office (and thus, if and under which conditions evidence might be used by policymakers) are familiar to political scientists. In fact, such discussions reignite long-standing (and still unresolved) tensions that have become known as the trustee – delegate dichotomy. The modern idea of trusteeship dates to 18<sup>th</sup> century England and Edmond Burke's now-famous speech to electors in Bristol. From Burke's worldview (1774), representatives should be familiar with the needs of the local areas they represent but should endeavor to come together with other representatives to pursue policies for the larger population toward which they all have responsibilities. However, one who rejects the idea of legislatures as deliberative bodies inclined toward the good of the whole rather than individual parts has little use for Burkes' views (Eulau et al. 1959) and instead requires an approach to representation that is more provincial in nature.

This historical debate between managers and elected representatives being "trustees" (that is, being entrusted by the public to do what they believe is in the best interests of those whom they represent) versus "delegates"<sup>18</sup> (meant to directly represent the popular will of those whom they represent) is likely to continue, but the rise of empirical research offers an opportunity to change the paradigm of the dichotomy: managers can now know what's best, and can use evidence to inform and teach the people they represent about how and why the decisions they are making can lead to the preferred outcomes. A predicted outcomes chart injecting the traditional trustee – delegate

---

<sup>18</sup> The delegate – trustee debate has also been referred to as the mandate – independence controversy.

dichotomy with a new layer of empirical knowledge posits ways in which evidence can change this debate.

*Figure 5: Predicted outcomes for trustee vs. delegate with think vs. know dynamics; An evidence-informed delegate vs. trustee dichotomy*

<b>Delegate who thinks- pushing through uncertainty and doubt:</b> “Even though I think this is the wrong policy, I am going to vote the way the people who sent me here want me to.”	<b>Trustee who thinks- a risk taker:</b> “Even though this goes against what the people who sent me here want, I think this change will be better for them and/or is the right thing to do regardless of the consequences that follow.”
<b>Delegate who “knows”- wrestling with conflicting values:</b> “I know I am supposed to represent the views of the people who sent me here, but am I doing them a service if I am pursuing policies that I am fairly confident that harm them – even if that’s what they say they want?”	<b>Trustee who “knows”- a more confident risk taker:</b> “I feel comfortable going against what the people who sent me here want because I have a high level of confidence that this is better for them and/or is the right thing to do, and I have the evidence to support my claim and convince them that this is the right action.”

While trustee and delegate models of representation demonstrate opposite approaches to governing, the reality is that most decision-making takes place somewhere between these two extremes. It draws parallels to the famous politics – administration dichotomy that has raged on for decades within public administration, beginning as far back as Wilson’s (1887) view on public administration. As Hannah Pitkin (1969) writes, “representation embodies a persistent tension between ideal and actual practice, between intention and institutionalization, between substance and form” (p. 22). Put another way, “there is not a single undisputed normative theory of political representation” (Thomassen and Schmitt, eds. 1999, p. 14). In real ways, the injection of evidence (especially quality evidence) into the policy making process could set the stage for a new way of approaching governance. Specifically, in instances where evidence is available

and trusted, it has the potential to shift legislators away from a delegate approach to lawmaking and more toward a trustee approach<sup>19</sup>.

To that end, the rise of evidence-based policymaking in states across the country *could* also be setting the stage for a renewed, revised, and robust discussion about the trustee versus delegate view of representation. It may no longer be a question of how a representative weighs doing what he or she *believes* to be in the best interest of constituents (a trustee worldview) against what constituents want elected officials to do (a delegate approach), but rather a question of how managers and officials reconcile an empirical body of evidence that allows officials to *know* better approaches against democratic norms and citizens who may not want officials to take certain actions, even if the evidence calls for it.

Do elected officials believe it is important to utilize empirical evidence when available? If leaders are concerned *only* with representing the views of their constituency (or what they believe to be the views of their constituency) or are focused on decisions that are best for their electoral concerns (or what they view to be best for their electoral considerations), the rise of evidence might not have much of an impact on this ongoing debate. In very real ways, elected officials are required to make the normative decision about where they fall in the context of this debate.

These considerations make monitoring how evidence impacts these ongoing tensions a worthwhile endeavor, as it could fundamentally reshape the way representation itself is thought about and carried out in the real world. Thus, it will be important to

---

<sup>19</sup> Some lawmakers will remain firmly aligned with the delegate approach to governing. However, I hypothesize that as evidence becomes more available and trusted by legislators, the most likely impact is that lawmakers are moved away from the delegate position and closer to the trustee position on the delegate – trustee spectrum.

understand this tension policymakers have long faced. The orientation on the Delegate – Trustee spectrum for elected officials plays a critical role in shaping how they approach decision-making. To better understand the role that this theoretical tension might play, one needs to understand what goes into the decision-making process for policymakers. The focus of this study is on evidence’s role, but evidence remains only one of many considerations facing elected leaders.

## **2.9 How elected officials make decisions**

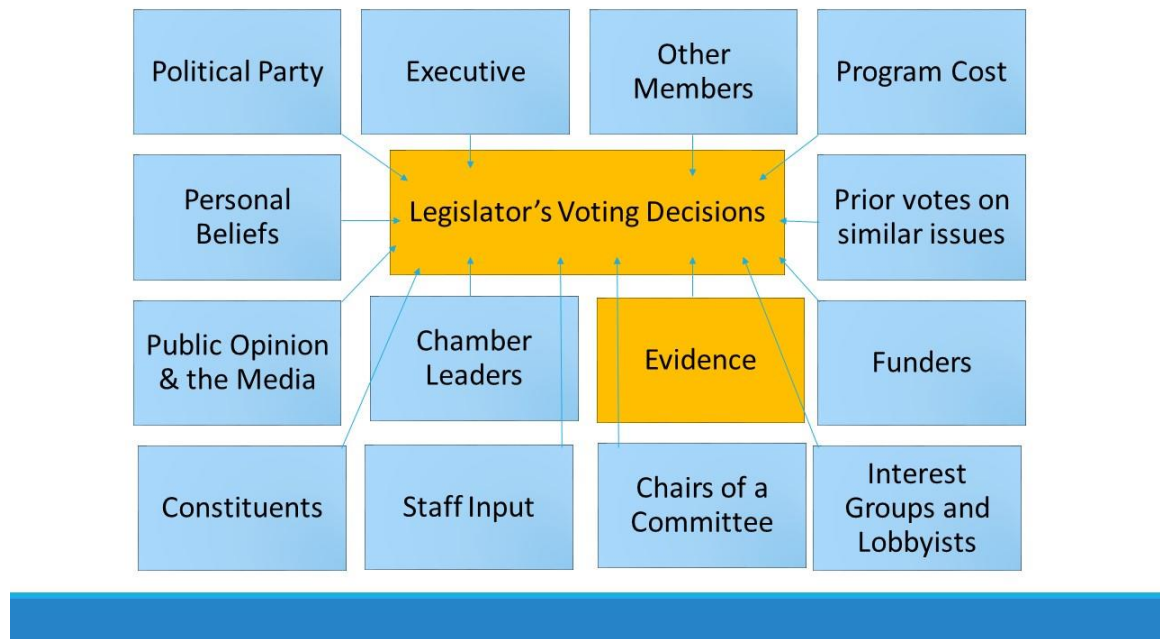
There are many inputs that factor into the decision-making process of an elected official. Some have studied these multiple inputs and channeled them into theories of how officials behave. Opinions developed from these reviews range from elected officials only being concerned with issues that might impact electoral outcomes to others which suggest officials not only seek re-election, but also might want to pursue good public policy, obtain power, or make a difference in the lives of others.

### **2.9.1 What goes into legislative decision making? A critical question**

Kingdon (1989) provides insights into many of the factors facing elected members of Congress in his exploration of voting decision behavior. Among the considerations facing congressmen Kingdon notes the independent variables measured in his study: constituents, political party leadership (for more on the influence on parties, also see Wright and Schaffner 2002) leadership of committees, interest groups, the executive branch, staff members, the media, and other members of Congress (on this point, see also Masket 2008). In addition to these considerations, other factors that have been identified as related to legislative decision-making include prior voting decisions,

fundings, personal beliefs, the cost of a program, and the focus of this research - evidence (see Craniold-Davis et al. 2009, Humane Society 2017). A full summary of these factors is included in figure 6 below.

*Figure 6: Factors that can influence decision making for elected officials, with evidence highlighted*



### 2.9.2 Single minded seekers of election?

For writers like Mayhew (1974), the trustee – delegate dichotomy depends less on how representatives choose to go about their elected task and centers more on how their decisions will impact their chances of re-election.<sup>20</sup> He found that the work of members of Congress primarily is geared toward securing re-election, which implies a set of decision-making calculations based less on a theoretical spectrum and more on a rational positioning within the electoral system. As he writes, “there can be no doubt that congressmen believe positions make a difference” (Mayhew 1974, p. 375). For this kind

<sup>20</sup> Mayhew claims that it makes sense to *assume* legislators strongly desire re-election. This would imply delegate-type behaviors; or at least, delegate-appearing behaviors.



of single-minded seeker of election victories, evidence based policy takes a backseat to electoral concerns. When convenient, elected officials might cite the findings of a study or be inclined to support the work of a program such as Results First. However, this cannot be assumed to be the case when an electorally popular course of action collides with what evidence suggests should be done.

The inclination to “fit” the processing of information (even unconsciously) to suit an end or goal for an individual has been referred to as “motivated cognition” (Mooney 2011). In a political context, an end goal of having one’s own personal beliefs affirmed (or, for that matter, the views of a constituency) motivates how mental operations are conducted. This could include impacting how one goes about “assessments of the weight and credibility of empirical evidence,” leading to biased information searches, biased assimilation, and identity-protective cognition (Mooney 2011).

#### 2.9.3 Willing to give up a seat: Not single minded, but a bounded rationality?

Other work has found that there are more interests to policy makers than re-election. One major example was Fenno’s (1973) *Congressmen in Committees*, in which he found that there are up to three goals that members pursue: re-election, power and influence within the chamber, and the pursuit of good public policy. Such a perspective acknowledges the tendencies of members’ pursuits toward re-election, but also suggests that such pursuits may not always be an end unto themselves. Fenno’s work leaves open the possibility that evidence could have a powerful impact on lawmakers, but only in ways that are bounded in relationship with other concerns. Assuming policymakers want to choose the “best” policies as often as they can, when and how often do the external

pressures of elections and internal struggles for influence within a chamber curb the pursuit of these policies?

Deborah Stone (2002) acknowledges that elected policy makers can and do have goals beyond those related to elections, but questions if those concerns can be trumped by politics. As she shares, “Politicians always have at least two goals. First is a policy goal [but] ... Perhaps even more important, though is a political goal” (Stone 2002, p. 2). These tradeoffs between political and policy goals reflect Behn’s (2005) accountability dilemma in public management in which he describes a tradeoff between accountability of finances and fairness against accountability of performance. Such a dilemma is all too familiar to government officials. In dealing with policy, there are often tradeoffs among accountability of performance (what constituents, or a political party, or another interest desires), accountability toward finances (which programs are the most efficient, get the best results) and fairness (such as marriage equality and other civil rights questions).

These types of tradeoffs lend support to the idea that elected officials operate within a bounded rationality, an approach first developed by Herbert Simon. While rational choice may lay much of the groundwork of the institutional analysis now so often associated with social science, practical experience suggests that the complex organizations and relationships in which actors operate make a simple model of individuals acting to maximize goals an oversimplification.

Simon (1999) discusses the ways in which the rationality of people is hampered by available knowledge, cognitive ability to access that knowledge, ability to project consequences of decisions, to imagine alternative courses of action, dealing with uncertainty, and deciding amongst many competing preferences. It questions the

economic worldview that suggests individuals will always seek to maximize their utility and preferences, suggesting that the ability to make perfectly rational decisions often proves elusive in practice. Similarly, Jones (2001) reminds us that “humans are goal driven, but not fully adaptable in our current decision-making activities” (p. xi). Such is not quite the incremental “muddling through” argument of Lindblom (1959) where power is more equitably distributed so that all the major interests have an opportunity to be heard. Rather, it acknowledges that power and the ability to impact governmental actions are unequally distributed (Forester 1984). In theory, such an uneven distribution places substantial limits on the extent to which rational policies (those that have the best outcomes, regardless of other implications) prevail in the political process.

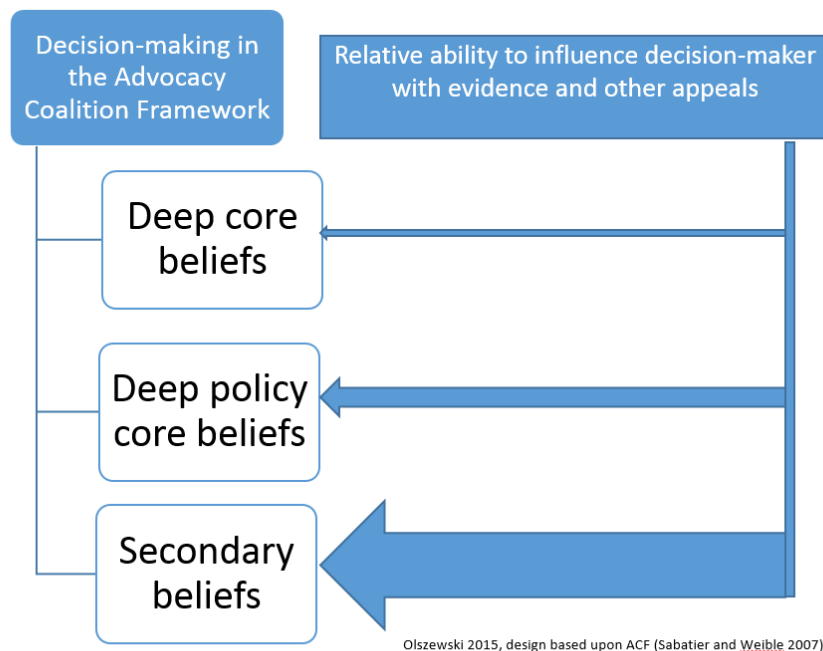
McCubbins and Schwartz (1984) found that Congressional oversight followed a similar type of rationality in practice: elected leaders have created institutions that allow citizens, interest groups, and others concerned with particular issues and oversight the opportunity to pull fire alarms if legislative action is needed. They likened their model to a fire alarm, as compared to ongoing and well-established police patrols. As they find, it is typically only “fire alarm” issues that are brought to the attention of Congressmen for action, while the daily patrol posts are left to other parts of the governmental structure. The fire alarm approach allows members of Congress to focus on other parts of their job (be they official or campaign-related) while still responding to issues as necessary. In some respects, it can be viewed as bounded rationality in practice.

The view of the world expounded by bounded rationality offers, then, at least limited opportunities to change policy based on evidence – when the environment is right. One example of this approach in action is the Advocacy Coalition Framework

(ACF). The ACF rests upon five principles, including having scientific and technical information playing a central role in the policy making process (Sabatier and Jenkins-Smith 1999). Built upon bounded rationality, it posits that varying levels of belief systems serve as the primary heuristic approach (as in Simon's theory) to decision-making (Buechner Institute for Governance 2014). Consequently, subscribers of the ACF focus on policy making subsystems and the quirks of actors, context, etc. They also focus most effort where it is likely to be most successful: at the secondary beliefs of policy makers, which are more empirical in nature and less so in the areas that deal with policy core beliefs and deep core beliefs (Buechner Institute for Governance 2014). While the advocacy coalition framework does not preclude the possibility of influencing policy makers on core personal or policy beliefs, it recognizes that there are differences between policy work in areas that may arouse great personal or policy passions (such as abortion) versus others that are likely to be influenced with more impact through advocacy and evidence (perhaps some educational-related issues). As Paul Sabatier (one of the ACF founders) and Christopher Weible (2007) write, "scientific and technical information may facilitate learning at the secondary level, but not the policy core ... because major change from within the subsystem is impossible, it must come from an external source" (p. 198). It may be that research can help to determine if the ACF hits or misses the point in this regard, both in understanding policy makers' overall views of evidence as well as their views when it relates to issues that might related to secondary, core policy, or deep core beliefs of lawmakers. In other words, how does evidence and the systems that produce evidence make themselves players in varying parts of the policy world? Is it worthwhile to produce evidence in the places where lawmakers are mostly likely to entrench

themselves with their own beliefs and/or the beliefs of their constituency? Do some systems of evidence production “do better” than others in these categories? In figure seven below, I have attempted to visualize the limitations of evidence’s ability to influence policy as expressed by Sabatier and Weible (2007).<sup>21</sup>

*Figure 7: Ability of evidence and other factors to influence decision-making in the ACF*



Jones (2001) points out that it is not enough to merely know the incentives facing policy makers, but that we must also know *how* elected leaders are thinking about the political process (*italics original*, p. ix). Can an evidence-based program (Results First) impact the “bounded” nature of rationality facing politicians? Does having access to information presented as empirical and unbiased change the decision-making landscape in all, some, or no instances?

<sup>21</sup> My visualization leaves open the possibility that evidence has some very small possibilities of impacting deep core and policy beliefs of lawmakers, but largely affirms the ACF position that change, including evidence-led change, is most likely to be achieved on issues considered to be “secondary beliefs” for lawmakers.

While it is argued here that the bounded rationality approach offers a thoughtful lens through which one can view the political process, it retains its share of critics. Stone (2002) claims that “the rationality project misses the point of politics” and that it is “an impossible dream” (p. 7). Depending on how one approaches such rational behavior, it could be argued that the bounded nature of bounded rationality *includes* provisions for Stone’s polis model of decision or policymaking. In other words, a bounded rationalist would accept that optimal policy solutions may be thwarted because of the limitations of the political process or as part of embracing the value and function that comes through deliberative democratic institutions and the normative values associated with them. Indeed, the purpose of exploring if there are ways in which evidence can become more salient within public discourse does not ignore the fact that there will almost always be some level of normative principles guiding governance behaviors. As Stone rightly points out, even if to an extreme, an extreme and full embrace of the rationalist model would eliminate the need for politics and elections altogether.

The literature suggests that elected officials might be concerned with issues beyond service in office, but questions the extent to which these other goals come into play when they might legitimately threaten re-election efforts. How legislators might resolve tensions between considerations of influence within a legislative chamber, a desire to support a colleague on an issue, and being presented with clear evidence about the value of a legislative program however, remains unclear.

#### 2.9.4 State legislative behaviors

The literature has found competing interests in the behavior of members of Congress in their service. It also explores a range of factors influencing state legislators.

In many respects, the same general competing interests exist as with Congress (re-election, influence within a chamber, pursuing good policy, for example), but research has also found that there are individual and institutional characteristics that influence how state legislators behave. These factors include but are most certainly not limited to race, gender, partisan dynamics, term limits, and legislative leadership.

Within state legislatures, women and people of color, for example, have been found to be especially attuned to other women and minorities – having those constituencies influence their legislative sponsorship agendas more than their male and white counterparts (Bratton and Haynie 1999). Women also tend to introduce more legislation dealing with women and children than their male counterparts (Thomas 1991). In addition to the tendency for women to pursue different legislation than their male counterparts, it has also been found that there has been an increasing disparity between female legislators in the two dominant political parties (Sanbonmatsu 2006), with a strong majority of women legislators being associated with the Democratic Party.

Term limits have been found to increase turnover among legislators (Moncrief et al. 2008) as well as making them feel less beholden to the constituency that elected them, allowing these legislators to become more attentive to other concerns (Carey et al. 2011). Beginning with Mayhew's *Divided We Govern* (1991), there has also been a robust debate about the extent to which legislative gridlock has been shaped though both intrabranh and interbranch rivalries (Binder 1999). On the other hand, the lack of partisan elections has been found to “break the policy linkage between citizens and their representatives in the statehouse” and the influence of parties are thus a key component

of coherence and accountability in a democratic system of government (Wright and Schaffner 2002, p. 377).

State legislators have also been shown to be more attuned to preferences of their voters as elections draw near and in situations where terms are shorter (Kuklinski 1978) and that the prospect of a competitive election forces incumbent politicians to be more interested in the potential negative reaction of voters to their behaviors and decisions while serving in office (Gordon and Huber 2007).

Within the institutional structures, the issue of leadership has been investigated and been found to depend, to a certain degree, on the context of the institution (Jewell and Whicker 1994). Depending on the state legislature, presiding officers and committee chairmen have anywhere from some to a great deal of power and authority over what gets accomplished in a state capitol. Legislative professionalism, on the other hand, has been found to have a link to legislative outcomes. Measuring this professionalism has proven wide ranging, with variations that include salary, length of session, amount of staff, and compensation for legislators, to name a few factors (Hamm and Moncreif 2012). To address these multiple components, scholars have attempted to consolidate them into categories and indexes, the most notable being the Squire Index. While there does not appear to be a one-size-fits all measurement tool of legislative professionalism, the Squire Index has been found to “accurately capture the core conceptual differences between citizens and professional legislatures” (Bowen and Greene 2014, p. 292). In a similar fashion, the National Conference of State Legislators has created a five-tiered structuring of types of legislators. These categories include full time, well paid, large staff; full-time lite; hybrid; part-time lite; and part-time, low pay, low staff (NCSL 2014).



Research suggests that legislative professionalism has changed a legislator's job (Moncrief et al. 1996), changed public opinion (Squire 1993), and had partisan implications (Fiorina 1994). Other work has found that less professional legislatures have different incentive structures for legislators than found in their more professional counterparts (Maddox 2004).

Finally, there is a growing literature surrounding the skepticism about science, especially among conservative Republicans. One of the most complete critiques of the GOP position was conducted by Chris Mooney in *The Republican War on Science*. In his book, Mooney (2005) argues that conservatives have not only taken to attacking policies favored by environmentalists, but the scientific information itself that is being used to justify arguments in favor of these policies. "To hear the modern Right tell it," Mooney writes (2005, p. 6), "you would think that environmental science, as conducted at America's leading universities, suffers from endemic corruption on a scale reminiscent of Tammany Hall." Even a study by Nisbet, Garrett and Cooper (2015) suggesting that liberals and conservatives alike can be biased against science that does not align with their political views, found that negative reactions to scientific pages challenging their views were four times greater among conservatives.

The state legislature is a complex organization (Hamm and Moncrief 2012). There are many factors shaping the ways in which legislatures ultimately reach the point of enacting public policies. In the same ways that the dynamics above shape legislative outcomes and electoral concerns, it is reasonable to believe that they could impact if and to what extent evidence utilization is undertaken. For these reasons, this research seeks to include, as appropriate, an accounting for these factors (as independent variables beyond

Results First) in its analysis of the data. But what lessons can be gleaned from the limited research about its role in the decision-making process? We turn to this question through original research in the following three chapters.

## **2.10 Chapter conclusion from the literature to new data**

The opening chapter introduced the use of evidence in policymaking and described in basic terms the efforts organizations such as Results First are undertaking. It explained the global importance of the topic and outlined the specific focus and methodology of this research.

This second chapter has provided both the more complete historical perspective on the evolution of evidence use in policymaking. In doing so, a light has been illuminated on the future direction and importance of the topic. More directly, the point has been made that significantly more original research is needed directly with a key constituency at the state policymaking level – legislators themselves.

What follows is the original research that seeks to further inform this future and meet that need, beginning with data on evidence utilization from state legislators themselves in chapter three, where the results of survey responses are unpacked and analyzed.

## Chapter 3: Evidence utilization among state legislators: the survey data

How do state legislators view evidence utilization? Using data obtained from a six-state survey, this chapter examines attitudes and practices as it relates to evidence utilization. The survey data speaks to how difficult it is for evidence to play a front and center role in the policymaking process, especially in the most controversial of topic areas. It also highlights considerable variation among evidence utilization by individual legislators, making it difficult for strong generalizations to be made across similar groups or categories. Finally, the results proved valuable in shaping the content of the qualitative interviews with legislators and staff professionals, reviewed in greater detail as part of chapter four.

First, this chapter includes a description of the survey instrument development, methodology and appropriate descriptive statistics. Next, an evidence utilization index – which serves as the key dependent variable – is developed. This measures how likely legislators are to incorporate evidence-based information into their policy considerations. Finally, this research examines how attitudes toward evidence utilization vary across key independent variables, including the existence of Results First.

### **3.1 Survey instrument development – preliminary interviews**

The data collection process for this project involved survey and interview methodologies. The process began with test interviews to help shape the questions on the survey that were distributed to all legislators in the study. Test interviews were conducted with a convenience sample of four members of the Maryland General Assembly. They

each represented a different local jurisdiction; three were men and one was a woman; three were Democrats and one was a Republican; three were Delegates and one a State Senator. Information from the test interviews and subsequent surveys from these Maryland legislators were not included in the survey analysis, but they were added to the case studies and insights from them are also included in final analysis regarding legislators and their uses of evidence in policymaking. As a former member of the Maryland House of Delegates, I was fortunate to have access and personal relationships with several colleagues that continue to serve.

These first participants were engaged in open-ended conversations about the use of evidence utilization in their work in Maryland. This narrative interviewing allowed these legislators to share their own stories and provide lengthy explanations about the type of information they utilize and how they utilize it (Jovchelovitch and Bauer 2000). The narrative interview guide can be found as Appendix D and was used to clarify issues such as the extent to which they and fellow legislators understood different evidence quality standards.

This pre-test mechanism proved valuable, helping to strengthen the design and implementation of the interview protocols while not contaminating the results from other states. While not directly considered in the survey data of this chapter, the conversations and survey responses were incorporated into qualitative lessons learned and next steps sections of this work (chapters four and six).

### **3.2 Survey methodology**

I reached out to a contact at the National Conference of State Legislators (NCSL), an organization that works with legislators across the country, who provided publicly

distributable contact data for the legislators in the six states (New York, Virginia, Massachusetts, Pennsylvania, Rhode Island and West Virginia) in which this project is focused. The state legislator survey consisted of 20 substantive questions, with an additional 15 demographic inquiries. A complete copy of the survey can be found as Appendix A. In pre-tests conducted by those associated with the survey design,<sup>22</sup> as well as with Maryland legislators, the average time to complete the survey lasted between 10 and 13 minutes. The survey was distributed via Qualtrics to every state legislator in the six states of interest for whom there was appropriate contact information available and that were in office as of January 30, 2015. Most legislatures welcome new members after the start of the new calendar year. Allowing a full month to pass ensured the sample allowed excluded departing members and included new members. When possible, personal email addresses were used from the NCSL database provided. If the contact field was empty on the NSCL database, official email addresses were added to the distribution list. In total, there were 1,038 of a total 1,049 legislators who were sent an email introduction and invitation.<sup>23</sup>

Compared to mail or phone-based surveys, leveraging the Qualtrics Internet-based option had the added benefit of being more quickly administered while also having been found to generate results that are largely representative (Fisher and Herrick 2012). The challenge to Internet-only approaches, as realized in this research, was response rates – which have been found to be lower than their mail-based counterparts (Fisher and

---

<sup>22</sup> Specifically, I completed a self-administration of the survey, as did dissertation Chair Dr. Roy Meyers and dissertation committee member Dr. Mileah Kromer.

<sup>23</sup> Four legislative seats were vacant at the time of survey distribution. There were also seven legislators who did not have publicly available contact information – either from the NCSL contact, on their formal government page or on a political page supporting their candidacy.

Henrick 2012). Compounding the challenge in this research was the “elite” nature of those being studied.

Especially with Internet-based surveys, reaching a legislative constituency is difficult. State legislators find themselves inundated with communication, with email being the primary form. Alperin and Schultz (2003) studied legislators in Minnesota and Wisconsin and found the average legislator in those states received 129 and 63 emails a week, respectively. In the same study, legislators indicated that they received significantly more communication via email than from other forms (letters, phone calls, office visits) combined, and that emails were a less effective form of communication than letters and phone calls. For example, one study comparing mail and Internet-based surveys found mailed surveys produced significantly higher response rates – the mailed survey response rate of 31.9 percent was more than double that of the Internet-based approach, which had a 12.5 percent rate (Fisher and Herrick 2009)<sup>24</sup>.

Considerable efforts were made to avoid the primary challenge of low response rates. Sending repeated requests to potential web respondents has been shown to be one of the most effective ways to increase response rates (Cook et al. 2000). For this project, respondents received one initial invitation as well as five follow ups. Most this activity occurred over the course of six weeks, with the email messages varying across the invitations (Dillman et al. 2014). For the duration of that time, the survey instrument

---

<sup>24</sup> In hindsight, I believe mailed surveys – in addition to the Internet-based approach, should have been more seriously considered as an option, especially considering the elite nature of my survey and interview subjects. There are, of course, tradeoffs in this decision; namely, the ease of administration with Qualtrics, less time and much less cost. While I remain confident that the data collected in both this and the next chapter adds a great deal of value to the literature, I believe including mailed surveys would have greatly improved response rates and helped greatly in addressing what is one of greatest limitations of this research.

remained opened. The specific timing for each message were loosely consistent with the model proposed by Dillman, Smyth, and Christian (2014).

There were four other methods utilized to help improve response rates. First, calls were made to the offices of each presiding officer for states included in the survey. In Massachusetts, a staff member to the Speaker of the House indicated he would encourage his members to participate. No other response was given from an office of a presiding officer. In practice, there was no tangible impact after these conversations.

Another approach utilized was to record a “robo call” that was delivered to all the phone numbers on file for state legislators. This recorded call outlined the project, its purpose, and reminded legislators that another email reminder would be in their inbox the next day. I also sent a cover letter with each survey informing prospective respondents that I am a former member of another state legislature conducting research in conjunction with this dissertation. However, given the huge information flow facing legislators, it is uncertain how many read and processed this information. The script for this call can be found in Appendix F and Table 6 includes the date each contact was made in the research process.

*Table 6: Survey contact calendar*

Survey Launch: Initial invitation to complete survey	Monday, February 1, 2016
Reminder email #1	Tuesday, February 9, 2016
Reminder email #2	Wednesday, February 17, 2016
Reminder email #3	Tuesday, March 1, 2016
Robo call distributed to member offices	Sunday, March 6, 2016 (voicemails)
Reminder email #4	Tuesday, March 8, 2016
Outreach to offices of Presiding Officers	Monday, March 16, 2016
Reminder email #5: final reminder	Thursday, March 24, 2016
Individual reminder emails	Various during open survey window
Survey Closed: Responses no longer accepted	Monday, April 4, 2016

Finally, legislators who opened and started the survey but did not finish within a week were sent personalized follow-up emails with a request to finish the survey.

When these efforts were exhausted, there were 116 individual legislators who opened the unique URL sent to them. Of these responses, 100 provided their informed consent to continue with the survey and two opted out. Among those giving consent to continue, there were 88 state legislators who responded to at least one of the questions posed in the survey.

Of the 88 representatives beginning the survey, there were 61 state legislators who made it to the final question (even if they skipped other, individual questions along the way). These numbers generated a “roll-off rate” of 31%. The most likely reason for the roll-off may have been associated with the survey itself: while easy to navigate, it was lengthy and included responses that required careful consideration. In other words, legislators were not able to “click through” the survey. Given this length and complexity, lawmakers who were not prepared to complete the entire survey would be unlikely to begin it (as observed in the biggest roll-off between informed consent and those answering at least one question) and may have either had difficulty or lack of desire to return to the survey later. The complete response rate picture is outlined in detail in Table 7:

*Table 7: Survey response rates*

<b>Type of Response</b>	<b>Frequency</b>	<b>Response Rate</b>
Opened their unique URL	116	11.2%
Provided informed consent	100	9.6%
Answered at least one question	88	8.5%
Provided enough data to be loaded for analysis	80	7.7%
Made it to full completion of survey	61	5.9%



These numbers were less than desired – in particular, the 7.7% response rate for sufficient data to be included in the analysis. The rate was several points lower than the approximately 12% response rate found in other studies focused on state legislators.<sup>25</sup> Thus, one could argue there is limited power in the findings due to the small N size of this study. The belief in the “law of small numbers” often provide people with erroneous beliefs about the laws of chance – considering a sample drawn from a population as highly representative (Tversky and Kahneman 1971). While this belief can be problematic even in randomly drawn samples, the skepticism is especially worth discussion when survey respondents were self-selected, as the case in this research.

To address this concern head-on, there remain many reasons to be optimistic about the value of feedback from the survey and the complete research of this project. First, as discussed above, Fisher and Herrick (2009) found that, despite the internet surveys having lower response rates compared to a mail survey, the two approaches have generated samples of similar representativeness. Moreover, the characteristics of the respondents to this survey strongly reflected the diversity of the full population of interest, at least on demographic measures. Finally, Hird’s (2005) example of a robust state legislative study had closer to a 10% response rate<sup>26</sup> -- a rate much closer that the one realized in this survey. To that end, there remains valuable information that can be gleaned from a review of this quantitative research effort.

---

<sup>25</sup> On the other hand, one technical measure of the response rate for this survey (opening of the URL) tracks with the 12% rate found in other studies of state legislators.

<sup>26</sup> There were 773 responses (Hird 2005, p. 132), with surveys sent to legislators in the 48 contiguous states within the US. This makes the denominator 7383, for an overall response rate of 10.4%. This tracks closely with the number providing informed consent for this project.

Moreover, the next chapter further supplements the survey data with in-depth conversations with elected officials and legislative staff members alike, adding a richness and depth to the data that can help affirm or question findings from the survey. These interviews include more than 20 from among the survey respondents, as well as additional interviews with four Maryland legislators, five staff members from Maryland, New York and West Virginia and one interview with the Director of the Results First Program. The research takes full advantage of the limited survey data available and tackles the limitation challenges head-on with the help of a demographically reflective data set and deeper qualitative insights that allow for deeper questioning and triangulating of the survey information provided.

Beyond response rate, this survey research sought to address the challenge of eliciting honest answers. On that front, there are several obstacles that could have influenced the results. The first potential challenge for obtaining honest responses is the tendency for respondents to provide socially desirable answers. Socially desirable responding has been labeled “one of the most pervasive response biases” in survey data (Mick 1996, p. 106), as self-reported data includes respondents who are providing answers that aim to make them look good (Paulhus 2002, Tourangeau and Yan 2007). Legislators are among those most likely to want to look good – and thus there is a risk that they responded in a way that exaggerated their own evidence utilization. However, several actions were taken to fight against this concern. This includes the confidential nature of the survey and triangulation with follow-up interviews as well as discussions with staff members. In providing a safe environment to answer questions and conducting follow up conversations on the issues covered in the survey, this work retains a strong

level of confidence that much of the risk in this space has been mitigated, even if it can never be entirely abated.

One final challenge facing the survey lie in interpretation issues – specifically, that legislators read a question or instruction to mean one thing while the intended reading was entirely different. To that end, the survey instrument worked hard to set clear definitions of evidence and specific examples, when appropriate, on policy issues. As with the concern about socially desirable answers, the additional qualitative research with follow-up legislator and new staff interviews also helped to address this challenge.

### **3.3 Independent variables: demographic data and beyond**

The survey considered many points of interest that, based on prior literature, are believed to influence how lawmakers behave while in office. These demographic traits included: gender, party, ideology, educational attainment, if the legislator faced a competitive primary election, if the legislator faced a competitive general election, race, if the legislator considers themselves in leadership, if the legislator plans to run again, if the legislator considers themselves a generalist or specialist, legislator age and years of service in the legislature.

The frequency distributions for each category are below: Of the survey respondents who made it to the demographic section of the survey, exactly two thirds were men and one third were women. The average age for legislators among respondents was 55.9, with a median of 57. Just under seven percent of the survey population was African-American, while approximately eighty-eight percent were white. 58.4 percent of survey respondents had a degree beyond a bachelor degree, and many more Democrats (66.7) responded than Republicans (31.7). There was, however, a more even distribution

of ideology: 31.7 percent described themselves as liberal, 43.4 claimed to be moderate, and 25 percent said they were conservative. Tenure tended to skew toward the shorter end of service, with only 35.7 percent of respondents reporting that they have been in office for more than 10 years.

*Table 8: Survey demographics, individual level characteristics; N = 60*

	Frequency	Percent
<b>Gender</b>		
Male	40	66.7
Female	20	33.3
<i>Total</i>	<i>60</i>	<i>100</i>
<b>Age</b>		
Mean = 55.9		
Median = 57.0		
<b>Race / Ethnicity</b>		
Black / African-American	4	6.8
White / Caucasian	52	88.1
Hispanic / Latino	1	1.7
Multiracial	2	3.4
<i>Total</i>	<i>59</i>	<i>100</i>
<b>Education</b>		
Less than Bachelor degree	7	11.7
Bachelor degree	18	30
Master degree	19	31.7
J.D., Ph.D., or equivalent	16	26.7
<i>Total</i>	<i>60</i>	<i>100</i>
<b>Party</b>		
Democratic	40	66.7
Republican	19	31.7
Independent	1	1.7
<i>Total</i>	<i>60</i>	<i>100</i>
<b>Ideology</b>		
Liberal	19	31.7
Moderate	26	43.3
Conservative	15	25
<i>Total</i>	<i>60</i>	<i>100</i>
<b>Tenure</b>		
1 to 9 years	36	64.3

10+ years	20	35.7
<i>Total</i>	<i>56</i>	<i>100</i>

In addition to personal characteristics, the survey contemplated demographics related to the electoral considerations facing the legislators in this study. When it comes to competitive elections, 18.6 percent of respondents indicated they faced a “very” competitive primary election – the exact same number who said they faced a “very” competitive general election. The least amount of competition came in the primary, with a strong majority of those completing the survey (69.5 percent) indicating they faced *no competitor* in the primary election. Likewise, respondents reported they were more likely to face no competition in the general election (44.1 percent) than either to have a somewhat or very competitive challenge (35.6 and 18.6 percent, respectively). Nearly 90 percent of those surveyed indicated their intention to seek re-election, while only one person definitively ruled out election to the same or a new office in the next election cycle. Sixty-four percent of respondents did not identify themselves as a part of their state’s leadership team. Of those who are in leadership positions, those self-identified as committee leaders were largest share (28.8 percent). Nearly 64 percent considered themselves to be legislative generalists, and there was a 58.6 to 41.4 percent split between legislators who had other professions and those who served as full-time legislators.

*Table 9: Survey demographics, electoral characteristics; N = 58*

	Frequency	Percent
<b>Primary Competitiveness</b>		
Very competitive primary election	11	18.6
Somewhat competitive primary election	6	10.2
No competition in the primary election	41	69.5
Don't Know / Refuse	1	1.7
<i>Total</i>	<i>59</i>	<i>100</i>

<b>General Competitiveness</b>		
Very competitive general election	11	18.6
Somewhat competitive general election	21	35.6
No competition in the general election	26	44.1
Don't Know / Refuse	1	1.7
<i>Total</i>	59	100
<b>In Leadership</b>		
Presiding Officer	1	1.7
Committee leader	17	28.8
Party leader	3	5.1
Not in leadership	38	64.4
<i>Total</i>	59	100
<b>Running Again</b>		
Re-election for same office	52	89.7
Election but new office	1	1.7
Do not plan to run in next election	1	1.7
Unsure of plans	4	6.9
<i>Total</i>	58	100
<b>Generalist or Specialist</b>		
Legislative generalist	37	63.8
Legislative specialist	20	34.5
Don't know	1	1.7
<i>Total</i>	58	100
<b>Full Time Legislator</b>		
Other Profession	34	58.6
Full Time Legislator	24	41.4
<i>Total</i>	58	100

### 3.4 Comparing population demographics to sample demographics

On many of the metrics of interest, the respondents of the survey appear to be very representative of the larger pool of legislators from which they come. Specifically, there are close ties on age, educational attainment, race and political party. Men are slightly under-represented in the survey data relative to their numbers in the states of interest, but not by overwhelming numbers.

The table below displays demographic data from individual states (as of February 2015) and pools that data to be compared against the demographic data of the survey.

*Table 10: Survey demographics versus population demographics*

	<b>Percent Survey</b>	<b>Percent Population<sup>27</sup></b>
<b>State</b>		
MA	21.3	19
PA	20.0	24
NY	13.8	20.2
RI	13.8	10.7
VA	18.8	13.3
WV	12.5	12.7
<b>Gender</b>		
Female	33.3	21.3
Male	66.7	78.6
<b>Age</b>		
Average Age	55.9	54.9
<b>Party</b>		
Republican	31.7	43.2
Democratic	66.7	56.1
Other	1.7	0.7
<b>Education</b>		
Less than bachelors	11.7	2.7
Bachelor degree	30	30.2
Advanced Degree	57.7	47.7
No Data on degree	0	19.3
<b>Ethnicity</b>		
Black/African American	6.8	7.3
Hispanic/Latino	1.7	2.7
Multiracial/Other	3.4	2.2
Total Minority	11.2	12.7
White/Caucasian	88.1	87.3

The fact that, overall, the survey respondents corresponds nicely when compared to the full population of interest in the six states studied makes it more reasonable to

<sup>27</sup> Source for comparisons: NCSL (September 2015).

begin to advance lines of inquiry and to hone the hypotheses set forward in the dissertation proposal.

### **3.5 Survey results**

The survey instrument included six distinct question sets, referred to as “blocks” within the Qualtrics system, which were broken down by issue area. It also collected demographic information at the conclusion. The questions asked, as well as their complete phrasing and question set assignment, is in Appendix F. What follows are the results and analysis of the results of the survey data, by question set.

#### **3.5.1 Question set one: factors that influence decision making**

Legislators were first asked to rate various factors that typically influence their decision making as it relates to two policy issues, bail reform and gun control. The text of question one is below:

There are many factors that go into the decision-making of legislators. Below are several of the factors typically given as a reason for supporting or opposing a specific piece of legislation. Thinking about a proposed change to your state’s bail requirements-- and using a scale of 1 to 5, where 1 means ‘not important at all’ and 5 means ‘very important’-- how heavily does each of the following factors influence your decision making?

The list of possible responses, randomized for each survey, included: personal opinion, constituent opinion, legislative staff expertise and research, presiding officer’s preference, committee chairman’s preference, how my colleagues view the issue, position of my political party, expected impact on crime rates, governor’s position on the issue, position of my campaign funders, expertise and research from an outside organization or university and how I have voted in the past.



Question two asked about the importance of the same set of factors provided as choices in question one, but asked legislators to focus their answers instead in the context of the issue of gun control.

Now, using the same set of factors and thinking about a proposed change to your state's gun laws, and using a scale of 1 to 5, where 1 means 'not important at all' and 5 means 'very important,' how heavily does each factor influence your decision making?

These two issues were selected to create a stark contrast identified in the Advocacy Coalition Framework (ACF). In this circumstance, questions about bail were meant to be representative of the "secondary beliefs" structure where evidence and other efforts are more likely to substantially influence policy outcomes. In comparison, using gun control as the other policy area of interest was meant to represent the "deep core belief" category of the ACF, where evidence and outside influence is believed to have limited or no impact on policy outcomes. Table 11 below includes the percentages of respondents who registered a response at the higher end of this scale only, suggesting that the factor was a high priority for them as they made decisions on that legislative issue.

*Table 11: Frequency distribution; factors that influence decisions – bail & gun control. Results in percent and indicate a 4 or 5 on the 5-point scale; N = 88 (bail), 80 (gun)*

	Bail	Gun Control	Difference
<b>Evidence Based Factors</b>			
Legislative staff expertise and research	55.3	39.0	-16.3
Expertise and research from an outside organization or university	58.7	49.4	-9.3
<b>Personal/Political Based Factors</b>			
Personal opinion	57.7	63.8	6.1
Constituent opinion	71.8	73.4	1.6
Presiding Officer's preference	10.0	6.9	-3.1
Committee Chairman's preference	14.5	10.4	-4.1
How my colleagues view the issue	20.5	14.4	-6.1
Position of my political party	13.1	11.8	-1.3
Governor's position on the issue	9.5	6.5	-3.0
Position of my campaign funders	5.6	5.4	-0.2
How I have voted in the past	50.0	49.4	-0.6

Expected Impact			
Expected impact on crime rates	77.7	73.1	-4.6

The most influential factors upon legislative decision making in this sample are expected impact on crime rates, constituent opinion and personal opinion. All three of these factors were ranked as influential by well over 50 percent of the sample in both scenarios. The *only* factors to increase influence on legislative decision making when transitioning from a question on bail reform to gun control were the opinions of constituents and the opinions of the legislators themselves.

Especially telling in this question is the *strength of negative differential* experienced by *all* evidence-based factors as it relates to shaping legislative positions when the question moves from bail reform to gun control. Legislative staff expertise suffered the biggest drop in influence, with expertise and research from outside research organizations and universities also experiencing a significant drop. Even *the expected impact on crime* itself as an issue mattered less when the topic at had switched from questions about bail reform to those on gun control.

The data suggests that on gun control issues, policy positions are informed as much by constituent opinion as expected impact on crime rates. In fact, a slightly larger number of legislators indicated that the views of constituents (73.4%) was very or extremely influential on their position on gun control than those (73.1%) who said a known impact on crime rates would impact their vote. These kinds of numbers support the hypothesis presented by the ACF. Specifically, it affirms the belief that some policy areas are more amenable to change driven by evidence than others.

To examine these differences even further, I completed a paired sample *t* test, examining the differences in means between the bail and gun control scores as outlined in Table 11 above.

*Table 12: Test of significance between bail and gun control scores on factors influencing legislative decision making*

	<i>t</i> score	df	Sig. (2-tailed)
<b>Evidence Based Factors</b>			
Legislative staff expertise and research	2.357	73	0.021 **
Expertise and research from an outside organization or university	1.581	73	0.118
<b>Personal/Political Based Factors</b>			
Personal opinion	-1.242	77	0.218
Constituent opinion	-0.11	76	0.912
Presiding Officer's preference	3.134	67	0.003 **
Committee Chairman's preference	2.474	74	0.016 **
How my colleagues view the issue	2.614	74	0.011 **
Position of my political party	-0.903	72	0.369
Governor's position on the issue	2.382	72	0.02 **
Position of my campaign funders	0.575	69	0.567
How I have voted in the past	-1.438	68	0.155
<b>Expected Impact</b>			
Expected impact on crime rates	0.567	73	0.572

*\*\* significant at the .05 level*

This analysis found several statistically significant impacts. Namely, there was a statistically significant difference with the following factors on legislative decision-making: the expertise and research of legislative staff, the preference of the presiding officer of a legislative chamber, the position of the chairman of a standing committee, the views of a legislator's colleagues and the position of the governor. The statistically significant decline of reliance on legislative staff when moving from bail to guns further affirms the literature that says evidence is most likely to play a role in lower profile and less politically charged issues. Similarly, a statistically significant decline for the influence of governors, presiding officers, committee chairs and the views of peers

speaks to the ways in which high profile and issues of core policy beliefs are more difficult to influence. The fact that these political and personal factors, ingrained into the typical lawmaking processes, also experience a statistically significant decline in importance, affirms how important the issue at hand is for those seeking to influence legislation. In other words, if one is looking to change laws, the easiest way is to focus on issues that are under the radar. Finally, I would note that personal and constituent opinion, the only factors experiencing an overall increase of meaningful reliance in Table 11, were not statistically significant – a result I interpret a function of how important these factors are regardless of the issue. That importance is further reinforced in the next chapter.

#### 3.5.2 Question set two: evidence, personal values, and constituent opinion

Evidence as a stand-alone concept is something for which individuals and legislators alike are much more likely to express support. However, a truer measure about the belief in evidence can be identified when support for evidence is directly contrasted with other major influences on policy-making decisions; namely, personal opinion and constituent feedback. As seen in the responses to the questions in question set one, evidence-based factors are among the highest rated influences on legislative decision-making. So too were non-evidence based factors such as personal opinion and constituent feedback.

To get at this truer measure, the survey was structured to create more of a forced choice between reliance on evidence and these other decision-making influences on policy making. The survey asked questions about whether respondents would tend to rely more upon evidence or personal values, constituent opinion and policy preferences,

respectively, if the evidence conflicted with these other values. Specifically, questions one, two and three of this section were presented to survey respondents in the following way:

Information: For the purposes of answering the remainder of questions, please consider ‘evidence’ as the best available information related to the development of a particular policy and its implementation.

Thinking about policy decisions on criminal justice issues, if the available evidence were to conflict with your personal values, in your decision-making you would tend to rely more upon [sliding scale between evidence and personal values] when making decisions on criminal justice policy?

Thinking about policy decisions on criminal justice issues, if the available evidence were to conflict with your general policy preferences, in your decision-making you would tend to rely more upon [sliding scale between evidence and general policy preferences] when making decisions on criminal justice policy?

Thinking generally about policy decisions on criminal justice issues, if the available evidence were to conflict with what you are hearing from constituents on the issue, in your decision-making you would tend to rely more upon [sliding scale between evidence and constituents] when making decisions on criminal justice policy?

Respondents were again asked to use a 1 to 5 scale, with 1 indicating they would completely rely upon their personal values, policy preferences, or constituent feedback and a 5 meaning they would rely entirely upon the evidence presented. Table 13 reports the extent to which legislators reported a reliance on evidence relative to these three other considerations.

*Table 13: Frequency distribution; Legislators relying on evidence compared to select influences.*

*Results in percent indicate a 4 or 5 on the 5-point scale; N = 75*

	Percent
Evidence utilized more over ...	
Personal Values	46.7
Constituent Opinion	52.7
Policy Preferences	54.1
Evidence utilized equally with ...	
Personal Values	40.0

Constituent Opinion	32.4
Policy Preferences	35.1
Evidence utilized less than...	
Personal Values	13.3
Constituent Opinion	14.9
Policy Preferences	10.8

As shown in the table above, approximately 50 percent of legislators self-reported a willingness to rely more on evidence in their policy-making decisions over their own values, constituent opinion and their policy preferences when evidence directly contradicted those considerations. Between 32.4 and 40 percent indicated that the evidence and these influences were weighed equally when in conflict with each other and approximately 15 percent continued to rely upon personal values, constituent opinions and policy preferences over evidence when they were at odds with what the evidence suggested the best path forward. Thus, it can be said that nearly half the legislators in this data set acknowledge that they continue to rely at least as much on personal values and constituent opinion *even when there is evidence directly contradicts those positions*. Based on results in question one, these numbers could be exacerbated if they were related to more controversial issues, rather than asked as a generic question regarding criminal justice. Put more directly, the open-ended nature of this question did not force legislators to make, even in theory, a difficult decision regarding choosing the weight of the evidence over more practical political considerations. While not tested, I suspect there would be a *much* greater reliance on non-evidence factors reported in a forced-choice scenario if the choice was between, say, the preferences of constituents and evidence on a highly partisan and salient political issue – abortion, gun control, the death penalty.

Still, there is no way to entirely prove this suspicion based on the structure of the survey and this information highlights the challenges of interpreting survey responses in a vacuum. It begs additional questions that can be partially answered by a deeper dive into the issue and that are partially answered in chapter four: What are the reasons legislators rely more heavily on factors other than evidence? Are there common characteristics of legislators who express a preference toward reliance on evidence? There are some of the issues that are explored further in individual conversations with the legislators responding to this survey. From my prior experience as a legislator, these numbers are unsurprising and may even reflect an over-estimation regarding the norms of a typical legislator relative to evidence. While many legislators espouse their policy decisions are driven by sound evidence, the reality is that (and other literature has confirmed) there are real and pressing constraints requiring decisions be made for normative and/or political purposes. Thus, legislators have always and many will continue to rely on personal and constituent opinions at least as much as evidence – even when the evidence directly contradicts those positions. The fact that so many (nearly half of all) legislators who self-selected into this study are willing to admit that in general is the case, one must wonder how much higher the number is for controversial issues as well as for those who did not join the research effort.

Still, the survey data has established that evidence is reported to be an important factor in legislative decision-making. The next set of questions seeks to better understand where legislators find their information. It also seeks to determine the relative influence each information type has upon survey respondents.

### 3.5.3 Question set three: finding evidence and its influence by type

Where legislators find their evidence says something about its quality. At the beginning of this question set, legislators were instructed to consider evidence as “the best available information related to the development of a particular policy and its implantation” in their work. This was done to frame the thinking of respondents so that they considered both where they obtained their information and so they believed the information they were accessing was only of top quality.

On a scale of 1 to 5, with 1 being ‘never’ and 5 being ‘all the time,’ where do you typically find evidence that relates to legislation that you are considering?

The evidence sources legislators were asked to rank, which were randomized on each survey, were legislative staff, state or federal agencies, legislature’s policy organization, think tanks, other legislators, institutions of higher education, lobbyists or interest groups, Results First reports, news media and governor’s office. By far, legislative staff is the most consulted form of information among those responding to the survey, with 49.3 percent of legislators ranking staff as a resource that is accessed most or all the time for legislative activity. In addition to legislative staff, policy organizations of the legislature as well as state and federal agencies are identified as the most regularly consulted sources by legislators in the sample.

*Table 14: Sources regularly consulted by legislators for legislation they are considering. Results in percent indicate a 4 or 5 on the 5-point scale; N = 74*

	Percent
I regularly find evidence from...	
Legislative staff	49.3
State or federal agencies	36.1
Legislature’s Policy Organization	33.3
Think tanks	29.7
Other legislators	24.7
Institutions of higher education	24.3
Lobbyists or interest groups	23.9



Results First reports	18.0
News media	13.9
Governor's office	7.1

---

Nearly 14 percent of legislators responding to this survey indicated that news media is a significant source of trusted evidence in their work. Lobbyists and interest groups were considered a source of evidence most or all the time to approximately 24 percent of the survey respondents – about the same number of legislators in the sample who turn to other legislators and institutions of higher education. The Governor's office was listed as the lowest of all sources for evidentiary quality, with only just over seven percent of respondents reporting that they frequently utilize it in their work. In contrast, a governor's *position* on a bill could be much more relevant in the policy process, where for example his or her veto could be impactful or if there were other sources of influence at his or her discretion.

These numbers are interesting in the context of the qualifier provided at the beginning of this question set regarding what “evidence” should be considered. In other words, 14 percent of legislators not only implied the news media is a source they regularly consult, but also that it is a place offering the “best available information” related to their legislative endeavors. Based on follow up interviews in the next chapter, there is reason to believe these numbers accurately reflect the feelings of legislators as the news being on par with other, more rigorous information outlets.

Another way to put these numbers in context are to compare them with the 2005 legislative survey from Hird, which was similar in scope but larger in size. Hird's overall work explored nonpartisan policy research organizations (NPROs) in depth, providing insights as to “how power affects expertise” as well as “how expertise effects power”

(2005, p. xv). In addition to survey data from the NPROs among the 50 states and follow-up interviews with NPRO agency directors, there was also legislative survey data collected from over 750 state policymakers. There was especially strong alignment between this and Hird’s research regarding questions about sources from which legislators most rely upon to reach policy decisions. Table 15 below calculates the mean scores for both the questions in this research as well as an adjusted mean for the Hird research. As converted to a five-point scale, a score of “1” refers to a source where legislators “never” find evidence related to legislation they are considering, while a score approaching “5” means “all of the time.”

*Table 15: Comparison of Olszewski to Hird (2005, p. 139) means<sup>28</sup>*

Source	Olszewski Mean	Hird Mean - Adjusted
Constituents	NA	4.23
Think Tank	2.84	2.98
Legislative Staff	3.36	3.37
Results First	2.52	NA
State/Federal Agencies	3.13	3.11
Lobbyists/Interest Groups	2.82	3.2
Higher Education	2.93	3.16
Legislature Policy Organization	2.9	3.49
Governor's Office	2.11	2.89
Other Legislators	2.95	3.75
News Media	2.42	2.56

The comparison of means produces relatively consistent findings between the two research efforts. The consistency in scores, more than a decade apart, speaks to a well-established set of norms vis a vis sources of legislative information outlets – and to how slowly change might be realized. The consistency also provides a layer of confidence to

---

<sup>28</sup> Hird means were calculated by dividing the original score by seven and then multiplying that percentage by five (to equalize to the Olszewski scale score. Constituents were not measured by Olszewski and Results First was not measured by Hird. For a complete list of other adjustments on this table, see appendix item 3E.

the validity of survey instrument used in this research. While Hird's highest rated source, constituents, was not measured in this question, prior questions in this survey (and subsequent information in chapter four) support the idea that legislators are likely to turn to those they represent as a critical information source. Ironically, such a source is among the least likely to have their information grounded in sound evidence.

Legislative staff is another area meriting further discussion. Hird's (2005) research found, beyond constituents, nonpartisan legislative staff were judged to be far more important than most other sources of information (a 3.76 out of 5 when converted to the scale I deployed). While the legislative staff ranks high and is consistent in the comparison chart, Hird's analysis separated partisan and non-partisan staffing – and there was a considerable gap between how important legislators viewed each in his or her work. Given how closely the other numbers tracked, it is likely there would have been a similar gap if legislators had the opportunity to differentiate between partisan and non-partisan staff options in this research. Similarly, the generic “legislative policy organization” that served as an organizational placeholder for groups which might house Results First could have likely scored higher if, as in Hird's survey, it was instead named legislative special committees and/or task forces. In either case, outside of constituents (and a legislator's own views), legislative staff is claimed to matter as an important source of information in both this and Hird's research – an issue explored more deeply in the next chapter.

On the topic of Results First, this question set was also the first real test of the strength of the program. Even though only half of the states included in this sample had a Results First program active within their state, 18 percent of respondents claimed to

regularly rely upon their reports for evidence production. Looked at it another way, the percentage of legislators actively turning to Results First as a trusted source of evidence remains only a few points above the news media. Taken from the former perspective, Results First could be over-performing relative to other sources within the states where it currently operates; from the latter, it remains a niche program with whom a select few legislators engage and utilize in the policymaking process.

Regularly consulting evidence might suggest information carries greater influence in legislative deliberations; however, it could also be an entirely different question. To better understand how influential varying types of information are to legislators, survey respondents were next asked about how influential several types of evidence are to them.

Thinking about how evidence is used in your decision-making process, on a scale of 1 to 5, with 1 being ‘not influential at all’ and 5 being ‘extremely influential,’ how influential would you say each type of evidence is in your decision-making process?

The information sources respondents were asked to rate included program cost / fiscal impact, formal program evaluations, information program evaluations, professional literature on the issue, scientific studies on the issue, letters, e-mails and phone calls, benefit-cost evaluations, testimony on the issue, data from performance reporting systems and Results First Reports. They were randomized on the surveys and the responses are listed the table below, broken down into evidence based factors as well as non-evidence based factors.

*Table 16: Influence of evidence by type. Results in percent indicate a 4 or 5 on the 5-point scale; N = 70*

	Percent
Higher Level Evidence Based Factors	
Formal Program Evaluations	44.9
Professional Literature	54.3
Scientific Studies	68.5

Benefit-Cost Studies	67.6
Performance Management Data	58.5
Results First Reports	29.8
<b>Non or Lower Level Evidence Based Factors</b>	
Program Cost	61.5
Informal Program Evaluations	16.2
Letters, e-mails, calls	32.9
Testimony on the issue	51.4

Overall, scientific studies, benefit-cost analysis and program cost were the top three information sources rated to have higher levels of influence in legislative decision making. While highly influential, benefit-cost studies take considerable time and effort and are not available for the thousands of legislative decisions made annually. In fact, it is unrealistic to expect that either robust benefit-cost or Result First evaluations would be able to be conducted but for a few issues. Even those issues would need to be identified far in advance of contemplated legislative action – a task that is often difficult to accomplish. Performance management data could be a more regular, ongoing feedback loop for legislators, but even here questions remain about the extent to which lawmakers have access to this information. Moreover, performance management data would not exist for a program not-yet created, compounding the challenges of finding quality evidence to inform decision-making.

While program cost makes sense as a highly rated factor influencing legislative decision-making (61.5 percent of respondents report they are highly influenced by cost), money alone is a poor determinant of program or policy quality and is no guarantee of effectiveness. The fact that so many legislators rank program cost alone as a critical factor – nearly as much as benefit-cost studies, which more holistically examines both the expenditures and savings/benefits of policies – is telling. It makes sense lawmakers want

to know how much they are spending on the programs of the government (after all, budgets are fixed and deficit spending is not a standard practice at the state level), but it could also be true that decisions based on cost alone could be wasteful and cost more in the long run. It might also be true that spending more now might, in the long run, turn into budget reductions or new revenue which can offset new expenditures.

Testimony was also strongly relied upon by more than half of the legislators responding to the survey, even though this type of information could be more troublesome than program cost numbers. This is because there are no guarantees testimony on the issues facing a legislature are grounded in solid evidence, let alone the truth. Indeed, most individuals offering testimony typically have a position they are trying to advance or defend, and their verbal and written testimony is almost always structured in such a way as to only support their case. Letters and emails are rated about 20 percent lower (32.9 percent of legislators are regularly influenced) than testimony, even though often the content remains the same in both formats.

Results First reports were the lowest rated influence factor among all the evidence-based options and near the bottom of all factors respondents were presented. It is worth noting that Results First purports to utilize and integrate findings from some of the other, higher-level sources that legislators report have more value. Thus, these results could reinforce the idea that Results First, even if working, is doing so as a niche program: it offers a limited scope of policy choices from which they work and take-up in only selected states, among other reasons. It might also be partially (as seen in other questions) due to the limited number of legislators who have had exposure to the work

and a general unawareness of the work the program conducts, which is a different kind of way to describe the program as niche.

#### 3.5.4 Question set four: personal values, constituent opinion, and policy area

In question set two, nearly half of the survey respondents reported they would be willing to rely upon the evidence above the views of their constituents, their personal beliefs and their personal policy preferences. Question set four of the survey delves deeper into these tensions, this time by focusing on specific criminal justice policy questions. This question set provides insights that go beyond the general issue of criminal justice and contrasts the two most difficult considerations to ignore, personal and constituent opinion, against the weight of the evidence in three specific areas of criminal justice.

Legislators were asked about the extent to which they were willing to rely upon the evidence produced by organizations with formal ties to a legislature, relative to personal and constituent beliefs. The same question was again asked in three policy areas: drug treatment funding levels, changes to parole and probation eligibility, and changes to drug sentencing.

Introduction: For each issue below, relative to your personal beliefs and those of your constituents, how important would evidence produced by people with formal ties to the legislature be for each of the following topics? Formal ties to the legislature means evidence produced by legislative staff and/or outside organizations that have formal partnerships with the state legislature.

Relative to your personal beliefs and those of your constituents, on a scale of 1 to 5 with 1 being 'evidence is not important at all' and 5 being 'evidence is very important,' how important do you think evidence produced by organizations with formal ties to the legislature would be on legislation that would change your state's drug laws:

Relative to your personal beliefs and those of your constituents, on a scale of 1 to 5 with 1 being 'evidence is not important at all' and 5 being 'evidence is very important,' how important do you think evidence produced by organizations with formal ties to the legislature would be on legislation that would alter parole and probation eligibility standards:

Relative to your personal beliefs and those of your constituents, on a scale of 1 to 5 with 1 being 'evidence is not important at all' and 5 being 'evidence is very important,' how important do you think evidence produced by organizations with formal ties to the legislature would be on legislation that would change funding levels for drug treatment programs, but neither create nor eliminate existing programs?

These three areas were kept intentionally vague; for example, none of them suggested changes would either increase or decrease sentences, eligibility or funding.

*Table 17: Reliance on evidence over personal or constituent opinion by policy area. Results in percent indicate a 4 or 5 on the 5-point scale; N = 67*

	Percent
<b>Criminal Justice Policy Area</b>	
Changes to Drug Laws	47.8
Parole and Probation Eligibility Standards	60.6
Drug Treatment Funding Levels	62.7

As expected, the rising level of issue controversy corresponded with a decline in reliance on evidence to inform policy making. While 62.7 percent of lawmakers in the survey were willing to rely more on evidence than personal and constituent opinion on drug treatment funding levels and 60.6 percent for parole standards, the number significantly dropped (to 47.8 percent) when asked to consider a fundamental change to drug laws themselves.

These responses generally track with those these rates to the “general criminal justice” preferences expressed in question set two. There, evidence was reported by respondents to typically trump personal values by 46.7 percent, constituent opinion by 52.7 percent, and general policy preferences by 54.1 percent. It makes sense that less



controversial issues (such as funding levels) track higher than these 50-percent range numbers, while the more controversial changes to drug laws drops below the numbers shared in question set two.

In all cases, I specifically asked about evidence produced by organizations with formal ties to the legislature as these are the organizations which most directly resemble – or directly rely upon results produced by the Results First program. Whether there is an active Results First presence, the contextualization of this question allowed legislators from all states to weigh in on how much they rely upon what should be the absolute best evidence coming from their most trusted sources – those directly and formally connected with the legislature. It also affirmed that, the more controversial the issue, the less likely evidence is to matter within legislative deliberations.

#### 3.5.5 Question Set five: evidence and legislative work

Respondents were also presented with various statements regarding the evidence produced by organizations with formal ties to the legislature and how that evidence applied to their work. These questions were broken down into “agree” or “disagree” categories, although legislators were also given options to “strongly” agree or disagree. The totals below reflect both agree/strongly agree as well as disagree/strongly disagree. Percentages totaling less than 100 reflect the balance of respondents indicating they did not know if they agreed or disagreed with each statement.

*Table 18: Evidence produced by organizations with formal ties to the legislature. Percentages reflect those both agree/disagree as well as strongly agree/disagree; N = 62*

	Agree	Disagree
Is relevant to my work	83.9	14.5
Provides intellectual background or framework to my work	12.9	3.2

Creates an opportunity for common language to work with	93.5	4.8
Helps me reach across partisan divides	74.2	22.5
Makes it easier to go against what my constituents want	16.1	51.7
Makes it easier to go against what my party leaders want	46.7	48.4
Makes it easier to go against what legislative leaders want	53.2	41.9
Makes me challenge my own views on legislation	83.9	14.5

The response with the most legislators agreeing was the statement about evidence produced by organizations with formal ties to the legislature creating an opportunity for common language to work with. High scores were also associated with the degree to which respondents believed this type of evidence was relevant to their work (83.9 percent agreed), made them challenge their own views on legislation (83.9 percent also agreed with this statement) and helped them reach across partisan divides (74.2 percent agreed with this statement).

There remained a dramatic drop when it came to going against what constituents, party leaders and legislative leaders want. When it comes to constituents, there were very few legislators willing to rely on evidence produced by organizations with formal ties to the legislature to go against what their constituents want – only 16.1 percent replied such evidence helps in that regard. This is consistent with the survey responses of legislators in earlier question sets and the findings of Hird’s research.

Continuing a line of questioning about organizations with formal ties to the legislature, survey respondents were next asked about if they felt the evidence from these organizations were of greater quality, equal quality, or lesser quality when compared to other forms of evidence.

Do you view the quality of evidence produced by organizations with formal ties to the state legislature as less than, equal to, or greater than the quality of other types of evidence?<sup>29</sup>

The responses to this question are outlined below in Table 19.

*Table 19: Quality of evidence produced by organizations with formal ties to the legislature as compared to other forms of evidence; N = 62*

	Percent
Evidence quality is –	
Less than other forms of evidence	8.1
Equal to other forms of evidence	67.7
Greater than other forms of evidence	21.0
Don't know	3.2

The clear majority of respondents to this question felt evidence produced by organizations with formal ties to the legislature was equal to other forms of evidence. As individual legislators, there are limited opportunities to influence the makeup and design of these organizations with which they collaborate. As a collective, legislatures could entirely refashion these organizations into those where evidence production is viewed as second to none. The lack of changes internally may suggest that there is limited interest in or political capital available to make these changes possible. Consequently, only 21 percent of legislators in the survey said they felt organizations associated with their legislature produced better quality evidence than other types. More than eight percent of legislators felt that the evidence quality was *worse* when coming from organizations affiliated with the legislature. Collectively, more than 75 percent of respondents feel the quality of evidence coming from organizations with formal ties to their legislature are no better than or worse than other forms of evidence in their quality. For policy scholars,

<sup>29</sup> Respondents were reminded (again) at the top of this question set about which definition of evidence to use. “Formal ties” to the legislature was also defined, meaning “evidence produced by legislative staff and/or outside organizations that have former partnerships with the state legislature.” This was intended to include organizations such as Results First.

these numbers raise some red flags; if legislators question the quality of the evidence being produced by their *own* organizations relative to other outlets, there are legitimate questions about the quality and source(s) of the information they are seeking. The doubt surrounding legislative policy organizations could be tied to myriad issues, making it even more important to understand where policymakers think “quality” evidence comes from and even what constitutes quality evidence.

Having a better sense of how legislators view information from organizations with formal ties to the legislature, another question asked the idealist question of how often evidence *should* play a prominent role in policy decision-making.

On a scale of 1 to 5 with 1 being ‘never’ and 5 being ‘all of the time, how often do you think evidence should play a prominent role in the decision-making process on public policy issues?

*Table 20: How often evidence should play a prominent role in policymaking; N = 61*

	Percent
Evidence should play a prominent role in policymaking	
Never or rarely (rating of 1 or 2)	0.0
About half the time (rating of 3)	13.1
Most or all of the time (rating of 4 or 5)	86.9

No lawmakers responding to this survey suggested evidence should not play a prominent role in policymaking (there were no scores of 1 or 2 provided for this question) and most, in fact, suggested that it “should” play a role most or all the time – suggested by nearly 87 percent responding with either a 4 or 5 to this question. The “should” aspect of this question may explain why the scoring was so high on when evidence should play a role in policymaking, but how significantly lower scores were realized when prior questions pitting evidence directly against other factors. In this case, the ideal does not appear to be aligned with the reality.

I also asked a question about when evidence is used by legislators within the legislative process. The options presented to legislators included research, bill drafting, committee work, floor debate and/or policy implementation. Respondents had the option to check all situations that applied to their personal circumstance.

*Table 21: Use of evidence in the legislative process; N = 56*

	Percent
<b>When Evidence is Utilized in the Legislative Process</b>	
Research	90.3
Bill Drafting	80.1
Committee Work	82.3
Floor Debate	74.2
Oversight of Policy Implementation	51.6

As seen in Table 21, evidence is used heavily in the research process among respondents, with more than 90 percent indicating it at this step in the legislative process. As the process continues, evidence begins to play an increasingly less influential role, dropping to 82.3 percent during committee work, about 80 percent for bill drafting, 74.2 percent during floor debate, and just over 50 percent for policy implementation.

#### 3.5.6 Question set six: Results First

Question set six explicitly asked legislators about their experience with the Results First program. Earlier in the survey process, more legislators expressed use of and reliance upon the efforts of Results First than the number of those expressing an awareness of the program even being active within their state.

The first item in this question set sought to determine how knowledgeable legislators were with the Results First Program. Respondents were asked to rate their familiarity with Results First, with a score of 1 meaning “not knowledgeable at all” and a ranking of 5 meaning that they are “very knowledgeable” of the program.

*Table 22: Knowledge of Results First program; N = 57*

	Percent
Knowledge of Results First program	
1 - "Not knowledgeable at all"	73.7
2	19.3
3	5.3
4	1.8
5 - "Very Knowledgeable"	0.0

As the table suggests, respondents have very limited knowledge of the Results First program. No one who responded suggested they were very knowledgeable about the program, and 73.7 percent admitted that they were not knowledgeable at all. These data stand in strong contrast to the larger number of elected officials who claimed to rely upon Results First Reports in an earlier question set. The crosstabulations on this question affirmed that a lack of knowledge was true among respondents from Results First and non-Results First states alike. While there was a very slight increase among scores of “2” and “3” in Results First states (three additional “2” scores and one additional “3” score), the overwhelming majority of respondents were “not knowledgeable at all” – 82.8 percent from non-Results First states and 64.3% from Results First states.<sup>30</sup> None of the cross-tabulations returned any results of statistical significance.

Respondents were asked next about the impact of the Results First, with a score of 1 meaning the program has “no impact at all” and a score of 5 meaning it has a “significant impact.”

*Table 23: Impact of Results First program; N = 32*

	Percent
Impact of Results First program	
1 - "No impact at all"	62.5
2	9.4

<sup>30</sup> See appendix H for a full split in responses between legislators from active and non-active Results First states. For all crosstabulation calculations, there was almost an exact 50/50 split between Results First and non-Results First state respondents.

3	18.8
4	6.3
5 - "Significant Impact"	3.1

While only 26.3 percent of respondents admitted *any* knowledge of the Results First, most of which included a limited knowledge of the program, 37.6 percent reported that it nonetheless impacted legislative activity within their state. Moreover, even though only 1.8 percent of respondents had a significant knowledge of Results First, 9.4 percent reported the program represented a more significant impact. At first glance, these combined results suggest Results First remains a largely unknown commodity amongst state legislators – but also that its impact is significantly larger than knowledge of the actual programmatic footprint. Most telling in the crosstabulations for this question were the approximately 30 respondents<sup>31</sup> that skipped the question. While I am ultimately unsure why this occurred, it is possible that legislators would have preferred a “do not know” option considering the limited knowledge of the program identified in the previous question.

I also wanted to know if those responding knew if Results First was active in the states of those responding. In the context of prior survey data, it came as no surprise that most of legislators in the survey population did not know if the program was active within their home state.

*Table 24: Is Results First active in your state? N = 59*

	Percent
Results First active in state	
Yes	1.7
No	22.0
Don't Know	76.3

<sup>31</sup> For context, at this point in the survey, 30 respondents represents approximately half of the survey population.

Reviewing the crosstabulations for this question, there were six legislators who correctly identified that their state was not a Results First state, one who correctly said their state was a Results First state, and seven who *incorrectly* reported that their state was not active, even though the Results First program is active in their state.

Finally, I asked if survey respondents personally utilized evidence produced by Results First to formulate a policy position. This was asked in such a way that it could have only helped influence and not be the only factor.

*Table 25: Personally utilized Results First to formulate policy position; N = 58*

	Percent
Personally used Results First to formulate policy position	
Yes	3.4
No	74.1
Don't Know	22.4

Despite the view that the program has an influence larger overall than the knowledge of the programmatic footprint, individual legislators within the survey largely did not identify Results First as a source to which they turned to formulate their own policy positions. In fact, nearly three-quarters of legislative respondents said they either did not rely on Results First (74.1 percent) or they didn't know if they personally utilized it (22.4 percent). A mere 3.4 percent acknowledged personally relying on Results First to formulate a policy position. Even more notable, the two legislators representing this 3.4 percent are lawmakers from non-Results First states.

### **3.6 Dependent variable: evidence utilization index**

Attempting to quantify “who uses” evidence should be done with abundant caution. Still, there are ways to begin sorting out those who are, at least from self-reporting tools such as this survey, higher and lower users of evidence in their policy-



making processes. For purposes of this project, an **evidence index** was created to help make meaning of the responses provided by the state legislators from the states of interest. To begin with the development of such an index, emphasis was placed on questions within the survey instrument that reasonably relate to the likelihood a legislator prioritizes evidence utilization in their own policy deliberations.

For example, answers to questions such as “when the available evidence conflicts with my personal beliefs, I tend to...” that skewed toward a reliance on evidence over personal opinion earned a legislator one “point” toward a total possible score. Similarly, high scores on questions about obtaining information from organizations inclined to produce credible, quality evidence (such as state evidence policy shops) awarded a “point” to legislators toward the evidence index score. Conversely, high scores to other information sources, such as newspapers and the Governor’s office, did not include the ability to earn a point towards the evidence index score.

Because all questions considered for the evidence index (and for that matter, nearly all questions in the survey) were based on a 1 – 5 scale, there were natural lines of differentiation. In the above example, legislators who used a 4 or 5 to indicate they used evidence over personal beliefs in conflict were given a point, while legislators who indicated a 50/50 balancing (with a 3 score) or greater reliance on personal beliefs (using a 1 or 2) were not awarded a point toward their evidence index score.

There were 16 total questions on the survey instrument which would have enabled legislators to indicate a preference for evidence over other considerations. These 16 questions are identified as contributors to the evidence index score with an (EIS) notation next to the relevant question in the appendix.

By allowing both 4 and 5 responses (essentially, equivalents to moderately and strongly support evidence above other considerations), the evidence index score goes beyond a strict interpretation about utilizers of evidence in the policy process, as such absolute approaches are extremely rare in the policymaking arena<sup>32</sup>. There were 41 individual legislators who answered all of the questions enabling the earning of a “point” toward the evidence utilization score; thus, only these respondents were included in the frequency distribution below so as to maintain a level playing field among scores.

*Table 26: Frequency distribution, evidence utilization index – liberal interpretation; N = 41*

	Frequency	Percent
<b>Evidence Utilization Index Score</b>		
1	1	2.4
2	1	2.4
4	4	9.8
5	6	14.6
6	1	2.4
7	2	4.9
8	4	9.8
9	2	4.9
10	4	9.8
11	3	7.3
12	4	9.8
13	4	9.8
14	1	2.4
15	2	4.9
16	2	4.9
<i>Total</i>	41	100

<sup>32</sup> In fact, using only “5” scores on the questions addressing evidence utilization over other considerations was examined. In that process, nearly the entire set of respondents were rated on the very low end of evidence utilization. While noteworthy and speaking directly to the challenges of applying evidence as the leading consideration in the public policymaking sphere, insufficient variation eliminated meaningful capacity to conduct comparisons between those more inclined to utilize evidence and those less include to utilize evidence. To that end, I viewed such an absolute standard for reliance on evidence as too strict of an interpretation.

As seen by the chart, the Evidence Index score produces a wide distribution of policymakers and their relative engagement with evidence in the policy-making process. The average evidence utilization score is 8.98, nearly identical to the median score of 9.0. The mode for the evidence utilization score is 5.0.

There are six legislators at the lowest end of the evidence utilization spectrum (scoring either a 1, 2, or 4) and five legislators at the highest end (scoring a 14, 15, or 16). These legislators are given special attention in attempt to conduct follow-up interviews, to be sure that a full stretch in viewpoints are provided in that format.

### 3.6.1 Index correlation

To further test the reliability of the evidence utilization index, I also calculated a Cronbach's alpha score to determine how strongly correlated the variables making up the index are to each other. On the 16-question index, the Cronbach alpha score was .823. Considering that many methodologists recommend a minimum  $\alpha$  coefficient of at least 0.65 up to 0.8, the utilization index created for the purposes of this research passes muster in this area<sup>33</sup>.

### 3.6.2 Alternative constructs of the evidence utilization index

One unfortunate consequence of the 16-question index was that several respondents with missing values were eliminated from the computation of the EIS. One alternate construct of the utilization index would be using only the first six questions of the 16 that inform a legislator's EIS score. Making this change would significantly

---

<sup>33</sup> I also calculated Cronbach's alpha for a smaller version of the index (see "alternative constructs of the EIS"), one that only utilized the first six questions toward which the 16-question scale was created. In that circumstance,  $\alpha = 0.763$ , which is above the 0.65 floor but shy of the 0.8 target.

increase the number of respondents with an EIS score (the N would increase from 41 to 71). However, because of missing demographic data, the number of “usable” comparisons would only represent an increase from an N of 41 to 52. Considering the robust nature of the 16-question EIS (confirmed by a larger Chronbach’s alpha score) and the minimal overall increase in usable respondents, I proceeded with analysis using the 16-question evidence utilization index / score.

Another possible approach to such an index would be conceptualizing the evidence utilization score as average score on the five-point scales from all available data for respondents who answer a minimum number of questions. This approach was explored with both the 6-question subset as well as the full set of 16 questions comprising the evidence utilization index. This approach showed essentially no variation because most of the scores centered around the average value of “3” on the Likert scale. Such an outcome reinforces the decision to design the “point” system for answers of “4” and “5” on the relevant questions, as these demonstrate a truer commitment to evidence utilization in the policymaking process. An average score of “3,” on the other hand, allows respondents to provide a safe answer that neither chooses the side of evidence nor the myriad other policy considerations. Presser and Schuman (1978), for example, found using a split-ballot experiment a highly significant increase in the “middle category” when offered explicitly as an option. As it relates to this research, they found an even larger middle category increase on liberal/conservative self-identification, with an increase of as much as 22 to 39 percent (Presser and Schuman, 1978). In light of this literature and my experiences, both as a state legislator and throughout this project, I am

confident that the original evidence utilization score more accurately reflects a legislator's commitment to evidence.

### 3.7 Crosstabs of interest between evidence index and select independent variables

Ideally, this research would have had the ability to incorporate multivariate analysis between the evidence index (the dependent variable) and the myriad independent variables that could influence the extent to which lawmakers are utilizing evidence in their policymaking work. However, a lower than expected N on the survey responses eliminated the possibility of completing this analysis. In the future, similar work that can obtain sufficient response rates should consider deploying these more sophisticated statistical techniques to the comparisons between key demographics and evidence utilization remains an important goal.

Even without multivariate analysis, there remained opportunities to dive deeper into the data and glean insights and help inform the context of interview conversations that take place with legislators regarding their experiences with evidence in their work. To that end, this work executed a comparison of means between the evidence utilization index and many of the key independent variables available. The results of that analysis can be found in Table 27 below.

*Table 27: Comparison of means between dependent and independent variables*

	Mean	Std. Deviation	p value
<b>Gender</b>			.427
Male	8.74	3.91	
Female	9.9	3.87	
<b>Party</b>			.185
Democratic	9.22	3.37	
Republican	9.31	4.48	
<b>Ideology</b>			.613
Liberal	9.92	3.45	

Moderate	8.4	4.01	
Conservative	9	4.37	
Full / Part Time Legislator			.263
Full Time Legislator	8.29	3.27	
Other Full Time Profession	9.77	4.13	
Generalist vs. Specialist			.337
Legislative Generalist	8.81	4	
Legislative Specialist	10.2	3.36	
Legislative Tenure			.485
1 – 9 Years	9.52	4.18	
10 or More Years	8.57	3.46	
Education			.837
Less than Bachelor degree	8.4	4.73	
Bachelor degree	9.79	3.62	
Master degree	8.46	3.72	
J.D., Ph.D., or equivalent	9	4.66	
Results First State			.030**
Not a Results First State	7.67	3.93	
Results First State	10.35	3.69	
Pew Trichotomy of States			.998
Leading	8.91	3.62	
Mixed	9.00	4.24	
Trailing	9.00	4.31	

\*\* Significant at the .05 level.

To put the differences in the above chart in context, a gap of “one” between means is the equivalent to the average person answering one more question in such a way to suggest they rely upon evidence in the execution of their policy work. In other words, they would be one more step towards fully relying on evidence as it is measured on the evidence utilization scale I created as a part of this research. Of all the categories tested, each tended to include some sort of variation. The only *statistically significant* result was for legislators who reside in a Results First state. This result helps to further validate interest in examining the relationship between the program and evidence utilization within a state.

Among the independent variables which had the least amount of variation among the survey demographics were education, tenure and political party. In regards to education, while there was a significant jump from less than a bachelor degree to those with a bachelor degree, that difference dropped back down as educational attainment reached the Master level, and slightly up again at the terminal degree equivalent level. Across all levels of educational attainment, however, there was relative stability (all the means were within approximately one point of each other), suggesting education was relatively unimportant as it relates to the dependent variable. To the extent a point could be made about educational attainment and evidence utilization, one could point out that the lowest mean score for all the categories, certainly suggesting there might be some relationship between education and evidence use.

Legislative tenure's mean scores were within one point of each other, with newer legislators (those serving less than 10 years) having a higher score of 9.52 and more tenured colleagues (those serving 10 or more years) earning an average evidence utilization score of 8.57.

Most clear in its lack of variation in means was political party. Here, I found there was essentially *no difference at all* between Democrats and Republicans, with average evidence utilization scores of 9.22 and 9.31, respectively. On the other hand, ideology (a similar but different measure than party) did show considerable variation in mean scores. These differences were equally focused, however, on *both ends* of the ideological spectrum when compared to the political center of the survey sample. While self-described moderates posted an average evidence utilization index score of 8.4, liberals in the sample scored at 9.92 while conservatives posted a mean score of 9. One

interpretation of this would be that conservatives are slightly more likely to rely on evidence than political moderates, although liberal legislators are even more likely than both their conservative and moderate colleagues to depend on evidence in their legislative efforts.

There was also clear variation between the genders, with women on average scoring more than a full point higher than men, 9.9 compared to 8.74. There was a similar spread between full-time and part-time legislators, with part-time legislator's average evidence utilization score of 9.77 being more than a full point higher than of their full-time counterparts at 8.29.

The range between average scores for self-described legislative generalists and specialists were among the larger gaps. Here, the average mean score on the evidence utilization scale was almost one and a half points higher for specialists (at 10.2) than the score for legislative generalists (who averaged a score of 8.81).

Finally, a key independent variable of interest for the purposes of this project was whether a state was a Results First state at the time of the research. Interestingly, this comparison of means testing the variation between Results First and non-Results First states produced the greatest difference in all the categories – and the only statistically significant difference. On average, non-Results First states were almost 3 full points lower than Results First states, with means of 7.67 and 10.35. Moreover, the average utilization score for Results First state participants was the highest of **all** the variables measured. On the other hand, the “ranking” that a state had from Pew (leading, mixed, trailing) on the use of evidence in policymaking does not appear to have a direct connection to evidence utilization scores.



### **3.8 Conclusion: from surveys to interviews**

Of course, the areas of interest regarding evidence utilization are the places where the scores are significantly higher and/or lower than the mean and median score of approximately 9. Considering the above review, these would include higher scorers such as women (9.9), having a liberal political orientation (9.92), part-time legislators (9.77), legislative specialists (10.2), and the impact of being in a Results First state (10.35). I did not find these results especially surprising.

I would first note that these scores, while technically above the median score on the evidence utilization scale, were not substantially higher. Still, these results seem consistent with literature and my experience. Most striking is the connection between legislators in Results First states and the highest average utilization score. While it is still close to the average score of 9, the 10.35 average score for Results First legislators does begin to validate the hypothesis that the program's existence may marginally contribute to evidence use by legislators<sup>34</sup>. If there is a growing literature that Republicans and conservatives are turning away from science, I expected and identified modest, positive associations between a liberal worldview and self-reported evidence utilization. The fact that variation centered more on ideology, and not on party affiliation, reminds me not to necessarily associate the two categories: while a Republican is more likely to be a conservative than a Democrat, the two are not always one in the same. Also of little surprise was the higher scores for legislative specialists, who tend to get into much greater detail during their policy work.

---

<sup>34</sup> Correlation versus causation in this regard is not formally tackled in this project, however there are arguments made elsewhere that Results First tends to establish itself in those environments most hospitable toward evidence utilization – making the distinction even more difficult to discern.

If there was a surprise among the higher evidence utilization scores, it is the number produced by part-time legislators. Having less time to consider the topics facing them than their full-time counterparts, it was striking that they nonetheless had the higher average evidence utilization score. Some possible explanations of this outcome: part-time legislators rely on experience from other professions (teaching, science, etc.) that might be more grounded in evidence; they may rely more heavily on staff or other resources for information and evidence because they have less time; or statistical variation of the sample.

On the other side, low scores worth examining in more detail include being a full-time legislator (8.29) and the impact of not being in a Results First state (7.67). As discussed above, the connection of not being in a Results First state to the lowest average evidence utilization score further affirms the possibility that the hypothesis regarding the role of Results First holds true. As for full-time legislators, they may score lower than their part-time counterparts as a consequence of stronger desire to seek re-election (it is, after all, their only job), a desire the literature identifies to be a strong motivator among elected officials in their behavior.

Another hypothesis put forward in chapter one, based on literature such as the advocacy coalition framework, suggested that lawmakers would be more inclined in certain circumstances to deploy evidence in their policymaking than in others. Perhaps the strongest example affirming these differences of circumstances can be found in tables 11 and 12, where there are substantial and several statistically significant variations between the use of evidence-based factors depending upon the subject matter at hand. There were also differences noted among non-evidence based factors between these

subjects, though those differences were much smaller in size and scale than those associated with evidence.

One way that would help in better understanding the dynamics discussed in this chapter is through in-depth interviewing of state legislators. Doing so also bolsters the strength of this research, as the case for combining research methodologies, particularly when both quantitative and qualitative methods are present, is a strong one (Gable 1994). In the next chapter, individual legislators take time to more fully discuss their experiences with evidence utilization, information gathering and understanding of the Results First program. Through this process, I am able to more clearly make statements regarding the hypothesis of this research.

I begin with this more in-depth understanding of low and high utilizers alike in chapter four while also paying attention to other lessons learned by combining the insights of the survey research with interview data from responsive legislators and staff members alike.

## Chapter 4: Evidence utilization among legislators: in their own words

When studying the beliefs and behaviors of legislators as it relates to evidence utilization, it is instructive to understand the many constraints they face. Chapter three presented, through self-reporting, useful context about the background in which legislators operate. These included reminders about the importance of personal and constituent opinion in policymaking decisions, the high use of staff and colleagues as sources of information and low knowledge and use of the Results First program. It was also shown that there is a wide distribution of legislators on an evidence utilization index and that there are many factors negotiating the degree to which evidence can be used to inform policy work.

This chapter first shares the results of interview data from legislators with the highest and lowest scores on the evidence utilization index. Those results are followed with the common themes that developed from interviews with the lawmakers and staff members who participated in the qualitative data collection process. Next, specific feedback about the Results First program is discussed, as are concerns about the program, incorporating feedback from the program's Director. Finally, the chapter concludes with ideas from legislators about ways in which evidence can become more relevant in their policy work. Much like chapter three, the data continued to shed light on the many challenges facing legislators and making it more difficult for evidence to be prominent in the policymaking process.

## 4.1 Interview methodology

Following the survey data collection and analysis of chapter three, research in this chapter moved toward semi-structured interviews, conducted across the six states of interest and in Maryland, the pre-test state. The survey data was leveraged to identify legislators that are the highest and lowest-level self-reported users of evidence. I specifically chose to dedicate time exploring the highest and lowest self-reported users of evidence (per the evidence utilization scale developed in chapter three) to compare how lawmakers on the extreme ends respond similarly or differently to the qualitative inquiry on the use of evidence. The N in this circumstance is very small (one high and two low score legislators were interviewed), but the extreme nature of the evidence utilization score suggests that these are the legislators to be *the most or least* inclined to incorporate evidence into their decision-making process.

These interview insights help fill any gaps from chapter three, which would have (with additional data) included analysis beyond the descriptive statistics and comparison of means provided – more directly measuring the correlations between individual and institutional characteristics (including the existence of Results First) and self-reported evidence utilization. The assumption in tackling these extremes is that the survey instrument accurately captured these very different perspectives, allowing the interviews to capture an even more granular level of detail<sup>35</sup>.

Extreme cases are only part of the story, however, as most survey respondents ended up somewhere in the middle of the distribution on the evidence utilization scale.

---

<sup>35</sup> While there was room for improvement with survey instrument, I believe it effectively captured the big picture on self-reported views of evidence utilization by policymakers. Given the inherent difficulty of using a structured survey to capture such a complex subject, there is likely to be some level of error in the survey.

To conduct a deeper investigation, interviews went beyond the highest and lowest users of evidence. This broader range of interviews was from across the utilization spectrum, with case study interview recruitment efforts that focused on reflecting the diversity of the participants across the population of interest. To the extent practical, subjects for legislative interviews were selected in such a way as to have participants be reflective of a state's partisan balance, gender balance, geography balance, and racial balance. These interviews help to better understand a state's experience with evidence-producing organizations, and how lawmakers engage with them, if at all.

Within each state, the research originally aimed to conduct approximately six interviews with elected state leaders from a variety of backgrounds. The intention was that the broadly cast survey mechanism would provide an adequate pool from which to draw subjects. This produced an initial goal of 36 interviews. However, as discussed in earlier, overall response rates were slightly lower than anticipated – leaving a smaller than expected interview pool from which to draw. Despite this limitation, 18 legislators that responded to the survey were still engaged in a one-on-one, semi-structured interview process.

To further supplement the knowledge gained from interviews, insights from the four Maryland legislator interviews conducted prior to entering the field with the survey were added to the data. I found these conversations to be especially informative of the strong personal relationships I have with the interview subjects. These interviews were also conducted in person, further contributing to an environment of openness and trust.

Information from six legislative staffers and one representative from Results First<sup>36</sup> was also included. Staff insights were likewise useful, offering an “outside view” of the self-reported information provided by state legislators that was still close enough to the work being conducted that it could effectively validate legislator input.

In all instances, the staff I sought out to interview were those listed with some connection to either the Results First program (such as the New York staffer and Director of Results First), connection to a state’s program producing output similar to Results First without the cooperation of Pew (such as the Virginia staffer), and/or who worked with Pew or the legislative committee that addresses criminal justice (such as the Maryland staffer) – the focus issue area around which research was conducted. In all instances, I utilized the name, email and phone number of appropriate staff members as listed on government websites to request interviews. In the three states where I did not obtain a response (Massachusetts, Rhode Island and Pennsylvania), I attempted a similar outreach for a second possible staff but was again unsuccessful. Combined, there were a total of 29 legislators and staffers included in the interview process<sup>37</sup> for this chapter – nearly meeting the original base goal of 36 interviews while providing a more rounded set of perspectives in the process. A full review of the questions presented, as well as the state demographics of these interview subjects, is included in Appendix I and J of this chapter.

In each interview, subjects were asked if they would be willing to participate in an interview, providing their informed consent in the process. The semi-structured questions

---

<sup>36</sup> Results First was not originally intended to be included in this study. I ultimately decided that it would be preferable to allow them to answer directly questions about their program than merely relying upon other sources.

<sup>37</sup> Two of the staff interviews were Maryland interview data from a paper I completed in 2012, described in the works cited. A third interview was also conducted with another Maryland staffer in 2017 consistent with the conversations held with other staff members.

inquired about if, when, and how the relevant policymakers utilized evidence while conducting their policy decision-making. Where appropriate, questions were also formulated to ascertain the history and value of the Results First program. Questions were also asked about what values policy makers believe their colleagues place upon the program and evidence. The high-level staffers included in this data collection were utilized as a “check” on the data obtained from state elected officials. On average, interviews lasted about 30 minutes each.

Travel and scheduling considerations made in-person interviews impractical to conduct for all but the Maryland legislative and staff conversations. Thus, phone interviews were used as an alternative. Case study design considerations were modeled on Yin’s (2009) approach to case study design and methods and incorporated elite interviewing strategies such as those suggested by Berry (2002) and Beamer (2002). Nearly every call, 25 out of 29, as well as every in-person interview was recorded with the permission of the interview subjects. I placed the recordings on my Apple S4 audio application. Detailed notes were taken during the interview process for all conversations, whether recorded or not. In addition to notes, recorded interviews were transcribed to guarantee the accuracy of notes and to assist with coding efforts.

Legislative respondents were told that their responses would remain confidential (as they were on the survey response) and that nothing would be shared to identify them personally. Staff members reached outside of Maryland were willing to share their feedback “on the record” and as such are quoted with attribution in this work. Staff from Maryland requested that I keep their identity confidential. Following the completion of interviews, every subject that needed to be made confidential was put into an Excel file



and randomized in order. From there, they were assigned a generic description, such as “Rhode Island House member 1.” This allows quotes to be properly attributed but also allowed for a way to connect quotes to respondents. Interviews started with the Maryland pre-survey discussions, taking place between October and December of 2015. The remainder of the new legislative and staff interviews took place between January of 2016 and May of 2017.

These case study interviews afforded the opportunity to delve deeper into the idiosyncrasies of a state relative to its participation with the Results First program. This deeper exploration of the “why” helped provide more meaningful insights into the “what” identified in the online survey instrument. To examine these idiosyncrasies, I turn first to the highest and lowest users of evidence in the survey sample.

## **4.2 High evidence utilizer: in his own words**

Legislators with three of the highest evidence utilization scores were sought after specifically because of their score. Two Democratic House members – one from Massachusetts and another from Rhode Island – had the highest scores on the evidence utilization index, with each member scoring a “16” on the scale. There was also a Republican legislator from Pennsylvania with one of the highest scores (15). Of the three legislators, only the legislator from Massachusetts with the evidence utilization score of 16 agreed to be interviewed. Insights from that conversation are shared next.

### **4.2.1 Massachusetts House member, score of 16**

This representative consistently produced survey results that suggested that they were inclined to use evidence – and indeed, it appeared from their interview that they

followed through on this preference *as much as practically possible*. These limitations follow closely with the constraints faced by all legislators. The limitations discussed were political as well as practical, and he particularly honed in on paying attention to one's district and votes on conscience. As he described it:

When it comes to voting, there are three buckets. First one is 'how does this impact my district?' I am truly here to represent you and I need to be aware of this especially on higher profile issues. The second is no impact on my district and I may have to vote how I'm asked by leadership or will engage of a quid pro quo of sorts, where I'm with you on this one but expect you to be with me on the next one. Finally, there are conscience votes, like abortion, same sex marriage and the death penalty. Here I vote what I think is right (MA House member 2, 2017).

It was in this second bucket that the Massachusetts legislator indicated that his preference for good data mattered most, consistent with the advocacy coalition framework discussed in earlier chapters. He includes evidence in this category, but recognizes that it is not a standalone issue.

I love data and research. Give me the data and give me the research (MA House member 2, 2017).

I am a firm believer that I can never have too much information. It's just that most of my decisions can't only be based on that, no matter how much I believe in it (MA House member 2, 2017).

Much of this lawmaker's appreciation for information and evidence was attributed to the fact that his professional background included training in the sophisticated use of evidence. As he explained during his interview, *research and evidence is an essential part of his background*. Even though he has an appreciation for the ways in which data and evidence can inform the policymaking process, he suggested that the very best arguments will deploy all the possible approaches available to an advocate.

If someone can hit on all of the components, I am inclined to be really supportive (MA House member 2, 2017).

An example that this legislator provided was the debate over “Romney-care,” where Massachusetts was the first state to move in the direction of what later became known as the Affordable Care Act or “Obamacare.”

As the Congress considers repealing the Affordable Care Act, the best arguments against it here are being able to talk about the number of people who will lose healthcare because of the repeal of the ACA, the budget impacts for the state, and then also being able to tell a personal story. Like my sister or wife has cancer and would not have had coverage or would lose it as a pre-existing condition. That’s the ideal scenario (MA House member 2, 2017).

Not only did this legislator admit that multiple methods of advocacy worked best when people are trying to convince him, but he made another observation worth repeating: the best advocates know their audience extremely well, and will focus on the methods by which legislators are most likely to be swayed. Some legislators are not swayed by data and evidence, so why bother worrying about selling them on the merits of it? In the end, policy advocates of all stripes have a goal of enacting (or defeating) a proposed change – it makes sense to focus on the ways most likely to realize that preferred outcome.

Some get into spreadsheets. Others like studies and some get into personal stories. I prefer them all but different people have different preferences (MA House member 2, 2017).

There’s a joke we have in the legislature. When we are all sworn in, we are overwhelmed by the building and the gravity of our position and we ask ourselves ‘how did I get here?’ After a little bit of time, we will find ourselves asking the question, ‘Wow, how the hell did *they* [some colleagues] get here?’ (italics added, MA House member 2, 2017).

But I guess they reflect their districts. In some places, being an auto body shop owner is a reflection of the people they represent. Some people are more data wonks like I am and some people just aren’t (MA House member 2, 2017).

People will make their cast in the most effective methods possible. That is different for every legislator. One size does not fit all (MA House member 2, 2017).

This legislator stressed the importance of “credibility” from those sharing information – whether empirical evidence or personal testimony. He seeks to incorporate “better” forms of data whenever possible, but also takes a very realistic approach to governing and recognizes that there are many conflicting points of interest involved in the process.

In this conversation, I found a lawmaker committed, in principle, to the widespread deployment and use of evidence in their policy-making endeavors. However, in peeling back into the details, I also discovered that even this “high flyer” recognized the limitations he faces. As outlined in earlier chapters, these constraints include constituents, party and legislative leadership structure and practical considerations such as limited time. While each of these constraints are discussed in greater detail along with the dozens of other interviews with state legislators and legislative staff, it is especially noteworthy to point out the ways in which this lawmaker, who is particularly committed to the use of evidence, also struggled with the obstacles to effective evidence-based policymaking. If those most committed to advancing evidence in policymaking admit to these struggles in their own work, how much more difficult of a climb is it for those less interested in, aware of, or inclined to personally support such an approach? With such a realistic approach being taken by some of the highest self-reported utilizers of evidence, I turn next to those who were among the lowest users for maximum contrast.

#### **4.3 Low evidence utilizers: in their own words**

Like the effort with legislators with a high evidence utilization score, there were also successful efforts to engage with the two legislators scoring lowest on the evidence

utilization score. These included a West Virginia House member<sup>38</sup> with an evidence utilization score of one as well as a member of the New York House (with a score of two) that agreed to participate in a follow-up interview.

#### 4.3.1 West Virginia House member, score of 1

While there was limited time with the West Virginia legislator who had a low evidence utilization score, the rationale behind his rating did not take long to identify.

I don't know what you're hearing from others, but I can tell you that I'm here to represent my constituents (WV House member 1, 2017).

What do I think about when I am deciding on how to vote on laws? Simple: 'What do the folks back home want me to do (WV House member 1, 2017)?

This legislator had a keen focus on the needs, desires and interests of those whom he represented. Above all, it was about his district and a very distinct belief in the responsible to be a delegate that conveys the will of those who elected him to the legislature. In fact, even when the conversation could turn to the role of evidence in the legislative decision-making process, this legislator scoffed at the idea – saying that it might be for other lawmakers, but that it was *not* for him.

Do I read other stuff, sure? But the most important feedback is the number of calls or letters I have coming into my office (WV House member 1, 2017).

I know other people will be impressed by fancy studies and big-wig lobbyists. That's fine for them, but it's not really for me (WV House member 1, 2017).

You know what's right. You know what folks back home want you to do. Trust your own compass and follow it (WV House member 1, 2017).

In a way, it was reassuring that this legislator was so focused on the views of his constituents as another affirmation of the evidence utilization scale. The conversation

---

<sup>38</sup> This member was only able to conduct an unscheduled, abbreviated interview. While not as thorough as other interviews, there was great value including his feedback.

with the other lowest scoring member further affirmed the connection between the scale and legislative practices of a member as it relates to evidence.

#### 4.3.2 New York House member, score of 2

Like the member of the West Virginia House, this legislator also – though not as directly – placed a premium on the views of his district. Almost the polar opposite of the high-scoring legislator from Massachusetts, who seemed to want to integrate data into decision-making whenever practical, this lawmaker seemed convinced that nearly every study available has a bias and/or special interest behind it – limiting the value of including it in the debate. Instead, this New York lawmaker placed a high priority on the views of his constituents and, in some circumstances, his personal belief system. When it came to voting with his district, this representative was clear.

One of the most important tools a legislator has is to listen. I spend a lot of time gauging the pulse of my district. They are first and foremost in all of the decisions that I make. My district comes first (NY House member 2, 2017).

In other circumstances, especially those which he considered matters of conscience, this legislator with a political science background indicated a clear willingness to go against the popular will of the legislative district. Specific examples cited were marriage equality and New York's version of the DREAM Act.

Marriage equality is one issue that comes to mind where I had strong personal opinions. I am not sure that public opinion was quite there yet. I just can't be a weathervane in issue areas that involve individual and civil rights (NY House member 2, 2017).

Like the Massachusetts representative, healthcare legislation was also discussed as a real-life example. In New York, legislation on an all-payer system is currently being considered. Rather than sharing ways the evidence on the proposal might influence his

vote, this lawmaker instead spoke about the ways in which he found it difficult to believe *any* of the information presented to him on the issue.

The sponsor's information looks great. Until I am visited by the Health Insurance Underwriters and then it's a disaster. I want to know who is paying for these studies. It seems like everything is paid for by a stakeholder, so you need to just use your own theories and minimize the outside influence (NY House member 2, 2017).

The healthcare example spoke to a larger mistrust that this representative had with research. He indicated that, because it is so hard to find "truly good data," he gives it limited credibility and does not use it much to inform decision-making. He reported being more comfortable taking in feedback from stakeholders, providing examples that suggest he and fellow lawmakers were most likely to follow lines of arguments aligned most closely with groups most important to them, whether they are the farm bureau, a business group or the teacher's union.

The thing legislators look at more than the data, including me, is who is in favor or opposed to this bill (NY House member 2, 2017).

Trying to find truly objective information these days on a particular issue can be difficult to locate. It seems like every study is paid for by somebody with a point of view ... I've become a bit jaded (NY House member 2, 2017).

Finding information is easy. It's just hard to know what is credible. Everything seems to have some sort of outside influence (NY House member 2, 2017).

It's not very hard to find information on a topic, what is hard is doing the critical thinking so as to separate the wheat from the chaff. To find the information that actually makes sense and is applicable and not biased (NY House member 2, 2017).

If anything, I will compare what we are trying to do with other states, to make sure it works (NY House member 2, 2017).

Finally, as discussed in a separate section concerning overall insights, this legislator discussed the importance of trust and relationship – especially emphasizing

using people with whom they have a relationship in the legislative district as a source from which to draw their information.

You develop relationships and find out who you can you trust (NY House member 2, 2017).

I am more likely to call an expert in the district over anyone else. I have cultivated a good base of experts in the district – such as a group of superintendents for education issues and a group of business owners for issues related to homeowner’s insurance (NY House member 2, 2017).

For both low-scoring legislators, constituents were key. In the case of the West Virginia delegate, he appeared outwardly disinterested or dismissive of evidence and analysis. In the case of the New York legislator, there was an acknowledgement of the available information, though he spent time discussing clashing perspectives within evidence and inevitable biases in the evidence production process. The New York legislator also expressed a willingness go against his constituents, but only in the area personal opinion – it was not because of what evidence suggested is sound policy. Instead, he suggested that political influences were a better cue to follow than evidence. The only real mention the lawmaker made to evidence was the ways in which he felt it was confusing, hard to separate out what is real from fake and fueled by special interests.

Having spent time getting to know the “best and worst case scenarios” of the respondents to the survey as it relates to evidence utilization, I turn next to a hard look at the more overarching themes drawn from the dozens of interviews completed.

#### **4.4 Unifying themes**

In addition to exploring the feedback provided by some of the highest and lowest scores on the evidence utilization index, there were numerous other interviews with legislators across the spectrum on the evidence index. Feedback provided from staffers in



several states were also included in the study to provide new voices to the conversation and an important new perspective from within the legislature. The insights of these interviews tracked closely with much of the survey data as well as the literature, discussed in earlier chapters. For these sections, relevant feedback from the 29 interviews conducted are clustered by issue area, with further breakdowns and commentary within each theme. For example, in the first section on the importance of constituent viewpoints, the analysis is broken down between respondents who are solidly committed to voting what they believe to be aligned with their district's politics and those who allow for some gradation of that belief system, however minor a variation.

#### 4.4.1 Constituents matter most

Unsurprisingly, legislators ranked the views of their constituents above all else: This tracks with the data identified in the quantitative research of chapter three, where it was reported that 71.8 percent of respondents heavily weighted the views of their constituents on issues of bail while 73.4 percent ranked it as critical on gun control issues. Interviews confirmed that legislators held the views of those they represented above all other considerations. In both issue instances, constituent opinion was reported as the single most important factor influencing decision-making, consistent with Hird's (2005) findings.

Legislators and staff consistently reinforced the importance of appealing to the interests and desires of the constituents served by legislators, as evidenced by much of the feedback from the interviews below. Several legislators explicitly shared their intentions to make the politics of their district the primary focus in their legislative decision-making:

clearly I think about who I've heard from in my district – that I value above all else ... especially looking at people who have supported me in the community or during a campaign (MD House member 2, 2015).

It is often not good enough just that you vote pro-life, it's that you're perceived as being pro-life. That means voting your district and going above and beyond so they know you are supporting their issues (PA House member 3, 2016).

I need to lay down a marker to cover my bases in the event I vote for and am seen as accepting a compromise. I don't want anyone in the progressive community to doubt my goals (VA House member 1, 2016).

I am very keen about the desires of my constituents. They are the bosses and you kind of have to cave into their desires because that's what they want (MA House member 3, 2016).

Your first duty is to be a trustee and a servant of the people you serve. Listen to your constituents and not your party leaders (VA House member 2, 2016).<sup>39</sup>

I don't really take into consideration what's 'good policy.' That is not as important to me on big-ticket items, making sure you are in line with where your constituents want you to be (MD House member 1, 2015).

You know who you represent. You know what's important to them and you know what's going to get you in trouble if you disagree with them or choose to take a personal point of view in a public policy issue (MD House member 1, 2015).

First and foremost it's about the district I represent (MA House member 1, 2016).

Obviously concerned about how the district will feel. That's my primary decision-maker (PA House member 1, 2016).

The opinion of my constituency is very important to me ... not just what they say but how they say it to me (RI Senate member 1, 2017).

Any legislator who has not told you that their constituents and the feedback they get from them is top of mind is either lying to you or not concerned about re-election (MD staff member 3, 2017).

This certitude about following the politics of one's district was strong and resounding consensus among members. A few others leave open the possibility to, on occasion, stray from those whom they represent when casting votes:

---

<sup>39</sup> The use of "trustee" here is actually used as an equivalent to the earlier description and use of the term "delegate."

I am 95-percent attuned to my district. When it seems like there might be a disagreement, it usually is because they do not understand the issue (MA Senate member 1, 2017).

It's not just the evidence but it's also about the politics of my district. If there's leeway back home I can more fully consider pushing the evidence (MD House member 3, 2015).

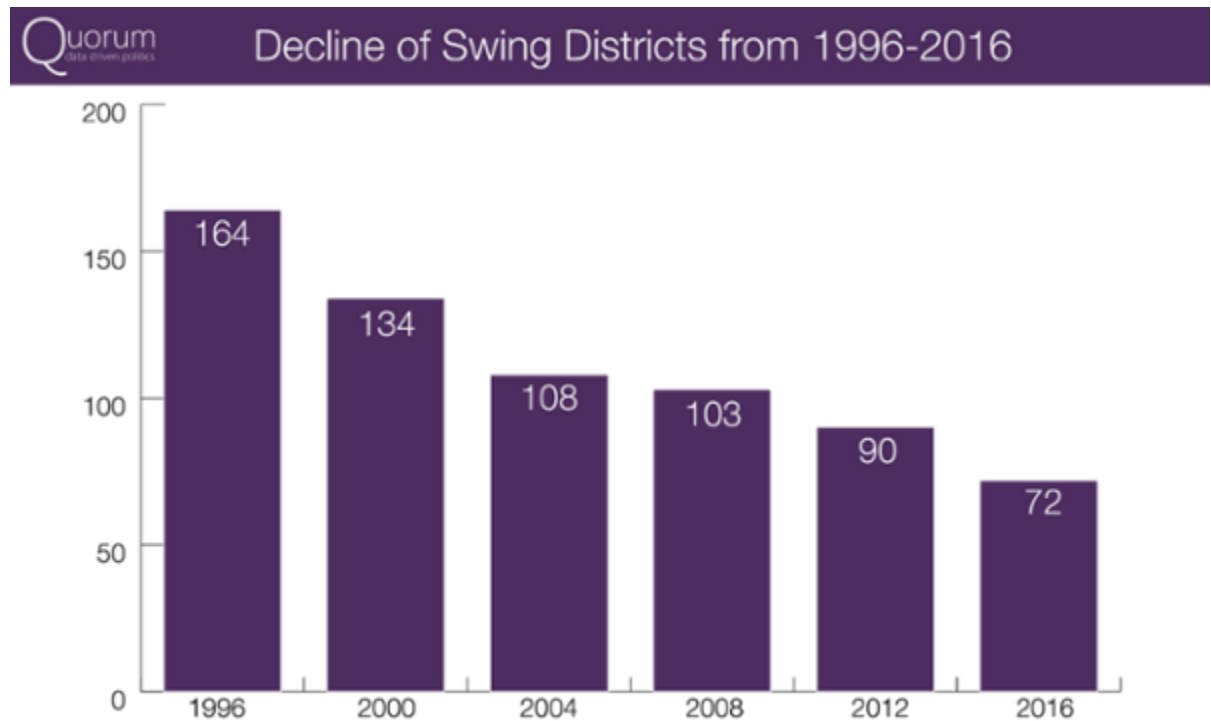
I absolutely talk to my constituents about how to vote, they are educated and have good ideas. Sometimes I will vote differently than what they want (VA House member 1, 2016).

Legislators were sensitive to the need to stay closely aligned with the ideology of their district, whether to execute on their trustee viewpoint of governing and/or for the more practical desire to remain in office. In the political science of legislatures, there is a big issue of how legislators view their district. Is it majority opinion, or a particularly intense minority that overpowers a quiescent majority in its election impact? Fenno (1977) offered one conceptual model, that of concentric constituencies – where one's view begins with his or her closest supporters and moves out, like concentric circles, to party activists, the primary voting electorate and finally the general electorate. Based on the feedback in this chapter, Fenno's model appears to fit well – legislators were most focused on the opinions and feedback of those in the smallest of their concentric circles.

Compounding this imperative for legislators are the extent to which election districts are also becoming increasingly partisan. Not only are elected officials trying to keep their residents happy with their voting record, they are attempting to do so in districts that are increasingly drawn to match up with an extreme ideology. "Swing" districts, whether congressional or state legislative, are becoming increasingly scarce across the country – making it more likely that the personal beliefs of a representative are aligned with those whom they represent. For example, the 2017 Cook Partisan Voter Index (PVI) shows that the number of congressional "swing" districts has reached a 20-

year low, declining from 164 seats in 1996 down to fewer than 70 this year (Quorum Analytics 2017). Note the decline of such districts in the figure below.

*Figure 8: Decline of swing districts over the last 30 years*



Source: Quorum Analytics (2017).

Shor has produced similar findings at the state legislative level through the use of spatial mapping of rollcall votes beginning in the mid-1990s. As Shor and McCarty (2011) write, “the states appear to follow the national pattern of high and growing polarization” (p. 549-550). The lack of swing and moderate districts tend to push lawmakers – and the electorates they represent – to one ideological extreme or another. These extreme worldviews, in theory, are less inclined to pursue policies entirely rooted in evidence. “Base” voters tend to be partisan – more interested in ideological purity and less interested in evidence-based policymaking than moderate and swing voters. Unless ideological purity includes an emphasis on evidence-driven policymaking, the ever-

increasing number of partisan districts presents an additional obstacle to enacting laws based on evidence.

#### 4.4.2 Personal beliefs not far behind constituent feedback

In chapter three, survey findings were that legislators often relied upon their personal beliefs to help drive some of their decision-making process. In fact, it was the second highest rated influence on decision-making, with 57.7 percent of legislators weighing personal opinion heavily on bail legislation and 63.8 percent doing so on legislation regarding gun control. While there were fewer who explicitly touched on this topic during the interview process, it was nonetheless an important topic and legislators clearly use personal beliefs to inform their voting:

As a person of faith and an evangelical Christian, I try my best to make informed decisions on the evidence but also listen to my heart on the basic right and wrongs of an issue (WV House member 2, 2016).

I will absolutely allow my personal experiences to guide my decision-making, especially on the issues I know more about (MD House member 2, 2015).

As time has gone on, I've become more reliant on my gut instinct (VA House member 1, 2016).

There are plenty of legislators on both sides of the aisle who will lead with their personal beliefs – be it abortion, climate change, guns ... you name it, and I can point out the Ds and Rs who will have their minds made up before the debate even begins.” (MD staff member 3, 2017).

One lesson particularly noteworthy was the extent to which legislators went beyond personal belief systems under the personal opinion umbrella. In fact, the feeling of certitude on legislation that came with personal opinion could be more accurately described as personal experience. During interviews, lawmakers would often cite experience from a background or profession (such as law) as a credential that made them better qualified to understand the topic than “other professionals.” Not only did personal

experience made legislators feel more comfortable in their own decision-making, but it gave them confidence when relying upon legislative colleagues for voting guidance. In other words, an insurance broker would feel like an expert in all things related to insurance and would be largely comfortable relying on an educator (of the same party) to inform their voting decisions about education policy. In both circumstances, there was a sense that less searching need be done to identify relevant evidence because there is a lot of “already knowing” from lived experience.

I have my own biases so I can’t say that I treat all information equally – I need to make decisions on my own gut reaction to things, keeping in mind that sometimes you need to evaluate what the constituents say and what the experts say and be able to explain your vote (PA House member 2, 2016).

I mean from personal experiences ... using my personal experience especially on issues that I know more about I certainly rely on that a lot more (MD House member 2, 2015).

My area of expertise matters a lot ... There are not many lawyers, so many non-lawyers are forced to vote on things which they are not very well aware of (PA House member 3, 2016).

I trust people with real experience, such as parents more than I do consultants on education issues (MA Senate member 1, 2017).

There are some areas where I have a great deal of expertise. There are others that I have no clue. In those areas [where I have no clue], I am more likely to rely on staff (NY House member 2, 2017).

Remember I’m a lawyer and lawyers know how to figure out everything (PA Senate member 1, 2016).<sup>40</sup>

Given this feedback, the irony of having close personal experiences with a topic in the legislature might mean that lawmakers close to an issue are actually *less likely* to seek new or cutting-edge evidence in their deliberations. Instead, they rely on their own

---

<sup>40</sup> Literature also exists on the difficulty that the courts have had dealing with the findings of social science (see Horowitz 1977).

views and perhaps the information or evidence that is created by and/or shared within their professional and personal circles.

#### 4.4.3 Institutional leadership may matter more than elected officials admit

It is much easier for an elected official to claim to or to stand behind the positions of those whom they represent than to go against these constituents for any reason, including evidence. In fact, accountability to a lawmaker's constituency is a bedrock of our democracy, and as discussed in earlier chapters, a major tension in trying to include evidence into the decision-making process. In a similar way, many people can understand that legislators might want to be viewed as standing on their personal principles while serving in a public capacity.

However, having one's constituents know that a legislator is following institutional leadership is an idea that is not as easy to defend. This helps to explain the low scores in the survey research. For example, the opinion of the Presiding Officer was listed as being a major influencing factor on bail reform and gun control by a mere 10 and 6.9 percent of survey respondents, respectively, while the same percentages for Committee Chairmen were 14.5 and 10.4 percent. No legislator wants to be perceived as following legislative leadership, particularly over their constituents, conscience or even evidence. However, based on the qualitative follow up discussions, the self-reported numbers in the online instrumentation did not align with the day-to-day realities faced by many legislators.

Oh yeah, absolutely lots of votes you know are going to go party line. It's easier to just go along with the flow than to go against it (MA House member 3, 2016).

Obviously I try to work with the leadership. If the issue is benign or does not really impact my district, I try to go with leadership (MA House member 1, 2016).

If it [legislation] is not in my committee, I generally trust the Chairman of the committee and follow their lead (PA House member 3, 2016).

Even party-line votes seem to allow some room to maneuver for legislators, assuming that they are not in legislative leadership and the legislation being considered is not a true legislative priority for leadership. As a few legislators described:

The position of leadership on policy matters a great deal, but in my experience, the times when leadership really pushes all of its members on legislation is pretty rare. But when they do weigh in and push you in a certain direction, there is very little wiggle room on those policy issues (MD Senate member 1, 2015).

I know when I have the chance to influence a bill on my own, and I also know when the Chair or the Speaker wants something a certain way and I do not have that leeway. Rarely are they entirely prescriptive, but when they are you better respond and other times it may just be ‘pass this bill’ and make whatever changes you want (MD House member 3, 2015).

The constraints of leadership appeared especially acute from discussions with members of the New York House. Members from that chamber shared the following:

We don’t really have an ability to do more than vote up or down. Even then, votes against a Chairman is considered sacrilegious. You do not have a lot of flexibility, if you’re going to vote against a bill, you had better have a good reason and give a heads up (NY House member 1, 2017).

I don’t know why you want to talk with me about evidence. You know it’s all bullshit and we do exactly what the Speaker and leadership tells us we are going to do (NY House member 3, 2016).

These sentiments from New York House member (3, 2016) were provided in the absolute shortest of all the interviews conducted. In fact, this member took just a few minutes to share that “I would be happy to talk to you about how nice it would be if we could make evidence used in our work, but the truth of the matter is that we will say that it was whenever we can and we will have a convenient excuse if there are circumstances where we do not” (NY House member 3, 2016). His response was reminiscent of the indictment against New York politics laid out by Lachman and Polner (2006) in *Three Men in a*



*Room.* Seymour Lachman, the former state Senator in the Empire state, claims that “it takes just three men [the Governor, Senate President and Speaker of the House] in a room to cause devastating harm to a democratic system of governance” and have made the state’s legislature “rotten to the core” (2006, p. 1 and 3).

Members of legislative leadership expressed similar constraints as the representatives from the New York legislature. The Maryland House’s “second reader rule” is particularly instructive into the influence of legislative leadership on member decision-making:

If you felt like the second reader rule was tricky as a member, try being a member of leadership where the rule expectation extended to just about every vote from every committee (MD House member 3, 2015).

Now typically I went with my party when I am less knowledgeable on an issue. Then again, that was a few years ago before I joined leadership. Now I kinda have to go with the party on just about every issue (MD House member 3, 2015).

Maryland’s “second reader rule” will not be found in a rule book for the House of Delegates. Instead, it is an unwritten expectation of the Speaker’s leadership team (which includes committee chairs, committee vice chair, majority leaders and majority whips) that they vote “with” committees as legislation is moved to the floor for amendment – essentially guaranteeing that *any* proposed amendment will not pass, unless accepted as a “friendly” amendment from a presiding chair.

Making the second reader rule even stronger, this unwritten expectation is extended to rank and file members of the majority party so that they will not vote “against” the committee on which members serve when amendments are offered. This expectation applied *even in circumstances* where legislators voted *for* an amendment in committee – on the floor, members were expected to either vote no or be missing from the vote to prevent amendments from becoming enacted. It also applied to amendments

that were not considered within a committee before legislation was brought to the floor of the House of Delegates, with the argument being that members had the opportunity to offer amendments as a member of the Committee and should have taken that opportunity prior to the committee vote.

In my personal experience, some committees thwarted even majority members from introducing amendments in the committee process. If they were introduced, the committee leadership would often ensure the amendment would fail. On the other hand, there were other committees within Maryland that were open to amendments, regardless of party – and that were also willing to entertain occasional floor amendments as friendly if they made sense and were consistent with the spirit of the legislation. But in all circumstances, the second reader rule remained.

The insistence on loyalty associated with Maryland’s second reader rule stands in stark opposition to the goals of evidence-based policymaking. It provides strong incentives for voting along party lines – regardless of whether or not a proposed amendment is rooted in evidence, regardless of one’s personal opinion and even regardless of the positions of a constituency that is represented by a legislator.

In Massachusetts, one member sought to rationalize their support of leadership in circumstances similar to Maryland’s second reader rule, suggesting that it was a compromise that allowed them to be in the room for decisions “when it really matters.”

I definitely feel the pressure of ‘wow the leadership really wants me to vote a certain way on this even though my district or the evidence is the other way.’ But you try to be honest with yourself and try to figure out what’s best ... you gotta go with lots of rationalization ... Like I said, I may vote with leadership but that allows me to influence policy more often behind closed doors when it really matters (MA House member 2, 2016).

However, what “really matters” is subject to interpretation. For some members, just being “behind closed doors” may be a sufficient justification to follow the lead of leadership in their voting – even if they do not truly influence the outcome of legislation. For others, it may be that their leadership position is such that they have influence on proposed laws. This might be in an area as a committee chairperson or across a broader spectrum of legislation in another position. What is clear is that, whether being “behind closed doors” equates to the ability to impact legislation or not for a member, there is an expectation of adherence to voting with the leadership team to be there.

Thus far, legislators have expressed a desire to be responsiveness to constituent opinion, shared how their personal beliefs (and experiences) shape their worldview, and discussed how legislative leadership can command their vote on issues. While one might wonder how there is room for additional considerations, it is important to also explore how evidence can play a role in legislative decision-making.

#### **4.5 Evidence: what is it, how it’s used, good versus bad evidence**

In chapter three, we learned that lawmakers rated legislative staff expertise and research (55.3 percent rated as important) as well as expertise and research from an outside organization or university (58.7) about as importantly as their personal opinion (57.7) on the issue of bail reform. These factors were used in the survey as proxies for evidence based factors, and while they scored high as a factor for legislation related to bail, they also experienced the most dramatic decreases when the subject was changed to gun control (-9.3% for universities and outside organizations and -16.3% for legislative staff research). What accounts for these changes? Clearly, gun control is a much more salient issue and one in which personal opinion, constituent opinion and leadership would

all weigh in upon. Given prior insights from earlier sections, that may help to explain part of the story – but is there more to it?

#### 4.5.1 Evidence matters, but there are varying views of what “evidence” means

A critical step in communicating a legislator’s perspective on using evidence in their job is to get a better understanding of exactly they believe evidence to be.

Legislators’ definitions of evidence had considerable variance during the interview process, with responses ranging from science, study and data to a simple “Googling” of information on the internet.

In chapter three, we asked several questions that provide some insight as to what lawmakers believe evidence to be. Legislators, for example, reported that scientific studies, benefit-cost studies and program cost were among the most influential forms of evidence. They reported to find evidence most often from legislative staff and state or federal agencies. Hird (2005) data suggests that constituents are also an incredibly important form of “evidence,” which is not surprising and consistent with the feedback earlier in this chapter. However, there is little comparison in the quality of evidence that will be prepared from professional staff as compared to the average constituent. There were some legislators who, in principle, shared a higher standard for what should be considered evidence:

When I hear the word evidence I go to a great big rulebook and I think of what I have to deal with in court ... Evidence is not hearsay; it’s not evidence unless it’s backed up by some sort of external validation” (MD House member 2, 2015)

Truly evidence based policy is grounded in study, science and data. It points us towards what we are working on for the long term, but it also is typically what gets thwarted by politics, or funding, or some other short-term consideration that prevents you from letting evidence completely drive the decision-making process (MD Senate member 1, 2015).

Without placing a value judgement on it, another legislator focused on how much of the analysis they get is based more on the law than on academics. This was confirmed by a Pennsylvania legislator and a Maryland staff person:

A lot of the analysis we get is from a legal standpoint rather than an academic standpoint. I don't really feel a strong input from statisticians or academics (PA House member 2, 2016).

We put a lot of pride into the fiscal notes we create, but given the constraints of session they often do tend to be pretty straightforward on the changes to the law with our best estimates on what the impact might be (MD staff member 2, 2012).

Many elected officials reported a desire to play the role of referee when it comes to evidence, knowing that they will have information thrown at them from multiple sources.

Evidence is really laying out the nature of the problem, providing why the proposed solution is the best one and acknowledging that there are potentially some drawbacks, maybe some negative externalities to the fix but are overall positive (MD House member 3, 2015).

Facts are you know the evidence provided by people and organizations I trust (MD House member 3, 2015).

You typically have two sides and they pretty much present whatever they consider to be evidence – and you have to weigh whether you consider that evidence to be credible or not (MD House member 1, 2015).

Others still had a much less sophisticated approach to what they considered as

“evidence,” opting for the most basic of methodologies:

I try to seek out evidence by doing my own research. I will do my own comparative analysis just by ‘Googling.’ The information and evidence I find on my own is less biased – when a group comes to me they have an agenda and a bias (MA House member 3, 2016).

When looking at overall policy, I seek out evidence, such as a Google search of per capita spending per pupil on education spending as it compares to other states. Evidence is not the right word for it. It's data and information (MA House member 3, 2016).

Very often I'll just get my ideas from newspapers (NY House member 1, 2017).

I represent a very conservative district, so I seek out a lot of my evidence from conservative sources and conservative groups (PA House member 1, 2016).

Speaking to the low-quality of information gathered and utilized by lawmakers, one staff person shared, “It’s amazing to me what some legislators consider evidence” (MD staff member 3, 2017). These concerns were shared by legislators and staff alike:

You’ve got legislators that will pull up something with no rigor from the internet during a hearing and they will give that as much weight as a report you or your colleagues have worked on for months (MD staff member 2, 2012).

There is not much difference between sources as far as legislators are concerned (RI House member 2, 2016).

Even the legislator scoring highest on the evidence utilization score profiled earlier from Massachusetts struggled to articulate a good working definition for evidence in the legislative process.

That’s a tough one, even for me. I would say there is no hard and fast rule, but it is something more than ‘this is a good idea.’ I personally place hard data above experience, but they all matter: personal experience, hard data and the experience of others (MA House member 2, 2017).

Until there are clear standards for evidence – established by legislatures but informed by academics – this hodgepodge of “evidence” utilization is likely to continue. The lack of a clear definition and standards for evidence within the legislative sphere allows legislators who “Google” information feel as if they are among the highest users of evidence, falsely confident in the decisions they are making along the way. It is an environment where a member of the legislature who scored near zero on the evidence utilization scale feels comfortable criticizing fellow members for being “scared of their own shadow.”

#### 4.5.2 Trying to identify “good” evidence that can be trusted

If the case is such that legislators are struggling to differentiate good evidence from other sorts of information, what can be done? Are there ways to help identify

“good” evidence that can and should be trusted more than others? Based on the feedback from the lawmakers who agreed to interviews, creating an environment where there is clearly “good” evidence that can be wholly trusted remains far on the horizon.

If there was all the best research that really looked at an issue and did it right and presented it in a way that people could understand it, I do think it would improve policymaking, but I don’t think it would drastically change it. There are people where it just doesn’t matter – it is only about values and it is only about I like you (MD House member 3, 2015).

Information is something you have to take with a grain of salt and you have to look at where it is coming from. Sometimes information provided to you is not data driven, it is often perspective driven or ideology driven (RI House member 1, 2017).

In legislative settings people play fast and loose with the facts. You are not always getting a true insight into what real evidence would be ... You have to be somewhat skeptical in this job about what people say is evidence (MD House member 1, 2015).

I would say that there is a big gap between political science and politics (PA House member 3, 2016).

Politics isn’t science, it’s art ... What gives the academics the right to tell us what to do (PA House member 3, 2016)?

This skepticism was especially prevalent among self-rated conservative lawmakers among the interview subjects. However, there is insufficient evidence in my research to fully support the claims raised by writers such as Mooney (2005), who argue that modern conservatives pose a significant challenge to both science and evidence. As he writes, “if we care about science and believe that it should play a critical role in decisions about our future, we must steadfastly oppose further political gains by the modern right. This political movement has patently demonstrated that it will not defend the integrity of science in any case in which science runs afoul of its core political constituencies” (Mooney 2005, p. 269). This research failed to identify any statistically significant differences between political parties and ideology, but did find several

examples through interview data to support Mooney's claims. I believe this could have something to do with the area of focus against which the evidence utilization index was created. As discussed in more depth in chapter five, there is a strong consensus around evidence around the criminal justice area, a consensus that does not exist in other areas researched by Mooney such as climate change. This research has identified that there are significant barriers to enacting evidence-informed policies is challenging enough. Having legislators that seek to proactively debunk and discredit reputable evidence (as argued by Mooney and partially validated by this research) is downright concerning.

#### 4.5.3 Social science research: "It's a huge black box."

Part of the challenge in differentiating quality evidence from inferior evidence may be the inability of legislators to understand how information is gathered and evidence generated. While not measured on the survey instrument, interview subjects consistently referred to social science methods as a mystery that they could not understand:

It's a huge black box (PA House member 1, 2016).

I think you would get blank stares from most legislators. I personally would not be comfortable (PA House member 2, 2016).

I think that it is definitely a black box and I even have taken statistics (RI House member 2, 2016).

[Social science methods] is not language or a topic you hear a legislator talk about (PA House member 3, 2016).

Too often it is a black box, and the press takes the executive summary of a study on an issue that may be politically charged and it loses all the nuance when it is reported out." (MA Senate member 1, 2017)

Quite frankly, no, we have found very little appetite for understanding the nuts and bolts of the project (Schabes 2017, New York staff).



There are maybe one or two Senators that could explain what is being done (RI Senate member 1, 2017).

Even those who concede that advanced social science methods might have a place in the legislative debate, if done correctly, still leave doubt because of the ways in which some might approach the efforts with particular agendas, or that the work being produced remains “over the head” of those for whom it is being created. This includes work produced by institutions of higher education:

if you look at other individual representatives, some look at higher education information as being agenda driven (RI House member 1, 2017).

Academics need to know that lawmakers come from varying backgrounds. I might have two colleagues who come from a research background and can understand more technical details (MA House member 1, 2016).

If lawmakers do not understand the process, it is possible that it becomes easier to challenge or question the outcomes. As we will see in a later section, however, lawmakers also have neither the time to consume detail nor the interest to get into these weeds, preferring instead executive summaries – leaving evidence (and those who produce it) between a rock and a hard place. When discussed with legislative staffers, there was generally a consensus that legislators often struggled to understand the methodology behind most of the more sophisticated social science techniques – or chose not to understand them.

Most legislators do not have a deep understanding of the more sophisticated analysis. Maybe fortunately, a lot of what we produce is pretty straightforward and relatively easy to understand (MD staff member 3, 2017).

A focus towards evidence is definitely not the status quo (Dube 2017, Results First director).

I think bottom line most of the decisions around here are ‘does this feel right, does this make sense, what are other states doing?’ (Baker 2017, WV staff).

They are all very interested in what their peers are doing, regardless of what the evidence says. If we come up with something that is counter to their beliefs, they will question the methodology or find someone else to do it for them ... Sometimes we struggle with how much of a level of detail to go into when we meet with them (Brown 2017, VA staff).

There was feedback from Maryland staff sharing that they believed legislators were effective at understanding more straightforward metrics, such as program cost and individuals served – metrics that certainly fall into a category of “evidence” but are not as rigorous as many academic inquiries of benefit-cost efforts.

They [legislators] understand and use program data to inform their arguments. It is not always the most advanced argument but it works for them (MD staff member 2, 2012).

We find that our reports and audits on programs are often used as the basis for Delegates and Senators to draft their legislation (MD staff member 1, 2012).

This differentiation among Maryland staff suggests a fault line between more of a performance management approach and a wholly integrated evidence-based policymaking. Certainly, there are elements of performance management that provide relevant data points to inform policymaking. As discussed in earlier chapters, however, the two are not one and the same – and if these staffers are correct, I propose that there are policymakers content to only use performance management data (and not include more robust evidence) while others may not know the difference between the two, and still others may choose not to use evidence of any variety.

#### 4.5.4 Evidence rarely overruling other considerations

Legislators *consistently* reinforced the idea that, above all, constituents were the most important factor influencing their decision-making. They also discussed in great length the extent to which personal opinion (and experiences) and legislative leadership directs their legislative decision-making. With that context, there was little surprise when

legislators of all backgrounds consistently shared the ways they and their colleagues opted to actively vote against evidence:

Sometimes we do legislation because one person said something, regardless of what the experts say. I find some of my votes very painful (PA House member 2, 2016).

Sometimes it is a ‘pick your battle’ and determining if it is worth it – I go by a case by case basis, honestly (PA House member 2, 2016).

That’s a really tough thing to do (WV House member 2, 2016).

I’ve voted against the evidence. There have been lots of times where the district feels strongly about an issue but I am personally ambivalent. Which way do you think I’d go? (RI House member 2, 2016).

I suffer through those tensions; I absolutely suffer through those tensions. Sometimes I’ve gone with the data and sometimes I’ve gone with my feelings. I wish I could give you a solid reason why I do that, but sometimes I just can’t get past my feelings regardless of what the evidence says (RI Senate member 1, 2017).

Data is important but honestly, what is more moving for legislators are personal stories (MA House member 1, 2016).

This is my 14<sup>th</sup> year and more and more I realize sometimes it doesn’t matter what the numbers say if the politics are against you, and it is frustrating (RI House member 2, 2016).

There is an element of compromise and political factors that go into the legislative crafting process. It would be foolish to say there are not times when either select evidence or evidence as a whole gets lost in the process (Schabes 2017, New York staff).

More troublesome to creators of quality evidence are the ways in which it and all other information is not sought after as an independent arbiter on political issues of the day, but instead political fodder that is only deployed when it affirms the belief system of someone within the political system. Such a tendency toward confirmation bias is a major challenge for human reasoning and thus a real impediment toward robustly evidence-informed democratic institutions. “If one were to attempt to identify a single problematic aspect of human reasoning that deserves attention above all others,” writes Raymond

Nickerson (1998), “the *confirmation bias* would have to be among the candidates” (italics original, p. 175).

More often than not I see other people using evidence to just verify their own confirmation bias (RI House member 2, 2016).

A lot of people choose not to use evidence, they have their minds pretty much made up – especially the many who are ideological (MA House member 3, 2016).

People for and against issues all come forward with their own data. I am not saying it’s inaccurate, but it seems like people have a conclusion in mind and then try to find the data to support it (RI Senate member 1, 2017).

The information is there. Lawmakers will use it when they want to – most often to confirm what they already believe.” (MD staff member 3, 2017)

Similarly, several legislators recognized and expressed a frustration about the pre-determined outcome of legislation.

By the time legislation reaches the floor of the House, you know the debates are pretty well scripted (MD House member 3, 2015).

We have been at this long enough to know what gets passed isn’t just based on what the evidence shows (Schabes 2017, New York staff).

Sometimes that rock hard data just goes up against feeling or emotions and it gets watered down or completely ignored (RI Senate member 1, 2017).

There are folks whether it is their actual belief or otherwise they are going to deny the science and the facts. Even people who believe in what the science says, they just feel the need to say that they don’t believe in it, so the politics will get in the way of any real progress (RI House member 2, 2016).

Less evident in the frustration on pre-determined outcomes was a disappointment that evidence was lost in the process. While much of the known outcome on votes was attributed to a leadership call, legislators were resigned to the fact that that there is often legislation that will not be determined by quality evidence or any other objective measure of what constitutes “good” or “the best possible” policy outcomes. A few, as seen in quotes above, recognized that evidence was being left behind in these outcomes, but none actively expressed disappointment.

#### 4.6 Colleagues do not use evidence, “but I do”

One reason legislators might be comfortable with evidence being pushed to the wayside is they simply do not or do not care to use it in their work. However, one of the most interesting insights was the degree to which legislators explained how most of their *colleagues* struggled to put evidence to use in their daily legislative activity yet that *they*, on the other hand, were among the rare few that successfully incorporated evidence.

I think there’s a continuum of people between deep diving on every single issue but a majority really only want to get into bigger issues when it impacts their district, truly impacts their district, and those who will always just do what is safest and politically expedient (MD House member 2, 2015).

I would say I am definitely more of an exception when it comes to using evidence ... Out of [my entire chamber], I would say less than a third of us use hard evidence and are swayed by that (RI House member 2, 2016).

I think I’m pretty unusual [compared to my colleagues], but I do review the bills coming up on the calendar for next week ... I look at the bill analysis (PA House member 3, 2016).

My quantitative training puts me in a better place than most everyone I work with. I would say others take less advantage of data than me, partly through capacity and partly through interest (MD House member 3, 2015).

Maybe I am biased, but I am one of the most thorough when it comes to evidence in the legislature. I have seen colleagues make decisions that make me gasp. I am disappointed that people are much less thorough than I am (VA House member 1, 2016).

There is a huge gap between those who are thorough and those who barely do any work. And those people concern me. (VA House member 1, 2016).

Everyone thinks they use evidence so much more than other lawmakers (MD Staff member 3, 2017).

These data create a disconnect of sorts between a generally acknowledged and accepted lack of evidence in the policymaking process against a backdrop where individual

legislators nonetheless persist in interviews that they continue to make it a priority in comparison to their peers.<sup>41</sup>

#### 4.6.1 Use of colleagues for answers

The fact that legislators so universally had limited confidence in the use of evidence by their colleagues made the next common theme even more confounding: lawmakers trust and frequently turn to their fellow legislators for cues on how to vote, especially in those areas where their colleagues have an educational or professional expertise.

If I am not informed about an issue I will turn to a colleague I trust and sometimes my caucus informs me as well (PA House member 2, 2016).

I rely on colleagues I trust. For example, I lean on two physicians in our chamber for health and medical issues – one from each party – and I essentially give their feedback equal weight (WV House member 2, 2016).

On health care stuff, I have colleagues who will come to me for the cliff notes. On tax stuff, I may go to the CPA. We can't all be experts on everything, so we have ask 'who is my in-house expert who can educate me?' when I need it (MA House member 2, 2017).

The legislature has ADD – jumping from one subject to another. We are an inch deep and a mile wide, which forces us to rely on other people's judgement a lot (PA House member 3, 2016).

Everyone is focused on their own area – you have to trust colleagues in whom you believe in their work (MA Senate member 1, 2017).

Colleagues are the most important signal for me to follow (VA House member 1, 2016).

In another twist, one lawmakers reported an aversion to upsetting colleagues by going against their proposed legislation.

---

<sup>41</sup> It is possible that “low users” of evidence rolled off from the survey portion, leaving a skewed interview pool for interviewees. However, in both the survey and interview research, the sample represented the full range of scores on the evidence utilization index.

You have to remember you are voting against your friend if you vote no – and you want them to vote for your bills. So most of the time I will vote for the bills that come out (NY House member 1, 2017).

If the average person would be surprised to know how much sway the colleagues of legislators have over each other in the voting process, they might be shocked to learn how strongly lobbyists and special interest advocates have affixed themselves into the evidence production, deliberation and vote influencing process.

#### 4.6.2 Lobbyists as a proxy for evidence – “You start to rely on lobbyists”

Based on several interviews, it appears that many state policymakers allow the lobbying corps to be an integral part of the lawmaking process. Lobbyists work alongside state legislators, acting as providers of “evidence” and ultimately influence the votes of legislators. La Pira (2008) speaks about the extent to which lobbyists hold a special advantage in impacting legislative change. Writing about members of Congress and their staff, La Pira’s comments also apply to state legislative bodies. As he shared, lobbyists and the interests they represent “meet a critical legislative demand for timely and accurate information” (La Pira 2008, pg. 5). As such, he argues that the interactions with lobbyists is an area in which exploring how information processing behavior impacts public policy decision-making is a promising endeavor (La Pira 2008). Based on feedback from interviews, it appears that there is in fact a similar meeting of demand for timely and mostly accurate information at the state legislative level, as well. Legislators in Pennsylvania seemed especially comfortable with this arrangement:

I would also give a plug to lobbyists and special interests, they do know what they are talking about in their area. The best ones are credible and honest and you trust them until they lie to you ... *You start to rely on lobbyists*, special interests and associations – *you may even trust them more than the House staff*. (Italics added, PA House member 3, 2016).

Some people say PA is too liberal allowing lobbyist access. I have no issue with it because I want the straight scoop and I'll make the decision based on what you tell me. Lobbyists know some of the integral details that you won't see in a newspaper article (PA Senate member 1, 2016).

I go back to the lobbyists for much of my information – the ones I trust. (PA Senate member 1, 2016)

If the Pennsylvania legislators included in this study were wide open in their comfort of the lobbying corps, many other state legislators indicated at least an openness to include them in the lawmaking process.

I like to invite advocates in ... I think we would make more mistakes with them not being there (MD House member 3, 2015).

Often times our information is provided by the advocates on either side. They have time to do the gathering or research that legislators don't (NY House member 1, 2017).

People don't lose points for being advocates. What is important is if they provide you with stuff you can rely on (NY House member 2, 2017).

Some of my constituents look at lobbyists the same way they look at used car salesmen. But I have no issue dealing with lobbyists. In some situations, it has actually been very helpful (MA House member 2, 2017).

After all, lobbyists can only contribute up to \$200 towards my campaign. The public can contribute up to \$1,000. They both have some influence, the only difference is that the public is not registered (MA House member 2, 2017).

The success with which lobbyists can infiltrate state legislative chambers has not gone unnoticed by the several staff members interviewed for this project. Their responses reflect a mix of annoyance and frustration – both by the levels at which paid advocates can influence the process as well as the degree to which they can become a substitute for them. Staff members believe that they would largely be picking up much of the slack of resource and information requests that may currently be directed toward lobbyists.

Lobbyists end up having large amounts of influence [in West Virginia]. I think a lot of this has to do with the fact that we don't have an independent budget finance office. On non-budget policy, usually what happens is there is some



organization or entity interested in a topic and lobbying for it – that’s where legislative information comes from (Baker 2017, WV staff).

It would be nice to have good policy experts who are not paid lobbyists (Baker 2017, WV staff).

I would say for me personally there are definitely lobbyists that you know are completely lecherous. And you know that ... How much are they willing to shrill for their client? People develop reputations for being fair and honest brokers, and people develop reputations for not being fair and honest brokers (MD staff member 1, 2012).

Another Maryland staffer took a more nuanced view on the subject, stating their appreciation for the ways in which they were better equipped to do their work when lobbyists would share information to which the staffer might not otherwise be privy.

If it’s someone that I respect and feel that there is a two way street in terms of information sharing and work too – sometimes they will come to you and expect that you are going to go lobby the members [of the legislature] on their behalf (MD staff member 2, 2012).

For me, the lobbyists that I like to work with and try to develop a relationship with are the ones who are most prepared, know their issues, and they bring information to the table that you might not have otherwise. That I really appreciate (MD staff member 2, 2012).

Across the board, it was fairly evident from interviews that lobbyists have done a good job becoming entrenched into the daily operations of most legislative functions.

Legislators (and even some staff) have come to rely on these lobbyists for information, insight and guidance, to the point where lawmakers will defend the work conducted by these lobbyists in confidential interviews.

#### **4.7 Results First unknown to legislators – even in states where it exists**

If lobbyists were a known value to legislators, the program and variable of interest – Results First – was quite the opposite. Legislators were completely unaware of the Results First program during interviews, so much so that many of the questions originally

slated for the semi-structured interviews were never put into production. This was aligned with the findings in chapter three, which informed us of this gap in awareness on the part of legislators. Nearly 74 percent of all respondents indicated that they were not knowledgeable at all and only 1.8 percent had an above average familiarity – with absolutely zero legislators considering themselves very knowledgeable about Results First. A full 76.3 percent of respondents did not know if Results First was active in their state and a few even incorrectly identified that the program was not active in their state (when in fact, it is).

For legislators without Results First active in their state, it was less of a surprise when they reported during an interview not knowing about the program. While there is a chance lawmakers without an active Results First effort could have learned about it elsewhere, the likelihood of that is not especially high.

No, I have not heard of them. Most of the information we get is from one side or the other (PA House member 3, 2016).

However, legislators from states *with* an active Results First were also unable to articulate any real working knowledge of what the program was or how it operated.

I have not heard of Results First. We get lots of flashy documents (MA Senate member 1, 2017).

I have never heard of Results First (RI House member 1, 2017).

I don't think so, it does not sound familiar (MA House member 3, 2016).

It sort of rings a bell. (MA House member 1, 2016).

Nope, I'm not familiar. But I think it's great [the goals they have]. The more people can put into our hands, the better (MA House member 2, 2017).

One legislator from Rhode Island could describe the *work* being conducted by Results First, but was unable to make the connection between what he was describing and what the program delivers. This legislator was familiar with the Washington State Institute for

Public Policy and expressed an interest in “having it created” in within his state, sharing concerns that it would be perceived as a creature of the Democratic leadership. The program, of course, is already in place in Rhode Island.

Pew has a good reputation, but I have not heard of Results First. I’m in a state where there is this bipartisan research council funded and created by the legislature that has Republicans and Democrats. It apparently has a pretty strong standing from what I have heard, and I think it was a Washington state model. I’d like to see if there is something we can do here to be more like that. I think Pew or someone at an entity like that could really help demonstrate an unbiased sort of place and folks could count on that, it would be helpful. It would help provide some sort of *trust* in what is being said (italics added, RI House member 2, 2016).

While Results First was virtually unknown to all the legislators with whom I spoke, there was more knowledge and awareness of the Pew Charitable Trusts.

I’m drawing a blank on that one. I am familiar with Pew but not Results First (RI Senate member 1, 2017).

I’ve heard about Results First but I don’t know a lot about it. I don’t know enough to have any feeling about it (PA House member 2, 2016).

I can’t really tell you about Results First. I know they have some sort of presence here in West Virginia, but that’s about it (WV House member 2, 2016).

This lack of program knowledge may not change anytime soon. As discussed in more detail in a “in their own words” section on Results First, the program’s intentions are not focused on becoming known and supported by rank and file members of the legislature.

Sara Dube (2017), the Executive Director of Results First, noted that the average legislators are “honestly not always the target ... For us, they are generally seen as a secondary audience.” If Results First is influencing members of a legislature and making them greater users of evidence in their policy work, that impact is probably happening without their knowledge – and certainly without making any connection to the Results First program.

## **4.8 Results First: in their own words**

In addition to speaking with legislators, an interview was also conducted with Sara Dube, Director of the Results First Program. Much of this conversation centered on allowing the program to speak to possible concerns and limitations regarding Results First. These limitations, mentioned in earlier chapters, included state variation, program replicability across content areas, selection bias, implementation standards and time to value. I believed it was important to allow the program to directly address these topics rather than to leave them completely unexplored or discussed without the benefit of insights from those directly managing the program. The interview also presented an opportunity to dig deeper into the ways in which the program's leader feels Results First has practically been influencing policy outcomes. Questions also covered successes and challenges working with state lawmakers, operational questions (such as model deployment and implementation data collection).

### **4.8.1 Variation among the states**

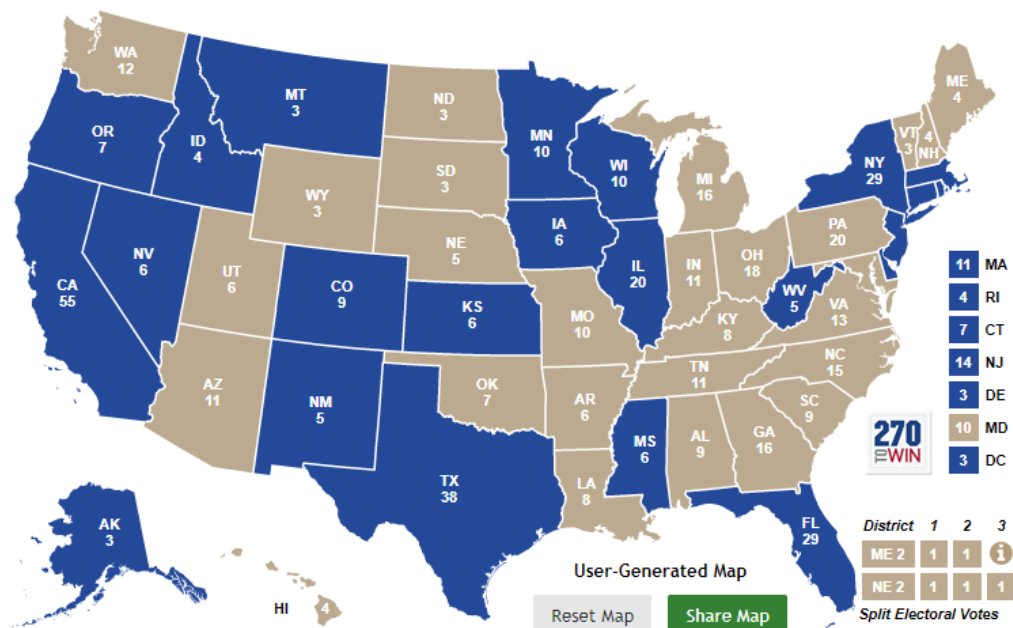
There are lots of variation among the states within the program. This variation presents an opportunity to measure how effective the Results First model is, with states from across the country creating a diversity in just about any desired metric. The largest exception to this diversity, discussed next, is the potential staff capacity or implementation ability. The extent to which the statistical models underlying the Results First analysis accurately compensate for the known differences between the states in which Results First operates is also unknown. During an interview, the program director

only shared that “the underlying math is the same,” with a “methodology that is pretty set” and that states are “not really able to change assumptions” (Dube 2017).

It is a benefit-cost model built on WSIPP. It was generated 20 years ago and it has been refined over time. In 2010 there was an ‘experiment’ with a large uptake in other states around the country. We have taken that underlying model and made it much more robust (Dube 2017).

This may or may not present concerns regarding the operation of the program. If, for example, the proprietary information of the model was made available for a thorough review, there could be a better confidence that it addressed the demographic variation of the states being served. Considering the wide distribution of states serviced by Results First (see the figure below), a better understanding of the underlying model would provide additional confidence that it meets the unique circumstances of each state.

*Figure 9: States in the Results First program, those in blue are Results First states)<sup>42</sup>*



Graphic Source: [www.270towin.com](http://www.270towin.com) (2017).

<sup>42</sup> Appendix K is a complete listing of all states and local governments currently partnering with the Results First program.

#### 4.8.2 Ability for the model to “travel”

Model availability would likewise provide some clarity on the extent to which the model works in varying fields of public policy. There is an open question both about how effectively the underlying WSIPP model has been adopted for individual states within the Results First program as well as regarding how effectively the Results First criminal justice module, which has been the most thoroughly developed, applies to the other areas where Results First is attempting to expand with states, such as behavioral health. Despite over a decade of overall experience, Dube (2017) conceded that “most of our experience has been with adult and juvenile justice. It’s where most states have done their work and there is quite a bit of literature in those fields.” I asked Dube directly about the state’s ability to tweak the model’s assumptions if they had an interest and the talent to do so and was informed that the “methodology is pretty set” and access to make changes is highly controlled and limited (2017). The same is true for reviews by outside organizations – something that has never been allowed in the name of protecting intellectual property. While Results First rightfully considers its work proprietary, the complete shutdown of external model validation and inability for individual states to consider model changes threatens to stifle the very rigor the program seeks to promote.

Just as the secrecy associated with the Results First model makes it very difficult to determine how well the program accounts for state-level variation, it also makes it difficult to address questions about how well the model fits with varying content areas. Finally, would-be evaluators might also rightfully be frustrated by the lack of access to confidence levels of Results First analysis – another metric not shared externally by the program.

#### 4.8.3 Selection criteria / selection bias / implementation standards

As discussed in earlier chapters and above, there are selection criteria required for entry to the Results First program that could exclude states. These include an unwillingness or an inability to join the partnership – for political or practical purposes, such as staffing resources. If such an exclusion exists, one could argue that there is a selection bias associated with the program, possibly skewing the nature of the results. When confronted with this question, Dube (2017) said “I think there’s some of that. There are a variety of reasons why the other 26 states are not working with us now.”

New York implemented their own quality assurance program, recognizing that implementation challenges can derail even the best intended of programs. “For us, a tipping point of the work was when we realized that to make this the most effective program possible, we really needed a fidelity and quality assurance system to be in place. We wanted to confirm that programs were implemented as proposed in the field. We needed to close the loop, otherwise it creates a substantial unknown” (Schabes 2017, New York staff). New York’s activity in quality assurance may not as much signal skepticism regarding Results First as much as it affirms the importance of efforts following policy adoption. Results First does encourage partner jurisdictions to set up an implementation framework to follow through on the data produced, suggesting a group comprised of technical staff to collect and analyze data and another as an advisory group to review assumptions and make other operational decisions (Dube 2017). Whether this approach is successful and all states follow (or plan to follow) New York’s lead in this space remains an open question – as is if there is a selection bias such that the states with which Results First works has greater capacity to implement programmatic changes.

#### 4.8.4 Time to value

The program admits that there can be a considerable gap between the standing up of a Results First effort and the realization of value. This is an area where Director Dube admits it “can be tricky” to know the benefit, especially for legislators who often are no longer serving in the seven to 10 years it can take before the tangible benefits are accrued (2017).

Another issue related to the time to value concept is the potential for wildly inconsistent tracking of results among states. Individual states are left entirely on their own to determine if, how and when to measure their own results. “We are capacity building only,” Dube (2017) shared, adding “We don’t do that tracking and monitoring other than pieces of legislation and how much money has been moved.” This lack of consistent tracking measures makes a comprehensive measurement of the Results First program problematic.

#### 4.8.5 Leading as enlightenment or through power channels

In addition to the above questions regarding how Results First works, this research has shed some light into the ways in which the program has influenced policy outcomes. This is more of an academic inquiry regarding the practical nature of how the program works and less of a measure of effectiveness. There are at least two ways in which social research such as Results First can influence the policy process. The first, described by Carol Weiss (1977), suggests that the primary use of such research is not applying specific data to specific decisions. Instead, she argues, research is used indirectly as an enlightenment function, providing ideas and orientation (Weiss 1977).



The other opportunity to influence outcomes is via power channels. In this space, Barnard's (1938) argues influence stems from those in leadership while Dahl (1957) speaks to the ways in which agents can utilize social influence through the manipulation of resources available to them.

Between these two approaches, the path taken by Results First is clear: power channels are the path to policy change. Legislative leadership and a governor's office are the primary audience for Results First *both* as it relates to the decision to bring the program to a state in the first place *as well as* efforts to implement use of Results First's cost-benefit analysis, once established. If an enlightenment approach is influencing outcomes through the Results First process, it is only done so as a secondary consequence of the work. This is consistent with remarks shared earlier from Dube (2017) that average legislators are a "secondary audience" for the Results First program and "probably not always the target." "Leadership needs to be bought in ... the top brass needs to at least be generally supportive" (Dube 2017). Kevin Baker, the lead staff member tasked with implementing the project there, agreed with this sentiment.

most of this is leadership driven (Baker 2017).

In West Virginia, it was absolutely the Majority Leader of the State Senate who got interested in getting Results First going here ... we had just gone through a justice reinvestment project and had some momentum already. With the leadership pushing too, it made it pretty easy (Baker 2017).

The risk of using power channels is that power can work both ways. A new election result quickly turned the tide and momentum of Results First in West Virginia.

In the 2014 election, the Democrats lost control and in the wake of the election change, the program just sort of got ignored. The Speaker is not a fan of the program and not a fan of Pew and there was no real leadership in the Senate that continued to carry the ball, so the program has sort of fallen to the wayside (Baker 2017).

It was supposed to be less susceptible to political change, and that sort of didn't work out in West Virginia (Baker 2017).

Despite that risk, it makes sense that Results First goes after leaders to deploy their model within states. Support from leadership can ensure that a Results First system is established, makes it more likely that the program will remain in place, and more likely that legislators will follow the lead of the results.<sup>43</sup> Rank and file legislators can feel good about also pursuing evidence-based policies, but in the end, it remains more of a byproduct of a process enacted by and carried out by the power channels within legislative and executive branches. Power channels can include a take-up of Results First recommendations through specialist channels such as staff, cue-taking from legislator specialists, leadership buy-in, and lobbyists using report information in their arguments, among other methods.

While Results first may leverage legislative power channels to enable their work, enlightenment through policy research may still have played a role enabling the progress. To wit: Results First has primarily operated in the criminal and juvenile justice policy areas, and it has been argued that policy enlightenment in these areas has equipped legislative leaders with the leeway to allow findings from the Results First program (and programs like it) to drive the policy agenda. Even the most conservative of lawmakers, who have historically been strongly opposed to any law changes that appear “soft on crime,” have largely bought into the value of justice reinvestment efforts – with policy enlightenment arguably playing an important role in this transition. This is a dynamic that I explore in more depth during the next chapter.

---

<sup>43</sup> This assumes, of course, the legislative leadership *wants* to enact the results of the research.

The lack of such “enlightenment” in other areas may be one reason why Results First has been slow to expand into other policy areas. Instead, the growth has focused primarily on increasing the number of states and local governments involved in the program. This is not to suggest that Results First does not deliver value in the states and subject areas in which they work. Quite the contrary, chapter five details how Results First and other partners in the criminal justice space have played a critical role in the development of more evidence-informed policy. These concerns do, however, provide a set of questions that state (and local governments) can and should be asking of Results First or any organization seeking to partner with them. Moreover, considering the tens of millions of dollars (and possibly more) in public investment being driven by the Results First program, it would not be unreasonable for Result First states to ask to see and perhaps be involved in a robust assessment of the effectiveness of the program itself.

This interview adds an important perspective to the operation of the Results First program. I did not endeavor to include the program directly because I thought there would be new insights related to the chapter themes; for example, I had no reason to believe anything about the program’s standard operating procedures encouraged or discouraged its use. I did, however, believe it was important to provide a first-hand discussion of the potential challenges and limitations of the Results First program. It remains unclear how worried state officials should be about the limitations of the Results First program, but I was at least able to validate that limitations exist and should be kept in mind while pursuing the program. Having done that, I return next to consistent themes in the interview data about ways in which evidence can most successfully be integrated into the policy process.

## **4.9 Embarking on change – how to push evidence**

These more granular interviews of this chapter had the benefit of allowing legislators to share details on the obstacles that were most likely to prevent their use of evidence. Similarly, it afforded them the opportunity to volunteer ideas about what would be most likely to help them access evidence in their work. Key themes in this section included a lack of time, difficulty in accessing evidence, a desire for executive summaries, the need for credible messengers and the importance of having strong relationships.

### **4.9.1 The lack of time**

It comes as little surprise to anyone who has spent time near legislative bodies to know that lawmakers are tremendously busy and that their time is a scarce and valuable commodity. This was a sentiment that was consistently shared across interviews.

Come session, we really do not have a lot of time to really get into the details of an issue (MD House member 2, 2015).

When I get long reports, I ask ‘really, seriously?’ It’s nearly impossible to have enough time to read (PA House member 2, 2016).

Time is the enemy of a thoughtful process (PA House member 2, 2016).

I don’t have the time or the energy to pursue every piece of research that’s cited (MA Senate member 1, 2017).

I may have seen a bill a day or two before I vote on it. There is not much time to evaluate what we are voting on (VA House member 1, 2016).

I can’t meet with every constituent and interest group individually, I just don’t have the time. With 6,000 bills each session, I can’t spend all of this time just on your issues (MA House member 2, 2017).

We don’t have a ton of time to do research, so if it isn’t falling into our lap, we probably aren’t seeking it out (NY House member 1, 2017).

One Maryland legislator also professed concerns for the time constraints facing the professional services organization providing feedback on legislation – not only reinforcing the time challenge faced by those within the legislative process, but serving as a reminder why lobbying and special interest groups have an opening to fill a void and establish credibility. With focusing on only one or a few niche issue areas, lobbyists can spend all their time generating information on a topic to distribute *and* cultivating relationships, a separate topic covered more in a later section.

DLS [Department of Legislative Services] writers are given very little time to collect the information and provide an analysis. They are under a lot of stress and sometimes have only a few hours to put together a complex fiscal note. So I would say it's hard to trust their analysis as much as, say an advocacy group that has a year to go out and put their evidence together (MD House member 1, 2015).

This time concern was shared by JLARC staff in Virginia. As Justin Brown (2017) from JLARC shared, “Virginia is a very part time legislature so being timely even to them can be difficult. They give us three to five topics a year to research. It is hard to say how long something is going to be relevant.” He added that studies can but done as quickly as four to six weeks but could also take anywhere from six to nine months to complete. These staff timing issues are a challenge because often reports may not fall into a perfect rhythm with legislative sessions, especially if there are new or extremely pressing issues before a legislature.

#### 4.9.2 Making evidence more available

Despite time constraints facing legislators and staff members alike, several lawmakers shared that it would be helpful if evidence were more readily available to them. Many suggested they were aware that considerable information and evidence existed, but they also were not inclined to go and find it on their own.

The depth of knowledge is really shallow, unless it's an issue you or your constituents really care about (PA House member 2, 2016).

The bottom line is that most of the decisions around here are based on does this make sense, does this feel right, and what are other states doing? (Baker 2007, WV staff).

The question is how do academics get our attention? How does one find information if you're not looking for it (MA Senate member 1, 2017).

It is not because I am against it, but the evidence just does not present itself to me. They have to find me, I don't have time to seek it out (VA House member 1, 2016).

We don't have enough at our disposal to really delve into these difficult topics we are faced with (RI House member 1, 2017).

I would say that a lot of the types of information you are referring to would be much more useful to the bill sponsor. It wouldn't be common for me to use someone's research to help a colleague fix a bill. But I am also not sure the sponsors would be seeking this information out (NY House member 1, 2017).

There were some mixed messages on this point, as some other legislators shared that they were overwhelmed by data and information. While not directly shared in conversation, it is possible that the difference here is that legislators do believe that "academic evidence" is limited in their work even as there is an abundance of "general information."

#### 4.9.3 Executive summaries over complicated stories

One way that academic information / quality evidence could be made more readily available to legislators is through executive summary documents. As discussed above, interview data included reports that general information being presented to legislators was overwhelming. Not only is information often presented from multiple sources, but that information does not always come with a short review. Legislators indicated that they preferred getting as much information as they can in the shortest amount of time possible and seemed especially appreciative to have complicated evaluations honed into shorter documents.

Unfortunately, anything that's long it's going to be hard for me to get through an entire document (PA House member 2, 2016).

Maybe they will just read the synopsis or something like that, but I wouldn't expect much more than that (RI House member 2, 2016).

It is always wise for them to have an executive summary. There are some people who just like the executive summary – I would say less than half our members would ever get into the numerical data (VA House member 2, 2016).

We are just flooded with so much information as a legislator. It's hard to keep track of it all and impossible to be an expert. I am going to spend the least amount of time as possible to understand an issue (MA House member 1, 2016).

We are all just drowning in information and data points (RI House member 2, 2016).

The mix of legislators calling for more evidence while sharing they are overwhelmed by information is, at first appearance, a contradiction. However, looking at the issue in more depth and in context, legislators seem genuinely interested in having more quality evidence to consider, if it can be presented in a concise and sleek way that does not add significant time or effort to an already-full plate.

#### 4.9.4 Credible messengers

The message provided by evidence producers is one issue. The messenger is quite another, but equally important subject to consider. As it relates to academic and nonprofit organizations, a Pennsylvania representative shared that there is often a credibility gap with these messengers, asking rhetorically “What gives the academics the right to tell us what to do?” (PA House member 3, 2016). Based on the interview responses from legislators, they are most likely to turn to and accept evidence from what I call *credible messengers*: those who have “done their homework,” do not have an agenda (or admit upfront what their agenda is) and who have external validators that matter to the legislator being approached.

Certain indications will tell me that you haven't really done your homework. You need to be able to tell me why this is important, important to you (MA House member 2, 2017).

There will never be one person who is truly independent. I think you have to talk to both sides and let it all balance itself out ... You have the committee staff who is honestly the people I find to be most helpful far and away (MD House member 2, 2015).

My experience is that a lot of legislators don't actually trust experts (PA House member 2, 2016).

I think people around here would listen if evidence was presented to show something did work. I think our members could be persuaded by that even if their instinct is to go against something, but they need to have data they can trust (Baker 2017, WV staff).

I do not think there's much research that does not start with a pre-determined outcome. (MA Senate member 1, 2017)

Do not come across as a know it all and just bombard members with numbers and statistics. Don't be condescending: I'm in the real world trying to balance a lot of things and not in academia only trying to meet one objective (RI Senate member 1, 2017).

If the source is something random sounding, members are less likely to be attentive to it (RI House member 2, 2016).

Having the press communicate the importance of the studies is a good start and really important. Constituents leveraging their relationships with representatives are also likely to have attention paid to these issues. Either way you need a validator (MA Senate member 1, 2017).

Good data shows and instructs the process but I am not so naive to think it is the only thing. Data is relevant but it is all about politics. It is all about values. Know the difference and know when to back off and say 'all right' (MD House member 3, 2015).

There are certain groups that I have learned who are 'straight shooters' and ... they're not really cooking it or making it sensational or kind of using it to suit their needs. They have earned credibility (MD House member 1, 2015).

Above all else, a credible messenger is one whom has relationship ties with legislators.

Familiarity requires relationship and it is from relationship where trust is fostered. It is *the* key ingredient in advancing any change in a legislature.



#### 4.9.5 The power of relationships

The most consistent finding in all the interviews conducted are the ways in which strong personal relationships are the key foundational ingredient in successful policy outcomes – whether those outcomes are evidence based or not. It might seem intuitive, but it is useful to have gone through the exercise to confirm the importance of relationships and relationship building – and ultimately, its impact on the legislative process. Not surprisingly, the foundations of strong relationships (trust, respect, communication, etc.) in society are similarly evident in the best relationships legislators have with outside interests. A finding about the importance of relationships and the ways in which they are cultivated confirm my own experiences and interactions in my former role as a state legislator. It also affirms prior work by Hird (2009), who concluded that the “most important dimension of policy advice is personal connections” (p. 533).

While relationship building served as one of the overarching themes, it is important to note that the other themes serve as mutually-reinforcing mechanisms that interact with relationship building. For example, the strongest relationships were based on an open and honest exchange of information, credibility and fair dealing by participants, and an investment in the long-term. The strongest of relationships will blur professional and personal lines and include regular, recurring communication. Legislators were clearer on this point than perhaps any other in the research.

I don't find national people to be helpful unless they've spent a lot of time doing real work and research in [my state] (MD House member 2, 2015).

I think it would be helpful to have the experts come in for a hearing ... having them come in to lobby and speak with us personally is helpful. Unless we start educating people in this country, we will continue to just go by the seat of our pants (PA House member 2, 2016).

How do you break through? I don't know, but I will say personal relationships or a major splash in the media would go a long way in helping (MA Senate member 1, 2017).

Sometimes outside groups don't realize that their analysis is irrelevant in general, and it certainly is without the personal touch (MD House member 3, 2015).

If a social scientist were to contact me directly, I would be more likely to consider (VA House member 1, 2016).

I don't think we get much from academia – it's not like they come to me and ask me to put a bill in on something they have studied. Maybe they should try that (NY House member 1, 2017).

It is all about relationships. There are some lobbyists and legislators who would come here and beg and plead and I don't think we would help them at all (MD staff member 1, 2012).

I think the people that have the best ability to impact this office's views are people who have come out of state government ... because you develop a relationship with them outside of them lobbying you as a contract lobbyist. So you know their strengths and weaknesses because those things become inherent when you dig through the trenches with someone (MD staff member 1, 2012).

Oh yeah, there is totally an intersection between politics and personal relationships (MD staff member 2, 2012).

I will turn to the people that I trust – If I'm not an expert on a subject, there are people around me that I will lean on (MA House member 1, 2016).

It would be interesting to have academic institutions come to us early in the session and ask 'what do you need' as opposed to providing research on what they think we need. It is all about relationships, and the same is true with institutions of higher education (MA House member 1, 2016).

Contact me at other times than when you have an ask. Build that credibility, build a relationship with me. Become someone I know and with whom I've interreacted and you're more likely to get what you're asking out of me. I think it really is the personal contact that makes all the difference (MA House member 2, 2017).

Hird (2005) writes that “Virtually all think tanks go to great lengths to be relevant by publishing short summaries of scholarly research, testifying on Capitol Hill, promoting events where researchers and policymakers interact, and generally making their work more accessible” (p. 113). This approach stands in contrast to the role that

many state legislative policy shops often find themselves playing. While legislative policy organizations may get the most amount of face time with legislators, what they do not enjoy is significant “personal” time nor many opportunities to develop relationships in such a way that think tanks, lobbyists and special interest organizations go out of their way to cultivate. Lobbyists make the most of their time with legislators, and make their interactions regular and ongoing. In Maryland, for example, lobbyists and their clients will regularly invite members of the legislature to join them for elaborate dinners in the capital at locations such as Ruth’s Chris Steakhouse. One journalist opened his reporting on the subject by saying “It’s political party time in Annapolis as the Maryland General Assembly cranks up its 2017 legislative session” (Donovan 2017)<sup>44</sup>. Layer on the ability to contribute to a representative’s campaign, and the differences in dynamics are very striking – and help to understand the information reported above.

Public policy schools themselves go out their way to train policy analysts as independent arbiters of information. Many focus on the idea that policy analysts play the role of client to legislatures (Hird 2005). In truth, they are; however, unless there are more concrete steps taken to allow for personal ties and bonds to develop between state policy analysts (and/or academics and/or outside organizations pushing for more evidence-informed policy), there will continue to be an uphill climb against other well-entrenched forces that typically have little use for evidence unless it advances their own agenda. Considering the ways in which legislators already must navigate their own political terrain within the district and leadership structure of the legislature, any additional hurdles do not bode well for the use of more evidence in policymaking.

---

<sup>44</sup> I can validate these dinners personally from my experience in the legislature as can a recent *Baltimore Sun* article on the subject at <http://www.baltimoresun.com/news/maryland/sun-investigates/bs-md-sun-investigates-protocol-calendar-20170121-story.html>.

#### **4.10 Transition to chapter five – lessons learned, next steps, etc.**

This chapter has taken a deep dive with the actors most involved with evidence utilization. First, it provided an overview of the methodology and process for collecting data. It then analyzed survey data from most of the legislators scoring the highest and lowest on the evidence utilization score developed in chapter three. Next, it reviewed key findings and common themes from across all 30 of the semi-structured interviews. Following that, data were shared with staff members most familiar with Results First to address key questions raised about the program. Finally, the chapter ended with insights shared by legislators and staff on ways to better infuse evidence into the policymaking process.

As part of that exercise, there have been important insights that continue to inform the hypothesis set forward at the beginning of this research, affirming the considerable hurdles that remain to having evidence significantly influence the policy-making process. Specifically, these interviews affirmed that constituents matter the most to the legislators with whom I spoke. Personal beliefs were also very highly rated, and the strength of these findings recall the analysis of chapter three indicating a similar strength of influence across the entire sample of lawmakers.

On many of the elements described by legislators as fundamental to advancing the use of evidence in policy, Results First can point to considerable but incomplete successes. In the most important category, relationships with legislators, they are incredibly strong in connection to key leaders, but they fail to relate to rank and file lawmakers. The program, to the extent it is known, is viewed as a credible producer of information and they also effectively produce summaries in addition to lengthy analysis.

Results First does not make evidence readily available beyond the normal deliberative process in the legislature, making the most informed those lawmakers most connected to the Results First governance or policy reviews. This gap in evidence distribution mirrors the discrepancy in overall relationship levels that Results First manages with leaders versus back-benchers. The final element, time, is one over which Results First has no ability to influence.

I also found that, while lawmakers were not often proud of quick to point to the influence of institutional leadership, it nonetheless remained an important factor – again affirming findings from chapter three. The fact that chamber leaders, committee chairs and governors continue to wield such outsized influence in the legislative process in many ways makes the approach from Results First and Pew very wise, actively bringing together improved evidence with the power channels that can pass legislation based on its findings. If there happens to be enlightenment that occurs during the legislative process, it is a bonus. The focus is clear: leverage leadership to both establish the program and make the evidence-based results drive governance. What does that look like in practice?

Chapter five focuses specifically on that subject, and in a policy area in which Results First has, to date, focused the bulk of its time and resources: criminal justice. Specifically, chapter five drills down more deeply on the area of justice reinvestment to provide additional context into the role Results First plays in policy outcomes. I focus on this as an area in which Pew has provided considerable advice and influence on a policy topic, using justice reinvestment as a comparison to the efforts of the Results First project.

In the meantime, the biggest takeaways from this chapter as it relates to the hypothesis are the extent to which Results First is largely an unknown program but that it may still also, at times, meaningfully impact the enactment of significant legislation. Data from this chapter also further affirm findings from the advocacy coalition framework regarding the varying circumstances under which evidence is likely to impact legislating—less so on issues related to core personal beliefs and more so on issues where there is limited personal, constituent and legislative leadership interest.

## Chapter 5: Prison Break and justice reinvestment: a survey of legislative activity showing success with limitations

This chapter takes a brief but deeper dive into an issue that has become especially salient in recent years: justice reinvestment (JRI). States across the country are turning to justice reinvestment as a process to curb escalating costs of corrections and lower the rate of offender recidivism while also maintaining public safety. Lawmakers of all variety, including conservatives usually opposed to efforts that are “soft on crime,” have largely bought into the reforms associated with JRI. Such widespread buy-in from legislators suggests that, if there were a policy area in which evidence utilization could thrive, it would do so with justice reinvestment. There are willing legislative participants, ample supplies of quality evidence around which consensus has been formed, and significant resources being deployed to move toward policy change.

With Justice Reinvestment serving as an ideal example of evidence informing policy, one should expect to see many of the promises of evidence-based policymaking to be realized. Specifically, legislation would be proposed that delivers the best return on investment, as determined by detailed benefit-cost analysis. One would also expect to see proposed legislation enacted in ways that reflect only small changes to maximize the return on investment. JRI is straightforward enough of an issue on the cost and the return that it is hard, objectively, to argue against its enactment. However, as found in the case studies of this chapter, not only were JRI reforms argued against, they had been thwarted for several years running in some states. In some of the states where JRI was enacted,

there was first significant policy changes that weakened the projected return on investment expectations.

I would note that the subject of interest is not JRI per se, but rather Pew's approach to providing advice and influence on this policy topic – an approach that differs from the Results First model. Justice reinvestment is a useful policy case study because it is similar to gun control and bail reform, issues explored at much greater detail in earlier sections. Thus, I explore the following in this chapter: What do justice reinvestment efforts look like in practice? How do they align to the ideals of evidence-based policymaking? What lessons might be drawn from the differences between the ideal and reality? The intent of these questions (and this chapter) is not to expand the scope of inquiry of the overall research project, but to provide additional context to the information presented in the earlier chapters.

In a perfect world, the questions explored in this would have been fully triangulated with chapters three and four with detailed histories of criminal justice policies in the six study states. This *might* have better revealed the direct influence of Results First: not through general knowledge of its efforts and not even through top legislative leaders, but rather through committee chairs and issue specialists. These relational connections make sense to further explore, as chapter four identified that strong personal relationships are the most important element to influencing policy outcomes, affirming literature such as Oliver et al. (2013) discussed in chapter two. With insufficient time to accomplish this, the JRI / Results First comparison presented in this chapter is an attractive alternative.



## 5.1 Justice reinvestment defined

Optimally, JRIs are data-driven reforms that involve the following strategies: developing policies that re-allocate existing resources to generate future savings without compromising safety; re-investment of a portion of criminal justice savings into programs that further reduce offender recidivism and prevent crime; and measuring the fiscal and criminal justice impacts of these reforms (Lawrence 2017). Other definitions include:

Justice reinvestment is a data-driven approach to improve public safety, reduce corrections and related criminal justice spending, and reinvest savings in strategies that can decrease crime and strengthen neighborhoods. The purpose of justice reinvestment is to manage and allocate criminal justice populations more cost-effectively, generating savings that can be reinvested in evidence-based strategies that increase public safety while holding offenders accountable. State and localities engaging in justice reinvestment collect and analyze data on drivers of criminal justice populations and costs, identify and implement changes to increase efficiencies, and measure both the fiscal and public safety impacts of those changes (United States Bureau of Justice Assistance, 2016).

Justice reinvestment is taking money from the correctional system and using it to fund community programs that will end the prison pipeline (Ella Baker Center for Human Rights, 2014).

For many, justice reinvestment is truly about reinvestment of funds saved through policy change. A lesser form of JRI focuses only on diversions from incarceration and shorter jail sentences to save on prison construction and operating costs. In many cases, “once savings were realized, legislators either wanted to divert the recouped funds to shore up sagging budgets in other areas or pay for more correctional infrastructure” (Ella Baker Center for Human Rights, 2014).

As a component of criminal justice reform, it not only aligns with a core competency of Results First but it has also been a top focus area for state legislators across the country. Pew has become heavily involved in this policy space – choosing to engage in this topic through direct, one-time relationships with states. They have

developed an alternate delivery mechanism created specifically for this issue area, focused on either Pew alone offering technical assistance to states or providing support in partnership with the Council of State Government's (CSG) Justice Center and the United States Department of Justice's Bureau of Justice Assistance. The partnership has served nearly 20 states on justice reinvestment issues (Rhode Island 2015).

That is a departure from the Results First model because, while reinvestment work may continue, the support to states is *not* an ongoing process and it is focused on a singular issue. It is also different in that Pew and their partners provide resources and analysis to the state rather than training state staff to conduct the work themselves. However, the underlying principles of benefit-cost analysis and evidence-based policymaking (at the core of the Results First program) *is* intended to shape the decision-making in this space. While not directly measured, this could be an example of providing evidence to states that are more relationship-based.

Heavy involvement and investment from Pew and others is one reason why this topic area makes for a strong case study. Another reason is that, while these cases are related to the Results First project, they remain different. Finally, these cases make sense to explore considering the extent to which the issue has become prominent on the policy agenda across the country. The criminal justice issue area – and justice reinvestment in particular – has become fertile ground for those positioning evidence as a driver of policy. Many have argued that it has undergone Weiss (1980)-style political enlightenment: where the influence upon legislation comes not from research and data, but through an ongoing and slow-moving medium of ideas. She describes the journey as

one of percolation, whereby research slowly makes its way through filters and into the minds and imaginations of policy-makers.

A strong and recent affirmation of the ways in which enlightenment has happened in the criminal justice space is outlined in *Prison Break: Why Conservatives Turned Against Mass Incarceration*. In that literature, Deagan and Teles (2016) outline example upon example of leading conservative figures, from Newt Gingrich to Grover Norquist to Rick Perry, who along with their “successors, are now lining up to challenge the value of incarceration and express sympathy for those behind bars” (p. xi). The change is a stunning change of position from what once was a key conservative doctrine. The book goes on to outline the ways in which a reduction in public anxiety over crime allowed reform-minded conservatives and Republicans the ability to slowly change the minds of fellow conservatives and Republicans (Deagan and Teles 2016).

As the book outlines, change in this policy space was neither swift nor easy. Quite the contrary, *Prison Break* argues that changing minds on policy is a difficult project even when the policy environment is or should be amenable to change – it has taken both a long time (nearly two decades) and a combination of resources, strategy and organization to shift conservative positions on issues of crime (Deagan and Teles 2016). Specific Reasons cited in *Prison Break* for movement on justice reinvestment issues include a huge spike in prison costs, high recidivism rates and massive racial disparities (Deagan and Teles 2016).

In addition to the strong, available scientific information that has been calling for changes within criminal justice, Deagan and Teles (2016) argue that the emotional reactions of individuals needed to be satisfied (and seen as non-threatening to individual

identity, such that other like-minded individuals feel the same way), an objective typically accomplished through a process called “identify vouching.”<sup>45</sup> This identify vouching began in smaller circles, expanding over time, and was empowered through strong personal relationships (see more on relationships in the conclusion). In addition to identity vouching, change was encouraged along by a more receptive public (with lower anxiety about crime) and the extent to which criminal justice reform efforts within states have diffused to others. Such diffusion continues today with the topic of JRI, where dozens of states have adopted the reforms and others have either failed to pursue reinvestment strategies or failed to adopt changes that have been proposed.

While liberal legislators and organizations have been more naturally oriented towards policy changes associated with reinvestment (such as the elimination of mandatory minimum sentences and reductions in penalties for crimes related to drugs), even some of the most conservative organizations have come to support these changes, too. For example, the American Legislative Exchange Council (ALEC) has been one of several think tanks that has helped turn similarly-minded legislators onto the idea of reforming the way criminal justice is conducted. In 2010, it adopted a resolution in supporting justice reinvestment, calling the status of many state prison systems “untenable” (ALEC 2017). One reason for the focus on reinvestment by ALEC is the incredible amount of resources in play, with corrections accounting for more growth than almost all other budget line-items over the last two decades, accounting for over \$50 billion in spending nationally (2017). Conservative and liberal organizations supporting justice reinvestment both speak to the ways in which resources saved and reinvested help

---

<sup>45</sup> Identity vouching typically is conducted by individuals of high esteem from the viewpoint of a group or culture. These opinions are especially credible when they are not reinforcing one’s own self-interest.

lower crime and keep our communities safe, but it also affirms the extent to which the data supporting these claims have been allowed to become mainstream across the political spectrum.

It is worth noting that, while groups like ALEC and Pew are pushing for similar changes in the JRI policy space, the politics of the two actors are typically very different. As discussed at length, Pew's Results First / WSIPP / benefit-cost model focuses on supporting the policy changes that yield the strongest return on investment. ALEC, on the other hand, operates with a process that produces "model legislation" generated by task forces, typically staffed with a mix of corporate representatives and lawmakers and each with equal voting rights as to which pieces of legislation to advance (Right Wing Watch 2011). In such an environment, "what works" is less likely to be the driving force it is under the Pew model. In fact, ALEC has been accused of pampering "lawmakers while providing them the opportunity to collaborate on legislation often previously researched and introduced by the policy shops of its corporate members" (Zwick 2010).

The Pew foundation, on the other hand, identifies policy topics which align with their mission and have gotten onto the agenda. Pew then spends heavily to ensure that these topics make it over the goal line. Pew tends to support more progressive policy initiatives while ALEC typically produces conservative draft legislation.

The Pew approach is described well by David Kirp in *The Sandbox Investment*, a book that describes the movement surrounding the pre-school movement and "kids first politics." It is an approach far less wonky than the Results First movement modeled on the WSIPP benefit-cost approach. In that book, Rebecca Rimel, the President and Chief Executive Officer of the Pew Charitable Trusts, is quoted saying "We (Pew) pick the

issues that are ripe, where the facts are clear and where we can bring change in a short time” (qtd. in Kirp 2007, p. 157). In practice, Pew is advocating for what they see as public goods, with a model of implementing change that is “unapologetically corporate” (Kirp 2007, p. 157). These methods include the full deployment of modern political tools: polling, a messaging campaign, grassroots organizing, focus groups, and the identification and cultivation of key, influential supporters on the subject. In football, a team on the offensive side of the ball is considered to have a high probability of scoring when they enter the red zone, because it is a short distance to score a touchdown and the team is also within range of a field goal for most kickers. The approach employed by Pew might be thought of as a “red zone offense,” where they enter the policy game in the places where they are most likely to put points on the scoreboard. Pew is careful in the topics they choose to engage with, but once they enter a space, they play to win – and the odds of success are always greater when there is a shorter distance of the field to travel.

Of course, Pew and Results First could argue that the progress seen over the decades in places like Washington and the WSIPP have also made it easier for organizations like ALEC to make the case for justice reinvestment among those with whom they work. Notwithstanding the different politics of Pew and ALEC, the convergence on the issue has certainly been helpful in allowing Pew to push for policy change in JRI. Indeed, the criminal justice space in which Results First is realizing its growth and success has been, in part, associated with the ways in which the landscape has been cleared by the work of others. In other words, having groups like ALEC help bring conservative lawmakers around to the ideas and changes associated with justice reinvestment makes it easier for Pew and Results First to propose interventions within

criminal justice policy, and especially within the more focused issue area of justice reinvestment.

One true test of the staying power of the Results First model will be the extent to which it can successfully replicate in other issues and across states. Can Pew and Results First replicate its success in policy spaces across states when there has not been widespread policy enlightenment across the partisan divide? Or has the criminal justice space become an exception rather than the rule for evidence-informed policymaking – whether for Results First or other organizations interested in having evidence at the center of policy debates? These questions are not answered in this chapter, but do merit further investigation.

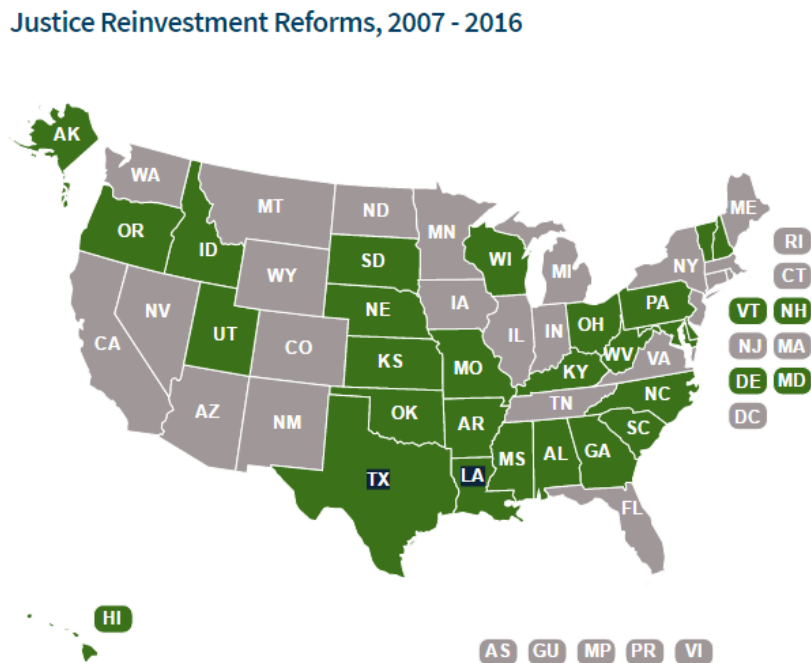
## **5.2 Justice reinvestment overview: states discussed in this study**

Of all the states included in this study, only two – West Virginia and Pennsylvania – successfully enacted new legislative reforms on the core criminal justice issue of justice reinvestment, according to Alison Lawrence (2017) with the National Conference of State Legislatures. In that same review, she shows that just over half – 27 out of 50 – of the states embarked on these types of efforts between 2007 and 2016. Even more striking: none of the three Results First states from this study were included in that number.

Legislative enactment of policy reforms is a common-sense and appropriate standard lens through which success can be measured in this policy arena. That being said, I also considered the experiences of states where there were unsuccessful efforts to

enact justice reinvestment policies<sup>46</sup>. Whether or not the study states adopted JRI, I also found that, with the exceptions of New York (which has yet to undertake these reforms) and Virginia (which made similar changes before the window of Lawrence’s study), Pew was engaged in the effort – either alone or through their partnership with the CSG. In the graphic below, “successful” efforts to enact justice reinvestment reforms over the last decade (between 2007 and 2016) are identified by those states shaded in green:

*Figure 10: An overview of JIR reforms between 2007 and 2016*



Source: Lawrence (2017)

With this backdrop in mind, I conduct a brief survey of activity / case study in each state around the subject of justice reinvestment, utilizing both primary (such as legislative

<sup>46</sup> This research includes and explores efforts that have helped make the issue high on a legislative agenda, even if they have not technically been “successful” from the perspective of passing legislation.



records and agency documents) and secondary (news reports and interviews) sources. The review but includes some information on every state. To complete this work, there were two new, confidential informant interviews completed (a committee staff member in Pennsylvania, staff member to the Maryland Governor) as well as a non-confidential interview with Elizabeth Lyon from the Council of State Governments. These interviews were conducted between March and September 2017. In addition to these three new interviews, content from a prior, confidential interview with a West Virginia legislator and well as non-confidential content from both a West Virginia and New York staff member were utilized in the completion of this chapter. Finally, quotes from the Speaker of the House and the Senate President in Rhode Island as well as a state Senator from Massachusetts were included as content to be analyzed in this case study.

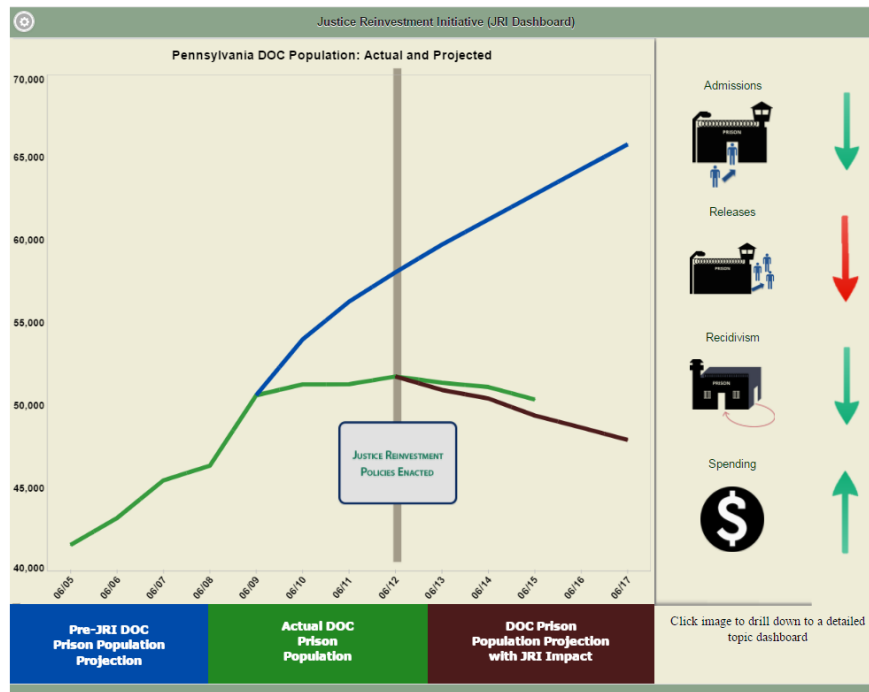
#### 5.2.1 Pennsylvania: action taken

Pennsylvania, a non-Results First state, successfully enacted justice reinvestment reforms beginning in 2012, with additional legislation enacted in 2012, 2015 and 2016 on the topic (Lawrence 2017). The first change helped reduce inefficiencies in the state's parole and corrections systems, changed the response to parole violations and transformed approaches to community corrections programs (CSG for PA 2017). These legislative changes reflected much of the core values of a justice reinvestment approach – and were enacted without having outside organizations actively engage in the policy transformation process.

Following these changes, the state realized a decrease in the prison population as well as cost savings. Pennsylvania has even adopted a state-specific dashboard where citizens can track the projected versus actual number on the prison populations and see

detailed information on admissions, releases, recidivism and state spending and reinvestment of funds.

*Figure 11: PA interactive dashboard for justice reinvestment performance measurement*



Source: PA Justice System Joint Performance Measures (2017)

As noted in the state’s JRI Dashboard, there is a considerable gap between the projected population in the Department of Corrections and the actual number – nearly 15,000 fewer individuals are in the corrections system as of June 2015 than were originally projected in 2009. A closer look at the chart, however, suggests that the impacts of reinvestment may have been over-estimated. Rather than the large spike in a corrections population beginning in 2009 as projected, the actual number sharply flattened out. To be sure, the population trend line is moving (for the first time in over a decade) in the negative direction, but the slope is not steep and the projected JRI

reductions have not been realized<sup>47</sup>. The state did report, however, over \$11 million in savings in 2015 alone (PA Justice System Joint Performance Measures 2017).

More recently, the state has re-engaged on the issue. Governor Wolf and legislative leaders invited the Justice Center from the Council of State Governments (and Pew) to conduct “a comprehensive analysis of extensive data sets collected from various relevant state agencies and branches of government to build a broad picture of statewide and local criminal justice trends in the state” (CSG for PA 2017). The working group released findings to the Governor and General Assembly for their consideration and legislative action in 2017 and beyond. The most recent ideas represent more than an estimated \$100 million of savings for future years, in partly realized by reducing the sentences of some non-violent offenders by five months (Esack 2017).

One Pennsylvania staffer close to the work shared that having extensive analysis affirming the outputs of these changes was important to the JRI effort. They also cast doubt that evidence was the driving force. Instead, they suggest, there were legislative leaders “who were clearly in charge and wanting to move an agenda forward” but who also “very wisely leveraged the available national resources to tell the story they wanted to anyone else connected to the process” (PA staff member 1, 2017). In many respects, the approach resembles the “Trustee who knows” described in Figure 5 of Chapter two – lawmakers willing to push evidence-supported legislation.

The work could have absolutely gotten done without all of the additional analysis that was brought in because there was enough buy-in to the process from top administration officials as well as legislative leadership. Still, having gone through the process was both an opportunity for the public – and skeptic legislators – to see that there were good reasons behind the push for these reforms (PA staff member 1, 2017).

---

<sup>47</sup> The June 2016 and June 2017 updates have also not been uploaded into the system.

In Pennsylvania, the push of the Governor and some key legislative leaders, paired with national resources, helped enact meaningful justice reinvestment reforms. Much like the Results First model where the organization only begins work upon the request of top leadership, Pennsylvania's push to justice reinvestment reform began with a similar request from nearly every major player on the issue: then-Governor Corbett, Chief Justice Castille and the chairs of the House and Senate Judiciary Committees of the Pennsylvania legislature (CSG for PA 2012). While it may not have been necessary to ultimately enact the changes, it set the stage for a larger conversation among lawmakers and ultimately the public as well – complete with a public-facing dashboard which tracked the progress on several key metrics following legislative passage. Pennsylvania enacted justice reinvestment reforms with support from the Council of State Governments and the Pew Charitable Trusts even though they are not a Results First state.

#### 5.2.2 West Virginia: justice reinvestment action taken

West Virginia has a similar story, with very limited Results First exposure but considerable activity in the justice reinvestment domain – again with the support of Pew and CSG. West Virginia has acted on justice reinvestment reforms, using Pew and the Council of State Governments as resources to staff their bipartisan efforts. These efforts began in 2012, without a Results First engagement, when the state had an adult prison population that grew by 50 percent in a decade and projections for another 24 percent growth by 2018 (Pew April 2015). In 2013, the state adopted reforms that required supervision of probationers and parolees based on risk and needs assessments, authorized the use of graduated sanctions for technical violations of supervision, and prioritized treatment and reporting resources for offenders assessed as highest-risk (Lawrence 2017).

The impacts from the 2012 change were experienced almost immediately, with an approximate five percent reduction in prison population in 2013 – the first decrease that had occurred in 16 years (Marra 2014).

The changes required not only legislative but also administrative alterations as well. In addition to the changes in law enacted in West Virginia, the state’s Supreme Court also “acted immediately” to mandate that “all convicted felons undergo a risk/needs assessment, and that the results of the assessments be provided to judges before sentencing” (Kabler 2013). As Kevin Baker (WV Staff, 2017) shared, this collaboration was important, as “having the buy-in from the West Virginia Supreme Court and other local judges was very helpful.”

If the buy-in was important, the causes of it are perhaps even more important. Baker shared further about the ways in which the overall presentations to the legislature helped drive reforms and were key in getting and keeping the interested parties on board.

I don’t remember a specific document or organization, but there were presentations from the folks with the justice reinvestment group that showed how our prison population has grown over time and what they would mean for the future. The presentations included comparisons of what could occur if the state adopted the justice reinvestment package and moved forward with it. I believe those presentations were critical in showing the members what the problem was and a potential solution (Kevin Baker, WV staff 2017).

If the testimony from experts helped get the attention of legislators, a “fellow legislator” may have been what helped to tip the scales.

I think the most impactful testimony was when a legislator from another state (maybe the Speaker of the House in Georgia?) came to West Virginia and said *this works*. He talked from experience about how getting the bills passed was challenging in his state, but that they have seen the impacts and that it is all positive. I think the legislators here in West Virginia were much more open to the idea when they heard from someone who had a similar approach work (Kevin Baker, WV staff 2017).

Hearing from other elected officials who shared my apprehension was more helpful than anything else in the process. Knowing that it worked elsewhere and had positive outcomes was important. It was also important to know I could still get re-elected and I wouldn't be standing out in the crowd (WV Legislator 2, 2017).

More recently, West Virginia passed comprehensive juvenile justice reform in 2015. These changes are expected to avoid at least \$20 million in costs over five years and lower the number of youth in residential placements by at least 16 percent (Lawrence 2017). Like the reforms passed in earlier years, the recommendations were based on a bipartisan Task Force (this one on Juvenile Justice), which again received technical assistance from Pew. Like Pennsylvania, West Virginia is a non Results-First state that has enacted, with the help of Pew and/or organizations like it, significant changes to the way in which that state addresses justice reinvestment. What about the status of the third Results First state in this study, Virginia? In that case, there have not been any significant efforts to reform policy within the Commonwealth, but there is evidence that changes enacted in the past might align with the “new” work being conducted elsewhere on the topic.

#### 5.2.3 Virginia: no recent activity – but plenty in the past

Virginia was not listed in the Lawrence survey as a state taking action in this area. Additional investigations into the state as it relates to justice reinvestment confirmed that there has not been any major reinvestment legislation introduced, considered and debated. However, the state can claim credit for enacting reforms in the mid-1990's that continue to impact criminal justice spending and prison populations. Front and center among these were new sentencing guidelines (abolishing parole and adopting felony

sentencing guidelines) that began in 1995. Virginia also became the first state, in 2002, to adopt risk assessment guidelines for circuit courts (Hickman 2014).

The results might partially explain why there has not been much additional legislative activity. In 1984, the Department of Corrections in Virginia accounted for 7.4 percent of the state's general fund appropriation; in 2014, it accounted for two percent less – 5.4 percent. While the total state general fund spending has increased 434 percent in those twenty years, spending on the Department of Corrections only increased 288 percent (Hickman 2014). In the years following, the prison population and prison spending trends within the Commonwealth have held relatively constant (Vera 2017). Elizabeth Lyon (2017) affirmed that there has been no relationship between Virginia and CSG adding that she was “unaware” of any legislation related to justice reinvestment passed in the Commonwealth.

If there has been recent (and in the case of Virginia, not so recent) action taken on the topic of justice reinvestment within non-Results First states, we would expect the same to be true within Results First states. Beginning with New York, I find that “action” has been taken in one form or another – usually through the introduction of legislation. However, I also investigate the ways in which these states, despite having an on-the-ground presence from Pew and the Results First program, have failed to successfully enact the legislative changes required to implement justice reinvestment.

#### 5.2.4 New York: action not taken

New York is a Results First state that has not completed any sort of justice reinvestment in the years covered by the NSCL review, nor was there any evidence of similar efforts completed in the years before. There has, however, been legislation

introduced for several years in a row that attempts to have the state begin this work. The legislation was first proposed back in 2009, but has yet to be enacted into law. Based on a review, it does not appear that there are any concentrated campaigns or organizations seeking to push the changes set forth in the proposed legislation. There also has been no significant legislative analysis conducted on the impact of the legislation, either by state legislative staff or an outside organization such as Pew or CSG.

That stands in sharp contrast with the statements of Marc Schabes, New York's primary contact for the Results First program, who claimed that "New York state plans to fully integrate Results First as a normal part of doing business" and that it will become "routine" to conduct exhaustive benefit-cost analysis in all their programming decisions (quoted in Pew, June 2015). Affirming this contrast, Elizabeth Lyon from CSG shared, from their perspective, "New York is not a justice reinvestment state" (2017). This is a laudable goal, but the failure of the legislature to embark upon justice reinvestment reforms could be viewed as a mark against this goal and, to some extent, the Results First program itself. If Results First was brought in to conduct work within a state and their expertise is especially strong in the criminal justice domain, would justice reinvestment not be a natural area of expansion?

That depends. While Results First is limited to adult criminal justice currently, the primary customer from the agency's perspective is the Governor's budget and state division of budget<sup>48</sup> -- not the legislature. "We had a unique timing experience," shares Schabes (2017), who coordinates the Results First effort as the manager of statistical

---

<sup>48</sup> In New York, the SAC will produce information for legislators as requested, but Marc Schabes (2017) reported that these requests are limited in nature.



services in New York's Statistical Analysis Center (SAC), a research agency of about 50 people.

Specifically, Schabes was referring to the pending expiration of the American Recovery and Reinvestment Act federal funding. The Results First benefit-cost model was solicited by the Governors' office as a way to understand the best way to respond to the loss of funding in the alternatives to incarceration space, so it deployed the model to show the state's division of budget and the legislature that picking up the tab with state money and avoiding a shutdown of diversionary programs would be cost-beneficial in the long run. The cost was \$3 million, and as Schabes (2017) shared,

Not only did we save the funding, we were actually given additional state funds over and above what were federal funding dollars to expand these alternatives to incarceration. It was because of this commitment to evidence-based programming that we could model and prove the long-run cost savings.

In other words, when policy decisions are be driven by evidence, success can be identified. New York is looking to expand their efforts into juvenile justice, building upon the criminal justice work commenced in 2012. However, given the limited translation the program has seen into legislative activity, one might question the capacity to realize that desired growth, especially as it relates to having legislative buy-in.

We try to provide modeling data to answer the 'what if we' questions to legislators... what if we expand or shut down a prison? What is we put more or fewer individuals on parole or probation? But all we can do is put the data out there. We can't force our legislators to use it (Schabes 2017).

Further suggesting challenges for expansion, Schabes (2017) added that "having the program housed in the SAC has been a limiting factor for expansion." Still, as discussed in chapter one, Schabes and his colleagues are doing something right: New York was listed as "leading the way" among states when it comes to cost-benefit analysis in the

most recent rankings released by Pew. In this case, it appears that impacts resulting from Pew generally and Results First specifically are more focused on the state's executive branch, with legislators having minimal exposure to and use of evidence that could help inform their policymaking efforts.

#### 5.2.5 Rhode Island: attempts at progress stymied

For its part, Rhode Island has been actively working on justice reinvestment efforts, but has struggled to get a final product across the finish line. Rhode Island Governor Gina Raimondo signed an Executive Order creating a Justice Reinvestment Working Group, who (like other states) worked in collaboration with the Council on State Government's Justice Center to determine data-driven analysis and policy options for state leaders (Rhode Island 2015). The state also passed legislation in 2008 that was not recognized in the Pew graphic created by Lawrence as sufficient to be considered justice reinvestment reform. Per the Rhode Island government (2015), passage of House Bill 7204 of 2008 enhanced parole discretion and standardized credits for time served, helping lead to a lower correctional population and recidivism rate.

Following the work of the Justice Reinvestment Working Group supported by CSG, 2016 saw the introduction of six bills aimed at enacting the proposal based on the recommendations of the Working Group established by Governor Raimondo. These bills were intended to, among other things, modernize sentencing and probation policies, focus supervision on high-risk, high need populations, and assess defendants to drive diversion opportunities (CSG for RI 2017). This legislation passed the State Senate unanimously but stalled and ultimately failed in the House – legislation that was again considered and

failed this most recent year – passing the Senate again but falling also falling short in the House (CSG for RI 2017).

Despite the high-profile efforts of Raimondo’s 27-member Justice Reinvestment Workgroup, and the support from CSG, justice reinvestment efforts remained stalled in Rhode Island. In the process, the typical political finger-pointing has taken place. House Speaker Nichols Mattiello, for example, has described the legislation as “too soft on crime” and that he was “never convinced Rhode Island should be a test case for a national model on criminal justice.” (quoted in Pina 2016). While the Speaker’s language conveyed that Rhode Island would be a “test case,” the reality is that many of the proposals included in the package have been successfully enacted in many of the others states, especially those that have partnered with Pew and CSG.

Speaker Mattiello’s hesitation to move the justice reinvestment legislation could be more accurately tied to an aide close to him, who confidentially shared that it was intra-branch tension that prevented movement on justice reinvestment. As the official shared, “I think we wanted some bills over there [the Senate], and they wanted some here [the House]. That’s how it works” (quoted in Pina 2016). Senate President Pavia Weed strongly condemned the actions of the Speaker. “They [the justice reinvestment bills] were a priority for the Senate. They were a priority for the governor. They were a priority for many members of the House, and several others. And they were a collaborative effort” (quoted in Pina 2006).

An additional Rhode Island legislator who was familiar with the situation further confirmed the role of political considerations in the breakdown of progress.

Of course, this was about more than just the Speaker having some concerns with the legislation. I mean, he’s the Speaker. If there was something in the bill he

really wanted to change, it's not a very hard thing for him to do. This was about the Speaker not getting his way on other legislation being considered in the Senate. It was just politics. (Rhode Island House Member 3, 2017).

While politics helped cause the breakdown of legislative progress, it was clear that there was no shortage of evidence in support of the legislative changes.

CSG came in and did a great job working with the Justice Reinvestment Group. There was pretty widespread support for the changes. I have to say that I was personally a little surprised when the Speaker decided he was going to go against these recommendations and not move the bill forward (Rhode Island House Member 3, 2017).

Rhode Island is a classic example of political considerations thwarting progress on legislation built upon evidence. Despite the evidence, the Speaker of the House's reservations about justice reinvestment efforts being "soft on crime" and other political considerations have thwarted legislative changes. This is also despite the strong support of the governor, Senate President and considerable work being conducted by Rhode Island and outside staff on the issue. Not far away, the justice reinvestment story is similar in neighboring Massachusetts. As in Rhode Island, there have now been several years of legislative consideration on the subject along with outside support from CSG and Pew, but there has not been formal legislative adoption of the proposed changes.

#### 5.2.6 Massachusetts: progress also thwarted

Finally, in Massachusetts, lawmakers began considering justice reinvestment options in 2015. Of all the states included in this dissertation, Massachusetts has the strongest starting point: the state boasts the second-lowest incarceration rate in the nation. Despite that strong statistic, leaders identified additional room for improvement in recidivism, as more than half of those leaving the Department of Correction facilities and two-thirds of people leaving the House of Corrections in 2011 were re-arraigned within

three years of their release (MA Governor website, 2017). The results here mirror the frustrations felt in Rhode Island: despite considerable effort being put into the justice reinvestment effort and the buy-in of important leaders, legislative passage remains elusive.

The Council for State Governments, along with Pew, provided intensive technical assistance for an interbranch steering committee comprised of lawmakers and officials from both sides of the political isle (CSG for MA 2017). This work began in August of 2015 and came at the request of Governor Charlie Baker, Senate President Stan Rosenberg, House Speaker Robert DeLeo and Supreme Court Chief Justice Ralph Gants. There were four areas of focus for the CSG working group. Policy options were developed around better aligning probation and parole supervision with best practices, improving access to treatment, making the parole process more effective and reducing the DOC population and increasing the number of individuals receiving post-release supervision (CSG for MA, 2017). After two years of failed efforts to change the law, legislation representing a consensus of this steering committee was again introduced in February of 2017 following a final report on justice reinvestment policy options, but has, to date, failed to become law (CSG for MA 2017).

As discussed above, legislative enactment has remained elusive in Massachusetts, despite the backing of all the major institutional leadership. In addition to the challenges of policy change, other legislators charge that the work of the criminal justice working group – spearheaded by CSG and supported by Pew – does not go far enough. Boston state Senator Sonia Chang-Diaz had harsh words to describe the effort:

In my eight years as a state senator, I've never actually been made nauseous by the handiwork of government – until yesterday (Chang-Diaz, quoted in Schoenberg 2016).

Yesterday I watched a room full of black and Latino demonstrators, who have been patient for the past two years, plead with an all-white panel of CSG working group members to say something or ask some questions about the devastating effects the criminal justice system has on their communities. The three-hour meeting continued with polite technical question-asking, none of which had to do with the cries for help from communities most impacted by crime (Chang-Diaz, quoted in Schoenberg 2016).

Our justice system is deeply broken, expensive and racist, and for the last two years legislators and grassroots activists alike have been baited into believing this 'Justice Reinvestment Initiative' was going to make a serious effort at fixing that. Yesterday we were placidly told, 'that was never part of our charge.' (Chang-Diaz, quoted in Schoenberg 2016).

Chang-Diaz's complaints were strong and even CSG staff could not say that they were without merit. In fact, CSG Senior Policy Advisor Steve Allen responded by saying "Justice reform is never as comprehensive as folks would like it to be. There's only so much any project can do at a time" (quoted in Schenberg 2016). To that end, Massachusetts follows in the Rhode Island example of being a Results First state that has conducted considerable evidence-based reviews on the subject and introduced legislation but has failed to change state policy. It is also a state in which, for at least one key lawmaker and activists, the changes being proposed remain insufficient to address the larger challenges. Despite the failure to pass legislation, Elizabeth Lyon with CSG remains optimistic for Massachusetts' prospects. As she shared, "they have a great example of legislators committed to evidence-based lawmaking. Every time we presented, it felt like we were preparing for a final exam. We are still looking to pass legislation, but no other group has been more focused on the evidence in my experience." (Lyon 2017).

### **5.3 Maryland: “A step in the right direction” with Pew**

In Maryland, a non-Results First state, justice reinvestment changes *were* enacted with the help of Pew. However, the final product also had advocates claiming that more work needed to be done. I explore that example next, where legislative leaders invited Pew to come and help work with them on JRI reform. They also have enacted some of the changes proposed from the review effort. Considering the extent to which staff and legislator feedback from Maryland meaningfully informed the data of chapter four and given the relatively easy access to appropriate interview subjects, I felt it was appropriate to include Maryland in this chapter as well.

Prior to embarking on policy changes, Maryland first established a Justice Reinvestment Coordinating Council, which was charged to review the state’s criminal justice structure and identify recommendations that would lower Maryland’s prison population, reduce correctional costs and reinvest savings into programs providing the most efficient outcomes in the criminal justice space. The Chair of the Council was the Director of the Governor’s Office of Crime Control and Prevention, with membership and subcommittee leadership from the Chairman of the Senate Judicial Proceedings Committee and Vice Chair of the House Judiciary Committee.

Following a review of the issue by the Justice Reinvestment Coordinating Council, with the assistance of Pew staff, legislation was considered during the 2016 session of the Maryland legislature. As first introduced, the legislation was projected to save \$247 million over the span of 10 years with an associated decrease in the prison of 14 percent (Mehu 2016). Walking into the legislative session, there appeared to be considerable consensus on the changes that would be likely enacted based on the work of

the Council. As one Maryland House member (4, 2017) familiar with the process shared, “We assumed before session started that there would be very few if any changes to the framework agreed upon and forwarded by the Reinvestment Council.”

The legislation initially progressed with little sign of trouble. However, toward the end of session, progress stalled and it became clear that whatever product emerged from the Conference Committee between the House and Senate would be the final word on the legislation.

There were a few key legislators that were not entirely on-board with the changes in the law but who did not want to be seen as totally against it. So, I think they slow walked the legislation to the end so that they could be sure to make the changes they wanted to see with no real time to un-do those changes (Maryland House member 4, 2017).

Consequently, a re-scored fiscal note estimated that the state would only realize a fraction of the savings initially projected, \$80 million, for reinvestment purposes. Moreover, some advocates were troubled about other changes made to the bill that were never part of the discussion in the first place – such as an *increase* in the penalty for second-degree murder (Dresser 2016). There was clearly disappointment with the final product. As a member of the Governor’s staff shared,

yeah, it sounded great. And the bill was on its way to really doing something. But then, of course, they go ahead and ‘Christmas tree’ it with all kinds of changes at the last hour. So, we claim victory on what we can and move on (Maryland staff member to Governor 2017).

These last-minute changes had nothing to do with whether Maryland was a Results First state. It also had nothing to do with Pew’s involvement with this issue. It did, however, have everything to do with the district, personal and political considerations outlined in earlier chapters.



After all that work, all it took was a Chairman and a few willing lawmakers on the Conference Committee to entirely undo what had been months of work on the part of others. I wonder if that was the plan all along (Maryland House member 4, 2017).

As frustrating as it was, we supported the bill and we show up for the bill signing. It could have been, should have been so much better, but sometimes you just need to take what you can get, assuming the good is better than the bad. In this case, I think that was true (Maryland House member 4, 2017).

The frustration was not limited to those most closely involved with the legislative activity. The Justice Policy Institute (2016) weighed in with the following statement following the passage of Maryland's Justice Reinvestment Act:

The final JRA bill is a far cry from the initial vision of the Justice Reinvestment Coordinating Council (JRCC) process that has been advancing for the past year...

JPI acknowledges that, even with a series of harmful Senate amendments to the original JRCC package, the enacted JRA is a step in the right direction for Maryland: Emphasizing a treatment approach to people incarcerated for low level drug offenses, mitigating the role technical violations play in people under community supervision, repealing some mandatory minimums and improvements in the parole process are steps in the right direction for Maryland.

To the degree that the changes here help modestly reduce the use of incarceration, it represents base to build from in future justice reform efforts. Next session, legislators and the executive should promote changes in law, policies and practice to more meaningfully reduce the use of incarceration in Maryland (Justice Policy Institute 2016).

Marc Schindler, the executive director of the Justice Policy Institute, said legislators "diluted the already thoroughly debated recommendations of the [Justice Reinvestment] Council, producing a bill that was a compromise of a compromise" (Dresser 2016).

While politics did not derail the justice reinvestment project as it did in Rhode Island, political considerations were the reason that significant, unplanned changes were made to the state's legislation in this space. Maryland's experience, like the experience of the other states studied, serves as an important reminder of the other influences impacting

legislative decision making discussed in earlier chapters. Even well-intended legislation that was grounded in evidence and the consensus product of a large and diverse working group was substantially changed before final enactment. These changes moved the law even further away from an evidence-based approach to policymaking. In Maryland and in all states, the many tensions facing legislators identified in the literature of chapters one and two were validated by self-reports in chapters three and four and shown to be true in practice here in chapter five.

#### **5.4 Conclusion: a focused intervention in Pew's wheelhouse with limited results**

Results First and Pew have shown considerable promise for influencing the areas in which they have spent considerable time growing a model, getting legislative and/or executive branch buy-in and demonstrating results. As discussed in this chapter, Pew is exceptionally involved in the justice reinvestment policy area across the country, but this involvement typically does not come directly through Results First. Pew instead, either directly or with their partners at CSG, has been or is actively involved in dozens of states supporting policy research with an on-the-ground investment. The level of effort and investment on the part of Pew and CSG are different in that *they too* are committing resources into the investigation and reporting process. This is a key difference from the Results First approach, where states are expected to use their own resources to conduct most of the analytical legwork. This difference could be meaningful in terms of the results that states experience.

Lawrence's (2016) review of the activity across the country offers a solid jumping off point from which to do an overall analysis of progress on the issue of justice reinvestment. Table 28 below adds additional detail to that graphic, including not only

whether there was bill passage on the topic, but also indications about whether bills were introduced on the subject (“bill hearings”), if there was Results First / Pew input on the subject and whether legislative staff within each state provided analysis on the legislation.

*Table 28: justice reinvestment activity and outcomes by state, 2007 – 2016*

<b>Justice Reinvestment Activity</b>	<b>MD</b>	<b>PA</b>	<b>NY</b>	<b>RI</b>	<b>MA</b>	<b>WV</b>	<b>VA</b>
Bill Hearings	Yes	Yes	Yes	Yes	Yes	Yes	No
Bill Passage	Yes	Yes	No	No	No	Yes	No
RF / Pew Input	Yes	Yes	Yes	Yes	Yes	Yes	No
Legislative Staff Analysis	Yes	Yes	Yes	Yes	Yes	Yes	No
Results First State	No	No	Yes	Yes	Yes	No	No

As this case study suggests and the above chart helps to confirm, Results First / Pew involvement has not been *required* to make progress on the topic of justice reinvestment. That being said, Pew has strongly inserted themselves into the state-level criminal justice policymaking discussion, having been found to be associated with just about every effort discussed in this chapter. That may be due to the “red zone offense” approach deployed by Pew, choosing to involve themselves in the policy areas ripest for change. Just as Kirp identified pre-school as an opportunity Pew embarked upon to swiftly and meaningfully change state laws, so it seems they are attempting to swiftly and meaningfully influence criminal justice reforms.

This case study chapter shows the ways in which Pew and their relevant partners are making an impact on evidence-driven justice reinvestment policy at the state legislative level. Pew and CSG have been helping with thought leadership and research in this domain. It makes sense that the Pew expertise being deployed is informed by the experiences of the Results First program – and has been applied in crafting the

recommendations put forward by the study groups and commissions described in this chapter. However, the review here also suggests that Pew's involvement may not always lead to complete or immediate enactment of policy changes. Such a finding on Pew's involvement is potentially significant and negative. However, it should also be taken with qualification: compared to other sections, there was limited research effort on this topic. Moreover, it does appear that there remains strong momentum toward policy enactment in the states where Pew has been involved but progress has, to date, been thwarted.

Still, it is worth noting that most of the legislative success stories in this case study come from those states that *were not* active Results First participants – New York, Rhode Island and Massachusetts. While these states were, in one way or another, pushing for justice reinvestment reforms, they were not the states that experienced passage of new laws to reform their systems. Instead, non-Results First states of Pennsylvania, West Virginia and Maryland were among those achieving JRI legislative successes. Thus, while difficult to make direct comparisons among the methods of evidence production, it can be said that it appears that the specific policy environment in which evidence is being produced (including the relationships on the ground between evidence producers and lawmakers) continues to matter a great deal.

Viewed more skeptically, data from this case study warrants further investigation into how far Pew/Results First might effectively influence other policy areas. While Results First claims the capacity to transform all types of human services, such claims *might* be overstated. If states working with Pew and Results First are experiencing challenges with JRI reforms, an extremely ripe area for policy change, it does not bode well for expansion efforts. As Marc Schabes (2017) from New York shared, “My

understanding is that Results First is only beginning to have robust evidence basing capabilities in these new spaces. If work is done through these new lenses, there may be some catching up work to do.”

From the examples of this chapter, the full promise of evidence-based policymaking was not fully realized. As discussed in the beginning of this chapter, realizing evidence’s full promise would mean legislation largely reflective of maximum returns from a robust benefit-cost analysis would be introduced and passed with relative ease and speed. To be fair, some progress on justice reinvestment has been achieved in most of the states included in this chapter’s review and Pew has had a hand in most of this work. However, the fullness of policy change – promised and driven by evidence-based policymaking ideals – has failed to come to fruition, either at all (as has been the case in Rhode Island and Massachusetts) or through a partial implementation of evidence-based practices (as in the case of Maryland). Such limited progress is occurring even though there is incredible alignment between Democrats and Republicans, liberals and conservatives around this issue. If evidence is prevented from fully prevailing in a space that is incredibly resourced and that benefits from significant policy enlightenment, what can be said of the prospects for policy areas where party and ideological divides remain? Or where resources are more constrained? Or where there are additional hurdles to overcome?

The most important takeaway from this chapter is simple but instructive: *even in what may be the most fertile of policy environments, the success of evidence-based policy* (meaning, the ability for policy to be based on the insights of the evidence) *was limited*. This lesson reinforces information gleaned from survey responses and interview data in

chapters three and four, data that suggested even policymakers most inclined toward leveraging evidence in their lawmaking activities face significant obstacles in doing so.

In the next and final chapter, I review our findings and contemplate the ways in which Results First is and is not helping to create a more evidence-informed culture within state legislatures. I also offer thoughts on ways in which we can move more toward such a culture, with or without Results First leading the charge.

## Chapter 6: lessons learned and cultivating an evidence-informed culture

In chapters three and four, this research has collected and analyzed information about the extent to which Results First has encouraged state legislators to more meaningfully incorporate evidence in their policymaking efforts. Here, I found very small impacts of Results First on evidence utilization. Chapter five considered others ways in which the Pew Foundation and like-minded organizations influence policy with the power of evidence, where I found mixed impacts. In this final chapter, I review findings from the survey, interview and third party data in more detail to discuss relevant insights into the role of evidence in policymaking including the ways in which Results First and similar organizations inform that space.

I conducted this work because if a program like Results First is successful in its mission, one can expect policy outcomes within a state to be more informed by evidence. More evidence-based policy could, in turn, lead to improved outcomes for citizens: in criminal justice (fewer people jailed, less crime), economic development (more jobs and opportunities), and improved educational and health outcomes (more students graduating, fewer people diseased). The list could, theoretically, extend to just about every area of policy. Similarly, evidence-based policymaking holds the promise of a reduction in wasteful spending, focusing expenditures on programs that work and freeing up money to be reallocated towards those programs that would lead to the improved life outcomes discussed above. Moreover, allowing evidence to meaningfully influence policy decision-making makes innovation in programming more likely and can help strengthen accountability mechanisms within governmental systems (Pew, November 2014).

Results First served as the focus of this investigation because it is the only large-scale deployment of a cost-benefit analysis-based effort that seeks to meaningfully drive policymaking with evidence. What started as an inquiry focused on the impact of Results First developed into a more thorough investigation into evidence utilization by state lawmakers – and one that considered work conducted by Pew beyond only the Results First program (in chapter five). The data presented add to the ways in which we know if and how state legislators use empirical evidence – especially as weighed against other priorities in making legislative decisions. The research provides within-state and across-state comparisons regarding not only evidence use but also attitudes toward evidence use. It also confirms existing literature with new examples about the ways in which and the reasons why lawmakers do – and do not – engage with evidence. While different in meaningful ways, the research I have shared provides an important new perspective to the work on legislative evidence utilization as it relates to non-partisan research organizations completed by John Hird over a decade ago.<sup>49</sup>

This exploration into evidence utilization by state legislators started with background on the Results First program, the structure of the research and relevant academic literature in chapters one and two. Quantitative and qualitative data regarding evidence utilization from six states – three that are active Results First states and three that are not – was reviewed in chapters three and four, with chapter four including additional insights from the state of Maryland. This original research was supplemented with additional insights regarding the work being done around the county in the justice reinvestment space by the Pew Charitable Trusts – a subcomponent of *the* key area of

---

<sup>49</sup> In addition to focusing on legislators over NPROs, another major difference between this research and Hird's is that his was a national survey population.



policy (criminal justice) serviced by Pew using the Results First model. Having completed that review, this final chapter returns to address the research questions posed at the beginning of this effort, leading with the fundamental question: is there a connection between Results First and the use of evidence among state lawmakers?

## **6.1 Does Results First increase the use of evidence?**

The question is simple but the answer is meaningful. A strong relationship between Results First and evidence utilization by policymakers and/or a connection between better state policies because of Results First analyses would have important implications for the use of evidence in state legislatures. As a program with a presence in approximately two dozen states and an ever-increasing portfolio, the time was ripe to embark on this investigation. I used the existing variation among Results First states as fertile ground for a deeper exploration of evidence use by state legislators. Considering the claims being made by Results First and the ways in which the program is expanding, the investigation is timely and important. The research may serve as the first truly public review of the impact in Results First states.

In chapter one, I hypothesized that Results First marginally encourages policymakers predisposed toward the use of evidence to further incorporate evidence into their legislative decision making. I found this to be *true*, but only marginally and with many caveats. The first of those caveats is that legislators must want evidence to be a part of the policy discussion. Having a sufficient supply of information or evidence is only useful if there is also sufficient demand for that information. Even though the data suggested overwhelming support for the idea of evidence playing a role in policymaking

in a question posed in chapter three, there were still over 10 percent of respondents from a heterogenous pool who thought, at least half the time, that it should not.

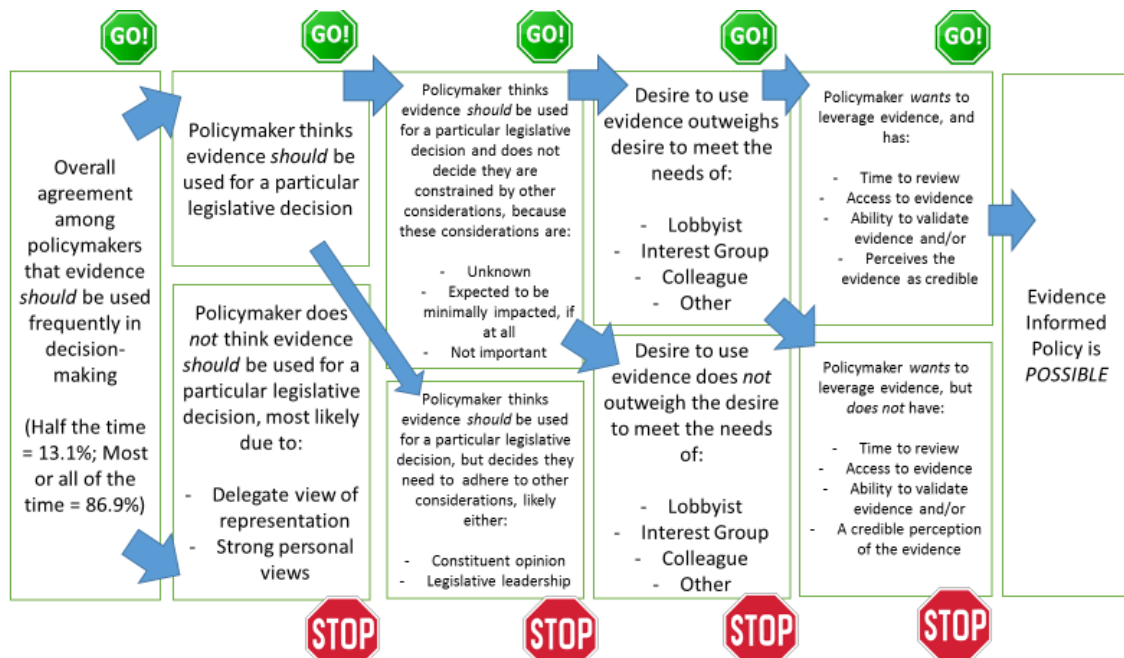
From there, effective evidence utilization also typically required issues to be largely free of both personal and constituent demands on a legislator. Based on research findings in chapters three and four, I add leadership demands to the list of major constraints. I also re-emphasize and expand the definition of “personal opinion.” First, there are matters of conscience upon which some lawmakers feel that evidence should not factor (or factor heavily) into their decision. As the research points out, there are many who are entirely opposed to evidence, largely because it conflicts with their worldview on their responsibilities as a legislator. These legislators view themselves so strongly as a delegate that there may be no circumstance in which it matters what the evidence says. This is a blend of personal opinion (it is their worldview, after all) and a more general need to be responsive to constituents – something to which all legislators, even those who consider themselves trustees, remain attuned.

If personal, constituent and leadership constraints are unknown, expected to be minimally impacted, or are not important to a legislator on a given topic, legislators are then required to weigh other considerations (such as lobbyists and their colleagues) against the use of evidence. Assuming the desire to have evidence inform a legislative decision is greater than these other considerations, legislators can move to an environment where evidence *might be sought*.

What remains is having the time to find the evidence, having the evidence be available and accessible and for lawmakers to be able to understand and validate what they are being presented with. If what has been described sounds long and arduous, it is.

Even if legislators make it “to the evidence” with definitive and clear to understand evidence, it is not guaranteed that their vote will align with the information with which they are presented. Having just enough doubt in the evidence or a strong enough pull from other factors influencing legislative decision-making could prevent an evidence-informed vote. The evidence-informed policymaking flowchart, below, is an attempt to simplify a very elaborate process and multiple considerations that often are being considered simultaneously. The fact that even this “simple” flowchart is cumbersome speaks to the challenges that evidence faces in playing a prominent role within policymaking.

Figure 12: Evidence-informed policy flowchart



Source: Olszewski (2017).

One could rightfully point out that advocates also can incorporate evidence into their arguments regarding their respective “needs,” bolstering the likelihood their opinion prevails in earlier portions of the flowchart – and prevails partially due to the influence of evidence. This gives appropriate consideration to the impact of evidence as explained

through the Weiss enlightenment argument – rightfully acknowledging that evidence is not merely considered at the end of a process, but instead that it can and occasionally does influence the entire policymaking process. The chain in Figure 12 does, however, help to reinforce the idea that legislation that is more authentically dependent on the weight of the evidence is a rare circumstance. While evidence certainly enlighten and informs throughout the process, it rarely is the critical influence on legislative outcomes.

This complicated legislative process reinforces the idea that the impact of Results First on the average legislator is limited, even among those predisposed to supporting evidence. While the data and research produced by Results First can be impactful upon legislative outcomes in the states where they conduct their work, I have shown that other considerations continue to also have significant influence on the outcome of laws that have passed.

In chapter two, I questioned if a new push for evidence in policymaking might enable a reimagination of the delegate versus trustee dichotomy. Having completed the research, I return to that discussion briefly to offer the following observations: to the extent evidence becomes more standardly integrated into the governing process, I still expect there will be some natural overall movement away from delegate and toward trustee on that continuum. However, I did not find significant movement within my own research. Data in chapter three suggested that more than half (52.7%) of legislators surveyed are *open to the idea* of utilizing evidence over constituent opinions. It also showed how much legislators rely on citizen input in their decision-making, with over 71 percent of respondents saying they heavily rely on citizen input on the issue of bail and even more (73.4 percent) saying it is a critical factor when considering gun control

legislation. This strong reliance on citizen input is further supported in chapter four with interview data. To more fully know the extent to which a shift on the trustee-delegate scale might be happening requires its own research project. It is also likely true that, as evidence production systems begin to become part of the policy process, it may be some time before evidence-informed policymaking trends are meaningfully noticeable.

As I have also shown, much of the work being done by Results First is done “behind the curtain.” In other words, the average legislator has little to no idea about the ways in which Results first is connecting evidence and policy because they have little to no idea about the Results First program. Thus, even in the places where Results First is driving incremental progress among evidence utilization by state policymakers, legislators themselves largely are unaware that Results First is the “thing” leading them towards additional evidence in their lawmaking. To be fair to Results First, the impact of their program is inherently difficult to measure – and the fact that political back-benchers might be unaware about whether or not Pew, Results First or similar organizations are driving change does not mean that these organizations might not still be leading important evidence-based reforms within a state. In fact, even Results First staff acknowledge that back-bencher legislators are not their target audience.

## **6.2 Other related questions**

Understanding the ways in which Results First does – and does not – increase the use of evidence required other questions to be examined and answered. Namely, if Results First can help overcome previously identified obstacles to evidence use; if the program makes legislators look more favorably upon evidence; and to what extent legislators understand the evidence production process.

6.2.1 Can Results First help overcome institutional, policy, political, and communication challenges identified by research in the past?

Yes. Results First (and Pew more globally) have played an important role in breaking down political barriers, especially in the criminal justice space. However, I have shown that Results First is not the only way in which to accomplish this end. For example, Virginia's JLARC produces similar output and has been associated with similar legislative gains in the criminal justice space and beyond. I have also shown, through examples in chapter five, that legislative successes based on evidence do not come easily even in the most fertile of policy environments.

6.2.2 Does Results First make legislators look more favorably upon evidence than in prior attempts to push evidence?

Based on this research, it is more accurate to say that there appears to be a correlation between evidence utilization and the existence of Results First than it is to say legislators look more favorably upon evidence *because* of Results First. Based on the research, I believe this may be due to the fact that the Results First program helps to further improve awareness of evidence by creating additional evidence for lawmakers to consider and by drawing attention to it through the support of governors and key legislative leadership. The data also suggests evidence is playing a larger role in legislative deliberations in the criminal justice space when Pew (as was the case for several states in chapter five) is involved in the process. However, the data also suggests that very few lawmakers make the connection between Results First (or even Pew) as the *reason for* the additional evidence with which they are being presented. Finally, it is hard to determine, and was not effectively measured in this research design, the extent to

which legislators might have changed their overall opinion about evidence *because* of the Results First program.

### **6.3 Evidence informed policymaking**

If there is one area, both through my personal experiences and research in this space, in which I feel most comfortable staking a claim, it is this: the term academics, scholars and news media should utilize in describing the area of evidence and policymaking is **evidence-informed policymaking**. I draw a sharp distinction between the terms evidence-based and evidence-informed, confident that even in the most ideal of circumstances evidence is, at best, informing our policy outcomes. Using the term evidence-based implies that the results of a study or several studies determine the outcome of policy. As was discussed and affirmed in the quantitative, qualitative and case-based reviews of this dissertation, we know this is simply not the reality on the ground.

Even proponents of evidence-based policy making concede that there are limitations to their reliance on that evidence. This ranges from considerations about leadership, personal beliefs, and constituent feedback. That is why I accept the limited interaction of policy and evidence proposed by Gow and Wilson (2014) where “public administration *tries* to base decisions on evidence, but also democratic deliberation” (emphasis added, p. 133). I believe that this research has been insightful on that subject. It specifically has further acknowledged that these boundaries not only exist, but also serve as critical parts of our democratic society, knowing the “boundaries” will give a better sense of where this type of bounded rationality begins – and ends.

Just as the Results First model presents a reconceptualization in policymaking that focuses more on outcomes rather than just inputs or outputs, so too has there been a push for a similar reconceptualization in the performance budgeting space. Phillip Joyce (1993), for example, has been calling for both a short-term emphasis on performance measures within agency management as he also advocates for longer-term reform that involves creating good measures of program results and connecting those results to the decision-making process when it comes to both budgeting and policymaking. We may be entering the time when those good measures sought by Joyce are available in many areas of government review, courtesy of programs like Results First. We finally have entered a time when evidence informed policymaking can be a reality.

Why evidence-informed policy rather than evidence-based? First, it's difficult to predict with much precision in which areas and to what extent evidence will be driving policy decisions in the years ahead. Moreover, labeling decisions as "evidence-based" could unfairly confuse the issue, as many so-called evidence-based policies today are only partially informed by the evidence – and not wholly based. In taking an approach that describes evidence's use as informative rather than entirely instructive, policymakers can have a more honest conversation that acknowledges normative concerns remain play and that there are interests beyond only what the evidence suggests.

There are also practical reasons why the policy process should not be entirely driven by evidence. As Marc Schabes (2017) from New York shared, "Don't take this to an extreme of a paradigm shift ... There are programs worthy of investment that may not show a net gain." Some of these treatment and intervention programs, in other words, may cost more than the "net value" on return. However, there are some in our society –



the most vulnerable, the most needy – worthy of this investment. Looking only at the numbers could create a dangerous environment where those most in need of our help are overlooked. To be fair, The Results First methodology does cover social benefits to some extent by quantifying the long-run (discounted) fiscal savings from interventions nor do they overtly argue that only the best benefit-cost ratios are where public dollars should be expended. However, it was unclear to me that the program provides any other sort of “social weighting” option in their calculations. In other words, the social benefit calculation appears to be solely a financial calculation under the WSIPP and Results First approach. In benefit-cost analysis run amok, lawmakers can lose sight of the fact that, children with special education needs are worthy of investment, even though such expenditures may not yield a positive benefit to cost ratio.

Such an investment in vulnerable populations speaks directly to an argument about equity in government. It is an argument that suggests the value of services goes beyond the simple test of governmental fiscal benefit versus cost (the Results First approach). Instead, I argue that there needs to be a metric that creates a value for intangible social benefit and cost in addition to pure fiscal benefit and cost. In school funding, for example, additional resources are necessary with some populations of students to ensure that they have something resembling an equal opportunity to be successful within the classroom. This might be additional costs for school breakfast and lunches, additional staff (be they more teachers for smaller class size or pupil personnel workers and counselors to help address concerns beyond the classroom), or resources for after-school and/or summer programs. These investments, like those in special education,

could in a pure fiscal cost-benefit approach, leave important classes of people behind in a way that I believe is contrary to the goals and objectives of our social democracy<sup>50</sup>.

Indeed, within democratic institutions, I neither expect nor aspire to reach a place where every decision is rigorously tested and evaluated before a verdict might be rendered. It is aspirational at best to believe randomized control trials and/or other rigorous benefit-cost analysis will be the foundation of *all* policy decisions. Nor should we: normative values deserve an important place in our democracy. But these normative values deserved to be tested against the best possible evidence out there – both for the legislators casting the votes as well for the public they represent.

In pushing for this evidence creation and use, it is important to recognize the unique circumstances of each state, and to recognize that there may not be one-size-fits-all application of evidence-based best practices and legislation. For example, the research in chapter four suggested that, in New York, the House leadership has strong control of the legislative agenda. Infusing evidence-informed policy there might look differently than in, say, Pennsylvania, who appeared to have a strong lobbying influence among legislators interviewed.

Academics should also take seriously the ways in which legislators view scientific method and research methods as a black box. By “black box,” I mean to say that the social sciences are, to legislators, a system of inputs and outputs with relationships working between them in such a way that lawmakers may understand what the inputs and outputs are, but do not understand well, or at all, the internal structures and processes that allow researchers and academics to draw statistical conclusions on relationships between

---

<sup>50</sup> On some level, legislators should take responsibility and argue that even programs that have negative fiscal return on investment should remain funded.

the variables, projections about the variables, and more. Lawmakers understand if an audit reveals that state employees are making inappropriate expenditures, or if projects run over budget. They typically do not understand logic models or statistical methods deployed to determine the extent to which a program is likely to generate a net positive return on investment. This lack of understanding could be a cause for concern for many reasons, including a reliance on evidence that is not grounded in the best possible methods and whose findings could be suspect, or plainly wrong.

For some states, that may mean Results First is a good fit, even if Results First itself should be open to adaptable to these differences and willing modify their approach across states. In other states, homegrown agencies such as JLARC might be appropriate. For others, there could be an entirely different approach.

Regarding evidence, we can and should do better. For all the flaws and limitation associated with program evaluations, system measures and testing of outcomes – “the biggest problem ... is that they are not [done] nearly enough” (*Economist* 2015). A single finding of significance or impact should not necessarily lead to policy change, nor should even multiple findings. However, the more clarity with which we approach the major policy issues of the day through quality evidence, the more likely we are to develop the best possible responses. I turn next to some practical recommendations on ways in which we can accomplish that goal.

#### **6.4 Recommendations for Results First**

Results First is a program worthy of being profiled and celebrated. They actively seek to cultivate an environment where evidence helps to inform policymaking decisions in legislatures across the country. In many respects, they are succeeding. But while the

output has been strong, the scope of the work has been somewhat limited. If Results First wants to truly expand (not only to new states and local governments, but also to new issue areas), I offer two substantive recommendations:

**Open the process:** If Results First continues in its role as a premiere provider of evidence-production services, they would serve the democratic institutions with which they partner and the program itself well by opening to more and complete reviews of their analytical processes. A system designed to measure other systems should, almost by definition, be among the most measured and validated. If there is going to be widespread use of the Results First model, leveraging the underlying benefit-cost analysis system developed by WSIPP, there also should be external validators who run the models and affirm them. While Results First *may* be working fine, the limited window into its inner-workings leaves more questions than should exist for anything driving so much of the country's public investment.

Ironically, to the best of my knowledge, this work may be the *only* publicly available review of the work being done by Results First. Even so, the review has been limited regarding program operation, focusing more on its impact on evidence utilization by state policymakers. This research has not investigated the underlying models that score policy nor has it evaluated the outcomes of all the policy changes enacted because of the program. To truly validate the work of Results First, its leadership should open itself to a thorough independent audit. The audit process should be ongoing, with thorough reviews re-occurring every few years. The findings of these reviews, removing any language that could compromise the intellectual property of the program's model, should be made available for public review. If public dollars and resources are being

utilized to support the Results First (or any other program), taxpayers deserve to know that the product their state leaders are investing in is delivering them the product they deserve. Only regular, ongoing and independent reviews that are made available for public consumption can ensure that is happening.

**More active outreach with legislators:** Results First should more proactively engage with legislators, and especially with those *not* in leadership, in states where they are operating. By doing so, lawmakers will be more inclined to understand the purpose of the program and how recommendations are made, and will be able to make a clearer connection between evidence production and strong policy outcomes. As was discovered in the research, vast majorities of state legislators are unaware of the existence of and impact of the Results First program – even in places where the program is operating and making an impact. To ensure the ongoing success of Results First and programs like it, legislators of all positions – from presiding officers to freshmen back-bench members, should be aware of and connected to the output. There can and should be a concerted effort to inform legislators on topic selection and the analysis process, followed by significant outreach to disseminate research findings of the program. Such distribution should include direct outreach (emails and paper copies distributed to each member) as well as clearly-identified sourcing if Results First findings are included in state-based legislative analysis. Finally, if Pew has the resources to do so, they should consider the creation of two new positions within the Results First structure: one person would specifically be assigned to partner with state legislative organizations to better connect Results First findings into the production of fiscal notes and other policy reviews; the other would be a Results First cheerleader and salesperson for the organization,

connecting directly with the lawmakers currently disengaged with the process, building awareness of Results First and helping to further expand its footprint.

## **6.5 Recommendations for increasing evidence-informed policymaking**

In addition to the recommendations for the Results First program, there are more global recommendations that should be considered for cultivating an evidence-informed policymaking environment in our state legislatures.

### **Change the language of evidence-based policy to evidence-informed policy:**

This is discussed in greater detail above and it is the top recommendation of this effort. Not only is the term more intellectually honest, it may also have the added benefit of making legislators feel like it is more approachable. If evidence is not presented as a “take it or leave it” proposition, it might be more embraced by lawmakers.

**Encourage evidence-informed policymaking through trust, credibility and relationships:** “Truth seekers,” be they members of the academic community or professional staff, need to find ways to break down the barriers that currently exist between themselves and lawmakers. As learned from interviews in chapter four, creating a personal bridge between those who produce evidence and those who consume it may be the single greatest need for change.

In some respect, that is a key strength of Results First: the program recognizes that, to have policy outcomes follow their evidence production, there needs to be a strong relationship. It just so happens that Results First focuses these relationships at the highest levels of government – House Speakers, Senate Presidents and Governors of states. That’s helpful in that it shows the power of relationships, and that it helps establish and maintain the program, but as shared in the Results First-specific recommendation,

relationships need to be extended to all lawmakers and more wholly institutionalized to cultivate a true sustainability and a growth plan for Results First and programs like it.

**Open the process (redux):** What is good for Results First should be good for any program that is meant to inform legislative decision-making. Regardless of the model utilized, if it is going to be used to drive investment in a state, there should be external validators who run the models and affirm them.

**Consider future opportunities for additional research on the subject:** There were new and important insights gleaned from this work, but there are also additional questions that remain in need of answering. First, it would be worthwhile to find a way to conduct more sophisticated statistical analysis of important demographic information of individual, electoral, state and institutional varieties against either this or a similar evidence utilization index. This can be accomplished by creating a survey instrument that more directly engages legislators on questions that get at evidence use versus other considerations.

There is also a need for more clarity on what constitutes evidence and perhaps what constitutes “good” evidence among legislators. There needs to be a way to distinguish between information that can be useful in helping to make decisions (for example, budget numbers) as well as more sophisticated studies and analyses that go beyond single data points (for example, a well-done, peer-reviewed cost benefit analysis to understand the full picture of a project that is being proposed).

Finally, now might be a good time to do a more expansive, nationally-representative sample that follows up on the initial work of Hird, expanding both his work on NPRO’s while also leveraging the new information shared here.

**The responsibilities of legislatures and legislators:** Without clarity of evidence origin and a robust review process, decision-makers are likely to be left with incomplete or poor information as they embark upon legislative decision-making. Thus, it behooves all variety of legislatures to seek ways to implement a standardized, nonpartisan and highly respected method of filtering evidence. It can be a frightening prospect for lawmakers. As Hird (2005) writes, “Information is powerful; therefore, NPRO’s can threaten other legislative institutions and legislators themselves because of their role in developing and controlling information” (p. 106). Perhaps that is why, “even as policymakers in developing countries harness the power of randomization, those in rich countries resist – especially for their most cherished schemes” (*Economist* 2015). Despite the ways in which legislators might feel threatened by research organizations, leadership is needed that will continue to push for investment in this area from the legislative side. Every state should have, or create, a well-staffed and independent / bipartisan department of legislative services. Legislative professionalism is a good idea and makes it more likely (though not certain) that outcomes will be grounded in evidence and supported by best practices.

## **6.6 Obstacles and limitations of this research**

Having presented my review of findings and providing recommendations for moving forward, I also wish to recognize the obstacles faced and limitation of the research conducted. The design of this research includes inherent analytical tradeoffs – the clearest tradeoff was a depth of understanding and variation on the issue in exchange for selection of issues and generalizability concerns. Perhaps the most practical limitation of this research relates to the issue of selection bias. As mentioned in earlier chapters,



there is a good chance that a selection bias exists among the states with whom Results First partners. Including states from the full spectrum on the use of evidence (as determined by Pew) in both the case study and survey work helped to address this, but the concern is nonetheless a valid one. Moreover, for a study with limited scope, examining the states most likely to have a selection bias seemed the most natural starting point for examining the impact of Results First and evidence utilization generally. If perceptions and actions relative to evidence are not changing in the most likely of places, it is increasingly doubtful that such a change would come about in states with less capacity and interest.

A lower than expected response rate, discussed in chapter three, was another challenge that called for additional attention to the question of the internal validity of the sample. There was a strong effort to drive up these rates with some success, although I do feel more comfortable knowing how closely, at least on demographic grounds, the ways in which the survey and full population reflected each other. I also believe the in-depth data collected during interviews helped to further validate the information presented.

It is also possible that the states examined as a part of this research may not have complete external validity to all other states. States choosing to engage with a program such as Results First may be the most likely to self-select into these types of programs and those most likely to have the administrative capacity to embark upon these programs. However, the ways in which this research proposes to examine both adopter and non-adopter states will capture a wide variance in several key indicators that can help to address some of the concerns that might be raised relative to selection. It can also provide useful base insights for future work in this area.

## **6.7 Final concluding remarks**

In all instances, states, legislatures and the citizens served will be better off because of these efforts to incorporate more evidence into the policymaking process. Continuing to understand the environments in which evidence-based policy shops can inform the legislative dialogue will be a critical step in bridging the divide between the ivory tower of academia and the hand-to-hand combat that often is politics and government.

Lawmaking has been compared to sausage-making, after all, for a reason. Many would be shocked to learn about what is happening behind the factory doors. With evidence-informed policymaking, we can make the factory work more effectively, even if only marginally – and even marginal improvement in our policymaking can have huge impacts for the people served by government. Evidence-informed policymaking holds the promise of a reduction in wasteful spending (freeing up money to be reallocated towards those programs that would lead to the improved life outcomes discussed above), can expand innovation in programming, and holds the potential to help strengthen accountability mechanisms (Pew, November 2014).

That potential is something that everyone needs to have a hand in if the full benefits of evidence-informed policymaking are to ever come to fruition.

## Appendices

### Appendix A: sample Results First report (Colorado)

The following screen grab is from an actual report from Colorado on benefit-cost determinations on varying programs related to juvenile justice. It provides an example of a typical benefit-cost reporting conducted under the Results First model.

Benefit-Cost Results of Evidence-Based Juvenile Justice Programs								
Benefits and program costs are rounded to the nearest dollar.								
Program Name	Department/ Organization	Total Benefits	Taxpayer Benefits	Non- Taxpayer Benefits	Program Cost (FY 2015)	Benefits Minus Costs (Net Present Value)	Taxpayer Benefits to Cost Ratio	Total Benefits to Cost Ratio
<b>Facility Programs</b>								
Aggression Replacement Training for Youth in State Institutions	Department of Human Services	\$13,599	\$3,202	\$10,397	\$1,948	\$11,651	\$1.60	\$7.00
Other Chemical Dependency Treatment for Juveniles (Non-Therapeutic Communities)	Department of Human Services	\$1,562	\$368	\$1,194	\$2,293	\$(731)	\$0.20	\$0.70
Sex Offender Treatment (Non-Multisystemic Therapy) for Juvenile Offenders	Department of Human Services	\$11,776	\$2,698	\$9,078	\$1,708	\$10,068	\$1.60	\$6.90
<b>Parole Programs</b>								
Functional Family Parole with Quality Assurance	Department of Human Services	\$7,032	\$1,465	\$5,567	\$2,583	\$4,449	\$0.60	\$2.70
Multisystemic Therapy for Youth with Problem Sexual Behaviors	Department of Human Services	\$27,699	\$6,548	\$21,151	\$12,673	\$15,026	\$0.50	\$2.20
Other Chemical Dependency Treatment for Juveniles (Non-Therapeutic Communities)	Department of Human Services	\$1,433	\$319	\$1,114	\$1,176	\$257	\$0.30	\$1.20
Sex Offender Treatment (Non-Multisystemic Therapy) for Juvenile Offenders	Department of Human Services	\$9,799	\$2,031	\$7,768	\$2,582	\$7,217	\$0.80	\$3.80
Victim Offender Mediation	Department of Human Services	\$5,831	\$1,338	\$4,493	\$526	\$5,305	\$2.50	\$11.10

## Appendix B: updated benefit-cost rankings

After the writing of the dissertation proposal and the completion of the bulk of the research data, an updated survey of the 50 states by the Pew Charitable Trusts (Dube and White 2017) generated a new range of categories related to state utilization of evidence-based policymaking. This new range placed states into one of the following categories: leading, established, modest, or trailing – with most states showing a modest level of evidence-based policymaking. As it relates to the states included in this research, the rankings are as follows:

State	Level of Evidence-Based Policymaking
Massachusetts	Modest
New York	Established
Pennsylvania	Established
Rhode Island	Modest
Virginia	Modest
West Virginia	Trailing

There were only five states identified as “Leading,” none of whom were included in the research. Several states were ranked as “Trailing,” although none of those formally included was listed at this level – except for Maryland, which provided only a few interviews towards chapter four.

These changes in rankings were likely caused by an adjustment in methodology. In the benefit-cost evaluation for the categories used in this research (leading the way, mixed and trailing behind), investigators sorted states by the number of comprehensive cost-benefit reports released by a state between January 2008 and December 2011. Leading states produced at least 11 reports, mixed states produced between three and 10, and trailing states produced two or fewer such reports. The new metric deployed a more complex evaluation process, seeking six actions related to evidence-based policy (define levels of evidence, inventory programs, compare costs and benefits, require action through law, report outcomes and target spending) across four functional areas (behavioral health, child welfare, criminal justice and juvenile justice). A full graphic for where every state in country landed on these new rankings is below:

Map 1  
Most States Show Modest Levels of Evidence-Based Policymaking



248

## Appendix C: NCSL Legislative professionalism definitions

The National Conference of State Legislatures (NCSL) has five separate categories when it comes to full versus part time legislatures. They are full time (green), full time lite (light green), hybrid (gray), part time lite (light gold) and completely part time (gold). Additional context for the green versus grey versus gold types of legislatures are below with information pulled from the NCSL website on the subject.

Source: NCSL June 2014.

### FULL- AND PART-TIME LEGISLATURES

CONTACT: BRIAN WEBERG



It seems like an easy question: Which legislatures are full-time and which ones are part-time? But with 50 different formulas for designing a state legislature, it's difficult to paint this issue in black and white. So we've done it in green, gray and gold.

Being a legislator doesn't just mean attending legislative sessions and voting on proposed laws. State legislators also spend large amounts of time assisting constituents, studying state issues during the interim and campaigning for election. These activities go on throughout the year.

Any assessment of the time requirements of the job should include all of these elements of legislative life.

Beyond that point, NCSL prefers to look more broadly at the capacity of legislatures to function as independent branches of government, capable of balancing the power of the executive branch and having the information necessary to make independent, informed policy decisions. To measure the capacity of legislatures, it's important to consider the amount of time legislators spend on the job, the amount they are compensated and the size of the legislature's staff.

NCSL has grouped the 50 state legislatures into three major categories: Green, Gray and Gold-and for those who want to know more, NCSL has provided some shading within those categories.

### Green Legislatures (Full-time, well-paid, large staff)

Green legislatures require the most time of legislators, usually 80 percent or more of a full-time job. They have large staffs. In most Green states, legislators are paid enough to make a living without requiring outside income. These legislatures are more similar to Congress than are the other state legislatures. Most of the nation's largest population states fall in this category. Because there are marked differences within the category, we have subdivided the Green states. Those in Green generally spend more time on the job because their sessions are longer and their districts larger than

those in Green Lite. As a result, they tend to have more staff and are compensated at a higher rate. Within subcategories, states are listed alphabetically.

## Gray Legislatures (Hybrid)

Legislatures in the Gray category are hybrids. Legislatures in these states typically say that they spend more than two-thirds of a full time job being legislators. Although their income from legislative work is greater than that in the Gold states, it's usually not enough to allow them to make a living without having other sources of income. Legislatures in the Gray category have intermediate sized staff. States in the middle of the population range tend to have Gray legislatures.

## Gold Legislatures (Part-time, low pay, small staff)

In the Gold states, on average lawmakers spend the equivalent of half of a full-time job doing legislative work. The compensation they receive for this work is quite low and requires them to have other sources of income in order to make a living. The Gold states have relatively small staffs. They are often called traditional or citizen legislatures and they are most often found in the smallest population, more rural states. Again, NCSL has divided these states into two groups. The legislatures in Gold are the most traditional or citizen legislatures. The legislatures in Gold Lite are slightly less traditional. States are listed alphabetically within subcategories.

Table 1 shows the breakdown of states by category. Table 2 shows the average scores for the Green, Gray and Gold states for time on the job, compensation and staff size. For 2017 legislator compensation figures, take a look at the [latest figures](#).

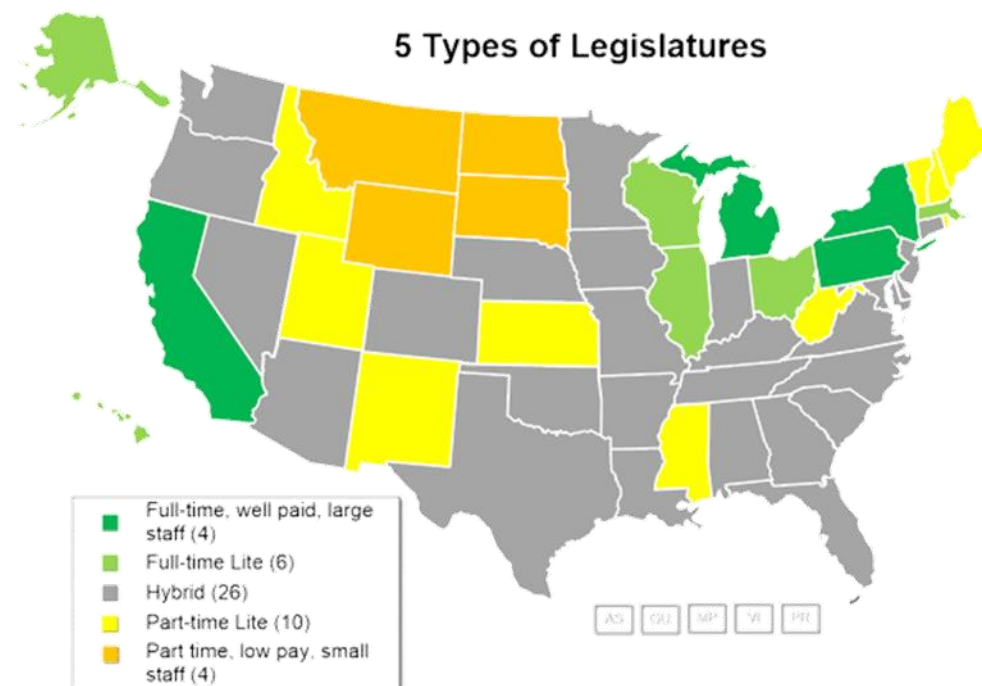


Table 1. Green, Gray and Gold Legislatures

Green	Green Lite	Gray		Gold Lite	Gold
California	Alaska	Alabama	Minnesota	Idaho	Montana
Michigan	Hawaii	Arizona	Missouri	Kansas	North Dakota
New York	Illinois	Arkansas	Nebraska	Maine	South Dakota
Pennsylvania	Massachusetts	Colorado	Nevada	Mississippi	Wyoming
	Ohio	Connecticut	New Jersey	New	
	Wisconsin	Delaware	North Carolina	Hampshire	
		Florida	Oklahoma	New Mexico	
		Georgia	Oregon	Rhode Island	
		Indiana	South Carolina	Utah	
		Iowa	Tennessee	Vermont	
		Kentucky	Texas	West Virginia	
		Louisiana	Virginia		
		Maryland	Washington		

Table 2. Average Job Time, Compensation and Staff Size by Category of Legislature

Category of Legislature	Time on the Job (1)	Compensation (2)	Total Staff (3)
Green	84%	\$82,358	1,250
Gray	74%	\$41,110	469
Gold	57%	\$18,449	160

**Notes:**

1. Estimated proportion of a full-time job spent on legislative work including time in session, constituent service, interim committee work, and election campaigns. *Source: 2014 NCSL survey of all state legislators.*
2. Estimated average annual compensation of legislators including salary, per diem, and any other unvouchered expense payments. *Source: NCSL 2014.*
3. Average number of staff--partisan and nonpartisan--working for the legislature. *Source: NCSL 2015.*



## Appendix D: Narrative interview script for Maryland pre-test, given on various dates

Thank you for taking the time to talk with me today. I am conducting research on how state legislators balance different types of factors when they are making policy decisions, including the role of evidence. Before I begin with my case studies in other states, I wanted to talk to some people in Maryland to make sure that I am asking the right questions on this issue. With that in mind, I have just a few questions to ask you and I would like to let you take as much time as you care to really working through them – even telling any relevant stories you might have about the topic.

Before we begin, do you give your informed consent to participate in this interview and research project? *Upon consent, begin interview.*

- 1) I wanted to start by asking if you could talk to me about the full range of factors you consider when you are deciding on how to vote on a piece of legislation?
- 2) One of the factors that *you mentioned / some legislators* consider is what the “evidence” says. What does the term “evidence” mean to you as a legislator?
- 3) Do you value some evidence more than other kinds? Can you explain why or when?
- 4) How important would you consider evidence relative to the other considerations you mentioned earlier?
- 5) Can you talk to me about what evidence you have available to you and what evidence you take advantage of as a legislator here in Maryland?
  - a. Probes: information that is part of the legislative process and information that is from sources outside the legislative process
  - b. Probe: do your colleagues take advantage of as much, about the same as, or less evidence than you?

Thank you again for taking a few minutes to talk with me about legislative decision-making and the role of evidence. Your feedback will be a big help to this research project the in preparing for the case studies associated with it.

## **Appendix E: Email recruitment scripts/prompts used in the data collection process**

*Email 1: Survey launch and Invitation to join survey on legislative decision-making, distributed on Monday, February 1, 2016*

Dear [Personalized Legislator Name],

As a former state legislator, I was constantly interested in the ways in which lawmakers can get information about effective programs. While I am no longer in the Maryland General Assembly, my interest in this issue has continued as a doctoral candidate at the University of Maryland, Baltimore County.

To that end, I am now collecting data for a PhD in public policy and public administration on this very issue: how legislators get and use information about programs and policies. The understanding I am attempting to glean is about how you – a former colleague – make decisions on important policy issues. It is an important issue to study from an academic perspective and it is a research area that has generated an interest from places such as the National Conference of State Legislatures (NCSL) in the results of my study.

I have created an on-line survey instrument that will easily collect information from legislators in such a way that your anonymity will be closely guarded and maintained. Your participation in this survey is very important to its success and would be greatly appreciated.

Thank you for taking the time to read this letter, and more importantly for your help in making this project a success through your participation – which you can do by clicking on the following link: [participation link]

Sincerely,

John Olszewski, Jr.  
Former Maryland State Legislator and  
PhD Candidate, University of Maryland Baltimore County

Follow the link to opt out of future emails:  
\${1://OptOutLink?d=Click here to unsubscribe}

*Email 2: First Reminder, sent on Tuesday, February 9, 2016*

Dear [Personalized Legislator Name],

Last week you should have received an e-mail from me seeking your participation in a PhD research project I am undertaking at the University of Maryland, Baltimore County.

The questionnaire has been kept intentionally short and should not take much of your time to complete. Your participation in this effort would be greatly appreciated and is critical to better understanding how legislators across the country get and use information about programs and policies.

I hope that providing you with a link to the survey website will make it easy for you to respond – you need only click on the following:

**Follow this link to the Survey:**

[\\${1://SurveyLink?d=Take the Survey}](#)

Or copy and paste the URL below into your internet browser:

[\\${1://SurveyURL}](#)

I would be grateful to have your participation in this project.

Many thanks,

John Olszewski, Jr.

Follow the link to opt out of future emails:

[\\${1://OptOutLink?d=Click here to unsubscribe}](#)

*Email 3: Second reminder, sent on Wednesday, February 17, 2016*

Dear [Personalized Legislator Name],

Recently, I sent you an e-mail asking you to participate in a survey about factors that go into the decision-making process of state legislators and specifically how legislators get and use information about programs. This research project is being done in conjunction with my PhD dissertation.

I would like to encourage you to take the time to complete this questionnaire, which should only take a little bit of your time to complete. Your response is voluntary and you can skip over any question you would prefer not to answer. Please be assured that all responses will be kept confidential and your name will not be associated with any data collected in the survey process.

**The participation link can be accessed here:**

[\\${1://SurveyLink?d=Take the Survey}](#)

Or copy and paste the URL below into your internet browser:

[\\${1://SurveyURL}](#)

If you have questions or comments about the project, please do not hesitate to contact me at 443-865-3370 or [johno1@umbc.edu](mailto:johno1@umbc.edu). I very much appreciate you considering this

request.

Regards,

John Olszewski, Jr.  
Former Maryland State Legislator and  
PhD Candidate, University of Maryland Baltimore County

Follow the link to opt out of future emails:  
\${l://OptOutLink?d=Click here to unsubscribe}

*Email 4: Third email reminder, sent on Tuesday, March 8, 2016*

Dear [Personalized Legislator Name],

About a month ago I contacted you to seek your help with a survey project that hopes to better understand legislative decision-making processes within state government. This has been an area of interest to me as a former state lawmaker and now a PhD candidate in Public Policy and Public Administration in Maryland.

I am reaching out to you because my ability to accurately describe how state legislators from across the country get and use information about programs depends on hearing from those who have not yet responded. I need your help to ensure that my results are as precise as possible. On average, completing the survey takes less than 10 minutes.

**To help me in this effort, you can click on the following web address:**

\${l://SurveyLink?d=Take the Survey}

Or copy and paste the URL below into your internet browser:

\${l://SurveyURL}

Respondents can be assured that their information will remain confidential. If you have questions about the survey, I can be reached anytime at 443-865-3370 or johno1@umbc.edu.

Thank you again for your help with this important project!

John Olszewski, Jr.  
Former Maryland State Legislator and  
PhD Candidate, University of Maryland Baltimore County

Follow the link to opt out of future emails:  
\${l://OptOutLink?d=Click here to unsubscribe}

*Email 5: Fourth reminder email, sent on Monday, March 16, 2016*

Dear [Personalized Legislator Name],

I am writing to follow up on the message sent to you last week asking for your participation in a survey about decision-making and the use of information by state legislators across the country. This will be one of your **final opportunities** to ensure that your responses are included with those from your colleagues.

**Follow this link to the Survey:**

[\\${l://SurveyLink?d=Take the Survey}](#)

Or copy and paste the URL below into your internet browser:

[\\${l://SurveyURL}](#)

I hope that you will add your input to this project – having your participation will make the research more robust and help provide new and important insights into your role as a legislator.

Thank you so much for your time and support of this important research project.

Sincerely,

John Olszewski, Jr.

Former Maryland State Legislator and

PhD Candidate, University of Maryland Baltimore County

Follow the link to opt out of future emails:

Follow the link to opt out of future emails:

[\\${l://OptOutLink?d=Click here to unsubscribe}](#)

*Email 6: Fifth and final Reminder, sent on Thursday March 24, 2016*

Dear [Personalized Legislator Name],

Back in February, I began reaching out to you, your colleagues and fellow legislators across the country to assist with my PhD dissertation research on legislative decision-making.

I am writing to you now because I have not heard back from you yet and am **scheduled to close this survey on Friday, May 6th**. That makes this reminder the **absolute last chance** for me to include your **completely confidential** responses in my work.

Your responses are critical for this research to be successful, so I appreciate your willingness to assist me with this effort. Since most legislatures have concluded their legislative sessions, I hope you might have a little more time to help me with this request.

**You can follow this link to the Survey:**

[\\${l://SurveyLink?d=Take the Survey}](#)

Or copy and paste the URL below into your internet browser:  
\${l://SurveyURL}

Again, please accept my sincere thanks for your help with this PhD dissertation research project.

Sincerely,

John Olszewski, Jr.  
PhD candidate, University of Maryland Baltimore County  
Former Maryland state legislator

Follow the link to opt out of future emails:  
\${l://OptOutLink?d=Click here to unsubscribe}

*Individual follow-up emails to legislators opening (but not completing) surveys, sent from personal email address and at varying times based on when legislators started a survey*

Dear [Legislator],

Thank you so much for your interest in helping with my PhD dissertation research on factors that go into legislative decision-making by starting the process of completing the online survey that I sent to you.

**I am writing to personally ask for your help -- will you return to the survey and complete it?**

As a former legislator in Maryland, I know there are many demands on your time so I am especially thankful for your help. To that end, I designed the survey to allow you to provide your input quickly -- on average, it takes approximately 10 minutes to complete.

Dozens of your colleagues have already completed the survey and I hope that I can add your feedback into my dissertation. You can be sure that your identity will remain entirely confidential.

If you have any questions about the survey collection process or the research itself, please do not hesitate to contact me at this email address or 443-854-3370. Thank you again for your willingness to help me with this important project!

Kind regards,

John Olszewski, Jr.  
PhD Candidate - University of Maryland Baltimore County (UMBC)

If you wish to no longer receive emails from me regarding this project, please reply to this email address with “unsubscribe” in the title or body of your response.

## **Appendix F: phone script for recorded call to legislators**

*Sent on Sunday, March 6, 2016*

Hi, this is John Olszewski, Jr. calling. I'm a doctoral candidate at the UMBC and I'm also a former state legislator in MD. I'm calling to personally ask for your help in completing a survey related to my PhD dissertation research. I'll be sending you an email reminder on Tuesday morning with a unique link to access that survey. As a former elected official, I know that your time is value, so I designed the survey so that it can be completed quickly. Your identity will remain entirely confidential in the process. If you have any questions about the survey collection or the research itself, you can reach me at [johnol1@umbc.edu](mailto:johnol1@umbc.edu) or 443-865-3370. Thanks so much for your help with this important project.



## Appendix F: final version of online survey

[note: EIS represents questions used toward a calculation of a legislator's evidence utilization score]

### *Introductory prompt and informed consent:*

Opening Prompt: This survey on legislative decision-making and the use of evidence is being conducted by John Olszewski, Jr., a former state legislator in Maryland and a Ph.D. candidate at the University of Maryland, Baltimore County.

You were contacted to participate because you were identified as a member of your state legislature. If you are not a current member of your state legislature, please do not complete the survey. If you have come to this site to complete this survey on behalf of a legislator, please ask the legislator for whom you work or are volunteering to fill out the survey.

A few things before you get started: 1. There are 20 total questions in addition to the collection of some demographic information. 2. All of your individual responses will be kept confidential, and nothing will be published that would identify you personally. 3. Once you answer a question and advance the screen, you will not be able to go back and change your responses. 4. A copy of the completed project can be sent to you upon request. To make such a request, please email me directly at [johnol@umbc.edu](mailto:johnol@umbc.edu).

Should you have any questions or concerns regarding the research project, you can contact me at 443-865-3370 or [johnol@umbc.edu](mailto:johnol@umbc.edu). You can also contact my advisor, Dr. Roy Meyers, at 410-455-2196 or [meyers@umbc.edu](mailto:meyers@umbc.edu).

This study has been reviewed and approved by the UMBC Institutional Review Board (IRB). A representative of that Board, from the Office for Research Protections and Compliance, is also available to discuss the review process or your rights as a research participant. Contact information for the Office is (410) 455-2737 or [compliance@umbc.edu](mailto:compliance@umbc.edu).

As a former state legislator, I know first-hand how many demands there are on your time. I truly appreciate you helping me with this research.

Thank you, Johnny

Informed Consent: Do you consent to participate?

- ☐ I hereby offer my informed consent and agree to participate in this research project. (1)
- ☐ I do not consent to participating in this research project. (2)

If I do not consent to participate is selected, then skip to end of survey

*Question set one – factors influencing decision making of legislators*

Introduction: As you complete the following questions, answer them within the context of your district, political considerations and desire for good government.

Q1: There are many factors that go into the decision-making of legislators. Below are several of the factors typically given as a reason for supporting or opposing a specific piece of legislation. Thinking about a proposed change to your state's bail requirements-- and using a scale of 1 to 5, where 1 means "not important at all" and 5 means "very important"-- how heavily does each of the following factors influence your decision making?

- \_\_\_\_\_ Personal opinion (1)
- \_\_\_\_\_ Constituent opinion (2)
- \_\_\_\_\_ Legislative staff expertise and research (3) [EIS]
- \_\_\_\_\_ Presiding Officer's preference (4)
- \_\_\_\_\_ Committee Chairman's preference (5)
- \_\_\_\_\_ How my colleagues view the issue (6)
- \_\_\_\_\_ Position of my political party (7)
- \_\_\_\_\_ Expected impact on crime rates (8) [EIS]
- \_\_\_\_\_ Governor's position on the issue (9)
- \_\_\_\_\_ Position of my campaign funders (10)
- \_\_\_\_\_ Expertise and research from an outside organization or university (11) [EIS]
- \_\_\_\_\_ How I have voted in the past (12)

Q2 Now, using the same set of factors and thinking about a proposed change to your state's gun laws, and using a scale of 1 to 5, where 1 means "not important at all" and 5 means "very important," how heavily does each factor influence your decision making?

- \_\_\_\_\_ Personal opinion (1)
- \_\_\_\_\_ Constituent opinion (2)
- \_\_\_\_\_ Legislative staff expertise and research (3) [EIS]
- \_\_\_\_\_ Presiding Officer's preference (4)
- \_\_\_\_\_ Committee Chairman's preference (5)
- \_\_\_\_\_ How my colleagues view the issue (6)
- \_\_\_\_\_ Position of my political party (7)
- \_\_\_\_\_ How I have voted in the past (8)
- \_\_\_\_\_ Governor's position on the issue (9)
- \_\_\_\_\_ Position of my campaign funders (10)
- \_\_\_\_\_ Expertise and research from an outside organization or university (11) [EIS]
- \_\_\_\_\_ Expected impact on crime rates (12) [EIS]

Q3 On a scale of 1 to 5 with 1 being "not important at all" and 5 being "very important," how important do you believe issues of crime and criminal justice are to your constituents?

- \_\_\_\_\_ Importance of Crime and Criminal Justice issues (1)

Q4 I serve on a legislative committee that hears and considers proposals relates to criminal justice issues:

- ☐ Yes (1)
- ☐ No (2)
- ☐ Don't Know / Refuse (3)

*Question set two – evidence versus personal values, versus policy preferences and versus constituents*

Information: For the purposes of answering the remainder of questions, please consider "evidence" as the best available information related to the development of a particular policy and its implementation.

Q1 Thinking about policy decisions on criminal justice issues, if the available evidence were to conflict with your personal values, in your decision-making you would tend to rely more upon ....

\_\_\_\_\_ What you rely upon in making decisions on criminal justice policy (1) [EIS]

Q2 Thinking about policy decisions on criminal justice issues, if the available evidence were to conflict with your general policy preferences, in your decision-making you would tend to rely more upon ...

\_\_\_\_\_ What you rely upon in making decisions on criminal justice policy (1) [EIS]

Q3 Thinking generally about policy decisions on criminal justice issues, if the available evidence were to conflict with what you are hearing from constituents on the issue, in your decision-making you would tend to rely more upon ...

\_\_\_\_\_ What you rely upon in making decisions on criminal justice policy (1) [EIS]

*Question set three – finding evidence and influence of evidence by type*

Introduction: The following questions deal specifically with your experience with the use of evidence in your work. For these questions, consider "evidence" as the best available information related to the development of a particular policy and its implementation in your work.

Q1 On a scale of 1 to 5, with 1 being "never" and 5 being "all of the time," where do you typically find evidence that relates to legislation that you are considering?

- \_\_\_\_\_ Think tank (1)
- \_\_\_\_\_ Legislative staff (2)
- \_\_\_\_\_ Results First reports (3)
- \_\_\_\_\_ State or federal agencies (4)
- \_\_\_\_\_ Lobbyists or interest groups (5)
- \_\_\_\_\_ Institutions of higher education (6)
- \_\_\_\_\_ Legislature's Policy Organization (7)
- \_\_\_\_\_ Governor's office (8)
- \_\_\_\_\_ Other legislators (9)
- \_\_\_\_\_ News media (10)

Q2 Thinking about how evidence is used in your decision-making process, on a scale of 1 to 5, with 1 being "not influential at all" and 5 being "extremely influential," how influential would you say each type of evidence is in your decision-making process?

- \_\_\_\_\_ Program cost / fiscal impact (1)
- \_\_\_\_\_ Formal program evaluations (2) [EIS]
- \_\_\_\_\_ Informal program evaluations (3)
- \_\_\_\_\_ Professional literature on the issue (4)
- \_\_\_\_\_ Scientific studies on the issue (5) [EIS]
- \_\_\_\_\_ Letter, e-mails and phone calls (6)
- \_\_\_\_\_ Benefit - cost evaluations (7) [EIS]
- \_\_\_\_\_ Testimony on the issue (8)
- \_\_\_\_\_ Data from performance reporting systems (9)
- \_\_\_\_\_ Results First reports (10) [EIS]

*Question set four – personal beliefs versus constituent belief by policy subject*

Introduction: For each issue below, relative to your personal beliefs and those of your constituents, how important would evidence produced by people with formal ties to the legislature be for each of the following topics? Formal ties to the legislature means evidence produced by legislative staff and/or outside organizations that have formal partnerships with the state legislature.

Q1 Relative to your personal beliefs and those of your constituents, on a scale of 1 to 5 with 1 being "evidence is not important at all" and 5 being "evidence is very important," how important do you think evidence produced by organizations with formal ties to the legislature would be on legislation that would change your state's drug laws:

- \_\_\_\_\_ Importance of evidence from organizations with formal ties to your legislature, relative to personal and constituent beliefs about drug laws (1) [EIS]

Q2 Relative to your personal beliefs and those of your constituents, on a scale of 1 to 5 with 1 being "evidence is not important at all" and 5 being "evidence is very important," how important do you think evidence produced by organizations with formal ties to the legislature would be on legislation that would alter parole and probation eligibility standards:

- \_\_\_\_\_ Importance of evidence from organizations with formal ties to your legislature, relative to personal and constituent beliefs about parole eligibility (1) [EIS]

Q3 Relative to your personal beliefs and those of your constituents, on a scale of 1 to 5 with 1 being "evidence is not important at all" and 5 being "evidence is very important," how important do you think evidence produced by organizations with formal ties to the legislature would be on legislation that would change funding levels for drug treatment programs, but neither create nor eliminate existing programs:

- \_\_\_\_\_ Importance of evidence from organizations with formal ties to your legislature, relative to personal and constituent beliefs about drug treatment funding (1) [EIS]

*Question set five – how evidence impacts legislative work*

Definition: For the purposes of answering these questions, please continue to consider "evidence" as the best available information related to the development of a particular policy and its implementation.

Q1 Next, you will find a list of statements concerning how you use a specific type of evidence -- that which is produced by organizations with formal ties to your legislature. Formal ties to the legislature means evidence produced by legislative staff and/or outside organizations that have formal partnerships with the state legislature. For each statement, please select if you "strongly disagree," "disagree," "agree," "strongly agree," or "don't know."

Evidence produced by organizations with formal ties to your legislature...

	Strongly Disagree (1)	Disagree (2)	Agree (3)	Strongly Agree (4)	Don't Know (5)
Is relevant to your work (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Provides an intellectual background or framework to your work (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Creates an opportunity for a common language to work with others (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Helps you reach across partisan divides (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Makes it easier to go against what your constituents want (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Makes it easier to go against what your party's leaders want (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Makes it easier to go against what legislative leaders want (7)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Makes you challenge your own views on legislation (8)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other: (9)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q2 On a scale of 1 to 5 with 1 being "never" and 5 being "all of the time," How often do you think evidence should play a prominent role in the decision-making process on public policy issues?

\_\_\_\_\_ How often evidence should have a prominent role in the public policy decision-making process (1)

Q3 Do you view the quality of evidence produced by organizations with formal ties to the state legislature as less than, equal to, or greater than the quality of other types of evidence?

☐ Less than (1)

- ☐ Equal to (2)
- ☐ Greater than (3)
- ☐ Don't know (4)

Q4 At what point in the policy making process do you utilize evidence? (check all that apply)

- ☐ Research (1)
- ☐ Bill drafting (2)
- ☐ Committee work (3)
- ☐ Floor debate (4)
- ☐ Policy implementation (5)

*Question set six – Results First*

Introduction: Now, changing focus to a specific type of evidence production called "Results First." As you may know, Results First describes itself as a program that works with states to implement an innovative cost-benefit analysis approach that helps state to invest in policies and programs that are proven to work.

Q1 On a scale of 1 to 5, with 1 being "not knowledgeable at all" and 5 being "very knowledgeable," how familiar are you with the Results First program?

\_\_\_\_\_ Results First familiarity (1)

Q2 On a scale of 1 to 5, with 1 being "no impact at all" and 5 being a "significant impact," when Results First produces evidence on an issue, to what extent does the evidence from Results First impact your decision-making on policy issues?

\_\_\_\_\_ Results First influence (1)

Q3 Results First is currently active in my state.

- ☐ Yes (1)
- ☐ No (2)
- ☐ Don't know (3)

Q4 I have personally relied upon Results First to help formulate a policy position.

- ☐ Yes (1)
- ☐ No (2)
- ☐ Don't know (3)

*Question set seven – demographics*

Introduction: To complete this survey, the last couple of questions seek to obtain some demographic information. This is a very important part of the data collection process.

Q1 Gender:

- ☐ Male (1)
- ☐ Female (2)
- ☐ Don't Know / Refuse (6)

Q2 I am a member of the following political party:

- ☐ Democratic (1)
- ☐ Republican (2)
- ☐ Independent (4)
- ☐ Other (8)
- ☐ Don't Know / Refuse (10)

Q3 I would describe myself as:

- ☐ Very liberal (1)
- ☐ Liberal (2)
- ☐ Moderate (3)
- ☐ Conservative (4)
- ☐ Very conservative (5)
- ☐ Don't Know / Refuse (6)

Q4 My level of education:

- ☐ Less than Bachelor degree (1)
- ☐ Bachelor degree (2)
- ☐ Master degree (3)
- ☐ J.D., Ph.D., or equivalent terminal degree (6)
- ☐ Don't Know / Refuse (7)

Q5 I would describe the competitiveness of my legislative district in the 2014 primary election as:

- ☐ I had a very competitive primary election (1)
- ☐ I had a somewhat competitive primary election (2)
- ☐ I faced no competition in the primary election (3)
- ☐ Don't Know / Refuse (4)

Q6 I would describe the competitiveness of my legislative district in the 2014 general election as:

- ☐ I had a very competitive general election (1)
- ☐ I had a somewhat competitive general election (2)
- ☐ I faced no competition in the general election (3)
- ☐ Don't Know / Refuse (4)



Q7 My race:

- ☐ Black / African-American (1)
- ☐ Asian / Pacific Islander (2)
- ☐ White / Caucasian (3)
- ☐ Hispanic / Latino (4)
- ☐ Multiracial / Other (5)
- ☐ American Indian (6)
- ☐ Don't Know / Refuse (7)

Q8 The following best describes my role relative to leadership positions within the legislature:

- ☐ In leadership, Presiding Officer (1)
- ☐ In leadership, Committee leader (2)
- ☐ In leadership, Party leader (3)
- ☐ Not in leadership (4)
- ☐ Don't Know / Refuse (5)

Q9 I plan to run for my current office or another office in an upcoming election cycle:

- ☐ Yes, I plan to run for this office in an upcoming election cycle. (1)
- ☐ Yes, I plan to run for another office in an upcoming election cycle. (2)
- ☐ No, I do not plan to run for this or another office in an upcoming election cycle. (3)
- ☐ I am unsure of my plans for the upcoming election cycles. (4)
- ☐ Don't Know / Refuse (5)

Q10 In the legislature, I would describe myself as more of a:

- ☐ Legislative generalist (1)
- ☐ Legislative specialist (2)
- ☐ Don't know (3)

Q11 I would describe my primary profession as:

- ☐ An attorney (1)
- ☐ An educator (2)
- ☐ A business professional (3)
- ☐ A health care professional (4)
- ☐ A full time legislator (5)
- ☐ A non-profit professional (6)
- ☐ Other (7)
- ☐ Don't Know / Refuse (8)

Q12 My age:

Q13 Total number of years I have served in the state legislature:

Q14 I am willing to schedule a follow-up interview to discuss factors of legislative decision-making in more detail to assist with this research.

☐ Yes (1)

☐ No (2)

If Yes Is Selected, then skip to Q15. If No is selected, then end survey and go to thank-you prompt.

Q15 I can be reached at the following phone number and/or e-mail address to set up a follow-up interview:

Following entry of this information, go to thank-you prompt.

## **Appendix G: Olszewski – Hird comparison of means**

As described in the footnote, the Hird mean scores were adjusted to the Olszewski scale of 1 – 5 by taking the Hird mean, dividing by 7, and then multiplying that percentage by 5. For example, university based research had a mean of 4.43 in Hird’s survey. Dividing that number by 7 (Hird’s scale was 1-7) gave a percentage of 62.14%. Thus, to translate to the Olszewski scale, the 4.43 score was multiplied by .6214 for an equivalent score of 3.16. Other notes and assumptions relevant to this comparison are:

- Constituents were not measured by Olszewski, nor Results First by Hird
- The Hird legislative staff score was an average of his partisan staff (2.97) and non-partisan staff (3.76) scores
- The Hird think tank score is an average of his independent think tanks (2.99) and partisan think tanks (2.97)
- The Hird “higher education” score is compared to the Olszewski “university based research” score
- The Hird “executive branch agencies” is compared to the Olszewski “state and federal agencies” score
- The Hird “legislative special committees / task forces” is compared to the Olszewski “legislature policy organization”
- The Hird “Governor” score is compared to the Olszewski “Governor’s office” score
- The Hird news media score is an average of his “print and broadcast media” score (2.70) and his “Internet / world wide web” score (2.42).

## Appendix H: Breakdown of responses on Results First questions

*Crosstabulation of Table 19 with Results First versus non-Results First status*

	Not a RF State	RF State	Total
Knowledge of RF   crosstab by RF state status			
1 - "Not knowledgeable at all"	24 (82.8%)	18 (64.3%)	42 (73.7%)
2	4 (13.8%)	7 (25%)	11 (19.3%)
3	1 (3.4%)	2 (7.1%)	3 (5.3%)
4	0 (0%)	1 (3.6%)	1 (1.8%)
5 - "Very knowledgeable"	0 (0%)	0 (0%)	0 (0%)
<i>Total</i>	29 (100%)	28 (100%)	57 (100%)

Pearson chi-squared = 2.992; p = .393

*Crosstabulation of Table 20 with Results First versus non-Results First status*

	Not a RF State	RF State	Total
Results First impact   crosstab by RF state status			
1 - "No impact at all"	13 (76.5%)	7 (46.7%)	20 (62.5%)
2	1 (5.9%)	2 (13.3%)	3 (9.4%)
3	1 (5.9%)	2 (7.1%)	6 (18.8%)
4	1 (5.9%)	1 (3.6%)	2 (6.3%)
5 - "Significant impact"	1 (5.9%)	0 (0%)	1 (3.1%)
<i>Total</i>	17 (100%)	15 (100%)	32 (100%)

Pearson chi-squared = 5.697; p = .225

*Crosstabulation of Table 21 with Results First versus non-Results First status*

	Not a RF State	RF State	Total
Results First active   crosstab by RF state status			
Yes	0 (0%)	1 (3.4%)	1 (1.7%)
No	6 (20%)	7 (24.1%)	13 (22%)
Don't know	24 (80%)	21 (72.4%)	45 (76.3%)
<i>Total</i>	30 (100%)	29 (100%)	59 (100%)

Pearson chi-squared = 1.260; p = .533

*Crosstabulation of Table 22 with Results First versus non-Results First status*

	Not a RF State	RF State	Total
Used Results First   crosstab by RF state status			
Yes	2 (6.7%)	0 (0%)	2 (3.4%)
No	21 (70%)	22 (78.6%)	43 (74.1%)
Don't know	7 (23.3%)	6 (21.4%)	13 (22.4%)

<i>Total</i>	30 (100%)	28 (100%)	58 (100%)
Pearson chi-squared = 2.034; p = .362			

*Note: none of the crosstabs resulted in statistically significant results.*

## **Appendix I: Sample interview questions mapped to research questions / hypotheses**

Thank you so much for taking time out of your very busy schedule to meet with me. As a former elected official in Maryland, I know how valuable your time is. As you may know, I am conducting research as part of my PhD dissertation project which seeks to better understand elected officials make decisions about policy questions confronting them. One of the issues I am especially interested in having additional conversations about is the role of evidence in your state. I have fourteen questions to ask you today. I also am leaving you with a document that outlines the details of this project, affirms that your identity will be kept confidential, and includes contact information for me and my dissertation advisor should you have any questions or concerns about the project. [Allow time to review document about project]. I am happy to answer any questions you may have. Once you are ready to proceed, please sign your name affirming that you give your informed consent to participate in this project.

Question 1: To start, can you tell me about your general philosophy when it comes to legislative decision-making? In other words, when making decisions, what are the things that you factor into your decision making?

[One of the issues you mentioned as a factor was evidence/information. OR: In addition to the issues you provided, others often time give evidence on the issue as a factor to consider].

Question 2: I would like to know a little more about the role you think that evidence should play in the legislative decision-making process. Specifically, I am curious to know what kinds of evidence you consider in your work?

Question 3: Is some evidence of better quality or more useful than others? Which kinds?

- Prompt: What kind of research do you believe produces quality evidence?

Question 4: When do you typically, if at all, consult evidence while conducting your legislative work? On which types of legislation are you most and least likely to use evidence?

Question 5: How do you reconcile when what the evidence says is different from what you believe and/or what your constituents want you to do?

Question 6: In your opinion, how do your colleagues typically consult evidence in their policymaking, if at all?

Question 7: Are there things that could be done that would make you use evidence more than you currently do? If so, what are those things?

### ***RESULTS FIRST STATES***

If Results First mentioned in question 6: You mentioned Results First in your response to the last question.

If Results First not mentioned in question 6: In addition to the places you've listed, there is also an organization called Results First that is working in and with your state on the issue of criminal justice.

Question 8a: Can you tell me anything about how Results First works and how they collect their data?

Question 9a: Do you have any feelings about the Results First program and how it operates?

Question 10a: Have you ever engaged with the work of Results First in this state? How so?

Question 11a: Do you consider the work of Results First to be similar or different than any other evidence presented to you as a lawmaker? If different, how so?

Question 12a: Can you tell me if or how it has had an impact on your decision-making?

Question 13a: Do you think you would have sought out the kind of evidence produced with Results First elsewhere if they were not in your state? If so, where?

### ***NON-RESULTS FIRST STATES***

In addition to the places you've listed, there is also an organization called Results First that is working in and with several other states on the issue of criminal justice.

Question 8b: Do you happen to know anything about how Results First works and how they collect their data?

Question 9b: Do you have any feelings about the Results First program and how it operates?

Question 10b: How about your colleagues, have they heard about it or do you have any sense of what they think about it?

Concluding question (Question 13 in RF states and Question 10 in non-RF states): To conclude our time together, I was wondering if there is anything unique about your state's use of evidence that we should be aware of as we consider this issue?

Thank you so much for your time. I am leaving you a sheet with my contact information should you have any questions or concerns about me or this work.

## Appendix J: Demographics of interview subjects by state

Below is a breakdown of the total number of interview subjects by state. The dissertation proposal sought to solicit six to eight legislators from each state as a part of the qualitative research. This would have yielded a total of between 36 and 48 interviews.

After considerable outreach, there were a total of 29 interviews, drawn from a mix of survey respondents, state legislative staff, Results First staff and Maryland legislators interviewed prior to the launch of the survey instrument. While the final number did not meet the range initially proposed, I am confident that the number, combined with survey data, is sufficient to support the analysis presented.

Initial goal for interview to be incorporated into the qualitative research:

State	Legislators
MA	6-8
RI	6-8
PA	6-8
NY	6-8
VA	6-8
WV	6-8
MD	0
<b>Totals</b>	<b>36 - 48</b>

While there were not Maryland survey data to compare the interview data against, it nonetheless made sense to keep the quality content generated through the pre-test interviews. Much of these conversations included the same topics covered with legislators in other states, making the insights equally as relevant.

In addition to expanding into Maryland, efforts were also made to inquire to legislative staff on their perceptions on the use of evidence by legislators within the legislative chambers. This was a recommendation of Dr. Laura Hussey during the defense process as a mechanism to validate responses and serve as a check to what is being shared by the legislators themselves. Adding staff interviews had the added value of driving up the number of total interviews into the range originally proposed. There was also one staff interview conducted with Sara Dube, the Director of the Results First program. Doing so provided important context from the program's perspective and allowed for questions to be asked that only the program itself was truly equipped to answer. Moreover, because of the decision to include Maryland legislator insights from the pre-survey interviews, it also made sense from a project as well as availability perspective to factor in staff interviews from the state.

If *all* Maryland interview surveys were removed (legislators and staff), there would still be a substantial number of interviews conducted, 22 in total. With the Maryland

interviews included, the number of interviews falls just below the total range initially proposed. The actual numbers are as follow:

Interview totals by state and legislator / staff

<b>State</b>	<b>Legislators</b>	<b>Staff</b>
MA	4	0
RI	3	0
PA	4	0
NY	3	1
VA	2	1
WV	2	1
MD	4	3
Results First	0	1
<b>Totals</b>	<b>22</b>	<b>7 (29 total)</b>



## Appendix K: States and local governments in the Results First program

### *States*

Alaska	Iowa	Oregon
California	Kansas	Pennsylvania
Colorado	Massachusetts	Rhode Island
Connecticut	Minnesota	Texas
Delaware	Mississippi	Vermont
Florida	Nevada	West Virginia*
Idaho	New Mexico	Wisconsin
Illinois	New York	

### *Counties*

Allegheny County, Pennsylvania	Santa Barbara County, California
Fresno County California	Santa Clara County, California
Kern County, California	Santa Cruz, California
Montgomery County, Maryland	Ventura County, California

Source: <http://www.pewtrusts.org/en/projects/pew-macarthur-results-first-initiative/where-we-work>

## Bibliography

- American Legislative Exchange Council (ALEC). 2017. "Resolution in support of justice reinvestment." Retrieved on 01 June 2017 from <https://www.alec.org/model-policy/resolution-in-support-of-justice-reinvestment/>
- Aos, Steve. 26 June 2012. "Improve Outcomes & Save Money: Switching from Lower to Higher Return-on-Investment Programs and Policies – Washington State's (Evolving) Approach." Presented at the Working Family Conference of the National Conference of State Legislatures. Denver, CO.
- Alperin, David J. and David Schultz. 2003. "E-democracy: Legislative-Constituent Communications in Minnesota and Wisconsin." In Philadelphia, PA.
- Barnard, Chester. 1938. *The functions of the executive*. Harvard University Press: Cambridge, MA.
- Baum, Erica B. 1991. "When the Witch Doctors Agree: The Family Support Act and Social Science Research." *Journal of Policy Analysis and Management*. Volume 10, Number 4. Pages 603 – 615.
- Behn, Robert. 2003. "Why measure performance? Different purposes require different measures." *Public Administration Review*. Volume 63, Number 5, pages 586 – 606.
- Behn, Robert. 2005. "The Accountability Dilemma." *Bob Behn's Public Management Report*. Volume 2, Number 7. Harvard University. Accessed 20 July 2014 from: <http://www.hks.harvard.edu/thebehnreport/March2005.pdf>
- Belluz, Julia and Steven Hoffman. 28 January 2015. "Inside the international effort to fund government programs that actually work." Vox. Retrieved from <http://www.vox.com/2015/1/28/7880203/evidence-based-policy> on 12 February 2015.
- Berry, Jeffrey M. 2002. "Validity and Reliability Issues in Elite Interviewing." *Political Science and Politics*. Volume 35, Number 4, pages 679 – 682.
- Black, Nick. August 4, 2001. "Evidence based policy: proceed with care." From the *British Journal of Medicine*. Volume 323, pages 275 – 279. Retrieved 12 December 2014 from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1120888/>.
- Binder, Sarah A. 1999. "The Dynamics of Legislative Gridlock, 1947–96." *American Political Science Review*. Volume 93, pages 519-533.
- Bovens, Mark. 2005. "Public accountability." In *The Oxfords Book of Public Management*, edited by E.Ferlie, L.E. Lynn, Jr., and C. Pollitt. New York: Oxford University Press. Pages 182 – 208.
- Bowen, Daniel and Zachary Greene. 2014. "Should We Measure Professionalism with an Index? A Note on Theory and Practice in State Legislative Professionalism Research." *State Politics and Policy Quarterly*. Volume 14, Number 3, pages 277 – 296.
- Bratton, Kathleen A. and and Kerry L. Haynie. 1999. "Agenda Setting and Legislative Success in State Legislatures: The Effects of Gender and Race." *The Journal of Politics*. Volume 61, pages 658-679.

- Bridgeland, John and Peter Orszag. 2013. "Can Government Play Moneyball?" *The Atlantic*. Retrieved from <http://www.theatlantic.com/magazine/archive/2013/07/can-government-play-moneyball/309389/> on 03 February 2015.
- Brookings Institute. 2015. "Brookings Institution Press: Systematic Thinking for Social Action." Retrieved from <http://www.brookings.edu/research/books/1971/systematicthinkingforsocialaction> on 02 January 2015.
- Buechner Institute for Governance. 2014. "Advocacy Coalition Framework Overview." University of Colorado Denver School of Public Affairs. Accessed 24 July 2014 from [http://www.ucdenver.edu/academics/colleges/SPA/BuechnerInstitute/Center s/WOPPR/ACF/Pages/ACFOverview.aspx](http://www.ucdenver.edu/academics/colleges/SPA/BuechnerInstitute/Center%20s/WOPPR/ACF/Pages/ACFOverview.aspx)
- Burke, Edmond. 1774. "Speech to the Electors of Bristol." *The Founders' Constitution*. Volume 1, Chapter 13, Document 7. Chicago University Press, 1987. Retrieved from <http://press-pubs.uchicago.edu/founders/documents/v1ch13s7.html> on 21 July 2014.
- Canfield-Davis, Kathy, Jain Sachin, Don Wattam, Jerry McMurty and Mike Johnson. 2009. "Factors of Influence on Legislative Decision Making: a Descriptive Study." *Journal of Legal, Ethical and Regulatory Issues*.
- Carey, J., M. Niemi, R.G. Powell, L.W. and Moncreif, G.F. 2006. "The Effects of Term Limits on State Legislatures: A New Survey of the 50 States." *Legislative Studies Quarterly*. Volume 31, pages 105 – 134.
- Chambers, Chris and Natalia Lawrence. 18 March 2014. "We aim to put research evidence on tap for UK politicians." *The Guardian*. Accessed June 05 2014 from <http://www.theguardian.com/science/2014/mar/18/research-scientific-evidence-information-service-politicians-eis>
- Coalition for Evidence-Based Policy. "Home Page." Retrieved 13 March 2014 from: <http://coalition4evidence.org/>
- Congressional Research Service (CRS). 2006. "Congress and Program Evaluation: An Overview of Randomized Controlled Trails (RCTs) and Related Issues." Clinton Brass, Blas Nunez-Neto, and Erin Williams. Library of Congress.
- Cook, C., Heath, F., and R. L. Thompson. 2000. "A meta-analysis of response rates in web-or Internet based surveys." *Educational and Psychology Measurement*. Volume 60, Number 6, pages 821 – 836.
- Council of State Government. January 2012. "Justice Reinvestment in Pennsylvania: Overview." The Justice Center – Council of State Governments. Retrieved from <http://www.cor.pa.gov/General%20Information/Documents/Justice%20Reinvestment%20Initiative/JR%20Pennsylvania%20State%20Overview.pdf> on 02 August 2017.
- Council of State Governments. 2017. "Justice Center: Massachusetts." Retrieved 06 June 2017 from: <https://csgjusticecenter.org/jr/ma/>
- Council of State Governments. 2017. "Justice Center: Pennsylvania." Retrieved 18 May 2017 from: <https://csgjusticecenter.org/jr/pa/>
- Council of State Governments. 2017. "Justice Center: Rhode Island." Retrieved 02 June 2017 from: <https://csgjusticecenter.org/jr/ri/>
- Dagan, David and Steven Teles. 2016. *Prison Break: Why Conservatives Turned Against Mass Incarceration*. New York, NY: Oxford University Press.

- Dahl, Robert. 1957. "The concept of power." *Systems Research and Behavioral Science*. Volume 2, Issue 3. Pages 201 – 215.
- Davies, Phillip. 1999. "What is evidence based education?" *British Journal of Educational Studies* Volume 47, Number 2, Pages 108-121.
- Department of Legislative Services (DLS). 2015. "Fiscal and Policy Analysis of Proposed Legislation." Maryland General Assembly. Retrieved from <http://dls.state.md.us/Content.aspx?page=84> on 03 January 2015.
- Deschenaux, Warren. Information provided by e-mail exchange on 31 May 2017.
- Dillman, Don, Jolene Smyth, and Leah Melani Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4<sup>th</sup> Edition. Hoboken, NJ: Wiley.
- Donovan, Doug. 21 January 2017. "Governor Hogan seeks to lift veil of secrecy on lobbyist parties for Maryland officials." *The Baltimore Sun*. Retrieved 31 May 2017 from: <http://www.baltimoresun.com/news/maryland/sun-investigates/bs-md-sun-investigates-protocol-calendar-20170121-story.html>.
- Downs, Anthony. 1967. *Inside Bureaucracy*. Boston: Little, Brown, and Company.
- Dreyfus, D. 1977. "The limitations of policy research in congressional decision-making." In C.H. Weiss (Ed.), *Using social science research in public policy making*. Pages 99 – 107. Lexington: DC Health.
- Dube, Sara and Darcy White. January 31, 2017. "Key Findings from 50-State Assessment of Evidence-Based Policymaking." The Pew Charitable Trusts. Retrieved from <http://www.pewtrusts.org/en/research-and-analysis/analysis/2017/01/31/key-findings-from-50-state-assessment-of-evidence-based-policymaking> on 15 February 2017.
- Economist*. 12 December 2015. "In praise of human guinea pigs." Page 18.
- Ella Baker Center for Human Rights. April 29, 2014. "What is Justice Reinvestment?" Retrieved from <http://ellabakercenter.org/blog/2014/04/what-is-justice-reinvestment> on 02 July 2017.
- Eulau, Heinz, John C. Wahlke, William Buchanan, and Leroy C. Ferguson. 1959. "The Role of the Representative: Some Empirical Observations on the Theory of Edmund Burke." *American Political Science Review* 53(3): 742- 756.
- Feldman, Andy. April 10, 2013. "Washington State, a leader in the use of cost-benefit analysis: An interview with Steve Aos, Director, Washington State Institute for Public Policy." GovInnovator Podcast. Episode 17. Accessed from <http://govinnovator.com/steve-aos/> on 02 February 2017.
- Fenno, Richard F. 1973. *Congressmen in committees*. Boston: Little Brown.
- Fenno, Richard F. 1977. "US House Members in Their Constituencies: An Exploration." *The American Political Science Review*, Volume 71, Number 3, pages 883-917.
- Fienberg, Stephen E. and Judith M. Tanur. 1989. "Combining Cognitive and Statistical Approaches to Survey Design." *Science*. Volume 243, pages 1007 – 1022.
- Fiorina, Morris P. 1994. "Divided Government in the American States: A Byproduct of Legislative Professionalism?" *American Political Science Review*. Volume 88, pages 304-16.
- Fisher III, Samuel and Rebekah Herrick. 2009. "Old Versus New: The Comparative Efficiency of Mail and Internet Surveys." Prepared for presentation at the APSA meeting Sept. 3-6, 2009. Toronto, Canada.

- Fisher III, Samuel and Rebekah Herrick. 2012. "Old versus New: The Comparative Efficiency of Mail and Internet Surveys of State Legislators." *State Politics and Policy Quarterly*. Volume 13, Number 2, pages 147 – 163.
- Forester, John. 1984. "Bounded Rationality and the Politics of Muddling Through." *Public Administration Review*. January / February 1984.
- Furubo, J.E. 1994. "Learning from evaluation: The Swedish experience." In F.L. Leeuw, R.C. Rist, and R.E. Sonnichsen (eds). *Can governments learn? Comparative perspectives on evaluation and organizational learning*. Pages 45 – 65. New Brunswick: Transaction.
- Gable, Guy. 1994. "Integrating case study and survey research methods: an example in information technology." *European Journal of Information Systems*. Volume 3, Number 2. Pages 112 – 126.
- Gordon, Sanford C. and Gregory A. Huber. 2007. "The Effect of Electoral Competitiveness on Incumbent Behavior." *Quarterly Journal of Political Science*. Volume 2, Pages 107 – 138.
- Governor Charlie Baker. 21 February 2017. "State Leaders Release Report on Criminal Justice Reform Measures." The Official Website of the Governor of Massachusetts. Retrieved on 04 August 2017 from <http://www.mass.gov/governor/press-office/press-releases/fy2017/report-released-on-criminal-justice-reform-measures.html>.
- Gow, J.I. and V. Seymour Wilson. March 2014. "Speaking what truth to whom? The uneasy relationship between practitioner and academic in public administration." *Canadian Public Administration*. Volume 57, Number 1, pages 118 – 137.
- Gray, Virginia and David Lowery. 2000. "Where Do Policy Ideas Come From? A Study of Minnesota Legislators and Staffers." *Journal of Public Administration Research and Theory*. 10 (3): 573 – 97.
- Green, Miranda. "The battle over science in the Trump administration." CNN Politics, 09 August 2017. Retrieved 08 September 2017 from: <http://www.cnn.com/2017/08/05/politics/trump-battle-science-epa-energy-climate/index.html>
- Greenberg, David, Mavin Mandell and Matthew Onstott. 2000. "The Dissemination and Utilization of Welfare-to-Work Experiments in State Policymaking." *Journal of Policy Analysis and Management*. 19 (3): 367-82.
- Gueron, Judith M. and Howard Rolston. 2013. *Fighting for Reliable Evidence*. New York: Russell Sage Foundation.
- Hamilton, Laura. 2005. "Lessons from performance measurement in education." In *High-Performance Government: Structure, Leadership, and Incentives*, edited by R. Klitgaard and P.C. Light. Santa, Monica, CA: RAND Corporation. Pages 381 – 404.
- Hamm, Keith E. and Gary F. Moncrief. 2012. "Legislative Politics in the States." *Politics in the American States: A Comparative Analysis*. Pages 163 – 207.
- Haskins, Ron. 2014. "Social Programs That Work." *The New York Times*. Retrieved from [http://www.nytimes.com/2015/01/01/opinion/social-programs-that-work.html?\\_r=0](http://www.nytimes.com/2015/01/01/opinion/social-programs-that-work.html?_r=0).

- Haskins, Ron, Christina Paxson, and Jeanne Brooks-Gunn. 2009. "Social Science Rising: A Tale of Evidence Shaping Public Policy." From the Future of Children. Princeton Brookings.
- Haskins, Ron and Greg Margolis. 2015. *Show me the Evidence: Obama's Fight for Rigor and Results in Social Policy*. Brookings Institution Press: Washington, DC.
- Haskins, Ron and Jon Baron. 2011. "Building the Connection between Policy and Evidence: The Obama evidence-based initiatives." NESTA.
- Heckman, J.J. and Smith, J.A. 1995. "Assessing the case for social experiments." *Journal of Economics Perspectives*. Volume 9, Pages 85 – 110.
- Heinrich, Carolyn J. 2007. "Evidence-Based Policy and Performance Management: Challenges and Prospects in Two Parallel Movements." *American Review of Public Administration*. Volume 37, Number 3. Pages 255 – 277.
- Hickman, Dick. "Crime and Corrections: NCSL Fiscal Analysts Seminar." Presented 30 August 2014 in Salt Lake City, Utah. Retrieved 05 June 2017 from: <http://www.ncsl.org/documents/nalfo/NCSLDickHickman.pdf>
- Hird, John A. 2005. *Power, Knowledge and Politics: Policy Analysis in the States*. Washington, DC: Georgetown University Press.
- Hird, John A. 2009. "The Study and Use of Policy Research in State Legislatures." *International Regional Science Review*. Volume 32, Number 4.
- Hoadland, William G., Edward "Sandy" Davis, Tim Shaw and Daniel Martins. May 2017. "Congress and Evidence-Based Policymaking: Creating a 21<sup>st</sup> Century Legislature." Bipartisan Policy Center.
- Horowitz, Donald L. 1977. *The Courts and Social Policy*. Brookings Institution: Washington D.C.
- Human Society Legislative Fund. 2017. "What Influences a Legislator?" Retrieved 03 September 2017 from <http://www.hslf.org/our-work/taking-action/be-a-citizen-lobbyist/what-influences-legislators.html?referrer=https://search.yahoo.com/#.Wa2VT7KGOM8>
- Institute of Education Sciences (IES). "What Works Clearinghouse." <http://ies.ed.gov/ncee/wwc/>
- Institute of Education Sciences (IES). "About IES: Connecting Research, Policy, and Practice." <http://ies.ed.gov/aboutus/>
- Jennings, Edward T. and Jeremy L. Hall. 2011. "Evidence-Based Practice and the Use of Information in the State Agency Decision Making." *Journal of Public Administration Research and Theory*. Oxford University Press.
- Jewell, Christopher and Lisa Bero. June 2008. "Developing Good Taste in Evidence: Facilitators of and Hindrances to Evidence-Informed Health Policymaking in State Government." *The Milbank Quarterly*. Vol. 86, No. 2. Pages 177 – 208.
- Jewell, Malcolm E. and Marcia Whicker. 1994. *Legislative Leadership in the American States*. Ann Arbor: University of Michigan Press.
- JLARC: Joint Legislative Audit and Review Commission. 2014. "About JLARC." Virginia General Assembly. Retrieved 24 July 2014 from <http://jlarc.virginia.gov/about.shtml>
- JLARC: Joint Legislative Audit and Review Commission. 2017. "JLARC authority in statute." Retrieved 29 May 2017 from: <http://jlarc.virginia.gov/authority.asp>



- Jones, Bryan D. 2001. *Politics and the Architecture of Choice*. University of Chicago Press: Chicago
- Jones, C.O. 1976. "Why Congress can't do policy analysis." *Policy Analysis*. Volume 2. Pages 251 – 264.
- Joyce, Philip. 1993. "Using performance measures for federal budgeting: Proposals and prospects." *Public Budgeting and Finance* 13, number 4. Pages 3 – 17.
- Joyce, Philip. 2011. "The Obama Administration and PBB: Building on the Legacy of Federal Performance-Informed Budgeting?" *Public Administration Review*. May/June 2011, pages 356-367. Retrieved 08 September 2017 from: <http://faculty.cbpp.uaa.alaska.edu/afgjp/PADM628%20Spring%202012/Obama%20Adm%20and%20PBB.pdf>.
- Jovchelovitch, Sandra and Martin W. Bauer. 2000. "Narrative interviewing." *LSE Research Online*. Retrieved 14 February 2015 from <http://eprints.lse.ac.uk/2633>
- Justice Policy Institute. "JPI Statement on passage of Maryland's Justice Reinvestment Act." 12 April 2016. Retrieved from <http://www.justicepolicy.org/news/10312> on 25 May 2017.
- Kabler, Phil. January 22, 2013. "New Study: Fix W. Va. Prisons this way." *Charleston Gazette-Mail*. Retrieved from <http://www.wvgazettemail.com/News/201301220093> on 01 August 2017.
- Key, V.O., Jr. 1961. *Public Opinion and American Democracy*. New York: Knopf.
- Kolata, Gina. September 2013. "Guesses and Hype Give Way to Data in Study of Education." *The New York Times*.
- Kingdon, John W. 1989. *Congressmen's Voting Decisions*. 3<sup>rd</sup> Edition. Ann Arbor: University of Michigan Press.
- Kirp, David L. 2007. *The Sandbox Investment: The Preschool Movement and Kids-First Politics*. Cambridge, Massachusetts: Harvard University Press.
- Kuklinski, James H. 1978. "Representativeness and Elections: A Policy Analysis." *American Political Science Review*. Volume 72, pages 165-177.
- Lachman, Seymour P and Robert Polner. 2006. *Three Men in a Room: The Inside Story of Power and Betrayal in an American Statehouse*. New York: The New Press.
- Lampinen, O. 1992. *The utilization of social science research in public policy*. Helsinki, Finland: VAPK-Publishing.
- La Pira, Timothy La Pira. 2008. "Information as Lobbying, or Lobbying as Information? Argument Quality, Group Credibility, and Heuristic Processing in Congress." Paper presented at the Midwest Political Science Association Annual National Meeting.
- Lawrence, Alison. 01 January 2017. "Justice Reinvestment: State Resources." Retrieved from the NCSL website on 30 May 2017 from: <http://www.ncsl.org/research/civil-and-criminal-justice/justicereinvestment.aspx>
- Larson, Judith K. 1980. "Knowledge Utilization: What Is It?" *Knowledge: Creation, Diffusion, Utilization* 1(3): 421 – 42.
- Leachman, Michael, Inimai M. Chettiar, and Benjamin Geare. 11 January 2012. "Improving Budget Analysis of State Criminal Justice Reforms: A Strategy for Better Outcomes and Saving Money." Center on Budget Priorities. Retrieved on 21 February 2015 from <http://www.cbpp.org/files/1-11-12sfp.pdf>.

- Lee, S., Aos, S., Drake, E., Pennucci, A., Miller, M., and L. Anderson. 2012. *Return on Investment: Evidence-based options to improve statewide outcomes*. (Document No. 12-04-1201). Olympia: Washington State Institute for Public Policy.
- Lewis, David E. 2007. "Testing Pendleton's Premise: Do Political Appointees Make Worse Bureaucrats?" *Journal of Politics*. Volume 69, number 4. Pages 1073 – 1088.
- Lindblom, Charles E. 1959. "The Science of 'Muddling Through.'" *Public Administration Review*. Volume 19, Number 2, pages 79 – 88.
- Lindblom, Charles E. 1980. *The Policy making process*. 2<sup>nd</sup> Edition. Englewood Cliffs, NJ: Prentice-Hall.
- Lupia, Arthur and Mathew D. McCubbins. 1994. "Who Controls? Information and the Structure of Legislative Decision Making." *Legislative Studies Quarterly*. Volume 19, Number 3, pages 361 – 384.
- Maddox, Jerome. 2004. "Opportunity Costs and Outside Careers in the U.S. State Legislatures." *Legislative Studies Quarterly*. Volume 29, pages 517 – 544.
- Marra, Ashton. 12 June 2014. "State Making Progress in Reducing Prison Population." West Virginia Public Broadcasting. Retrieved 15 June 2017 from: <http://wvpublic.org/post/state-making-progress-reducing-prison-population#stream/0>.
- Maynard, Rebecca. 2006. "Presidential Address – Evidence-based Decision Making: What Will It Take for the Decision Makers to Care?" *Journal of Policy Analysis and Management*. Volume 25, Number 2. Pages 249 – 265.
- Mayhew, David R. 1974. *Congress: The Electoral Connection*. Fredericksburg: Yale University Press.
- Mayhew, David R. 1991. *Divided We Govern*. 1<sup>st</sup> Edition. Fredericksburg: Yale University Press.
- Masket, Seth E. 2008. "Where You Sit is Where You Stand: The Impact of Seating Proximity on Legislative Cue-Taking." *Quarterly Journal of Political Science*. Volume 3, Pages 301 – 311.
- Matkin, David S. and H. George Frederickson. 2009. "Metropolitan Governance: Institutional Roles and Interjurisdictional Cooperation." *Journal of Urban Affairs*. Volume 31, Number 1. Pages 45 – 66.
- Maxwell, Lesli. December 21, 2016. "Head Start Advantages Mostly Gone by 3<sup>rd</sup> Grade, Study Finds." *Education Week*. Retrieved from [http://blogs.edweek.org/edweek/early\\_years/2012/12/head\\_start\\_advantages\\_mostly\\_gone\\_by\\_third\\_grade\\_study\\_finds.html](http://blogs.edweek.org/edweek/early_years/2012/12/head_start_advantages_mostly_gone_by_third_grade_study_finds.html) on 01 September 2017.
- McCubbins, Mathew D. and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols Versus Fire Alarms." *American Journal of Political Science*. Volume 28, Issue 1. Pages 165 – 179.
- Meha, Natasha. February 26, 2016. "Why 'Justice Reinvestment' Might Be The Biggest Issue of the Session." Maryland Association of Counties (MACo). Retrieved on 12 May 2017 from: <https://conduitstreet.mdcountries.org/2016/02/26/why-justice-reinvestment-mightbe-the-biggest-issue-of-the-session>
- Melkers, Julia and Katherine Willoughby. 2004. "Staying the Course: The use of Performance Management in State Government." IBM Center for the Business of Government.



- Merriam-Webster. 2015. "Empirical." Online dictionary. Retrieved from: <http://www.merriam-webster.com/dictionary/empirical> on 03 January 2015.
- Mick, David G. 1996. "Are Studies of Dark Side Variables Confounded by Socially Desirable Responding? The Case of Materialism." *Journal of Consumer Research*. Volume 23, September, pages 106 – 119.
- Moncrief, Gary, Joel A. Thompson, and Karl Kurtz. 1996. "The Old Statehouse, It Ain't What It Used to Be." *Legislative Studies Quarterly*. Volume 21, pages 57-72.
- Moncrief, Gary, Richard Niemi, and Lynda Powell. 2008. "Turnover in State Legislatures: An Update." Paper presented to the annual meeting of the Western Political Science Association, San Diego, March.
- Mooney Chris. 2005. *The Republican War on Science*. Basic Books: New York.
- Mooney, Chris. 05 May 2011. "What is Motivated Reasoning? How Does it Work?" Discover: Science for the Curious. Accessed 15 December 2014 from: <http://blogs.discovermagazine.com/intersection/2011/05/05/what-is-motivated-reasoning-how-does-it-work-dan-kahan-answers/#.VJCFLivF-o9>.
- Moynihan, Donald P. 2005. "Goal-Based Learning and the Future of Performance Management." *Public Administration Review*. Volume 65, Number 2, pages 203 – 216.
- Moynihan, Donald P. 2008. *The Dynamics of Performance Management: Construction information and reform*. Georgetown University Press: Washington DC.
- National Conference of State Legislatures (NCSL). 2009. "Size of State Legislative Staff." <http://www.ncsl.org/research/about-state-legislatures/staff-change-chart-1979-1988-1996-2003-2009.aspx>.
- National Conference of State Legislatures (NCSL). June 1, 2014. "Full- and Part-Time Legislatures." Retrieved from <http://www.ncsl.org/research/about-state-legislatures/full-and-part-time-legislatures.aspx> on 06 January 2015.
- National Conference of State Legislatures (NCSL). February 4, 2015. "2015 State and Legislative Partisan Composition." Retrieved from [http://www.ncsl.org/Portals/1/Documents/Elections/Legis\\_Control\\_2015\\_Feb4\\_11am.pdf](http://www.ncsl.org/Portals/1/Documents/Elections/Legis_Control_2015_Feb4_11am.pdf) on 15 February 2015.
- National Conference of State Legislatures (NCSL). September 2015. "Who We Elect: an Interactive Graphic." Retrieved from <http://www.ncsl.org/research/about-state-legislatures/who-we-elect-an-interactive-graphic.aspx#> on 15 December 2016.
- Newman, Kirsty, Antonio Capillo, Akin Famurewa, Chandirka Nath, and Willie Siyanbola. 2013. "What is the evidence on evidence-informed policy making?" International Network for the Availability of Scientific Publications.
- New State Ice Co. v. Liebmann*, 285 U.S. 262, 311 (1932).
- Nickerson, Raymond S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." *Review of General Psychology*. Volume 2, Number 2. Pages 175 – 220.
- Nisbet, Eric, R. Kelly Garrett and Kathryn Cooper. March 2015. "The Partisan Brain: How Dissonant Science Messages Lead Conservatives and Liberals to (Dis)Trust Science." *The Annals of the American Academy of Political and Social Science*. Pages 36 – 66.
- Nussle, Jim and Peter Orszag, eds. 2014. *Moneyball for Government*. United States: Disruption Books.
- Mooney, Chris. 2005. *The Republican War on Science*. New York: Basic Books.

- Oliver, Kathryn, Simon Innvar, Theo Lorenc, Jenny Woodman, and James Thomas. 2014. "A systematic review of barriers to and facilitators of the use of evidence by policymakers." BioMed Central. Retrieved 02 July 2014 from <http://www.biomedcentral.com/1472-6963/14/2>.
- Orszag, Peter. "Building Rigorous Evidence to Drive Policy." 2009. Office of Management and Budget. Retrieved 14 March 2014 from [www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidenceToDrivePolicy](http://www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidenceToDrivePolicy).
- Paulhus, Delroy. 2002. "Socially Desirable Responding: The Evolution of a Construct," in *The Role of Constructs in Psychological and Educational Measurement*, H.I. Braun, D.N. Jackson and D.E. Wiley, eds., Mahwah, NJ: Erlbaum, pages 49 – 69.
- Pennsylvania Justice System Joint Performance Measures. 2017. Retrieved 10 March 2017 from: <https://isra.hbg.psu.edu/paisipm/Dashboard/JRI/tabid/2042/Default.aspx>
- Petrosino, Anthony, Robert F. Boruch, Haluk Soydan, Lorna Duggan, and Julio Sanchez-Meca. 2001. "Meeting the Challenges of Evidence-Based Policy: The Campbell Collaboration." *The Annals of the American Academy*. Volume 579, November 2001, pages 14 – 34.
- Pew Charitable Trusts. 2017. "Pew-MacArthur Results First Initiative." <http://www.pewstates.org/projects/pew-macarthur-results-first-initiative-328069>. Accessed 31 October 2017.
- Pew Charitable Trusts. July 2013. "States' Use of Cost-Benefit Analysis: Improving Results for Taxpayers."
- Pew Charitable Trusts. 2014. "Achieving Success with the Pew – MacArthur Results First Initiative: A State Progress Report for 2011 – 2013." February 2014, pages 1-8.
- Pew Charitable Trusts. November 2014. "Evidence-Based Policymaking: A guide for effective government." Retrieved 02 December 2014 from: <http://www.pewtrusts.org/~media/Assets/2014/11/EvidenceBasedPolicymakingAGuideforEffectiveGovernment.pdf>
- Pew Charitable Trusts. April 3, 2015. "Public Safety in West Virginia." Retrieved on 10 June 2017 from <http://www.pewtrusts.org/en/research-and-analysis/fact-sheets/2012/09/28/public-safety-in-west-virginia>
- Pew Charitable Trusts. June 1, 2015. "Video: Results First in New York: Improving public safety through cost-benefit analysis." Retrieved 05 March 2017 from <http://www.pewtrusts.org/en/multimedia/video/2015/results-first-in-new-york-improving-public-safety>
- Pew Charitable Trusts. 23 June 2015. "Results First Clearinghouse Database." Retrieved from <http://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2014/09/results-first-clearinghouse-database> on 04 September 2017.
- Pitkin, Hanna Fenichel, ed. 1969. *Representation*. New York: Atherton Press.
- Presser, Stanley and Howard Schuman. 1980. "The Measurement of a Middle Position in Attitude Surveys." *Public Opinion Quarterly*. Volume 44, Issue 1. Pages 70 – 85. Retrieved 17 September 2017 from: [http://www2.amstat.org/sections/SRMS/Proceedings/papers/1978\\_008.pdf](http://www2.amstat.org/sections/SRMS/Proceedings/papers/1978_008.pdf).

- Quorum Analytics. 2017. "2017 Partisan Voting Index Reveals Increasingly Polarized Congressional Districts." Retrieved 28 May 2017 from: <http://info.quorum.us/2017-partisan-voting-index-reveals-increasingly-polarized-congressional-districts>.
- Rhode Island State Government. July 7, 2015. "Raimondo Launches Working Group to Improve Criminal Justice System, Reduce Costs." Retrieved 04 June 2017 from <http://www.ri.gov/press/view/25229>
- Riccucci, Norma. 2010. *Public Administration Traditions of Inquiry and Philosophies of Knowledge*. Washington DC: Georgetown University Press.
- Riccucci, Norma and Katherine Naff. 2008. *Personnel Management in Government: Politics and Process*. 6<sup>th</sup> Edition. New York: CRC Press.
- Right Wing Watch. July 2011. "ALEC: The Voice of Corporate Special Interests in State Legislatures." Retrieved on 10 July 2017 from <http://www.rightwingwatch.org/report/alec-the-voice-of-corporate-special-interests-in-state-legislatures/>.
- Rivlin, Alice M. 1971. *Systematic Thinking for Social Action*. Washington, DC: Brookings Institution Press.
- Sabatier, Paul A, editor. 1999. *Theories of the Policy Process*. 2<sup>nd</sup> Edition. Boulder, CO: Westview Press.
- Schoenberg, Shira. 22 December 2016. "Boston senator 'made nauseous' by criminal justice working group." MassLive.com. Retrieved 01 August 2017 from [http://www.masslive.com/politics/index.ssf/2016/12/boston\\_senator\\_made\\_nauseous\\_b.html](http://www.masslive.com/politics/index.ssf/2016/12/boston_senator_made_nauseous_b.html).
- Sanbonmatsu, Kira. 2006. *Where Women Run: Gender and Party in the American States*. Ann Arbor: University of Michigan Press.
- Seawright, Jason and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly*. Volume 61, Number 2. Pages 294-308.
- Shor, Bois and Nolan McCarty. 2011. "The Ideological Mapping of American Legislatures." *American Political Science Review*. Volume 105, Number 3. Pages 530-550.
- Simon, Herbert A. 1999. "Bounded Rationality in Social Science: Today and Tomorrow." *Mind and Society*. Volume 1, pages 25 – 39.
- South Carolina Legislature. 2013. "South Carolina Legislative Council." <http://www.scstatehouse.gov/council.php>. Accessed 08 December 2013.
- Stone, Deborah. 2002. *Policy Paradox: The Art of Political Decision Making*. New York: WW Norton & Company.
- Squire, Peverill 1992. "Legislative Professionalization and Membership Diversity in State Legislatures." *Legislative Studies Quarterly*. Volume 17, pages 69-79.
- Squire, Peverill and Keith Hamm. 2005. *101 Chambers: Congress, State Legislatures, and the Future of Legislative Studies*. Columbus: Ohio State University Press.
- Squire, Peverill. 2007. "Measuring State Legislative Professionalism: The Squire Index Revisited." *State Politics and Policy Quarterly*. Volume 7, Number 2, pages 211 – 227.

- Talbot, Colin and Carole Tablot. April 2014. "Sir Humphrey and the professors: What does Whitehall want from academics?" University of Manchester. Accessed 13 May 2014 from [www.manchester.ac.uk/policy](http://www.manchester.ac.uk/policy)
- Thevenau, Sealia. 27 January 2015. "Slider Questions – How Useful Are They and How Should They be Used?" Eval and Go. Retrieved 24 February 2015 from <http://www.evalandgo.com/this/>.
- Thomas, Sue. 1991. "The Impact of Women on State Legislative Policies." *The Journal of Politics*. Volume 53, pages 958-976.
- Thomassen, Jacques and Hermann Schmitt, Eds. 1999. *Political representation and legitimacy in the European Union*, pp. 3-21. Oxford: Oxford University Press.
- Tourangeau, Roger and Ting Yan. 2007. "Sensitive Questions in Surveys," *Psychological Bulletin*, 133 (5), pages 859 – 883.
- Tversky, Amos and Daniel Kahneman. 1971. "Belief in the Law of Small Numbers." *Psychological Bulletin*. Vol. 76, No. 2. Pages 105 – 110. Hebrew University Of Jerusalem.
- United States Census. 2015. "Quickfacts by state." Retrieved from <http://quickfacts.census.gov/qfd/index.html> on 04 January 2015.
- United States Committee on the Budget. 20 November 2014. "Murray and Ryan Propose Commission to Promote Evidence-Based Policymaking." House of Representatives. Accessed 22 November 2014 from: <http://budget.house.gov/news/documentsingle.aspx?DocumentID=393882>
- United States Department of State. 2008. "Program Assessment Rating Tool (PART). FY 2008 Performance Summary." Retrieved on 31 May 2017 from <https://www.state.gov/documents/organization/79676.pdf>.
- University of South Carolina (USC). 2013. "Governmental Research and Service Technical Assistance." <http://www.ipspr.sc.edu/grs/technical.asp>. Accessed 08 December 2013.
- Urahn, Susan K. "The Cost-Benefit Imperative." February 1, 2012. *Governing Magazine*. Accessed 25 January 2017 from: <http://www.governing.com/columns/mgmt-insights/col-cost-benefit-outcomes-states-results-first.html>.
- Vera Institute for Justice. May 2017. "The Price of Prisons: Examining State Spending Trends" Retrieved 02 August 2017 from <https://www.vera.org/publications/price-of-prisons-2015-state-spending-trends#prisonPopulationAndSpending20102015>.
- Weinberg, H. 1979. "Using policy analysis in congressional budgeting." In F.M. Zweig (Ed.), *Evaluation in legislation*. Beverly Hills: Sage Publications.
- Weiss, Carol H., Erin Murphy-Graham, Anthony Petrosino, and Allison Gandhi. "The Fairy Godmother – and Her Wants: Making the Dream of Evidence-Based Policy Come True." *American Journal of Evaluation*. Volume 29, Number 1. Pages 29 – 47.
- Weiss, Carol H. 1977. "Research for Policy's Sake: The Enlightenment Function of Social Research." *Policy Analysis*. Volume 3, Number 4. Fall 1977, pages 531 – 545.
- Weiss, Carol H. 1991. "Policy Research as Advocacy: Pro and Con." *Knowledge and Policy: The International Journal of Knowledge Transfer*. Volume 4, Numbers 1 & 2. Pages 37 – 55.

- Weiss, Carol H. 1989. "Congressional Committees as Users of Analysis." *Journal of Policy Analysis and Management*. Volume 8, Number 3. Pages 411 – 431.
- Weiss, Carol H. 1998. "Have We Learned Anything New About the Use of Evaluation?" *American Journal of Evaluation*. Volume 19, Number 1. Pages 21 – 33.
- Whiteman, David. 1996. *Communication in Congress: Members, Staff, and the Search for Information*. University of Kansas Press.
- Wilson, Woodrow. 1887. "The Study of Administration." *Political Science Quarterly*, reprinted in 1997 in *Classics of Public Administration*, 2<sup>nd</sup> Edition. Shafritz, J and Hyde, A. Chicago: Dorsey Press.
- Wolanin, Thomas R. 1971. "Congress, Information, and Policymaking for Postsecondary Education: 'Don't Trouble Me With the Facts.'" Pages 382 – 394.
- Wright, Gerald C. and Brian F. Schaffner. 2002. "The Influence of Party: Evidence From State Legislatures." *American Political Science Review*. Volume 96, Number 2, Pages 367 –379.
- Yin, Robert K. 2009. "Case Study Research: Design and Methods." 4<sup>th</sup> Ed. Washington DC:Sage.
- Zelio, Judy. 2008. "Five Actions to Enhance State Legislative Use of Performance Information." IBM Center for the Business of Government.
- Zwick, Jesse. October 2010. "The Secret World of ALEC's Hacks." *The Washington Independent*.



