Learning Disabilities: A Contemporary Journal 14(2), 121-142, 2016

Leveraging Evidence-based Practices: From Policy to Action

Ronnie Detrich
Randy Keyworth
Jack States
Wing Institute

Education is a public health issue. Poor educational outcomes are correlated with many health and social ills. To improve the quality of education, it will be necessary to take advantage of the leverage points of policy, evidence, and implementation science. The idea of evidence informing policy may be non-controversial, but a closer examination of the concept suggests greater complexity than is readily apparent. Even if policy is informed by evidence, it is necessary to utilize what is known from implementation science to assure that policies are actually implemented well enough to achieve the desired benefits. In this paper, we discuss how policy, evidence, and implementation science can be integrated to leverage the impact of evidence-based practices and bring about meaningful, systematic change to the educational system. The change in mental health services in Norway is used as an exemplar for occasioning change. We also review policy initiatives that failed to achieve outcomes because evidence or implementation science was not part of the initiative. Finally, suggestions are made about these three elements that can be applied to bring about change in teacher preparation programs.

Keywords: **Policy, evidence-informed policy, evidence-based policy, implementation science, evidence-based education.**

"*Without data, you are just another person with an opinion.*" *W. Edwards Deming*

Introduction

Educational outcomes are a public health issue. Poor outcomes have been associated with a number of adverse health conditions and social problems such as overall lower level of income and poverty that in turn brings about many social ills including increased risk of prison, higher rates of obesity, and increased risk of smoking (Biglan, 2015). Students in special education are a particularly vulnerable population because they are more likely to have poor educational outcomes and, as a consequence, experience many of the difficulties cited above.

If, as a society, we are to improve the educational outcomes for all students, including the most vulnerable, we must take advantage of all available leverage points. One of those leverage points is *policy*. Because of its reach, if developed wisely, it offers the opportunity to influence large numbers of educators and, in turn, improve outcomes for large groups of students. A second leverage point is *evidence* about what are effective practices that are most likely to improve student outcomes. These two leverage points can be combined to form evidence-informed policy. Well-constructed evidence-informed policy has the potential to have impact beyond the capabilities of

*Please send correspondence to: Ronnie Detrich, Wing Institute, 2102 Dennison Street, Suite B., Oakland, California, 94606, Email: rdetrich@winginstitute.org.

either policy or evidence alone. Policy without evidence is just a guess and the probability of benefit is likely to be low. Evidence without policy is information that is unlikely to have impact as it has limited reach. The goal of evidence-informed policy is well summarized in a memo from Jeffery Zients, Acting Director of the Office of Management and Budget, "Where evidence is strong, we should act on it. Where evidence is suggestive, we should consider it. Where evidence is weak, we should build the knowledge to support better decisions in the future" (Zients, 2012). The federal Department of Education is one of the government agencies that operates under the auspices of the Office of Management and Budget so this memo had influence on the policies and activities of the Department of Education. The memo reflects the emphasis the Obama administration has placed on evidence in guiding policy decisions. (Haskins & Margolis, 2015).

The development of evidence-informed policy is not sufficient to assure the benefits of the policy will be realized. Policies must actually be implemented well if they are to have impact. Many education policies have been enacted without any meaningful impact on educational outcomes. Often this was because there was no comprehensive, coherent plan for implementing the policy. *Implementation science* is defined as the study of factors that influence the full and effective use of innovations (National Implementation Research Network, 2015) and brings coherence to the implementation of policies. It is the third leverage point that can be utilized to turn policy into meaningful action, thus achieving desired outcomes. It is the bridge between policy, evidence-based practices, and improved outcomes for students. Without implementation science, the aspirations of evidence-informed policy, no matter how well intentioned, are not likely to result in benefit for students.

One of the largest efforts to improve educational outcomes for students with disabilities in the United States was Public Law 94-142 (Education for All Handicapped Children Act, 1975; EHA) now renamed, the Individuals with Disabilities Education Improvement Act (IDEIA, 2004). Prior to its passage, it was estimated that one million children were completely excluded from educational opportunities and another three million children had very limited access (U.S. Department of Education, 2010). It had incredible reach in the sense that for the first time all children with disabilities had access to educational opportunities. To facilitate implementation of the law, the federal government was to provide 40% of the costs to states to off-set the financial burden of educating special education students; however, the federal government has never fully honored that commitment (Council for Exceptional Children, ND). Lacking in the law and subsequent regulations was a comprehensive effort to facilitate implementation. The most recent (2013) federal data measuring how well states are implementing IDEIA show that only 19 of the 50 states and Washington D.C. met federal standards for adequate implementation while 27 states met the criteria for needing assistance in implementation for two consecutive years (U. S. Department of Education, 2015).

Since the law was initially passed in 1975, it has been reauthorized several times. The reauthorization in 2004 emphasized using research-based interventions to improve student outcomes (P.L. 108-446). Even with this increased policy emphasis on using research-supported interventions there is evidence to suggest that special

education teachers are not necessarily being prepared to use these practices (Begeny & Martens, 2006; Spear-Swearling, 2008; Walsh, Glaser, & Wilcox, 2006).

Given the high stakes for children and society, it is important that all three leverage points are used to their fullest to assure that every student receives an excellent education. In this paper we consider how these leverage points can work together to improve educational outcomes for all students. We briefly discuss the leverage points separately then describe the necessary interaction among the three leverage points to maximize the benefits of policy. Finally, we analyze successful policies as well as policies that failed to produce the expected outcomes.

The Leverage Points

Policy

One of the primary functions of policy is to guide action in specific ways to influence outcomes. At the federal and state level, policy usually comes in the form of laws and regulations. Policies can also be enacted at the organizational level, such as school districts and schools, and commonly take the form of guidelines, expectations, and directives. Even classrooms have policies, usually described as classroom rules. In the classroom, the goal is for rules (policies) to influence the behavior of those affected by them. The challenge is to transform the policies into effective action. Articulating a policy does not mean that action will automatically follow and the anticipated benefits of the policies will be achieved. In the classroom, it is necessary to get the rules off the walls and into the interactions between teachers and students. School-wide positive behavior supports has provided a process for turning rules into actions that influence student behavior (Sugai & Horner, 2006). In essence, it is necessary to teach the rules (assuring students know what to do), teach specific behaviors that represent rule following (assuring students know how to behave across different contexts), and provide positive and consistent feedback for rule following (assuring rule following benefits the student). All of this is done to insure that positive behavior is sustained.

Similarly, at larger system units, the challenge is to get the policies out of the manuals and into practice. There are two general approaches for influencing compliance with a policy. Policies can specify consequences for compliance or, more often, failure to comply. This was the approach taken in No Child Left Behind (NCLB, 2008). If a school failed to make adequate yearly progress (AYP) as measured by annual tests, then a series of escalating negative consequences were brought to bear. Ultimately, schools could be closed for failing to make AYP for four consecutive years. Essentially, this approach to enforcing policy is coercive.

An alternative approach to coercive policies is for policymakers to create incentives for voluntarily participating with the policy. If individuals or organizations choose not to participate, there are no negative consequences. One example of this approach was California's efforts to reduce class-size in grades K-3 (Borhnstedt & Stecher, 2002). This initiative allowed school districts to receive $650 per student for every K-3 classroom that had 20 or fewer students, provided the districts first reduced first grade class size, followed by second grade, and then either third grade or kindergarten. At the end of the second year of the initiative, almost all of the first and second

grade classes had been reduced to 20 or fewer students. By the end of the fourth year, reduction of class-size was complete for kindergarten and third grade (Borhnstedt & Stecher, 2002). The state of California spent approximately $1 billion in the first year (1996-97 school year) and through 2010 has spent $20 billion. In the last five years the costs have averaged $1.75 billion annually (Luckie, 2009). This approach resulted in virtually all school districts voluntarily participating in the class size reduction initiative. The empirical outcomes and some difficulties of implementing this large-scale experiment will be discussed in a subsequent section.

Evidence

The evidence-based practice (EBP) movement has had an increasingly significant influence in education in the last 20 years (Detrich & Lewis, 2013; Kratochwill & Shernoff, 2004). More recently, there has been increased interest in evidence-based or evidence-informed policy (Nussle & Orszag, 2015; Nutley, Walter, & Davies, 2007). Before considering evidence-informed policy we briefly review the concept of evidence-based practice, defined as the integration of the best available evidence, with clinical expertise, and client values and context (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). EBP can be thought of in two ways (Slocum et al., 2014). The first approach is to label a practice that has met rigorous evidentiary standards as evidence-based. The second way of thinking about EBP is as a decision-making framework. In this approach, evidence must work in conjunction with clinical expertise and consideration of client values and context as a basis for decisions. It does not allow clinical expertise to substitute for evidence of effectiveness. The best available evidence must inform decisions. The limitation of the first approach is when there are no relevant studies that meet rigorous standards then the practitioner has no evidence to guide decision-making. It is often the case in special education that there are no relevant studies that meet the evidence standards. In the second approach to evidence-based practice, all decisions can be based on the best available evidence even when that evidence is not rigorous.

The concept of best available evidence implies that there is a range of rigor with evidence and sometimes decisions must be based on evidence that is not of the highest rigor. Applying the idea that evidence-informed policy is a decision-making framework is consistent with the spirit of the Zients (2012) memo and allows flexibility so that practitioners can use their professional judgment as to how best implement the policy in a local context, giving consideration to community values and relevant contextual factors.

Policy is made broadly but implemented locally. Policy is generally made at a distance removed from the local context in which it is to be implemented and all of the differences across implementation settings cannot be anticipated. Features of local context play a central role in influencing implementation. Urban schools have different challenges than rural schools in implementing a policy, as do high poverty schools compared to more affluent schools. It has been argued that because of the complexities of differing contexts, the concept of evidence-informed policy is not realistic (Greenhalgh & Russell, 2009). The concern is that the research context is so different from the local context as to make research evidence irrelevant.

Evidence-informed policy can prescribe what to do, but not how to do it in a specific context. Those with the best understanding of that context are in a better position to make those decisions. At the local level decisions about how to best implement an evidence-informed policy requires professional judgment and a clear understanding of the values of the local community. Conceptualizing evidence-informed policy as a decision-making framework addresses many of the concerns about the feasibility of it being realistic to address issues of context (Greenhalgh & Russell, 2009).

Complexity in evidence-informed policy. The general assumption of evidence-informed policy is that if policymakers rely on high-quality evidence when making decisions, then students and citizens will benefit. The notion of evidence-informed policy may seem non-controversial, but a closer examination of the concept suggests there are complex issues to be solved if evidence is to influence policy. There is an implication that policymaking is a rationale process in the sense that if the evidence is available policymakers will act on it; however, the formulation of policy is influenced by a number of factors other than evidence. A challenge for those advocating evidence-informed policy is that policymakers bring their own political and personal biases to the task. In instances when evidence conflicts with political and personal preferences, preferences usually prevail and evidence is discounted (Gambrill, 2012). Kavale and Forness (2000) described the conflict between evidence and bias in the following statement "…when however, such conclusions encounter strong advocacy and intuitive appeal, it is difficult to dislodge the practice through evidence and reason" (p. 297).

One example of advocacy playing a significant role in policymaking is President Obama's maternal, infant, and early childhood home visitation initiative (Haskins & Margolis, 2015). The Obama administration was specifically interested in the Nurse-Family Partnership (NFP; Olds, 2006); however, there were other home visitation programs that had evidence of effectiveness and their developers effectively lobbied Congress and the officials in the Obama administration to broaden the scope of the initiative to include other visitation models (Haskins & Margolis, 2015). Even though all of the home visitations models had evidence of effectiveness, in some instances the evidence was not as strong as the evidence for NFP; however, there were features of these other programs that made them attractive to the policymakers. For example, NFP depends on nurses conducting the home visits while other programs allowed for visits to be carried out by individuals who were not nurses. NFP restricted participation to first time mothers, but other models were much more inclusive and allowed women who were not first time mothers to participate. Broadening the policy on home visitation program from the NFP to allow other evidence-supported models reflects this notion of EBP as a basis for decision-making. It gave latitude to local decision-makers to use their judgment when choosing the home visitation program that was the best contextual fit for that community. A part of making the decision required a clear understanding of the values of the local community and the available resources.

This example illustrates some of the difficulties of basing policy exclusively on evidence. The evidence for NFP is compelling, but some of the features may make it difficult to implement into a specific context. Chorpita and colleagues (2011) have

described programs that have similar goals and methods as treatment families and have suggested that any program within the treatment family should be considered an option for those who are responsible for adopting effective programs within a specific context. Thus, in this example, each of the specific home visitation programs would belong in the same treatment family, while a number of contextual variables might influence the decision about which to adopt.

If evidence is to play a central role in influencing policy, then the challenges of overcoming personal biases, political considerations, advocacy groups, and financial incentives must be confronted. The perspective here is that evidence is the best basis for determining policy. Evidence provides a more transparent basis for decisions than any of the other options, such as political or personal philosophy since anyone who is interested can evaluate the evidence that is informing policy.

Uses of evidence in policy. There are several types of evidence that can be useful to policymakers. Experimental evidence, such as randomized clinical trials, quasi-experimental designs, and single participant designs, inform policymakers about which programs and practices are likely to have the greatest impact. A limitation of experimental evidence is that one experimental study is never sufficient to definitively answer a question about what should be done and is a poor basis for formulating policy. If there is a body of literature, the common approach by education scholars has been to review the extant literature and make a reasoned judgment about what should be done. Policymakers are not necessarily prepared to conduct a review of the literature and come to reasonable conclusions about what should be done as a matter of policy. An alternative to the narrative type of review is a systematic review or meta-analysis that summarizes a body of research and can inform policymakers about the general effect of a practice. A significant advantage of meta-analysis for policymakers is that it provides a single score (effect size) that best estimates the strength of an intervention across populations, settings, and other contextual variables. Program evaluation is another type of evidence that is valuable to policymakers. It provides feedback about the effectiveness of a program or practice and can provide insights about how policies can be changed to increase benefit.

Nutley et al. (2007) have proposed four ways that evidence can be used in the development of policy. Evidence can be used in an instrumental sense in which research directly informs or answers a policy question about the best course of action. This is how most individuals think research is used to influence policy (Nutley et al.). Kavale and Forness (2000) have argued that meta-analyses can play an important role in developing policy. Similarly, large scale randomized clinical trials can make a significant contribution to formulating policy. One quasi-experimental or single participant design is less useful for informing policy because of the small number of participants and potential for bias because of the unrepresentativeness of study participants.

Research can also be used tactically to justify a position one has taken or to convince someone else about the best course of action. In this context, evidence is being used as a social influence tactic. It is often the case that individuals wishing to use evidence as a means of social influence will "cherry pick" studies that support a position and ignore evidence that contradicts that position. Well-conducted meta-analyses minimize this problem because it includes all studies on a topic

that meet inclusion criteria regardless of the outcomes; therefore, minimizing the confirmation bias.

A third use of research in the formation of policy is what Nutley and colleagues described as conceptual. In this instance, evidence is used to shape the way a problem is framed and discussed. An example of this is using evidence to change the discussion of school discipline policies from emphasizing reactive, negative consequences for rule violations to preventive and positive approaches to establish rule following. In this way evidence was used to move the discussion away from punitive zero tolerance policies.

The fourth use of policy is described as the process use. This refers to policymakers and practitioners using the methods of research to develop answers to their questions. An example of this might be for school administrators to use local evidence about when and where "acting out" behavior is most likely to occur to make decisions about how to allocate resources or use evidence to focus attention on "alterable variables" when trying to solve problems. In a sense this is practice-based evidence. Practice-based evidence is evidence that is developed at the local level. Even when an evidence-based practice is implemented with fidelity there is no guarantee that it will be effective in a specific context. Practice-based evidence is evidence about the effectiveness of intervention in a specific context. Program evaluation evidence is commonly used to evaluate the impact of an intervention in a local context.

Similarly, practitioners seeking answers to challenges they are facing can collect data about the frequency of occurrence, the contexts in which they are most likely to occur, and the differences between the contexts in which the problem occurs and does not occur. Practice-based evidence is the essence of data-based decision making (Ervin, Schaughency, Mathews, Goodman, & McGlinchey, 2007; Stecker, Lembke, & Foegen, 2008). Single participant designs are commonly used in data-based decision making. The unit of analysis can be an individual to determine if she is benefiting from an intervention and is common in response to intervention approaches (Barnett, Daly, Jones, & Lentz, 2004). The unit of analysis can also be larger such as a whole school. Practitioners of school-wide positive behavior support rely on single participant designs to make decisions regarding the effectiveness of whole school interventions (Ervin et al., 2007).

Maciolek (2015) has added "imposed use" as a fifth way that evidence can be used in policy. Imposed use is when others are required to use evidence to access funding or other resources. This was the approach that was used in NCLB and IDEA. Educators were mandated to use evidence to solve academic and social problems. In an evidence-based practice approach to education all sources of evidence are legitimate. The best available evidence construct assumes there is a hierarchy of evidence and the evidence that is the highest quality and the most relevant should be used at the local level to make decisions about which interventions to adopt.

Consequences of ignoring evidence. There are significant consequences for failing to consider evidence when making policy. Those who are intended to benefit from policy and those who fund the initiatives are often not well served when evidence is not considered. An example of the negative consequences of ignoring credible evidence is the widespread adoption of the drug education program, Drug Abuse Resistance Education (DARE). It is the most commonly implemented program of its

type in the U.S., mandated in law in some states, and it is estimated that it receives up to $1 billion annually in funding (Shephard, 2001). The policies mandating DARE preceded evidence of effectiveness. In this instance, policymakers were largely influenced by advocacy by DARE proponents. By doing so, these policymakers were not acting in the spirit of the Zients memo (2012). Several well-designed studies have subsequently demonstrated that DARE yields no benefit and, in some instances, may have negative effects (Dukes, Ullman, & Stein, 1996; Kanof, 2003; Perry et al., 2003).

In some instances, government and educational organizations have funded programs for years without knowing if they actually work (Liebman, 2013). Head Start, a $7 billion-a-year program that has been continuously funded since 1965, was never evaluated until 2002 (Liebman, 2013). Once the evaluation was completed, the results were largely disappointing, finding only moderately beneficial impact that disappeared by the end of first grade (Advisory Committee on Head Start Research and Evaluation, 2012). Further complicating the results, Head Start programs were highly variable in quality, which may have contributed to the limited findings about the effectiveness. If the program had been evaluated earlier, the variation in quality could have been identified and efforts made to improve the quality of individual Head Start programs and the overall quality of all programs. Improving the quality of Head Start could have altered the developmental trajectories of the participating children. If that were the case, there would have been a greater return on the investment of tax dollars in this vulnerable population that could have enormous societal benefits for decades.

Integrating evidence with policy. If evidence is to inform policy, then it is necessary to understand how it might best be integrated with it. There are two approaches for doing this. The first is policy mandating that educators implement practices that are supported by evidence, but leave the decision about which intervention to implement to the discretion of local decision makers (imposed use of evidence as described by Maciolek, 2015). As previously stated, this was the approach used in NCLB and the IDEA. The second approach is using evidence to inform policy.

Mandating evidence-based practices. There were a number of implicit assumptions in the use of this approach with NCLB (Detrich, 2008). First, it was assumed that there was an established body of evidence-based interventions. Secondly, it was assumed that educators were aware of the evidence supporting different practices. A third assumption was that educators had the expertise to implement a specific practice. A final assumption was that the necessary resources were available to support effective implementation. The experience with NCLB would suggest that these assumptions are not justified. When NCLB was enacted, there was no organized resource for educators that provided information about the evidentiary status of various interventions. More recently, there are a number of organizations that summarize and evaluate the evidence supporting educational interventions such as the What Works Clearinghouse and Best Evidence Encyclopedia.

There is evidence that those responsible for implementing NCLB and IDEA are not aware of the evidence-base for different interventions. For example, in special education, school psychologists are often responsible for identifying and developing interventions. In a survey of directors of school psychologist training programs, it was reported that 29% of directors of school psychology programs had no knowledge of a list of evidence-based interventions and 41% of these directors reported that stu-

dents in school psychology training programs received no training in the evidence-based interventions in the survey (Shernoff, Kratochwill, & Stoiber, 2003). If school psychologists are not aware of specific evidence-based interventions, then it is unlikely they will select them for intervention with the students they serve.

Reading is one of the foundational skills that all students must develop if they are to fully benefit from educational opportunities. NAEP data (2012) suggest that U.S. students have made little progress in the last 40 years in reading proficiency in spite of enormous expenditures. One explanation for these results is that educators are not using effective instructional practices. Walsh and colleagues (2006) in a review of teacher preparation programs reported that only 15% of reviewed teacher preparation programs taught all of the five empirically-supported elements of reading as determined by the National Reading Panel (2000) and 50% taught no more than one.

Similarly, formative assessment has been demonstrated to be a powerful tool for improving educational outcomes (Fuchs & Fuchs, 1986; Hattie, 2009; Yeh, 2007) and yet, it is not common for it to be taught in teacher preparation programs. In a review of special education teacher preparation programs, it was reported that only 14% of reviewed programs addressed formative assessment (Spear-Swerling, 2008).

As discussed above the punitive approach of NCLB for failing to make adequate educational progress could be considered coercive (Sidman, 1989). Coercive practices often have unintended consequences such as cheating to avoid the negative consequences. One result of the coercive policies of NCLB, educators not knowing what to do or how to improve test scores, began altering the answers on annual high stakes tests resulting in cheating scandals that occurred across many cities, particularly in high poverty, low performing schools (Lattal & Detrich, ND). If policies require educators to use empirically-supported interventions, it will be necessary to develop comprehensive implementation plans to assure that they have access to the evidentiary status of interventions, receive the necessary training to implement with fidelity, and have the resources required to support implementing those interventions. Failing to do so will result in the policy having little impact or negative unintended consequences that harm the educational process.

Evidence informing policy. A second approach to incorporating evidence in policy is for the policy to be informed by evidence. In this case, evidence is used to influence decisions regarding the content of policy (instrumental use as described by Nutley et al., 2007). Examples of this include warning labels on cigarettes, restaurants, bars, and office workspaces required to be smoke free, and everyone in a car required to use seat belts (CDC, 2011; Cohen & Einav, 2001; Hiilamo, Crosbie, & Glantz, 2014). In all instances, the evidence was clear that there were health and safety benefits for the citizens by enacting these policies even though there was opposition to these policies when they were introduced. In education, policymakers in the state of California relied on the evidence about the benefits of class size reduction to develop the statewide policy to reduce class sizes in grades K-3 (Bohrnsted, & Stecher, 2002). The National Reading Panel (2000) applied the procedures of meta-analysis to develop their recommendations for effective reading instruction. Similarly, the Institute for Education Science (IES) relies on a combination of meta-analysis and expert opinion to develop recommendations for their practice guides for educators across

a variety of topics such as reducing behavior problems (Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008). One of the interesting features of the practice guides is that for each recommendation the strength of evidence supporting the recommendation is provided. This is another example of relying on the best available evidence and expert judgment to inform decision-makers. Each practice guide is accompanied by the following disclaimer:

> This practice guide should be reviewed and applied according to the specific needs of the educators and education agency using it, and with full realization that it represents the judgments of the review panel regarding what constitutes sensible practice, based on the research that was available at the time of publication. This practice guide should be used as a tool to assist in decision-making rather than as a "cookbook" (e.g., Epstein et al., (2008, p. ii).

This disclaimer reflects the perspective that evidence informed policy is a decision-making process that requires the integration of best available evidence, professional judgment, and client values and context.

These two approaches to incorporating evidence into policy are best used in different circumstances. Policy mandating services based on research-supported interventions is best used when the mandate is broad, such as improving educational outcomes for all students in all schools as in NCLB and IDEA. Federal policymakers cannot anticipate all of the circumstances and contexts in which interventions will be necessary, so it may be most efficient to simply make policy that whatever the circumstances and contexts, decisions about the selection and adoption of interventions should be guided by evidence.

Even when policy is broad at one level (i.e., federal), at other levels closer to the setting in which the solution will be implemented, the decisions and circumstances become more defined and evidence can be used to inform the policy. For example, a school district must make decisions about which reading programs to adopt given the best available evidence. This is a much more defined set of decisions. The selection committee has to decide which reading curriculum is most appropriate for the relevant grade levels, characteristics of the students, and the resources required to implement the new curriculum. Under these constraints, the decision-makers can make use of existing research evidence to inform their policy about reading curricula so they can comply with the broader mandate to use evidence as a basis for policy decisions.

Implementation Science

One of the themes in this paper is that policy does not automatically lead to effective action. In this section we explore the stages of implementation science before turning to how policy, evidence, and implementation science can be integrated to produce important outcomes for students. Implementation science has been described as a systematic approach to transforming policy into meaningful action. Implementation science has been defined as the study of how an evidence-based practice gets translated to different contexts in the "real world" (Martinez-Beck, 2013). It is a multi-staged approach with specific actions at each stage that are necessary to facilitate implementation. The stages of implementation science are exploration and

adoption, installation, initial implementation, and full implementation (Blasé, Van Dyke, Fixsen, & Bailey, 2012).

Exploration and adoption is the phase in which all stakeholders (teachers, administrators, etc.) are involved in the decision-making in terms of defining the problem they are trying to solve and identifying possible solutions. Involving all stakeholders in the process increases "buy in" and commitment to an adopted program (Blasé et al., 2012). It is during this stage that all stakeholders have access to evidence of effectiveness, core features of the intervention, and resources required to implement effectively.

Once a specific program has been adopted, implementation moves into the installation stage. During this phase it is necessary to make a number of changes to the infrastructure of the school and school district before actual implementation can begin. Training and coaching has to be completed, policies and procedures must be changed to align with the new program, and the necessary resources to make high quality implementation possible must be purchased (Blasé et al., 2012). It is often the case that resources are not adequately allocated to the implementation to new initiatives (Schaffer, Nesselrodt, & Stringfield, 1997). Without careful attention to installation issues, initiatives will never get fully implemented. Common shortcomings during the installation phase are to train for less time than is recommended by the program developers and to hire individuals who do not meet the recommended standards for training and experience (Elliott & Mihalic, 2004).

The initial implementation stage is when the new intervention or program is introduced to students. During this phase, those responsible for implementation at all levels of the system are practicing new skills and figuring out how to solve the various problems that become apparent as actual implementation begins. No matter how well things work on paper, the reality is that there will be many different problems to solve. Usually initial implementation involves a small number of classrooms or schools within a district as a pilot project. Many of the obstacles that are identified during this phase can be solved before involving all classrooms, schools, or districts in the implementation process.

Full implementation occurs when the new initiative is introduced across the whole system. This is the phase in which the practices and procedures related to the initiative become the norm. Fixsen, Naoom, Blase, Friedman, and Wallace (2005) estimate that it takes two to four years to reach this stage. Full implementation can be accomplished only when all aspects of a system are aligned. If a district initiates a new reading curriculum across all elementary schools, to be successful, the school must acquire all resources that are necessary for implementation including textbooks and teachers' manuals; allocate the necessary time in the school schedule to allow the program to be implemented as recommended; arrange initial training as recommended by the developers; develop a system for ongoing coaching and support for teachers in their classrooms as they learn to implement the new curriculum; and establish instructional teams that review and revise implementation efforts as implementation of the intervention progresses.

The fundamental goal of implementation science is to make sure that at each phase of implementation the necessary steps are taken to assure that an intervention is implemented with integrity. No matter how effective an intervention is, if

it is not implemented with sufficient integrity, it will not result in anticipated outcomes (Noell, Gresham, & Gansle, 2002). In order to assure high quality implementation, it is necessary to have a system in place to monitor treatment integrity and give feedback to practitioners about how well they are implementing (Burns, Peters, & Noell, 2008; Duhon, Mesmer, Gregerson, & Witt, 2009; Sanetti, Fallon, & Collier-Meek, 2013). Not only is it necessary to have a system in place to monitor and provide feedback to those who are responsible for implementing an intervention for students, it is also necessary to measure and give feedback to those who are responsible for supporting the teachers who are implementing. If the plan to support teachers is not implemented well, then teachers are less likely to implement with integrity (Detrich, 2013). This logic extends to all levels of the system and makes the case for the importance of all aspects of the system being aligned to support implementation at the level of the classroom.

Integrating Policy, Evidence, and Implementation Science

It is abundantly clear that policy alone is not sufficient to improve students' academic achievement. Since the publication of *A Nation at Risk* (Gardner, 1983) there has been a steady stream of policy initiatives with the intent to reform the U.S. education system. In the time period covered by these various policy initiatives there is almost 50 years of data suggesting that academic performance in reading and math as measured by NAEP has not changed in any significant way despite all of the policies and money spent (Nations Report Card, 2015).

One has to ask, if policy has not been effective, what needs to be done? One part of the solution is to incorporate evidence about effective practices and implementation science with policy. Each is necessary to improve educational outcomes for all students, but alone none are sufficient to bring about the necessary improvements. Figure 1 shows the interaction among the three components and the effect the various combinations of these components is likely to have on educational outcomes. As can be seen in Figure 1, when policy, evidence, and implementation science are present, there is high probability of having impact at scales of social importance. When any one of the components is absent, the probability of impact or scale of impact is significantly limited. The challenge is integrating the three elements into a single multi-component intervention package.

Successful Implementation of a Policy Initiative

Perhaps the best example of integrating policy, evidence, and implementation science is the national effort in Norway to improve mental health services. As described in Biglan and Ogden (2008), after years of program evaluation data showing that Norwegian youth were not being well served by placement in residential treatment programs, the Norwegian Research Council hosted an expert conference in which a number of invited presenters described their research to a group of Norwegian practitioners, policymakers, and researchers (Biglan & Ogden). Following the conference, a governmental initiative to strengthen services, competence, and research was developed. The Ministry of Child and Family Affairs funded a research unit at the University of Oslo and gave priority to the adoption and implementation

of evidence-based practices. One of the programs that was adopted because of its strong research base was the Parent Management Training Oregon model (PMTO). The 19 county health directors were invited to participate in implementing PMTO in Norway. With the participation of the health directors, this became a national implementation project to improve the psychological and behavioral well-being of youth and families in Norway.

Figure 1. **Relationship among leverage points and outcomes.**

Policy	Evidence	Implementation	Outcome
√	√	√	Policy assures significant reach and broad impact; evidence and systematic implementation improves probability of high impact.
√	√		Policy assures significant reach; Evidence increases impact but lack of systematic implementation limits impact.
√		√	Policy assures reach; however, the impact is limited because policy not informed by evidence. Additionally, systematic implementation of non evidence-based practices is not likely to result in impact.
√			Policy assures reach; however, limited impact because there is no evidence to justify policy and there is no systematic implementation plan.
	√	√	Evidence based practices and systematic implementation likely to produce benefit but without the reach of policy the impact is likely to be limited
	√		Even with an evidence-based practice the impact is limited because there is no policy to increase reach and there is no systematic plan for implementation.
		√	Systematic implementation of a non-evidence-based practice is not likely to achieve benefit.

The PMTO project began with the developers from the Oregon Social Learning Center (OSLC) providing an 18-month long extensive training in the approach. At the end of the training period, 30 professionals were certified as PMTO practitioners. The next generation of professionals (83) was trained by individuals from the first generation cohort with coaching and support provided by the developers from the OSLC. All of the professionals were recruited from the regular service system so they could immediately begin providing PMTO to families.

These professionals came from organizations that had agreed to conditions to participate including providing equipment for training, paying the salaries for the trainees, and allowing for time off to participate in the training. To support the implementation process, the national ministries funded infrastructure development, training, and supervision. Practitioners from the first two cohorts trained a third generation cohort (Biglan & Ogden, 2008).

To maintain the fidelity of implementation and solve implementation challenges, a significant infrastructure was developed. The implementation effort was led by the National Implementation Team, which was comprised of five national consultants and six regional coordinators. Regional groups met eight times a year for one day each time. Groups of trainees met with a supervisor every other week throughout the training period. The Regional Supervisors met three or four times per year with their Regional Coordinator (Ogden, Forgatch, Askeland, Patterson, & Bullock, 2005). Finally, the national ministries funded a randomized evaluation of the PMTO effects and found that PMTO successfully reduced child externalizing problems, improved social competence, and enhanced parental discipline (Biglan & Ogden, 2008). From this description, it is clear that there was a systematic plan to bring PMTO to Norway. It involved policy, evidence, and implementation science. Program evaluation evidence about the poor outcomes served as motivation to consider alternatives to existing practices. This is an example of evidence being used to answer a specific policy question as suggested by Nutley et al., (2007). In this instance, the question was how well are current practices serving children and families and the answer was that they were not being well-served.

Once a decision was made to do something different, the initiative entered into the adoption and exploration stage of implementation (Fixsen et al., 2005). To guide the process, exploration was constrained by identifying those programs that had strong research support, reflecting the instrumental use of evidence in policy (Nutley et al., 2007). To accelerate the development and dissemination of the PMTO, the Ministry of Child and Family Affairs developed policy to develop the necessary infrastructure, beginning with the creation of the research center (Biglan & Ogden, 2008). Ultimately, the National Implementation Team was responsible for assuring that all levels of the system were aligned and supporting the implementation of PMTO (Biglan & Ogden, 2008). All of these efforts represent steps in the installation phase of implementation as described by Fixsen and colleagues (2005).

The next phase in the implementation process is initial implementation (Fixsen et al., 2005). In Norway this was exemplified by the developers of PMTO training the first generation of professionals who had volunteered to use PMTO in their practice. Policy supported the initial implementation by providing funding for infrastructure and training and supervision (Ogden et al., 2005). There was no mandate that agencies adopt PMTO and participate in the training, but all 19 county health administrators did participate and advocated for the initiative in their regions (Biglan & Odgen, 2008). Progress toward full implementation was made by the second and subsequent generations being trained by those who had been previously trained rather than by the developers of PMTO. This step was necessary to make it part of the institutional practices of national health agencies as well as local agencies

responsible for implementation in the cities and counties. Doing so increases the likelihood of the practices sustaining after the outside consultants leave.

Examples of Unsuccessful Implementation of Policy Initiatives

The experience of Norway has been an enormous success compared to some other initiatives such as class size reduction and the introduction of whole language frameworks in California. Comparing how those initiatives were implemented with the efforts of Norway may be informative for development of future educational initiatives. First, we will consider the California experience with class-size reduction. The initiative began because of falling test scores on standardized tests and a significant budget surplus. Small class size is always popular with many stakeholders in the educational process (Bornstedt & Stecher, 2002) and Tennessee had recently completed a project evaluating the effects of reduced class size in grades K-3 in which students who participated in classes no larger than 17 students and were served by experienced credentialed teachers had better educational outcomes in reading and math than students in larger classrooms (Mosteller, 1995). With the available funding, enthusiasm from stakeholders, and the evidence from Tennessee, California initiated class size reduction providing extra funding to districts that reduced K-3 class size to 20 or less. The initiative came out of the governor's office rather than the state department of education. When the initiative began, state and district officials were taken by surprise because the governor had not worked with them to develop the policy or an implementation plan (Detrich, 2013). This initiative created an immediate need for 18,000 classrooms and 12,000 new additional teachers the first year and an additional 15,000 over the next two years. The long term results of this initiative is that it resulted in negligible benefit for the students of California (Bornstedt & Stecher, 2002). By examining Figure 1, the sources of failure are clear. First, there was policy, but that policy was developed without the input of stakeholders from around the state. As a consequence, there was no systematic implementation plan. This is considerably different than the approach in Norway to bring PMTO to full implementation. An additional shortcoming in the development of the policy to reduce class size was that all available evidence from Tennessee suggested that class size should be 17 or less and the teacher should be credentialed and have experience. California reducing class size to 20 was without support in the available evidence so even with fully credentialed teachers, the effects may have been minimized. By creating such a great demand for additional teachers, districts were forced to hire many inexperienced teachers working on emergency credentials (going to school to complete the necessary coursework while working in a classroom). Again, there is no evidence that teachers with emergency credentials can achieve the benefits promised by class size reduction. When examined through the lens of necessity of policy, evidence, and implementation science, it is not surprising that California did not achieve the effects that Tennessee had. The state of California had a policy that was not supported by evidence or a systematic implementation plan.

Another California educational policy initiative that warrants discussion is the decision to adopt whole language reading frameworks as the state approach to reading. The whole language frameworks were adopted in California in 1987. In the eight years following the adoption of the frameworks, California fourth grad-

ers were scoring at the bottom of states on the NAEP reading tests (Stewart, 2000). The frameworks were adopted not because there was strong evidence supporting the ideas in the frameworks, but rather because of the power of anecdotal stories about how whole language had made New Zealand the most literate nation in the world. Subsequently, international testing revealed that many New Zealand children could not read and, in fact, lagged behind U.S. children in reading scores (Stewart, 2000). The obvious question is how whole language gained such a foothold in California. In the beginning, the Superintendent of Public Instruction organized a meeting of a select group of educators to discuss how to improve reading in California. By design, very few reading experts were invited to participate in the meeting, so the participants would not be influenced by the agendas of the experts. As a result the theoretical basis of whole language was accepted with very little critical evaluation. In this instance, policy was formed without any influence from evidence about effective practices. In fact, some whole language proponents reject the value of research arguing that the skills they are trying to teach cannot be measured (Stewart, 2000). Also, there was no systematic implementation plan, thus leaving each district to determine what should be done to implement the frameworks. In some instances, this resulted in schools abandoning the explicit instruction of fundamental skills such as decoding in favor of literature-based literacy (Stewart, 2000). The consequences of this type of policymaking fell to the students who were not learning to read and their families who sought out private tutors to remediate the reading difficulties.

Improving Education Through Evidence-informed Policy

If policymakers and educators are to avoid California's experiences with class size reduction and whole language reading frameworks, it will be necessary to more explicitly seek out evidence about effective practices. In this section, we will discuss teacher preparation as an example of how evidence can be used to influence policy and how that policy can influence educators and result in better outcomes for students.

There is compelling evidence that the classroom teacher is one of the most powerful influences over a child's academic achievement (Babu & Mendro, 2003; Rowan, 2004). There is also evidence about which instructional practices have the greatest likelihood of impact on students (Fuchs & Fuchs, 1986; Hattie, 2008; Marzano, Marzano, & Pickering, 2003; National Reading Panel, 2000; Wang, Haertel, & Walberg, 1997); however, many teacher preparation programs are not teaching those skills, so teachers entering the classroom are ill-prepared to provide high quality instruction (Begeny & Martens, 2006; Greenberg, Putnam, & Walsh; 2014; Greenberg, Walsh, & McKee, 2015; Spear-Swerling, 2008; Walsh et al., 2006). Furthermore, the training practices to prepare teachers to teach are often ineffective. (Joyce & Showers, 2002).

If students are to receive a meaningful education, then it will be necessary to rely on evidence to inform policy about what teachers should be taught and how they are to be taught in their preparation programs. Given the complexity of altering the content and process of teacher preparation programs a thoughtful, long term plan of systematic implementation will be necessary. One of the biggest challenges to changing teacher preparation is that most teachers are educated in university preparation

programs. These programs often operate outside any direct influence from policymakers and have been resistant to outside influences.

There are several ways in which evidence can be used to influence the content and process of teacher preparation programs. As discussed earlier, evidence can be used to serve multiple functions. One of the functions of evidence is to influence the perspectives of others (Nutley et al., 2007). There are currently a number of groups in the U.S. that are using evidence as a part of advocacy. The National Council on Teacher Quality (NCTQ) uses evidence about what teacher preparation programs are teaching and not teaching to increase public and political awareness about how, in many instances, teacher preparation programs are failing to adequately prepare teachers to be effective and ultimately, this is limiting the educational outcomes of public school students (Greenberg et al., 2015). The reviews that the NCTQ conducts on teacher preparation are based on their assessment of what are effective instructional practices. They assess the syllabi of teacher programs to determine if these practices are being taught within a specific teacher preparation program.

One of the advocacy methods that the NCTQ uses is widely disseminating and publicizing their reviews of teacher preparation programs. It is their perception that this form of public pressure is beginning to influence what teacher preparation programs are doing (Greenberg et al., 2015). If universities are to change their practices around teaching, it will be necessary to create a motivation for change. On their own, NCTQ reports only have so much influence. The evidence from these reports in the hands of policymakers can have a greater impact. As discussed earlier, policy can increase motivation to behave in particular ways by linking funding to adoption of a particular practice (Haskins & Margolis, 2015). Since public universities receive a significant amount of their funding from state budgets, legislators could allocate an amount of money over the traditional levels of funding to support the transition to a teacher preparation curriculum that more closely reflects what is currently known about effective practices. To avoid the negative effects of coercion, the allotted funds should be in addition to the usual level of funding. This is more likely to result in programs volunteering to participate.

As part of the initiative to change what is taught in teacher preparation programs, other agencies such as the teacher credentialing office should be involved. If the teacher credential requirements at the state level change to better reflect what is known about effective teaching, then universities will have to change their instructional practices so that their graduates are prepared to meet the credentialing requirements.

An additional source of influence that can be exerted on the universities to change their practices is local school districts. A school district can use evidence about effective teaching practices to shape their recruitment practices. If a significant number of school districts are explicit that they are more likely to hire graduates if they have had training in these specific practices and their job descriptions and actual hiring practices reflect this preference, then universities are more likely to alter their training practices so that their graduates are competitive in the job market.

There is precedent for these kinds of external influences altering what is taught in a university program. With the increased incidence rate of children diagnosed with autism and the recognition that applied behavior analysis is the treat-

ment of choice, there has been an exponential growth in universities offering training programs in applied behavior analysis for the explicit purpose of providing services for children with autism (Behavior Analyst Certification Board, ND). This growth was almost exclusively fueled by external demands for trained, qualified individuals to support children with autism. Part of the implementation infrastructure to assure that the programs maintained are of high quality, the Behavior Analyst Certification Board (BACB) approves course sequences offered by universities as preparing the students of these programs to sit for the certification exam. Being certified by the BACB gives those university programs a competitive advantage over other non-certified programs in recruiting students and is a financial benefit to the universities.

A Way Forward

We began this discussion with a quote from W. Edward Deming. If education is, as we have suggested, a public health issue, then more than opinions are necessary. One of the challenges in integrating evidence with policy is that policymakers and researchers are from two different cultures, speak different languages, and have different values (Baron, 2010). To cross the cultural divide, it will be necessary for researchers to function as cultural anthropologists so they can understand the language and values of policymakers and can interact more effectively. Even if policy becomes evidence-informed, it will not be sufficient to assure that these policies will affect change. Policies and practices must be implemented with sufficient integrity that improved educational outcomes are likely. Policies mandating the use of evidence are a good start, but it is necessary to develop the infrastructure to support all of the implications of that policy. State education agencies can become a clearinghouse for practices that are empirically-supported and provide training, coaching, and implementation guidance. As suggested here, universities can begin to train teachers and other educators to use these empirically-supported practices. Districts can set policies requiring that practices and programs have strong research support. The Norway example offers hope that it can be accomplished. With evidence, it is a difficult task; without evidence, it is impossible to bring about effective, systematic, and coherent change.

Improving educational outcomes, and by extension public health, will require a generation of policymakers, researchers, and practitioners working in partnership to change the educational system. Evidence-informed policy is a recursive process in which change is continuous and policies and practices change as the evidence dictates. The process of developing evidence-informed policy is not complete until there is effective implementation and the evaluation data show that the intended benefits have been achieved. Rather than isolating themselves in their individual silos of policy, research, and practice, stakeholders need to be involved throughout the entire process. It seems fitting to end with a quote from Henry Ford, "Coming together is a beginning; keeping together is progress; working together is success."

References

Advisory Committee on Head Start Research and Evaluation: Final Report (2012). Submitted to U. S. Department of Health and Human Services. Retrieved from http://www.acf.hhs.gov/programs/opre/resource/advisory-committee-on-head-start-research-and-evaluation-final-report.

Babu, S., & Mendro, R. (2003, April). *Teacher accountability: HLM-based teacher effectiveness indices in the investigation of teacher effects on student achievement in a state assessment program.* Paper presented at the Annual Meeting of the American Education Research Association, Chicago, ILL.

Barnett, D. W., Daly, I. I. I., Edward J, Jones, K. M., & Lentz, J., F. Edward. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *Journal of Special Education*, *38*(2), 66-79.

Baron, N. (2010). *Escape from the ivory tower.* Washington DC: Island Press.

Begeny, J. C., & Martens B. K. (2006). Assessing pre--service teachers' training in empirically-validated behavioral instruction practices. *School Psychology Quarterly*, *21*(3), 262-285.

Behavior Analyst Certification Board (ND). *Approved course sequences.* Retrieved from http://info.bacb.com/o.php?page=100358

Biglan, A. (2015). *The nurture effect: How the science of human behavior can improve our lives and our world.* Oakland, CA: New Harbinger Publications.

Biglan, A., & Ogden, T. (2008). The evolution of evidence-based practices. *European Journal of Behavior Analysis*, *9*(1), 81-95.

Blase, K. A., Van Dyke, M., Fixsen, D. L., Bailey, F. W., Kelly, B., & Perkins, D. F. (2012). Implementation science: Key concepts, themes and evidence for practitioners in educational psychology. In B. Kelly & D. E. Perkins (Eds.) *Handbook of implementation science for psychology in education* (pp.13-34). New York, NY: Cambridge University Press.

Bohrnstedt, G. W., & Stecher, B. M. (2002). *What we have learned about class size reduction in California. Capstone report.* Retrieved from http://eric.ed.gov/?id=ED471331

Burns, M. K., Peters, R., & Noell, G. H. (2008). Using performance feedback to enhance implementation fidelity of the problem-solving team process. *Journal of School Psychology*, *46*(5), 537-550.

Center for Disease Control and Prevention (April, 2011). Morbidity and Mortality Weekly Report: *State smoke-free laws for worksites, restaurants, and bars-United States, 2000-2010.* (Vol. 60. No. 15), 472-475.

Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., & Trent, L. (2011). Evidence based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice*, *18*(2), 154-172.

Cohen, A., & Einav, L. (2001). *The effects of mandatory seat belt laws on driving behavior and traffic fatalities.* Retrieved from the Harvard John M. Olin Discussion Paper Series website: http://www.law.harvard.edu/programs/olin_center/

Council for Exceptional Children (ND). *Funding, Policy & Advocacy.* Retrieved from https://www.cec.sped.org/Policy-and-Advocacy/Current-Sped-Gifted-Issues/Funding

Detrich, R. (2008). From policy to practice: IDEIA and evidence-based practice. In E. L. Grigorenko (Ed.), *Educating Individuals with Disabilities: IDEIA 2004 and Beyond*, (pp. 85-114). New York, NY: Springer Publishing Company.

Detrich, R. (2013). Innovation, implementation science, and data-based decision making: Components of successful reform. In M. Murphy, S. Redding, & J. Twyman (Eds.), *Handbook on Innovations in Learning* (pp. 33-49). Charlotte, NC: Information Age Publishing.

Detrich, R., & Lewis, T. (2013). A decade of evidence-based education: Where are we and where do we need to go? *Journal of Positive Behavior Interventions*, *15*(4), 214-220.

Duhon, G. J., Mesmer, E. M., Gregerson, L., & Witt, J. C. (2009). Effects of public feedback during RTI team meetings on teacher implementation integrity and student academic performance. *Journal of School Psychology*, *47*(1), 19-37.

Dukes, R. L., Ullman, J. B., & Stein, J. A. (1996). Three-year follow-up of drug abuse resistance education (DARE). *Evaluation Review*, *20*(1), 49-66.

Education for all handicapped children act of 1975. Public Law 94-142. *Federal Register, (August 23, 1977)*.

Elliott, D. S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science*, *5*(1), 47-53.

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). *Reducing behavior problems in the elementary school classroom*. Washington, DC: U.S. Department of Education.

Ervin, R. A., Schaughency, E., Matthews, A., Goodman, S. D., & McGlinchey, M. T. (2007). Primary and secondary prevention of behavior difficulties: Developing a data-informed problem-solving model to guide decision making at a school-wide level. *Psychology in the Schools*, *44*(1), 7-18.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: The National Implementation Research Network.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*(3), 199-208.

Gambrill, E. (2012). *Critical thinking in clinical practice: Improving the quality of judgments and decisions*, (3rd ed.). Hoboken, NJ: Wiley.

Gardner, D. P. (1983). *A Nation at Risk*. Washington, DC: US Department of Education.

Greenberg, J., Putman, H., & Walsh, K. (2014). *Training our future teachers: Classroom management*. Revised. Retrieved from http://www.nctq.org/dmsStage/Future_Teachers_Classroom_Management_NCTQ_Report

Greenberg, J., Walsh, K., & McKee, A. (2015). *Teacher prep review: A review of the nation's teacher preparation programs*. Retrieved from http://www.nctq.org/dmsStage/Teacher_Prep_Review_2014_Report.

Greenhalgh, T., & Russell, J. (2009). Evidence-based policymaking: A critique. *Perspectives in Biology and Medicine*, *52*(2), 304-318.

Haskins, R., & Margolis, G. (2015). *Show me the evidence: Obama's fight for rigor and results in social policy (2nd Ed)*. Washington DC: Brookings Institution Press.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hiilamo, H., Crosbie, E., & Glantz, S. A. (2014). The evolution of health warning labels on cigarette packs: The role of precedents and tobacco industry strategies to block diffusion. *Tobacco Control*, *23*(1), 1-18.

National Implementation Research Network (2015). *Implementation science defined*. Retrieved from http://nirn.fpg.unc.edu/learn-implementation/implementation-science-defined

Individuals with Disabilities Education Improvement Act of 2004. 108th Congress, Public Law 108-446. *Retrieved from idea.ed.gov/download/statute.html.*

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: ASCD.

Kanof, M. E. (2003). Youth illicit drug use prevention: DARE long-term evaluations and federal efforts to identify effective programs. Retrieved from http://files.eric.ed.gov/fulltext/ED473428.pdf

Kavale, K. A., & Forness, S. R. (2000). Policy decisions in special education: The role of meta-analysis. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds). *Contemporary special education research: Syntheses of knowledge base on critical instructional issues*, (pp. 281-326). Mahwah, NJ: Lawrence Erlbaum Associate.

Kratochwill, T. R., & Shernoff, E. S. (2004). Evidence-based practice: Promoting evidence-based interventions in school psychology. *School Psychology Review*, *33*(1), 34-48.

Lattal, D., & Detrich, R. (ND). *Analysis of the Atlanta public school cheating scandal: A culture of coercion*. Retrieved from http://aubreydaniels.com/institute/behavior-watch/special-article/part-i-analysis-the-atlanta-public-school-cheating-scandal-culture-co

Liebman, J. B. (2013). *Building on recent advances in evidence-based policymaking.* Retrieved from http://www.hamiltonproject.org/papers/building_on_recent_advances_in_evidence-based policymaking/.

Luckie, M. S. (2009). *California's class-size reduction program: Total spending since 1996.* [Interactive graph]. Retrieved from http://californiawatch.org/k-12/californias-class-size-reduction-program-total-spending-1996.

Maciolek, S. (2015). *Use of research evidence: Social services portfolio.* Retrieved from http://wtgrantfoundation.org/?s=Maciolek

Martinez-Beck, I. (2013). Introduction: Where is the new frontier of implementation science in early care and education research and practice? In T. Halle, A. Metz, & I. Martinez-Beck (Eds.) *Applying implementation science in early childhood programs and systems* (pp. xix-xxx). Baltimore, MD: Paul H. Brookes Publishing.

Marzano, R. J., Marzano, J. S., & Pickering, D. (2003). *Classroom management that works: Research-based strategies for every teacher.* Alexandria, VA: ASCD.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The future of Children*, 5(2), 113-127.

National Institute of Child Health and Human Development (2015). *National institutes of health nations report card.* Retrieved from http://www.nationsreportcard.gov/reading_math_2015/#reading?grade=4

National Reading Panel (US), National Institute of Child Health, & Human Development (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.*

No Child Left Behind Act of 2001, 20 U.S.C. & 6319 (2008).

Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *Journal of Behavioral Education*, *11*(1), 51-67.

Nussle, J., & Orszag, P. (2015). *Moneyball for government 2nd Edition.* Washington DC: Disruption Press.

Nutley, S. M., Walter, I., & Davies, H. T. (2007). *Using evidence: How research can inform public services.* Bristol, UK: Policy Press.

Ogden, T., Forgatch, M. S., Askeland, E., Patterson, G. R., & Bullock, B. M. (2005). Implementation of parent management training at the national level: The case of Norway. *Journal of Social Work Practice*, *19*(3), 317-329.

Olds, D. L. (2006). The nurse–family partnership: An evidence based preventive intervention. *Infant Mental Health Journal*, *27*(1), 5-25.

Perry, C. L., Komro, K. A., Veblen-Mortenson, S., Bosma, L. M., Farbakhsh, K., Munson, K. A., & Lytle, L. A. (2003). A randomized controlled trial of the middle and junior high school DARE and DARE Plus programs. *Archives of Pediatrics & Adolescent Medicine*, *157*(2), 178-184.

Rowan, B. (2004). Teachers matter: Evidence from value-added assessments. Research points. *American Educational Research Association, 2*(2), 1-4

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (Eds.). (2000). *Evidence-based medicine: How to teach and practice EBM*. Edinburgh, UK: Churchill Livingstone Inc.

Sanetti, L. M. H., Fallon, L. M., & Collier-Meek, M. A. (2013). Increasing teacher treatment integrity through performance feedback provided by school personnel. *Psychology in the Schools*, *50*(2), 134-150.

Schaffer, E. C., Nesselrodt, P. S., & Stringfield, S. C. (1997). *Impediments to reform: An analysis of destabilizing issues in ten promising programs*. Retrieved from http://files.eric.ed.gov/fulltext/ED408676.pdf

Shephard, E. (2001). The economic costs of DARE. Institute of Industrial Relations. *Research Paper*, 22). Retrieved from drugpolicy.org

Shernoff, E. S., Kratochwill, T. R., & Stoiber, K. C. (2003). Training in evidence-based interventions (EBIs): What are school psychology programs teaching? *Journal of School Psychology*, *41*(6), 467-483.

Sidman, M. (1989). *Coercion and its fallout*. Boston, MA: Authors Cooperative.

Slocum, T. A., Detrich, R., Wilczynski, S. M., Spencer, T. D., Lewis, T., & Wolfe, K. (2014). The evidence-based practice of applied behavior analysis. *The Behavior Analyst*, *37*(1), 41-56.

Spear-Swerling, L. (2008). Response to intervention and teacher preparation. In E. L. Grigorenko (Ed.) *Educating Individuals with Disabilities: IDEIA 2004 and Beyond*, (pp. 273-294). New York: Spring Publishing Company.

Stecker, P. M., Lembke, E. S., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing School Failure*, *52*(2), 48-58.

Stewart, J. (2000). The blackboard bungle: California's failed reading experiment. *LA Weekly*, pp. 22-29. Retrieved from http://www.kidsource.com/kidsource/content/whole.1.html

Sugai, G., & Homer, R. R. (2006). A promising approach for expanding and sustaining schoolwide positive behavior support. *School Psychology Review*, *35*(2), 245-259.

U. S. Department of Education (2010). *Thirty-five years of progress in educating children with disabilities through IDEA*. Retrieved from http://www2.ed.gov/about/offices/list/osers/idea35/history/index.html?exp=5

U. S. Department of Education (2015). *Determination letters on state implementation of IDEA*. Retrieved from www2.ed.gov/fund/data/report/idea/2015-ideafactsheet-determinations.pdf

Walsh, K., Glaser, D., & Wilcox, D. D. (2006). *What education schools aren't teaching about reading and what elementary teachers aren't learning*. Washington, DC: National Center on Teacher Quality.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). Learning influences. In H. J. Walberg & G. D. Haertel (Eds.). *Psychology and educational practice*, (pp.199-211). Berkeley, CA: McCutchan Publishing.

Yeh, S. S. (2007). The cost-effectiveness of five policies for improving student achievement. *American Journal of Evaluation*, *28*(4), 416-436.

Zients, C. P. O. J. (2012). *Memorandum for the heads of executive departments and agencies: Guidance on the use of challenges and prizes to promote open government*. Washington, DC: Office of Management and Budget.