From Fidelity to Fit: Enabling Successful Interventions in Complex Implementation Systems

Stephanie Moulton
Associate Professor
John Glenn School of Public Affairs, The Ohio State University

&

Jodi R. Sandfort
Associate Professor
Humphrey School of Public Affairs, University of Minnesota

November, 2015

WORKING DRAFT

For Presentation at the APPAM Annual Meeting, Miami, Florida November 12-14, 2015
Please do not cite without Permission of the Authors

This paper elaborates ideas at the foundation of our book *Effective Implementation in Practice: Integrating Public Policy and Management* (Jossey-Bass, 2015). We would like to thank Scott Allard, Jennifer Mosely, and staff and leaders at MDRC New York for their helpful comments on the ideas presented in this paper. A portion of this research was supported by the Family Self-Sufficiency Research Consortium funded by the Office of Planning, Research, and Evaluation in the Administration for Children and Families, U.S. Department of Health and Human Services. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

Successful policy and program implementation involves creating change in the target population, as well as change in the system of organizations and institutions carrying out the initiative. Both types of change have been discussed in varying degrees in prior literature, and are often the focus of evidence-based evaluations (Weiss et al. 2014; Hargreaves, 2010). While policy analysts and program evaluators are most often interested in whether or not an intervention improves the condition or behavior of the target population, sustainable interventions also require changes to the implementation system--the day to day activities of frontline staff, the operational processes within organizations, and the often contested dynamics in the macro level policy environment (Sandfort and Moulton 2015). In fact, a growing body of literature in implementation science focuses on the extent to which interventions become successfully integrated within implementation systems (May, et al 2009; Meyer et al, 2012).

In addition to changes in the system and target population, there is a third type of change that is often overlooked, downplayed, or even discouraged in evidence based evaluations: change in the public service intervention to be implemented. This is often because one of the hallmarks of scaling successful initiatives is "fidelity" to the original program design. Fidelity is presumed to be necessary to enable rigorous evaluations of the impact of the intervention. Conceptually, fidelity to the program design depends upon a precisely delineated program logic, a clearly specified implementation plan, and well-defined outcomes (Weiss et al. 2014). While perhaps beneficial for evaluating causality in a controlled field experiment, such precision does not reflect reality and may not be desired. Fidelity presumes that evaluators have been able to specify the causal mechanisms within an intervention that will lead to positive results. In fact, these mechanisms are often not

entirely clear a priori, and may differ depending on the context. Further, attempts to adhere to fidelity may impede an implementation systems' ability to adapt and, thus, to sustain the intervention over the longer term.

Rather than viewing fidelity as the end goal for successful implementation of an intervention, an alternative focus is to consider the fit between the intervention, target population and implementation system. From this perspective, implementation is more successful over the longer term when the intervention is adapted to fit the system and the needs of the target population. In this manuscript, we present a framework that integrates these three types of change as occurring simultaneously in a dynamic system. Instead of relying on program theory guide implementation (e.g., Chen & Rossi, 1989; Weiss 1987), this approach offers an alternative way to theorize how changes are made in implementation systems.

Building from our prior scholarship, we conceive of the implementation system as nested multi-level strategic action fields, bounded by sources of legitimate authority at each level that shape the core program intervention (Sandfort and Moulton 2014; Moulton and Sandfort 2015). Program implementers and evaluators – as field actors—can work together to resolve competing sources of authority as they make implementation decisions, thereby facilitating target population and program "fit." Such a framework can help inform future research agendas that bridge policy analysis, evaluation and management, and can help facilitate improved implementation practice within complex systems.

This paper proceeds as follows. First, we identify the need for a new framework that integrates knowledge of an evidence based intervention with behaviorally informed insights about the target population and an understanding of the dynamics within the

implementation system. Next, we describe the implementation system as multi-level sites of strategic action that enable and constrain changes to an intervention. Drawing from social theories, we then describe the processes of creating change within implementation sites—including changes to the intervention and changes to the system. Ideally, these changes are aligned to facilitate desired outcomes for the target population. We conclude with a worked example of the evaluation of a municipal financial empowerment intervention. Our aim is to provide a more systematic way to think about and enable changes not only in the target population (the desired outcome of the program), but also changes to the intervention and system that occur during implementation.

Interventions in Existing Implementation Systems: New Wine in Old Wineskins

Evidence-based interventions draw upon basic research or formal program evaluations to identify treatment methods that improve the behavior or conditions of a target group. The prevalence of such interventions has proliferated in the last twenty-years:[1] school curricula, adult mental health treatments, retirement savings, parent-development initiatives all have various intervention models that are either developed through consultation with basic social science research or rigorously evaluated post-hoc to ascertain impacts on their respective target groups. Two movements in this regard are worth noting: an increase in public support for behaviorally informed interventions, and increased prevalence of randomized impact evaluations.

First, by integrating insights from psychology, behavioral scientists identify and analyze systematic biases and heuristics that shape individual behaviors. They take these

[1] The theme of this year's APPAM conference "The Golden Age of Evidence-Based Policy" bespeaks this emphasis and achievement.

insights to inform understandings of how users experience programs, and direct implementers to design interventions that address biases in order to increase take-up and maximize behavioral change. Such behavioral insights result in interventions that are often not only better targeted but also less costly (Amir et al. 2005; Shafir, 2013). In light of this, policymakers and evaluators are increasingly advocating for behaviorally informed interventions, as indicated by the White House's recent executive order to "Use Behavioral Science to Better Serve the American People."[2]

Second, public support for and investment in experimental-design evaluations has increased. Randomized-control designs improve the likelihood of isolating the discernible impact of an intervention upon the target group members. Such designs emphasize a precisely specified treatment that is administered uniformly to the treatment group, and withheld from the control group. Both behavioral economics and impact evaluations use scientific rigor and analysis to identify how a particular intervention might change the conditions or behaviors of a target group. While these research developments have yielded more information about 'what works,' they have not created much insight about how to sustain and replicate these interventions outside of field trials.

In large part, this is because these new evidence-based interventions are rarely carried out by new organizations or delivery systems. Instead, initiatives are introduced into existing systems, where networks, organizations, and individuals have historical experiences, distinct core competencies, and varying accountabilities, professional norms, and beliefs about the intervention. This reality leads to what program evaluators perceive to be

[2]See https://www.whitehouse.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-american

problematic - varying degrees of treatment fidelity—where replication of an evidence-based program in a particular place differs from what was planned. In fact, when an intervention is being carried out across multiple sites, the degree of fidelity is likely different for each implementation site, which creates empirical challenges for evaluators trying to assess its impact (Weiss, et al, 2014). In practice, instead of one intervention, there are multiple interventions, which can often lead evaluators to be quite frustrated with the implementation process and the irrationality of program operators who deviate from replication plans (Werner, 2004). Well-established evaluation firms recognize this and some require randomized-control field trial sites to develop implementation plans in addition to their treatment plans (Weiss, et al, 2014).

A field of implementation science has evolved in response to these challenges, to better understand and enable system capacity to implement an intervention with fidelity (Eccles and Mittman, 2006). Like behavioral economics and program evaluation, the scientific foundation of this literature is notable; developed first in medicine, and broadened to community psychology, education, mental health and other human services fields, the research builds upon randomized control trials and systematic reviews of empirical studies (Fixsen, et al 2005). An often cited meta-analysis by Durlak and DePre (2008: 340) reviews over 500 quantitative studies and concludes there is "strong support for the premise that effective implementation is associated with better outcomes." According to their analysis, the magnitudes of effect sizes are two to three times higher when programs are carefully implemented. This finding and others motivated considerable attention among implementation science researchers and federal agencies to create predictive models to assess the "readiness" of an organization or system to adopt an evidence-based program

with fidelity (Aarons, Hurlburt, and Horwitz, 2011; Damschroder, et al 2009; Greenhalgh, et al, 2004; Horner, et al 2014; Meyers, et al, 2012; Tabak, 2012). Others have articulated that certain institutional levers – such as training, coaching, management team, or information systems – (Aarons, et al, 2012; National Implementation Research Network) are important in shaping this readiness. Still others advocate for understanding implementation of evidence-based interventions in phases (Fixsen, et al 2005; Aarons, et al 2011).

Overall, this research activity has yielded conflicting results. While predictive models proliferate, empirical studies fail to confirm that certain variables predict organization or system readiness for change to support replication of evidence-based models with fidelity (Aarons, Hurlburt & Horwitz, 2011; Supplee & Metz, 2015). Yet the activities of those engaged in this work is not without merit. Discussions about institutional levers or phases of implementation processes, or reports from federal funding agencies drawing attention to these issues, can provide heuristic devices useful for the actors trying to change implementation systems (Blasé, et al 2015; Aaron, et al 2011).

Interestingly, the development of implementation science research is largely divorced from long-standing scholarship in public policy, political science, and public management also concerned about the operation of and change in implementation systems (Nilsen, et al 2013; Roll, Sandfort & Moulton, 2015). Since the 1970s, scholars have examined the policy processes and political authority necessary to implement formal policy. For example, "top down" scholars emphasize the design of the formal policy (e.g., legislative mandate) as the most appropriate lever for improving policy outcomes, identifying variables that can be manipulated by policy makers to ostensibly improve outcomes (Palumbo and Calista 1990; Bardach 1977; Mazmanian and Sabatier 1989). A group of "bottom up" scholars critiqued

this approach and instead focused on assessing or mapping local contexts to appreciate incentive structures and behaviors on the ground (Berman, 1978; Elmore, 1980; Hjern & Porter, 1981; Lipsky, 1980).

Another approach to inquiry about system operation explores the governance of public action, drawing heavily from organizational theory and public management, focusing on the structures and coordinating mechanisms of governance and networks (Agranoff and McGuire 2003; Ingram and Schneider, 1990; Milward and Provan 2003; Provan and Kenis, 2007; Salamon, 2002). Many notable public management scholars, including Michael Hill and Peter Hupe (2008), Ken Meier and Larry O'Toole (2004), Carolyn Hill and Larry Lynn (2005) see program implementation in terms of multi-level governance. This provides a way to think about how policy decisions made in federal government resonate with the work of state agencies and service providers. However, because the focus of their interests is the institutional structures or coordinating mechanisms through which governance occurs, the policy and program content or concern for a particular evidence-based interventions is often downplayed or missing.

Thus, while program evaluators and implementation scientists may lack conceptual frameworks for understanding implementation contexts, those who study these contexts tend to operate at an abstract level that often overlooks a particular intervention. Moving implementation practice and research forward requires that we integrate these ideas, taking seriously the fit between existing systems, an intervention and the target population. Implementation of new evidence based practices takes place in existing systems. Like the parable, simply pouring new wine into old wineskins likely leads to poor outcomes (and a mess to clean up). But, social systems are capable of change, people are capable of change,

and interventions can be adapted. Skilled implementers and evaluators can learn new ways to help facilitate this change in intentional ways to bring about desired outcomes. A better understanding of the social processes underlying implementation systems offers a step forward.

Understanding the Sites and Social Processes within Implementation Systems

In developing a framework to understand implementation systems, we take seriously the work of governance scholars who highlight the important tasks of implementation that occur in multi-level systems (Hill and Hupe, 2014; Ostrom, 2011; Robichau and Lynn, 2009). We also draw upon other theories of social processes (Fligstein and McAdam, 2011; 2012) to understand how change and stability occur within these systems. When applied, we think this framework helps inform the factors that constrain and enable the integration of an intervention within a system, and the ultimate impact on a target population. It can provide a way to explore 'fit,' and to see changes in intervention models and implementation systems as desirable outcomes of program implementation.

First, it is important to conceptualize implementation systems as multi-level. While a number of the implementation science models (e.g. Aarons et al 2011; Greenhalgh, et al 2004) refer to an "inner context" (eg. organizational setting) and "outer context" (e.g. larger environment), it is important to more clearly identify the sites and processes important in shaping how interventions are changed during implementation. In fact, distinct actions take place at the policy field, organizational and frontlines of the system.

Policy fields are bounded networks among institutions and organizations carrying out a substantive policy and program area in a particular place (Stone and Sandfort, 2009).

They are made of up formal and informal relationships between various units of government, private corporations, and nonprofit agencies. Multiple government institutions at the national, state, and local levels engage in the same policy field at any given time, adding to the complexity. At the policy field level in an implementation system, actors possess and assemble resources and ideas that both enable and constrain a program intervention. At the very least, the introduction of a new intervention or changes to an existing intervention cannot conflict with the rules that are perceived to be in place at a given point time, stemming from existing legislation, funder requirements or other sources of authority. To enable larger-scale replication, existing policy or legislation might need to be adopted or informal rules about what is feasible or desirable altered at the policy field level.

An example from the housing field is the federal foreclosure initiative, the Hardest Hit Funds (HHF) program, authorized under the Trouble Assets Relief Program (TARP) and created by Executive Order in 2010. Essentially, through the HHF program, more than $7 billion was allocated to 18 eligible states. State housing finance agencies (HFAs) within each state were to develop interventions to distribute the funds to homeowners at risk of foreclosure. Evaluations of the effectiveness of HHF have been critical, noting a substantial delay in getting funds out to homeowners and lower than expected numbers of assisted.[3] However, the actions of state HFAs were constrained significantly by the existing rules in the policy field around mortgage servicing. Private lenders had to agree to accept the mortgage assistance on behalf of a homeowner, which often required a renegotiation of the existing mortgage contract. Before state HFAs could distribute funds, policy field action was required at the federal level to bring together regulators, secondary market participants (who had

[3]SIGTARP report 2012. http://www.sigtarp.gov/audit%20reports/sigtarp_hhf_audit.pdf

purchased the mortgages) and private lenders to clarify what was considered to be feasible given existing constraints—which may or may not have been the ideal intervention to stem foreclosures over the longer term. An incredibly important outcome to explore—that would be overlooked by most impact evaluations of the program-- is the extent to which there are changes to the housing finance system to allow for more behaviorally effective interventions.

At the ***organizational*** level, program interventions are operationalized and integrated into day-to-day activities. While legal constraints certainly apply, other concerns – such as existing organizational competencies, operating procedures, or financial constraints – seem more immediate. There are opportunity costs when creating new infrastructures to implement a new intervention. And often changes to procedures and practices are weighed against potential benefits. Practices that are already part of the formal or informal standardized operating procedures are often "sticky;" they are institutionalized and may shape managers' beliefs about how things should be done (Khademian, 2002). Changes to an intervention that require new infrastructure also require new resources and authority to make the case that it is a worthwhile investment; alternatively, when organizations merely change an intervention to fit within existing infrastructure because it is feasible, it might not be ideal.

An example of organizational level constraints can be found in the implementation of the National School Lunch program, and in particular, recent changes to increase the nutritional content of school lunches under the 2010 Healthy Hunger-Free Kids Act. There are a variety of behaviorally informed interventions that have been developed to help meet the program objectives, beyond simply changing the types of food served in the lunch room. For example, some school systems have integrated instruction about healthy foods into their

curriculum, others have engaged students directly in growing fruits and vegetables in community gardens, and still others have invited students and parents to participate in taste tests and contests to select the revised menus for their school systems (e.g. Hirschman and Chriqui 2013). However, the ability of a particular school system to carry out these interventions depends in large part on the capacity of the school and its staff. And, to the extent that the school is not financially dependent on the National School Lunch program to fund a large proportion of its students (e.g., students on free and reduced lunches comprise a small proportion of the total population), the schools may not have the financial incentives (or flexibility) to invest in these additional programming options. Under what conditions are local school systems likely to seek out additional resources or challenge existing norms to introduce new practices? Which elements of the implementation system enable or impede this type of change?

At the ***frontlines*** of implementation systems, interventions are enacted with the target population, either through direct interactions with frontline workers, or through indirect interactions with technology systems (e.g., online registration portals). Here, frontline workers respond to signals from the organization (e.g., rules, incentives, or performance targets), but they also exert agency by relying upon their own professional norms, or beliefs developed through experience (May and Finch, 2009). The interactions with the target population also often shape frontline workers strategies because they learn what seems to be effective in carrying out a program. It is at this level where direct conflict between evidence-based interventions and frontline practices come into focus. Based on their own experiences with the target population, frontline staff can question the legitimacy of the intervention mandates (Aarons, et al 2012; Lin, 2000; May and Finch, 2009).

For example, significant research focuses on developing and testing the efficacy of mental health interventions. Yet the delivery of the treatment depends significantly on the clinicians, the patients seeking care, and the interactions between them. Although mental health professionals are generally trained in professional schools that emphasize the importance of the research underlying specific interventions, national studies document a complex relationship between the frontline staff and their take-up of evidence-informed practice (Aarons, et al 2005; Aarons, 2004). While they might support the general attitude of 'evidence-informed care,' they feel quite legitimate in rejecting such a model if it is mandated by their employer or state agency. While they might voluntarily attend training in a new proven model, they also rely upon what they have learned from their own practice as equally legitimate. Investigations suggest that these attitudes are not held merely by individuals or maintained in isolation; the social context and culture among these frontline workers shapes their take up and incorporation of such interventions into their work (Aarons, et al 2012).

As these examples illustrate, implementation activities occur at various levels in a system and evidence-based interventions encounter various factors across these levels that blunt the likelihood of adoption with fidelity. While there is ample evidence that implementation processes alters interventions, we have few ways to understand this variation systematically.

The second part of our framework provides one such explanation. Because social dynamics lie at the heart of the coordinated activity taking place within implementation contexts, we turn to social theory and notions of social systems (Fligstein & McAdam, 2012; Moulton & Sandfort, 2015). In particular, the theories of strategic action fields and social

skill can inform how implementers work collectively to adapt the ideas underlying evidence-based interventions to 'fit' within existing contexts.

First, we find it helpful to think of each site in an implementation system as a distinct strategic action field (Moulton & Sandfort, 2015). For sociologists, a strategic action field is a social order created by people as they interact around "shared understandings about the purposes of the field, relationships to others in the field, and the rules governing legitimate action in the field" (Fligstein and McAdam, 2012:9). National housing finance systems, local elementary schools and group of clinicians working in mental health clinics can each be seen as strategic action fields. In these sites, people have a collective understanding of their purpose, know their relationships, and implicitly understand what actions are considered to be appropriate (and feasible) in that context. In any given implementation system, these social structures create boundaries around what types of change are possible in that setting.

As social systems, each strategic action field also includes particular technologies-- or bundle of processes and coordinating tools that are used to transform inputs into outputs in order to accomplish programmatic objectives (Scott, 2008; Orlikowski, 1992; Sandfort, 2010). Curriculum in use in elementary school classrooms, and the supplementary books, power point slides and worksheets provided to students are examples of such technologies. So are the intake processes and documentation procedures that are developed within mental health clinics to coordinate action. The program technologies currently in use in a particular setting may not be the most effective for creating change in an external target population, but rather are those that are considered feasible in that setting at a given point in time.

These dynamics have important implications for how institutions and individuals within an implementation site-- operating as a strategic action field-- adopt new

programmatic interventions, or makes changes to existing ones. The amount of effort exerted by these actors depends in part on how much the new or modified intervention differs from the status quo technology in use. In experimental interventions, this is sometimes referred to as the "treatment contrast," or the extent to which the new or modified treatment differs from the control group intervention (Weiss, et al 2014). Changes to an intervention must be either framed in a way that they fit with the existing social system or other sources of authority must be brought to bear. In our notion, authority derives from the perceived legitimacy by the entity or person issuing the directive (Blau & Scott, 1962; Suchman, 1995). Because there are often multiple and conflicting sources of authority in any strategic action field, how various sources are activated and reconciled is important for it determines what behaviors are carried out, and, given our interest here, what changes are made to an evidence-based intervention when it is taken up by other sites in an implementation system.

[Insert Figure 1 Here]

Figure 1 gives a visual depiction of the two main elements of this framework. It highlights that there is a central intervention of interest that is taken up by various sites in a multi-level implementation system. It also shows that distinct strategic action fields exist at these sites with competing sources of authority – any of which may influence the form and function of the intervention that is enacted in that place. And there are multiple sources of authority. As the visual depicts, there are multiple potential sources of authority that might appear in public implementation systems: "political authority," which defines what is legally permissible and mandated, stipulates which public institutions have legal responsibilities and the governance relationships; "economic authority," which stems from competition and

reliance on market mechanisms, including performance criteria or competitive activities; "professional norms," which are expressed through professional codes, white papers, trainings or other implementation resources (Hill, 2003) that come from non-state actors and communicate best practices about the interventions;[4] and, finally, "belief and values," the reflect even less formal aspects of culture such as the shared understandings that develop about target group members, local conditions, and daily experiences.

These potential sources of authority are in no ways deterministic. In one setting in an implementation system, a mandate associated with funding provides cause for managers to abandon a long-held practice in service quality, while in another context in that same system, the mandate would motivate another manager to seek supplemental funding or develop an advocacy strategy to push against the direction. Actors involved in strategic action fields at all levels in a system are actively engaged in co-creating the conditions within which implementation activities get carried out. Their significance evolves out of the particular people and conditions in the strategic action field.

This type of variation has proven troublesome for evaluators concerned about treatment fidelity or policymakers interested in mandating replication of evidence-based interventions. To provide a way to gather data about these social dynamics and their variation across strategic action fields in an implementation system, we have created a tool currently being field-tested (See Appendix 1). It provides a systematic way to gather perceptional data to better understand how these different forms of authority operate within the strategic action fields that shape core interventions.

---

[4] Promulgation of evidence-based intervention models through research papers or 'purveyor'-sponsored trainings is one form a professional norm; rather than assuming it will motivate others to replicate, our framework sees it as merely one potential source of authority available to field actors in a setting.

Facilitating Fit: Three Types of Change

One of the primary benefits of this framework is that it also provides a way of thinking about how the existing social dynamics within implementation settings might be altered to improve results in relation to both internalizing new system operations and improving target group conditions. According to strategic action field theory, changes to system dynamics are achieved through the use of *social skill* (Fligstein, 2001; 2008), defined as "the ability [of an actor] to induce cooperation by appealing to and helping to create shared meanings and collective identities" (Fligstein and McAdam, 2012:46). In the program evaluation and implementation science literatures, there is a resounding theme that "leadership" or presence of a "program champion" matters in efforts to successfully replicate research-based interventions (Blasé, et al, 2015; Weiss, Bloom & Brock, 2014). By thinking about implementation sites as strategic action fields, the activities of leadership that bring about successful change become clearer.

Effective implementers can employ social skill to make changes to the implementation system to fit the intervention, and to make changes to the intervention to fit the system constraints. Oftentimes, this requires framing and brokering changes in way that resonates with actors in the implementation setting (Fligstein, 2001; Fligstein and McAdam, 2012). Moving from current conditions to a setting in which the new intervention becomes normalized-- embedded in the standard operating procedures at the frontline-lines, organizational, and policy field levels-- is slow and pain-staking work (May, et al, 2013; Hall et al, 2014).

Yet, additionally, a carefully integrated and executed intervention that does not create the desired change in the target group is a poor outcome. This requires implementers who collect information about results—indicators of change in the behavior or conditions of the target population – and then use that information to make changes to the intervention. External evaluators may provide technical expertise about the extent to which outcomes for the target population are achieved, and offer recommended changes to the intervention to increase the probability of achieving desired outcomes.

But in order for the intervention to be sustained, the changes that are made to the intervention must be those that are feasible within the system. Hence, technically informed changes depend upon the social skills of actors within those contexts who can frame arguments and create commitments from those within a particular implementation context about what is possible within that setting. Altering an intervention so that it is more likely to lead to desired change in the target population requires social skill. Actors must know how to identify other key decision makers, activate appropriate sources of authority, and frame particular courses of action to get the acceptance of others.

Conceptually, fit for a particular intervention is achieved when all three types of change align, as demonstrated by segment #1 in Figure 2.

[Insert Figure 2 Here]

Unfortunately, there often are cases when no overlap occurs at all – interventions are designed without consideration of the target population or the system, systems change for political or institutional reasons despite the needs of the target population or individual programs, and target populations change due to factors outside of the control of the implementation system or program designers. Other times, partial alignment is achieved.

Changes to an intervention may be made to fit the needs of the target population, but without regard to specific system dynamics (segment #2 of Figure 2). This is often the case when evidence based interventions are developed based on evaluation results in one particular context, or in laboratory settings. The context is either presumed to be irrelevant or controlled for during implementation.

On the other hand, sometimes implementers emphasize the contextual fit of a pre-defined intervention (e.g., Horner, Blitz and Ross, 2014), where the objective is for the intervention to become part of standard operating procedures, with little adjustment for a particular target population (segment #4 of Figure 2). This may be appropriate in settings where the intervention is proven effective (e.g., medical treatments) and there is no need to adjust based on the responsiveness of a particular target population. But, this is probably not the case for many policy interventions, were an ideal treatment is unknown or even unknowable.

Finally, there are cases where the focus is on the system and the target population, but with little attention to the resulting intervention (segment #3 of Figure 2). In this case, there may be symbolic implementation of a generalized intervention but the factors that lead to success (or failure) operate behind the scenes, with no systematic identification of the mechanisms of the intervention that lead to change. If the program works, it is often attributed to a "secret sauce" or some combination of attributes of the people engaged in carrying out the program and the individuals participating. Replication and generalizability are severely limited.

Focusing on all three types of change provides a starting point for improving implementation, and in particular, improving the likelihood that evidence based practices

will be effective and sustained over time within the system. Enabling all three types of change requires more than technical expertise to monitor change in the behavior or conditions of target groups. It also requires socially skilled actors in each setting who can recognize and adjust the intervention in ways that align with current system constraints. Evaluators tasked with measuring and enabling the success of an "evidence based practice" can likely benefit from an understanding of the system dynamics in a particular setting that contribute to or deter from socially skilled action. The framework presented here provides a tool to inform this understanding. In the next section, we illustrate the applicability of this framework for evaluation with a worked example.

Example: LIFT-UP

The Local Interventions for Financial Empowerment through Utility Payments (LIFT-UP) is an intervention to align local financial empowerment services with municipal utility debt collection practices. The LIFT-UP intervention was first developed by the National League of Cities (NLC)—a national nonprofit organization that supports, develops and disseminates best practices to municipal governments across the U.S. Over the past decade, the NLC Institute for Youth, Education, and Families has worked closely with many of its member cities to develop financial empowerment (FE) strategies and services. FE programs support the efforts of low-income families to achieve financial stability and build assets by expanding their access to financial services, empowering residents to take advantage of

available benefits and tax credits, and providing tools to manage money more effectively and build assets to increase their financial stability.[5]

In ongoing consultation with member cities, the NLC recognized a dissonance in how municipal services—including FE services—were being provided to financially vulnerable households. The partner municipalities that had developed FE interventions to help their residents achieve financial stability were also issuing collection notices and terminating water utility services to the very same vulnerable households. Figure 3 provides a visual depiction of this dynamic in municipal services.

[Insert Figure 3 Here]

LIFT-UP was intended reconcile the duality, developing an intervention that connects residents who are struggling to pay for municipal debts, like water utility bills, with FE services. The ideal LIFT-UP intervention includes three components: (1) a streamlined identification and referral process to flag utility customers at risk of water shut-off; (2) restructured debt payment options to help at risk customers repay past due utility debt; and (3) connection of at risk customers to FE services, including but not limited to an initial financial counseling session. In 2013, the NLC secured funding from the Center for Financial Services Innovation and the Ford Foundation to design, implement and evaluate the impact of LIFT-UP through a pilot program to test the effectiveness of the intervention in five unique municipal environments. The success of the LIFT-UP model would be indicated by (1) a

[5]New York City is an earlier adopter of Financial Empowerment. City officials claim that "fully integrating financial empowerment and asset building strategies into core social service delivery" can produce a "super vitamin" effect whereby the FE can magnify the positive outcomes of an intervention, therefore maximizing the impact of scarce social service resources (Mintz, 2011). The Cities for Financial Empowerment Fund is a nationwide movement-- inspired by early successes in New York City-- that seeks to recruit cities and mayors to pioneer the assimilation of FE into government infrastructure. Three LIFT-UP cities, Newark, NJ, Savannah, GA and Louisville, KY are members of this coalition (Cities for Financial Empowerment Fund).

reduction in the rate of water-shut offs among delinquent utility customers (change in the behaviors of the target population), and (2) the sustained integration of the LIFT-UP identification and referral process into the municipal utility system. A rigorous (randomized) evaluation design to document impact was developed to test the pilot program.[6] Data systems were put in place to track customer take-up of LIFT-UP in each city, and to link treatment and control group customers to their ongoing utility payment data.

The LIFT-UP pilot launched in March 2014, when the first city mailed postcards offering the program to several hundred eligible utility customers, who would otherwise be at risk of losing water services. Enrollment across the five cities continued through October 2014. Throughout the pilot period, more than 3,000 customers had been offered the LIFT-UP program, with 307 enrolled in the program across the five cities. Table 1 summarizes preliminary program outcomes by implementation site (city). Keeping with the framework discussed here, it describes three types of change occurring during the pilot program implementation- change in the behavior of program participants (utility customers), change in the LIFT-UP intervention during implementation, and change and integration of the program within the utility collection system now that the pilot period has ended.

[Insert Table 1 Here]

As noted on Table 1, there were substantial differences in take-up across the five sites, ranging from a low of about 3% in Louisville to a high of 24% in Newark. A variety of factors contributed to these differences, including the amount of additional eligibility screening that cities employed and the methods of outreach to engage participants in the program. A high

[6]One of the authors of this paper is leading the evaluation team for the LIFT-UP pilot. The data on the LIFT-UP pilot shared here is preliminary and is not publicly available. The final evaluation report for LIFT-UP will be complete and made public in February, 2016.

take-up rate does not necessarily equate to a high success rate in the program. In fact, turning to outcomes, both cities at the extreme ends of the take-up distribution had difficulty tracking data through the pilot period for all customers. They had developed relatively customized programs in response to their unique target groups (moderate adaptation), but also had difficulty sustaining support within their cities for ongoing data collection and program monitoring.

For the cities that tracked outcome data, the results in Table 1 indicate the change in water shut-off rates for the treatment group during the first three months after program enrollment, relative to the rates of water shut-off for the control group in that particular city (a difference in difference approach). Statistically significant lower ($p<0.01$) rates of water shut-off were observed in two out of the five implementation sites, where water shutoff relative to the control group was reduced by 23-32%.

In addition to tracking outcomes in relation to the target group, Table 1 reports the degree of adaptation of the program intervention during implementation. While there were core components of LIFT-UP across all five cities, the funder provided each municipality flexibility and encouragement to tailor the intervention based on the capacity and constraints of their utility collection system, the nature of the financial empowerment services currently operating in their cities, and the needs of the target population in their municipality. Some of the cities adapted the program considerably in response to the needs of their customers, with the treatment deviating substantially from the control group conditions (the treatment contrast). For example, in St. Petersburg, the frontline collections officer called each of the 90+ customers enrolled in LIFT-UP on a regular basis to monitor their progress through the program. In Houston, the frontline utility staff completed training

in “financial coaching” and also provided individualized support to LIFT-UP customers, not typically available to customers. By contrast, in Savannah, while LIFT-UP customers were referred to external financial counseling they were not required to attend.

Table 1 also reports the extent to which the LIFT-UP program model became integrated into the ongoing utility collection practices after the end of the pilot program. Only one city (Houston) has maintained all components of the program and continues to enroll new customers on an ongoing basis. Two other cities have maintained at least some components of the program and planned additional waves of a modified LIFT-UP intervention with customers who are delinquent on their utility bills. Finally, two of the cities have reported that they are unable to continue the LIFT-UP program into the future due to lack of system capacity and support within their municipality.

While this example provides some preliminary indication that the LIFT-UP intervention can be successful at promoting behavioral change in the target population, it also illustrates that this is not always the case. Like most multi-site pilot programs, the impact is not uniform across implementation sites. And, it is not clear that the intervention can be sustained over time; only in one city are all of the components full integrated into the system, with partial integration in two additional cities.

What factors enabled or constrained the change observed in the cities? Clearly, there is variation across sites, and one approach might be to try to code various elements of each site, predicting the likelihood of success based on the presence or absence of specific variables. However, with only five sites, there is insufficient power for any rigorous analysis. And, moreover, the purpose of such analysis is typically to isolate the independence of specific factors on implementation outcomes (e.g., through a multivariate regression).

However, system factors do not operate independently; the presence of a factor in one city may interact with other conditions to enable change, while doing nothing in another city. It is not clear that such an analysis that seeks to isolate the impact of specific factors would be either appropriate or particularly informative.

Instead, we think it is useful to think more holistically about each site as a dynamic system, comprised of multiple strategic action fields in which actors consider the idealized intervention in relation to what is known about the context – its existing institutional infrastructure, the operational means of coordination, the beliefs and norms of key actors and frontline staff. Understanding these mechanisms found within the contexts provides a means to appreciate the realities and possibilities of fit, and the interactions to create observed outcomes. Table 2 compares the five sites based on the elements of the framework described here for illustrative purposes only (a full analysis of these dynamics for LIFT-UP is beyond the scope of this paper).

[Insert Table 2 Here]

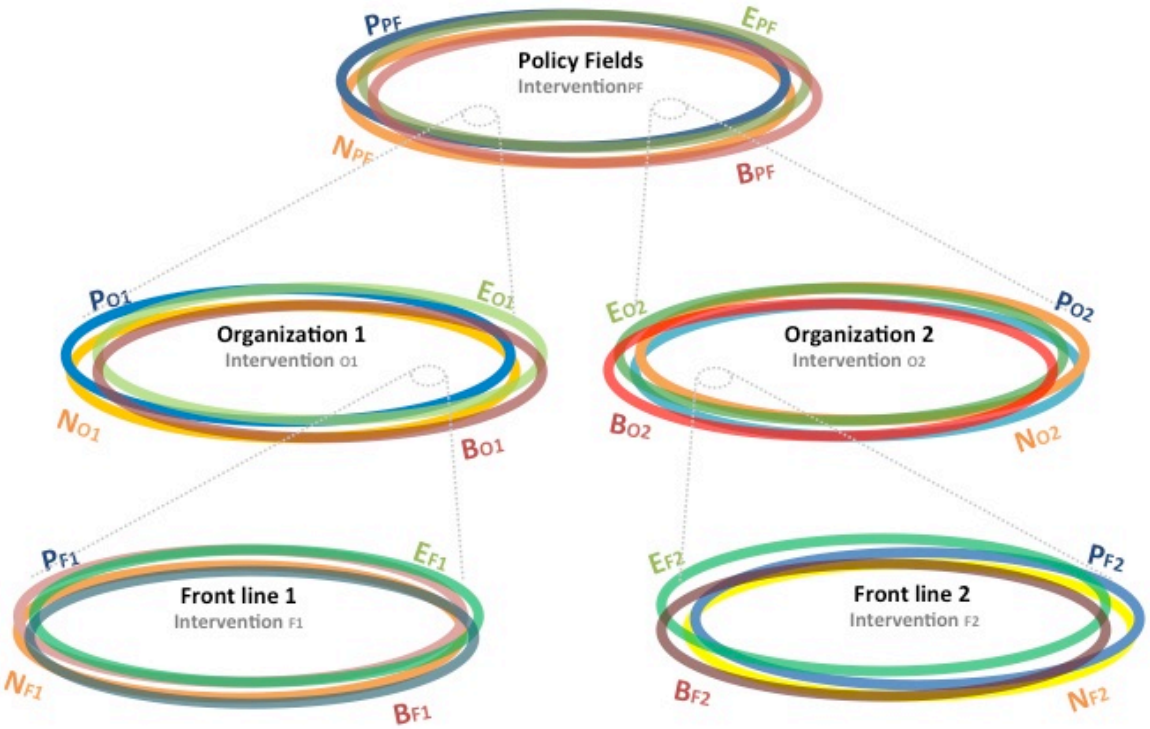
Across the five sites, Table 2 documents the three components of our framework – the operative authority in the context; the coordination of the core intervention, and the social skill demonstrated by actors to justify changes in terms of improving target group conditions. It focuses on some easily observed formal authority sources – the degree of support by the political leader (Mayor) and the rules bounding actions within the utility companies. A more complete analysis would use a detailed instrument (see Appendix 1) which would document more completely the mechanisms of authority. The Table also summarizes differences in technological capacity to coordinate activities around the intervention, such as the existing processes of collection on past-bills and reporting on

defaults, varied across these cities. Finally, it summarizes capacity of actors to mobilize others through social skill that supports the change of the pilot program. Again, a more complete analysis would probe evidence of social skill in various strategic action fields found in each system (as suggested by Appendix 1).

Conclusion

In this paper, we offer a way to think about how evidence based interventions can successfully—or unsuccessfully—become integrated into existing governance and service delivery systems. While fidelity to an intervention is often the goal to document impact on a target population, this goal may actually thwart the very impact that is being sought. Creating sustained change in target groups requires interventions that can be sustained within implementation settings. Here we suggest a framework which considers the fit between the intervention, target population and implementation system, arguing that sustainability of an intervention depends upon the alignment. We argue that actors within the system must use their knowledge of the context and their concern for the target population to appropriately adjust interventions.

FIGURE 1: MULTI-LEVEL STRATEGIC ACTION FIELDS IN IMPLEMENTATION SYSTEMS



KEY:
P= Political authority at different levels in the system
E = Economic authority at different levels in the system
N= Norms at different levels in the system
B= Beliefs & Values at different levels in the system

PF=Policy Field
O= Organization
F= Frontlines

FIGURE 2: TYPES OF INTERRELATED CHANGE

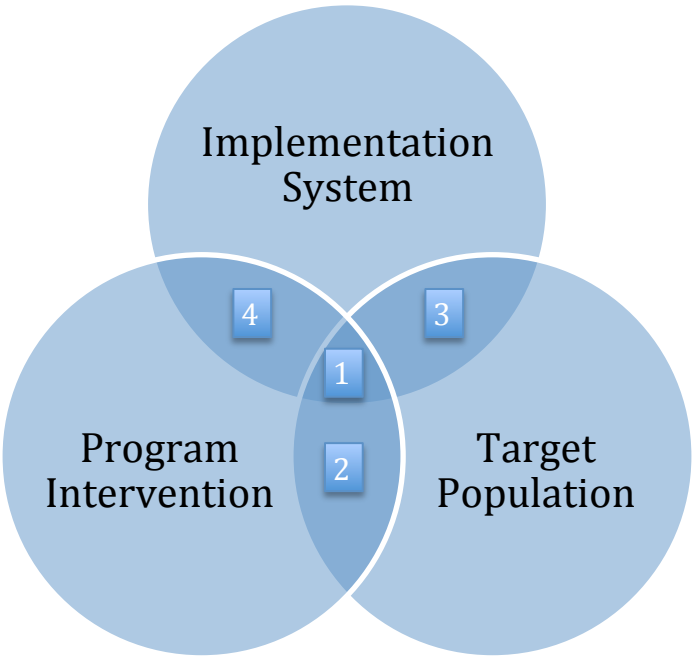


FIGURE 3: MUNICIPAL FINANCIAL EMPOWERMENT AND DEBT COLLECTION

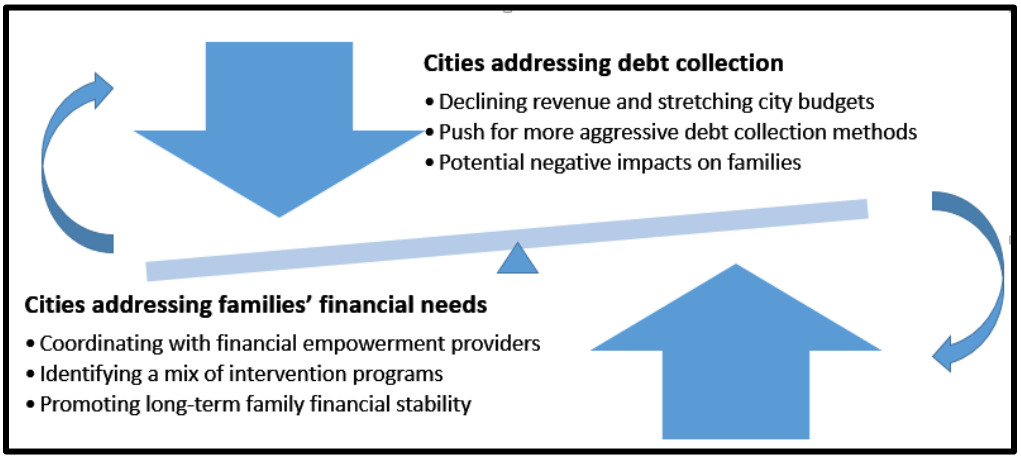


TABLE 1. LIFT-UP OUTCOMES: THREE TYPES OF CHANGE

	Savannah	**Newark**	**St. Pete**	**Louisville**	**Houston**
Change in Target Group Behavior					
Take Up Rate (Number Offered Services)	10.8% (900)	23.7% (266)	13.4% (680)	2.7% (730)	5.9% (615)
Change in Shutoff Rate (Significance)	-7% (NS)	Insufficient Data	-32% ($p<0.01$)	Insufficient Data	-23% ($p<0.01$)
Change in Program Intervention					
Degree of adaption to meet customer needs	Low adaptation	Moderate adaptation	High adaptation	Moderate adaptation	High adaptation
Change in Implementation System					
Degree of integration of LIFT-UP after pilot period	No integration	No integration	Partial integration	Partial integration	Full integration

TABLE 2. SYSTEM DYNAMICS IN LIFT-UP PROGRAM IMPLEMENTATION

	Savannah	Newark	St. Pete	Louisville	Houston
Operative Authority					
Strength of Support (Sources and direction)	Weak (City utility rules - ; Mayor +)	Moderate (Mayor +; Mayoral transition -)	Strong (City utility rules +; Mayor and city council+)	Moderate (Mayor+; city FE service providers +; city utility rules -)	Moderate (city utility transition -; city utility rules +)
Technological Capacity					
Flexibility of Existing Technology (Source and direction)	Weak (utility collection and reporting system -; flexibility in collections-)	Strong (utility collection and reporting system +; flexibility in collections+)	Strong (utility collection and reporting system +; flexibility in collections+)	Moderate (utility collection and reporting system +/- due to transition; flexibility in collections+)	Moderate (utility collection and reporting system +/- due to transition; flexibility in collections+)ate
Evidence of Social Skill					
Strength of Demonstrated Skill (Source and direction)	Strong (Strong champion with brokering and framing skills located outside of the utility system)	Weak (Champion left utility shortly after launch)	Moderate (Champion within utility; low conflict, lower need for social skill)	Moderate (Champion within city; some difficulty brokering with utility)	Strong (Champion within utility; used brokering and framing to secure resources)

REFERENCES

Aarons, Gregory et al. 2012. "The Organizational Social Context of Mental Health Services and Clinician Attitudes toward Evidence-Based Practice: A United States National Study." *Implementation Science* 7:15.

Aarons, Gregory A. 2005. "Measuring Provider Attitudes Toward Evidence-Based Practice: Consideration of Organizational Context and Individual Differences." *Child and Adolescent Psychiatric Clinics of North America* 14(2):255–71.

Aarons, Gregory A., Michael Hurlburt, and Sarah McCue Horwitz. 2011. "Advancing a Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors." *Administration and Policy in Mental Health and Mental Health Services Research* 38(1):4–23.

Agranoff, Robert, and Michael McGuire. *Collaborative Public Management: New Strategies for Local Governments*. Washington, DC: Georgetown University Press, 2003.

Amir, On, Dan Ariely, Alan Cooke, David Dunning, Nicholas Epley, Uri Gneezy, Botond Koszegi. "Psychology, Behavioral Economics, and Public Policy." *Marketing Letters* 16, nos. 3–4 (2005).

Aarons, Gregory A. (2004). "Mental Health Provider Attitudes twoard Adoption of Evidence-Based Practice: The Evidence-Based Practice Attitude Scale. *Mental Health Service Review* 6(2): 61-74.

Bardach, E. *The Implementation Game*. Cambridge: MIT Press, 1977.

Berman, Paul. *The Study of Micro and Macro Implementation of Social Policy* (Santa Monica, CA: Rand, 1978).

Blasé, Karen, Dean Fixsen, Barbara Sims and Caryn Ward (2015). "Implementation Science: Changing Hearts, Minds, Behavior, and Systems to Improve Educational Outcomes." National Implementation Research Network: University of North Carolina, Chapel Hill.

Blau and Scott 1962, *Formal Organizations: A Comparative Approach*. San Francisco: Chandler;

Daft, Richard L. *Organization theory and design*. Cengage learning, 2010.

Chen, Huey-Tsyh and Peter H. Rossi (1989). Issues in a Theory Driven Perspective," *Evaluation and Program Planning* 12(4): 299-306.

Damschroder, Laura J. et al., "Fostering Implementation of Health Services Research Findings into Practice: A Consolidated Framework for Advancing Implementation Science.," *Implementation Science* 4 (January 2009): 50.

deLeon, Peter, and Linda deLeon. "What Ever Happened to Policy Implementation? An Alternative Approach." *Journal of Public Administration Research and Theory* 12, no. 4 (October 2002): 467–492.

Durlak, Joseph a, and Emily P DuPre. "Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation." *American Journal of Community Psychology* 41, no. 3–4 (June 2008): 327–50.

Eccles and Mitlman, 2006.

Elmore, Richard F. "Backward Mapping : Implementation Research and Policy Decisions." *Political Science Quarterly* 94, no. 4 (1980): 601–616.

Fixsen, Dean L., Sandra F Naoom, Karen A Blase, Robert M Friedman, Barbara Burns, and Frances Wallace. *Implementation Research: A Synthesis of the Literature*. Tampa, Fl, 2005. doi:FMHI Publication #231.
Fligstein, Neil. "Social Skill and the Theory of Fields," *Sociological Theory* 19, no. 2 (2001): 105-125.
Fligstein, Neil, and Doug McAdam. *A Theory of Fields*. Oxford, England: Oxford University Press, 2012.
Fligstein, Neil, and Doug McAdams. "Toward a General Theory of Strategic action fields," *Sociological Theory* 29: 1 (2011): 1–26.
Fligstein, Neil. 2008. "Fields, Power and Social Skill: A Critical Analysis of the New Institutionalisms." *International Public Management Review* 9(1):227–53.
Greenhalgh et al., "Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations," *The Milbank Quarterly* 82, no. 4 (2004): 581–629.
Hall, Gene, Allison Dymnicki, Jennifer Coffey, and Melissa Brodowski. 2014. *Using Evidence-Based Constructs to Assess Extent of Implementation of Evidence-Based Interventions*. Washington D.C.
Hargreaves, Margaret B. 2010. *Evaluating System Change: A Planning Guide*. Princeton, New Jersey.
Hjern, Benny and David O. Porter, "Implementation Structures: A New Unit of Administrative Analysis," *Organization Studies* 2, no. 3964 (1981): 211–227.
Hill, Carolyn and Lawrence Lynn (2005). "Is Hierarchical Governance in Decline? Evidence from Empirical Research," *Journal of Public Administration Research and Theory* 15(2): 173-195.
Hill, Heather C. "Understanding Implementation: Street-Level Bureaucrats' Resources for Reform." *Journal of Public Administration Research and Theory* 13, no. 3 (2003): 265–282.
Hill, Michael and Peter Hupe. *Implementing Public Policy: An Introduction to the Study of Operational Governance*, 2nd ed. Thousand Oaks, California: Sage Publications, 2008.
Hirshmann, and Chriqui, 2013.
Horner, Rob, Caryn Blitz, and Scott W. Ross. 2014. "The Importance Of Contextual Fit When Implementing Evidence-Based Interventions." *ASPE Issue Brief* (September):1–16.
Ingram, Helen and Anne Schneider, "Improving Implementation through Framing Smarter Statutes," *Journal of Public Policy* 10:1 (1990): 67–88.
Khademian, Anne M. 2002. *Working with Culture: The Way the Job Gets Done in Public Programs*. Washington D.C.: CQ Press.
Lin, Ann Chih. 2000. *Reform in the Making: The Implementation of Social Policy in Prison*. Princeton, New Jersey: Princeton University Press.
Lipsky, Michael. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York, NY: Russel Sage Foundation, 1980.
May, Carl. "Towards a General Theory of Implementation." *Implementation Science* 43, no .3 (2009): 535–554.
May, Carl R. et al. "Development of a Theory of Implementation and Integration: Normalization Process Theory." *Implementation Science* 8:1 (2013).
May, C. and T. Finch. 2009. "Implementing, Embedding, and Integrating Practices: An Outline of Normalization Process Theory." *Sociology* 43(3):535–54.
May, Carl R. et al. 2009. "Development of a Theory of Implementation and Integration:

Normalization Process Theory." *Implementation Science* 4(29).
Mazmanian, Daniel A. and Paul A. Sabatier. *Implementation and Public Policy.* Lanham, Maryland: University Press of America, 1989.
Meier, Kenneth J. Laurence J. O'Toole, and Sean Nicholson-Crotty, "Multilevel Governance and Organizational Performance: Investigating the Political-Bureaucratic Labyrinth," *Journal of Policy Analysis and Management* 23, no. 1 (2004): 31–47.
Meyers, Duncan C., Joseph A. Durlak, and Abraham Wandersman. "The Quality Implementation Framework: A Synthesis of Critical Steps in the Implementation Process." *American Journal of Community Psychology* 50, nos. 3–4 (2012): 462–80.
Milward, H Brinton, and Keith G Provan. "Managing Networks Effectively." University of Arizona School of Public Administration and Policy, 2003.
Nilsen, Per, Christian Ståhl, Kerstin Roback, and Paul Cairney. "Never the Twain Shall Meet?--a Comparison of Implementation Science and Policy Implementation Research." *Implementation Science* 8, no. 63 (January 2013).
Orlikowski, Wanda J. 1992. "The Duality of Technology: Rethinking the Concept of Technology in Organizations." *Organization Science* 3(3):398–427.
Ostrom, Elinor. "Background on the Institutional Analysis and Development Framework," *Policy Studies Journal* 39, no. 1 (2011): 7–27.
Palumbo, Dennis J., and Donald J. Calista. *Implementation and the Policy Process: Opening up the Black Box*. New York: Greenwood Press, 1990.
Provan, Keith G. and Patrick Kenis. 2007. "Modes of Network Governance: Structure, Management, and Effectiveness." *Journal of Public Administration Research and Theory* 18(2):229–52.
Robichau, Robbie Waters and Laurence E. Lynn, Jr. "The Implementation of Public Policy: Still the Missing Link," *Policy Studies Journal* 37, no. 1 (2009): 21–36.
Roll, Stephen, Jodi Sandfort, and Stephanie Moulton. 2015. "A Comparative Analysis of Two Streams of Policy and Program Implementation Research." *Public Policy and Administration*.
Salamon, Lester. *The Tools of Government: A Guide to the New Governance* (Oxford: Oxford University Press, 2002),
Sandfort, Jodi R. and Stephanie Moulton. 2015. *Effective Implementation in Practice: Integrating Public Policy and Managment*. San Francisco, CA: Jossey-Bass.
Sandfort, Jodi. "Human Service Organizational Technology." *Human Services as Complex Organizations*, ed. Yeheskel Hasenfeld, 2nd ed. (New York: Sage, 2010), 269-290.
Scott, Richard W. *Institutions and Organizations: Ideas and Interests*, 3rd ed. (2008) Los Angeles: Sage Publications.
Shafir, 2013.
Stone, Melissa and Jodi R. Sandfort, "Building a Policy Fields Framework to Inform Research in Nonprofit Organizations," *Nonprofit and Voluntary Sector Quarterly* 38, no.6 (2009): 1054–1075.
Suchman 1995, "Managing Legitimacy: Strategic and Institutional Approaches," Academy of Management Review 20:571-610.
Supplee, Lauren and Allison Metz. 2015. "Opportunities and Challenges in Evidence-Based Social Policy." *Social Policy Report* 28(4):3–14.

Tabak, Rachel G., Elaine C. Khoong, David A. Chambers, and Ross C. Brownson, "Bridging Research and Practice: Models for Dissemination and Implementation Research," *American Journal of Preventive Medicine* 43, no. 3 (2012): 337–50.

Weiss, Carol Hirschon. 1987. "Evaluating Social Programs: What Have We Learned?" *Society* 25(1): 40–45.

Weiss, Michael J., Howard S. Bloom, and Thomas Brock. 2014. "A Conceptual Framework for Studying the Sources of Variation in Program Effects." *Journal of Policy Analysis and Management* 33(3):778–808.

Werner, Alan. (2004). *A Guide to Implementation Research*. Washington, DC: Urban Institute Press.