POLICY ESSAY

Scaling Effective Innovations

Dean L. Fixsen
Karen A. Blase
University of North Carolina

Amanda A. M. Fixsen
Invest in Kids

I mplementation and scaling evidence-based practices and other effective innovations is a challenge in criminology as it is in all social sciences. Yet, scaling is essential to achieving socially significant outcomes envisioned in state and federal legislation and recommendations for reform (Charles Colson Task Force, 2016). In their article, Christopher J. Sullivan, Brandon C. Welsh, and Omeed S. Ilchi (2017, this issue) provide one of the few attempts to estimate the social impact of known variables related to an evidence-based program. Their computer simulation model holds promise for estimating outcomes of scaling in human service settings. The computer model is useful for evidence-based interventions that have substantial data on implementation influences (impact of high or low fidelity) and on intervention influences (impact on offense rates or recidivism).

In this policy essay, we explore scaling more broadly with emphases on factors that should be considered, not just those that are available. *Scaling* is defined as the extent to which an innovation is used with good effect in the entire population of interest. It is argued that delivering the innovation as intended (with fidelity) and sustaining innovations in practice are essential to achieving social impact. *Innovations* refer to something new and previously

unused by potential users or recipients (Rogers, 1995). *Human services* include all forms of organized prevention and intervention efforts in corrections, health, education, and social services (Herzberg, 2015). *Implementation* is defined as the supports required to purposefully and reliably produce full and effective uses of innovations in practice (Fixsen, Blase, Metz, and Van Dyke, 2015). *Implementation capacity* refers to the availability of implementation teams with the knowledge, skills, and abilities to develop competencies and affect change with practitioners, organizations, and systems. *Social impact* is defined as the product of effective innovations used with fidelity (the numerator) at a sufficient scale to produce noticeable benefits for the entire population of interest (the denominator) (Perl, 2011).

In this conception, scaling innovations requires scaling implementation capacity. To illustrate this point, we compare atom-based innovations with interaction-based innovations and we highlight common components for producing and sustaining effective innovations at scale. Given the significant public and private investments in developing innovations that are intended to improve human services (Clancy, 2006), scaling the use of effective innovations in practice has global implications for the safety, health, and happiness of individuals in society.

As pointed out by Sullivan et al. (2017), information is becoming available to predict scaling outcomes. This allows a move away from relying on the broad conception that scaling must always take into account a "discount" or "penalty" of unknown magnitude (Tommeraas and Ogden, 2016). First, good progress is being made in developing effective innovations in human services (Abrams et al., 2012; National Institute of Corrections, 2004). The Nurse-Family Partnership (NFP) program cited by Sullivan et al. is an excellent example of an effective innovation. Second, key implementation factors related to the use of evidence-based programs and other innovations in typical practice settings are being operationalized and tested in implementation science (Glisson et al., 2010; Metz et al., 2014). There are now more than 30 implementation frameworks (Meyers, Durlak, and Wandersman, 2012; Tabak, Khoong, Chambers, and Brownson, 2012). With advances in research, frameworks are increasingly evidence based. For example, the active implementation frameworks (Fixsen et al., 2015) are based on a thorough review of the implementation evaluation literature (Fixsen, Naoom, Blase, Friedman, and Wallace, 2005) and systematic examination of practice knowledge from experienced developers and implementers (Blase, Fixsen, Naoom, and Wallace, 2005). Finally, there is growing recognition that current systems produce current results and systems must change if new and improved results are to be realized (Clay-Williams, Nosrati, Cunningham, Hillman, and Braithwaite, 2014). Although work continues on developing innovations and improving and extending the science of implementation, attention is now being given to scaling human service innovations in communities across the nation and the globe to produce social impact.

With scaling defined as the extent to which an innovation is used with good effect in the entire population of interest, *only* knowing how many individuals, organizations, or systems are benefiting from the use of an innovation means little as a scaling metric

(Fixsen, 2009; Horner et al., 2014; Spicer et al., 2014). The *denominator* and the quality of the *numerator* both need to be defined and operationalized when determining scale and assessing social impact.

The Denominator

The denominator for scaling is the entire population of intended beneficiaries. For example, the U.S. Bureau of Justice Statistics reported that more than 2 million adults were incarcerated in 2013 and nearly 5 million adults were on probation or parole. The Office of Juvenile Justice and Delinquency Prevention documented that, on average, more than 55,000 delinquent youth were incarcerated in Colorado on any given day in 2014. The National Institute of Mental Health estimated that more than 10 million adults have one or more forms of severe mental illness in the United States. The World Bank is concerned with alleviating extreme poverty for greater than 1 billion people in 188 countries living on US$1.50 or less a day. In education, approximately 60 million students attend 98,000 schools in the United States. An estimated 250,000 people in the United States have been diagnosed with pulmonary hypertension, a rare and fatal disease. As these examples illustrate, *the denominator for scaling varies greatly across domains, levels of systems, and specific populations of concern that could benefit.*

Scaling is essential to realizing the promised benefits of innovations in human services. Innovations cannot produce social impact unless they are used as intended in practice with the population of intended beneficiaries.

The Numerator

Although identifying and quantifying the population of intended beneficiaries (the denominator) is straightforward, determining the numerator in human service applications is complicated. The numerator is the number of members of the population who experience the innovation (with fidelity) and, therefore, have the opportunity to benefit. *The numerator for scaling is based on the number of recipients experiencing the innovation with fidelity.* The quality of the innovation as it is delivered in practice operationally defines the numerator.

Fidelity, the extent to which an innovation is used as intended, is an important determinant of the social benefit of innovations (Miller and Miller, 2015; Naleppa and Cagle, 2010). A case study of the importance of fidelity in scaling involves the inoculation of children with the Salk polio vaccine. Historically, some early batches of the newly developed Salk polio vaccine used by doctors to inoculate schoolchildren were substandard (low-fidelity production). Those substandard batches of polio vaccine *increased* the chances of children contracting the disease, which is not an outcome that benefits recipients or society as a whole and certainly not one to be scaled up. Those children who suffered the low-quality vaccine received an injection but did not receive a high-fidelity inoculation for polio. They would not count in the numerator related to scaling because they did not experience the intervention delivered as intended. Fortunately, medical personnel quickly detected and

corrected the problem, and the damage to children, although severe for infected children, was limited for the population (Bayly, 1956). Assessments of fidelity in human services are not common, and their absence presents problems for calculating the numerator for scaling.

Numerator Moderators

The quality of an innovation as it is delivered in practice is at risk because human services are *interaction based*—they involve a human provider and a human recipient, each with unique personal characteristics and changing emotional states and circumstances, that influence one another. In many fields, innovations are *atom based*. For example, when developing improved highway bridges (McNichol, 2006) and new components for rockets (Boyle, 2003), the innovations are atom based, meaning they remain static after they are developed and do not change from one engineer to the next, shift into new configurations overnight, or decide not to show up. Even though they may be complex, developing and scaling atom-based innovations present "tame problems," whereas interaction-based innovations present "wicked problems" (Rittel and Webber, 1973). Given the complexity of multiple humans interacting with and influencing one another, wicked problems are difficult to define, "fight back" when solutions are attempted, and require continual resolution (resolution) to maintain and improve impact. Wicked problems associated with interaction-based innovations impact the numerator for scaling because they impact quality/fidelity. When attempting to scale interaction-based innovations, fidelity is difficult to achieve and even more difficult to maintain as practitioners, supervisors/coaches, managers, and directors come and go in a sea of change. When attempting to scale interaction-based innovations, multiple challenges may add up to produce a negative effect that "is somewhat larger than the expected sum of the different factors" as indicated in the findings by Sullivan et al. (2017). This nonlinear effect further underscores the complexity of scaling up.

Given the complexity of interaction-based innovations, the quality of the numerator can be difficult to assess. A good fidelity measure assesses the extent to which the essential components of an interaction-based innovation are present and used as intended (with fidelity) in daily practice (Schoenwald and Garland, 2013). In addition, a good fidelity measure is highly related to outcomes—the better the innovation is delivered in practice, the better the outcomes for users and recipients. With a strong relationship between the two, if fidelity is known, then outcomes can be predicted (Bond and Salyers, 2004).

The complexity of interaction-based innovations also requires precise and persistent *implementation capacity* to produce high-fidelity uses of innovations reliably over time and across staff. For example, in the computer simulation run by Sullivan et al. (2017), arrest rates *increased* by 70% "when implementation factors are not fully considered." Implementation capacity is largely missing in human service systems (Schofield, 2004). The combination of lack of fidelity measures and the near absence of implementation capacity stand in the way of generating a numerator that meets and sustains a standard for quality (e.g., high fidelity) needed to produce socially significant outcomes.

Another impediment to scaling is *sustainability*. A numerator cannot continue to increase without attention to sustaining the quality of an innovation by those who already are using the innovation. Out of necessity, developers initially direct attention to expanding the number of new users ("adoption" of innovations). Over time, as the number of users increases, more attention must be paid to de-adoption and re-adoption (Massatti, Sweeney, Panzano, and Roth, 2008; McIntosh, Mercer, Nese, Strickland-Cohen, and Hoselton, 2015) while continuing to expand the numerator by developing new users. A container that is draining as fast as it is being filled simply maintains the same level. Likewise, if the number of users "de-adopting" an innovation equals the number of new users "adopting" an innovation, the numerator stays the same and scaling plateaus at that point. Any successful scaling attempt must attend to sustaining those who are already using an innovation while continuing to develop new users.

Scaling the Numerator

Given that the denominator (the population of interest) is somewhat fixed, *social impact* depends on scaling the numerator. In atom-based business environments, the production methods (manufacturing processes) for an innovation are essential to scaling (increasing the numerator). For example, the initial production of newly invented silicon-based integrated circuits was problematic with approximately one usable silicon wafer produced out of 100 attempts. When the engineers examined the low-fidelity wafers, they discovered that hair, flakes of dry skin, pesticides in the air (they were in the middle of citrus groves before Silicon Valley paved them over), and so on contaminated the manufacturing process. This led to the development of ultra-clean environments for producing silicon wafers and for assembling computer hardware and other electronics (Kaeslin, 2008). With improvements in production processes, manufacturing plants could send usable (high-fidelity) atom-based products out the door at an increasingly rapid pace (Ohno, 1988). Scaling is done by increasing the size and number of ultra-clean production facilities nationally and globally.

The job is not done when atom-based products leave the production plant. For sustained high-fidelity use, vaccines need to be stored properly, automobiles need to be serviced and repaired, computer software requires frequent updating (preferably automatically, to take users out of the process), and electronic parts are bound to fail under certain environmental conditions or after so many hours of use. Whole industries have been developed to maintain, repair, and improve atom-based products after they have been produced (sustain the numerator).

For interaction-based innovations, the challenges involved in production and maintenance are compounded many times over. "Production" depends on the skill of an implementation team (Metz et al., 2014) that does staff selection, training, coaching, and fidelity assessments in hospitable organization and system environments. "Maintenance" requires ongoing support from coaches and facilitative administrators and repeated assessments of fidelity and outcomes. Given the difficulties in reliably producing human competencies and

their sensitivity to many "contaminants" in human service environments that are anything but "ultra-clean," the shelf life of many interaction-based innovations is not long. In this sense, human interaction skills are "perishable" with a "use by" date that is short.

The development of implementation teams in human service organizations is tantamount to moving the production unit into the service environment (Metz et al., 2014; Schofield, 2004). Interaction-based innovations without the ongoing support of implementation teams stand little chance of high-fidelity use or sustained use in turbulent human service environments (Brunk, Chapman, and Schoenwald, 2014; Fixsen, Blase, Timbers, and Wolf, 2001; Tommeraas and Ogden, 2016).

If a human service system already had implementation capacity in the form of skilled implementation teams that were integrated and aligned to support the full and effective use of innovations, the social impact of effective innovations already would be at a high level and improving every year. Unfortunately, that is not the case (Guerino, Harrison, and Sabol, 2011; National Center for Education Statistics, 2013; Starfield, 2000).

Scaling Outcomes: Changing Systems

Any approach to scaling must (from the beginning) include plans to initiate and manage necessary system changes to increase implementation capacity and sustain high fidelity use of effective innovations. Human service systems are legacy systems (Ulrich, 2002) made up of an accumulation of fragments of past mandates, good ideas, beliefs, and ways of work that evolved over many decades as legislators, leaders, and staff came and went. Legacy systems typically are fragmented, siloed, and inefficient. To realize social impact, organizations and systems need to be designed on purpose to produce and sustain high-fidelity use of effective innovations. This requires changes in legacy systems, often in dramatic and transformative ways (Osborne and Gaebler, 1992), to produce, implement, sustain, and improve innovations. Transformation of systems may require only small changes that make a big difference (Lanham et al., 2013). For example, Ulrich (2002) found that transformed legacy systems retained approximately 80% of the original system. Only 20% had to change to produce dramatically improved performance. Nonetheless, distinguishing the 20% from the 80% presents problems when change must occur in human service systems that continue to function as-is during the time of transition. "Transformation zone" strategies and recursive practice-policy communication methods have been developed to discover the 20% that needs to change and then changing it in human service systems (Fixsen, Blase, and Van Dyke 2012; Ulrich, 2002).

In the scaling and system change literature, there is debate among researchers about the merits of top-down or bottom-up approaches (Klingner, Boardman, and McMaster, 2013; Pülzl and Treib, 2006), the wisdom of requiring fidelity or accepting adaptations in the use of innovations (Aarons et al., 2012; Szulanski and Jensen, 2008), and the value of policy-first or outcome-first strategies (Haskins and Baron, 2011; Manna, 2008). There also is discussion of scaling strategies that comprise more linear and planned approaches

(Winter and Szulanski, 2001) or complexity theory that emphasizes emergent forms and comfort with "not knowing" (Lanham et al., 2013; Manna, 2008).

In general, there is no "or" in scaling work, there is only "and." That is, top-down *and* bottom-up capacity development involves policy-first *and* outcome-first strategies where fidelity *and* adaption are required. A planned approach helps to operationalize the work to be done *and* complexity theory helps to make sense of these apparent contradictions and anticipates the emergent nature of the work.

The need for systemic change is apparent in an example from global health. Fenner, Henderson, Arita, JeZek, and Ladnyi (1988) recounted the 175-year history of the smallpox vaccine and the intensive global efforts that began in 1966 and led to the eradication of smallpox in 1979. The global mandate and billions of dollars invested in eradicating smallpox is an example of a "brute-force" approach (Christakos, 2011) to implementation and scaling. Fenner et al. (1988) and Foege (2011) described how the global health community had to identify and overcome quickly major impediments related to maintaining the quality and stability of the vaccine, supply of materials and qualified personnel in remote and low-resource areas, leadership and collaboration among nations, and shifting populations as a result of war and migration. Brute-force approaches may get the job done, but they are not focused on systemic change and do not leave improved systemic implementation capacity in their wake. Global health leaders eradicated smallpox but did not establish or sustain improved global or national health systems with the generalized capacity to help eradicate other diseases or health concerns in the subsequent decades (e.g., the AIDS epidemic that began in the 1980s).

Recommendations

The computer simulation model described by Sullivan et al. (2017) has the potential to guide investments in scaling effective innovations in criminology and human services. Given the importance of scaling, the simulation model holds great promise for policy makers and funders. A simulation approach is most useful when it considers inputs from (a) studies of the effectiveness of the innovation, and the relationship between fidelity and outcomes; (b) studies of the impact of implementation variables (e.g., staff selection, training, and coaching; organization supports; and leadership) on fidelity of the use of the innovation; and (c) studies of the influence of contextual factors (compliance monitoring, requirements for funding, support for training and coaching, and impact of competing initiatives) that may enable or hinder use of effective implementation methods and use of an effective innovation.

As discussed in this policy essay, scaling innovations in human services can be operationalized and social impact can be assessed in terms of the quality of innovations as they are used in practice. Knowing the denominator is a good place to begin. The power of scaling is in the ability to produce a numerator substantial enough to realize population benefits. The quality of the innovation as delivered in practice (fidelity) relies on purposeful use of

effective implementation methods. Scaled use of an effective innovation depends on first scaling implementation capacity to support, sustain, and improve high-quality replications of the innovation.

It makes little sense to expend resources in attempts to scale innovations that do not produce intended benefits. Pronouncements ("you have completed the workshop and we now declare you are using our innovation"), claims ("our group read your manual and we are using your innovation"), or policy statements ("we passed a law and funded an agency to take care of that") are not the same as using an innovation as intended (with fidelity), producing promised outcomes in typical practice settings, and sustaining the innovations and implementation supports so they become the new standard practices in changed human service systems.

Attempts to scale ineffective or harmful innovations are a waste of time, money, and opportunity. Yet, investing in strategies that do not work is the norm and not the exception as well-meaning legislators and leaders press for quick solutions in human service systems that are ill equipped to respond in ways that produce intended outcomes (Dodge, Dishion, and Landsford, 2006; Institute of Medicine, 2001; Lynam et al., 2004; Tyack and Cuban, 1995; U.S. Department of Education Institute of Education Sciences, 2010; U.S. Department of Health and Human Services, 1999, 2001). System administrators and managers simply do not have the levers for meaningful change available to them. It is like asking a scholar who only knows Roman numerals to calculate the solution to a complex math problem. The means are not available to achieve the ends. Investments in ineffective innovations and implementation strategies likely produce the jaded view of innovations and new initiatives often encountered in human service settings.

It is time for change in human services. Research to develop innovations is improving, and the number of evidence-based innovations is growing. Increasingly, we know WHAT to do in criminology. The availability of innovations that are effective but not used in practice has focused attention on implementation science. With rapid advances in implementation science, we are learning HOW to make effective use of innovations in typical human service organizations and systems (Tucker, Edmondson, and Nembhard, 2005). Now the task is to scale implementation capacity in the form of skilled implementation teams to produce on purpose social impact for the effective innovations at hand as well as those yet to be developed. Citizens of our nation and global community deserve no less.

References

Aarons, Gregory A., Amy E. Green, Lawrence A. Palinkas, Shannon Self-Brown, Daniel J. Whitaker, John R. Lutzker, Jane F. Silovsky, Debra B. Hecht, and Mark J. Chaffin. 2012. Dynamic adaptation process to implement an evidence-based child maltreatment intervention. *Implementation Science*, 7: 32.

Abrams, Elaine J., R. J. Simonds, Surbhi Modi, Emilia Rivadeneira, Paula Vaz, Chipepo Kankasa, Denis Tindyebwa, Ryan Phelps, Sara Bowsky, Chloe A. Teasdale, Emilia

Koumans, and Andrea J. Ruff. 2012. PEPFAR scale-up of pediatric HIV services: Innovations, achievements, and challenges. *Journal of Acquired Immune Deficiency Syndromes*, 60: 105–112.

Bayly, M. Beddow. 1956. The story of the Salk anti-poliomyelitis vaccine. Retrieved from whale.to/vaccine/bayly.html.

Blase, Karen A., Dean L. Fixsen, Sandra F. Naoom, and Frances Wallace. 2005. *Operationalizing Implementation: Strategies and Methods*. Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute.

Bond, Gary R. and Michelle P. Salyers. 2004. Prediction of outcome from the Dartmouth assertive community treatment fidelity scale. *CNS Spectrums*, 9: 937–942.

Boyle, Jon. 2003. *Working "Faster, Better, Cheaper": A Federal Research Agency in Transition*. Roanoke: Virginia Polytechnic Institute and State University.

Brunk, Molly A., Jason E. Chapman, and Sonja K. Schoenwald. 2014. Defining and evaluating fidelity at the program level: A preliminary investigation. *Zeitschrift für Psychologie*, 222: 22–29.

Charles Colson Task Force. 2016. *Transforming Prisons, Restoring Lives: Final Recommendations of the Charles Colson Task Force on Federal Corrections*. Washington, DC: Urban Institute.

Christakos, George. 2011. Implementation and technology. In (George Christakos, ed.), *Integrative Problem-Solving in a Time of Decadence*. Amsterdam, the Netherlands: Springer.

Clancy, Carolyn. 2006. The $1.6 trillion question: If we're spending so much on healthcare, why so little improvement in quality? *Medscape General Medicine*, 8: 58.

Clay-Williams, Robyn, Hadis Nosrati, Frances Cunningham, Kenneth Hillman, and Jeffrey Braithwaite. 2014. Do large-scale hospital- and system-wide interventions improve patient outcomes: A systematic review. *BMC Health Services Research*, 14: 369.

Dodge, Kenneth A., Thomas J. Dishion, and Jennifer E. Landsford. 2006. Deviant peer influences in intervention and public policy for youth. *Social Policy Report*, 20: 3–19.

Fenner, Frank, Donald A. Henderson, Isao Arita, Zdenek JeZek, and Ivan D. Ladnyi. 1988. *Smallpox and Its Eradication*. Geneva, Switzerland: World Health Organization.

Fixsen, Amanda A. M. 2009. *Defining Scaling Up Across Disciplines: An Annotated Bibliography*. Portland, OR: Portland State University.

Fixsen, Dean L., Karen A. Blase, Allison Metz, and Melissa Van Dyke. 2015. Implementation science. In (James D. Wright, ed.), *International Encyclopedia of the Social and Behavioral Sciences*. Oxford, U.K.: Elsevier.

Fixsen, Dean L., Karen A. Blase, Gary D. Timbers, and Montrose M. Wolf. 2001. In search of program implementation: 792 replications of the Teaching-Family Model. In (Gary A. Bernfeld, David P. Farrington, and Alan W. Leschied, eds.), *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs*. London, U.K.: Wiley.

Fixsen, Dean L., Karen A. Blase, and Melissa Van Dyke. 2012. *From Ghost Systems to Host Systems Via Transformation Zones*. Washington, DC: U.S. Department of Education Office of Vocational and Adult Education.

Fixsen, Dean L., Sandra F. Naoom, Karen A. Blase, Robert M. Friedman, and Frances Wallace. 2005. *Implementation Research: A Synthesis of the Literature*. Tampa: University of South Florida, National Implementation Research Network.

Foege, William H. 2011. *House on Fire: The Fight to Eradicate Smallpox*. Berkeley: University of California Press.

Glisson, Charles, Sonja K. Schoenwald, Anthony Hemmelgarn, Philip Green, Denzel Dukes, Kevin S. Armstrong, and Jason E. Chapman. 2010. Randomized trial of MST and ARC in a two-level evidence-based treatment implementation strategy. *Journal of Consulting and Clinical Psychology*, 78: 537–550.

Guerino, Paul, Paige M. Harrison, and William J. Sabol. 2011. *Prisoners in 2010* (Revised). Washington, DC: Bureau of Justice Statistics.

Haskins, Ron and Jon Baron. 2011. *Building the Connection Between Policy and Evidence: The Obama Evidence-Based Initiatives*. London, U.K.: National Endowment for Science, Technology and the Arts (UK).

Herzberg, Judith T. 2015. *Foundations in Human Services Practice: A Generalist Perspective on Individual, Agency, and Community*. Boston, MA: Pearson.

Horner, Robert H., Donald Kincaid, George Sugai, Timothy Lewis, Lucille Eber, Susan Barrett, Celeste Rossetto Dickey, Mary Richter, Erin Sullivan, Cyndi Boezio Bob Algozzine, Heather Reynolds, and Nanci Johnson. 2014. Scaling up school-wide positive behavioral interventions and supports: Experiences of seven states with documented success. *Journal of Positive Behavior Interventions*, 16: 197–208.

Institute of Medicine. 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press.

Kaeslin, Hubert. 2008. *Digital Integrated Circuit Design, From VLSI Architectures to CMOS Fabrication*. Cambridge, U.K.: University of Cambridge Press.

Klingner, Janette K., Alison G. Boardman, and Kristen L. McMaster. 2013. What does it take to scale up and sustain evidence-based practices? *Exceptional Children*, 79: 195–211.

Lanham, Holly Jordan, Luci K. Leykum, Barbara S. Taylor, C. Joseph McCannon, Curt Lindberg, and Richard T. Lester. 2013. How complexity science can inform scale-up and spread in health care: Understanding the role of self-organization in variation across local contexts. *Social Science & Medicine*, 93: 194–202.

Lynam, Donald R., Richard Milch, Rick Zimmerman, Scott P. Novak, T. K. Logan, Catherine Martin, Carl Leukefeld, and Richard Clayton. 2004. Project DARE: No effects at 10-year follow-up. *Journal of Consulting & Clinical Psychology*, 67: 590–593.

Manna, Paul. 2008. *Federal Aid to Elementary and Secondary Education: Premises, Effects, and Major Lessons Learned*. Williamsburg, VA: College of William and Mary, Department of Government and the Thomas Jefferson Program in Public Policy.

Massatti, Rick R., Helen Anne Sweeney, Phyllis C. Panzano, and Dee Roth. 2008. The de-adoption of innovative mental health practices (IMHP): Why organizations choose not to sustain an IMHP. *Administration and Policy in Mental Health and Mental Health Services Research*, 35: 50–65.

McIntosh, Kent, Sterett H. Mercer, Rhonda N. T. Nese, M. Kathleen Strickland-Cohen, and Robert Hoselton. 2015. Predictors of sustained implementation of school-wide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions*, 18: 209–218.

McNichol, Dan. 2006. *The Roads That Built America*. New York: Sterling.

Metz, Allison, Leah Bartley, Heather Ball, Dawn Wilson, Sandra Naoom, and Phil Redmond. 2014. Active Implementation Frameworks (AIF) for successful service delivery: Catawba County child wellbeing project. *Research on Social Work Practice*, 25: 415–422.

Meyers, Duncan C., Joseph A. Durlak, and Abraham Wandersman. 2012. The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, 50: 462–480.

Miller, J. Mitchell and Holly Ventura Miller. 2015. Rethinking program fidelity for criminal justice. *Criminology & Public Policy*, 14: 339–349.

Naleppa, Matthias J. and John G. Cagle. 2010. Treatment fidelity in social work intervention research: A review of published studies. *Research on Social Work Practice*, 20: 674–681.

National Center for Education Statistics. 2013. *The Nation's Report Card: Trends in Academic Progress 2012*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Institute of Corrections. 2004. *Implementing Evidence-Based Practice in Community Corrections: The Principles of Effective Intervention*. Washington, DC: National Institute of Corrections.

Ohno, Taiichi. 1988. *Toyota Production System: Beyond Large-Scale Production*. New York: Productivity Press.

Osborne, David and Ted Gaebler. 1992. *Reinventing Government*. New York: Addison-Wesley.

Perl, Harold I. 2011. Addicted to discovery: Does the quest for new knowledge hinder practice improvement? *Addictive Behaviors*, 36: 590–596.

Pulzl, Helga and Oliver Treib. 2006. Implementing public policy. In (Frank Fischer, Gerald J. Miller, and Mara S. Sidney, eds.), *Handbook of Public Policy Analysis*. New York: Dekker.

Rittel, Horst W. J. and Melvin M. Webber. 1973. Dilemmas in a general theory of planning. *Policy Sciences*, 4: 155–169.

Rogers, Everett M. 1995. *Diffusion of Innovations*. New York: The Free Press.

Schoenwald, Sonja K. and Ann F. Garland. 2013. A review of treatment adherence measurement methods. *Psychological Assessment*, 25: 146–156.

Schofield, Jill. 2004. A model of learned implementation. *Public Administration*, 82: 283–308.

Spicer, Neil, Dipankar Bhattacharya, Ritgak Dimka, Feleke Fanta, Lindsay Mangham-Jefferies, Joanna Schellenberg, Addis Tamire-Woldemariam, Gill Walt, and Deepthi Wickremasinghe. 2014. Scaling-up is a craft not a science: Catalysing scale-up of

health innovations in Ethiopia, India and Nigeria. *Social Science & Medicine*, 121: 30–38.

Starfield, Barbara. 2000. Is US health really the best in the world? *Journal of the American Medical Association*, 284: 483–485.

Sullivan, Christopher J., Brandon C. Welsh, and Omeed S. Ilchi. 2017. Modeling the scaling up of early crime prevention: Implementation challenges and opportunities for translational criminology. *Criminology & Public Policy*, 16: 457–485.

Szulanski, Gabriel and Robert J. Jensen. 2008. Growing through copying: The negative consequences of innovation on franchise network growth. *Research Policy*, 37: 1732–1741.

Tabak, Rachel G., Elaine C. Khoong, David A. Chambers, and Ross C. Brownson. 2012. Bridging research and practice: Models for dissemination and implementation research. *American Journal of Preventive Medicine*, 43: 337–350.

Tommeraas, Truls and Terje Ogden. 2016. Is there a scale-up penalty? Testing behavioral change in the scaling up of parent management training in Norway. *Administration and Policy in Mental Health and Mental Health Services Research*, 1–14.

Tucker, Anita, Amy Edmondson, and Ingrid Nembhard. 2005. *Contributions of Know-What and Know-How to Performance Improvement in Complex Service Organizations* (Paper 22). Boston, MA: Academy of Health Meeting.

Tyack, David and Larry Cuban. 1995. *Tinkering Toward Utopia: A Century of Public School Reform*. Cambridge, MA: Harvard University Press.

U.S. Department of Education Institute of Education Sciences. 2010. The condition of education. *NCES*. Retrieved from nces.ed.gov/programs/coe/.

U.S. Department of Health and Human Services. 1999. *Mental Health: A Report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services.

U.S. Department of Health and Human Services. 2001. *Youth Violence: A Report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services.

Ulrich, William M. 2002. *Legacy Systems: Transformation Strategies*. Upper Saddle River, NJ: Prentice Hall.

Winter, Sidney G. and Gabriel Szulanski. 2001. Replication as strategy. *Organization Science*, 12: 730–743.

Dean L. Fixsen is co-founder of the National Implementation Research Network and president of the Global Implementation Initiative. In 2005, he co-authored *Implementation Research: A Synthesis of the Literature* and has been active in developing implementation science as a professional discipline. From 1967 to 1992, he was co-director of the Teaching-Family Model research project that developed and replicated one of the few evidence-based residential care programs for youths in juvenile correction systems.

Karen A. Blase is co-founder of the National Implementation Research Network and emerita senior scientist at the University of North Carolina at Chapel Hill. She co-authored

Implementation Research: A Synthesis of the Literature and led efforts to operationalize and replicate implementation and scaling methods in human services. She was president of the Foster Family-Based Treatment Association and president of the Teaching-Family Association during the formative years when each group was setting standards for effective practice.

Amanda A. M. Fixsen is the director of implementation science at Invest in Kids with a focus on scaling The Incredible Years and The Nurse Family Partnership programs in Colorado. With a background in applied behavior analysis, social work research, and implementation science, she pursues her interest in how human and health services can be effectively implemented and scaled up in real-world contexts.