Science and Implementation

Dean L. Fixsen, Melissa Van Dyke, Karen A. Blase

Active Implementation Research Network

Active Implementation Research Network

Please cite as:

Fixsen, D. L., Van Dyke, M. K., & Blase, K. A. (2019). Science and implementation. Chapel Hill, NC: Active Implementation Research Network. www.activeimplementation.org/resources

Science and Implementation

Dean L. Fixsen, Melissa Van Dyke, Karen A. Blase

Active Implementation Research Network

Implementation science is being developed for use in human services and other domains (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Meyers, Durlak, & Wandersman, 2012; Tabak, Khoong, Chambers, & Brownson, 2012; Winter & Szulanski, 2001). While implementation science has been developing for some time (Saetren, 2005), in the last few decades it has begun to develop rapidly. For example, on December 15, 2018 a Google Scholar search for "implementation research" yielded 80,600 returns, 23% of them since 2014. A search for "implementation science" produced 46,100 returns with 40% of them since 2014. Rapid development can be attributed to recognition of the role of implementation in closing the science to service gap: the gap between substantial investments in developing effective innovations and the continuing need for significant improvements in health, education, and social services (Bryk, 2016; Kessler & Glasgow, 2011; Perl, 2011).

Given the increasing attention to implementation in a variety of contexts, the purpose of this article is to explore implementation science as a science. The idea of science, and the use of the scientific method to test predictions and hypotheses to advance science are discussed in the practical contexts faced by implementation scientists. Theory and the ability to establish implementation independent variables for study are cited as two areas for development.

What is Science?

The Merriam-Webster dictionary (https://www.merriam-webster.com/) defines *science* as the "system of knowledge covering general truths or the operation of general laws especially as obtained and tested through scientific method." The *scientific method* is defined as "principles and procedures for the systematic pursuit of knowledge involving the recognition and formulation of a problem, the collection of data through observation and experiment, and the formulation and testing of hypotheses." In his book *The Invention of Science*, Wootton (2015, p. 393) says, "What makes it science is not that it provides an explanation but that it provides reliable predictions."

Note that research can be done without contributing directly to science. A person may be curious about something and conduct research that subsequently adds to the knowledge base about that topic. The implementation field is overwhelmed with surveys and studies that add to the knowledge base without necessarily contributing to science. It is rare to find research that explicitly makes a theory-based prediction about an implementation variable then tests that prediction in practice (for a good example, see Seers et al., 2018).

Science-based research must outpace curiosity-based research if the interest in implementation is to further a science of implementation.

What is Implementation Science?

The Merriam-Webster dictionary defines *implement* as "to equip" and as a "device used in the performance of a task." In this sense, implementation science is the science of equipping people to perform tasks. After decades of study, the Active Implementation Research Network (Fixsen, Blase, Naoom, & Wallace, 2009) defines *implementation science* as "the study of factors that influence the full and effective use of innovations in practice. The goal of implementation science is not to answer factual questions about what is, but to determine what is required (mission driven)." Wootton points to the mission-driven nature of science by noting that "One must always proceed from wonder to no wonder; that is, one should continue one's investigation until that which we thought strange no longer seems strange to us." (Wootton, 2015, p 299, citing the Dutch philosopher and scientist Isaac Beeckman, 1626). The mission of implementation research is "to proceed from wonder to no wonder" as implementation knowledge is developed.

Theory

There are many things that potentially may have some bearing on how to improve the impact of proven or developing interventions. Unfortunately, "may have" does not equate to "does" in science. "Scientific inquiry" demands more than speculation; science requires a clearly stated theory that generates predictions (if-then) and testable hypotheses (Popper, 1963, 2002). There are multiple implementation frameworks (Meyers et al., 2012; Tabak et al., 2012) that could serve as a source of predictions (if-then) and hypotheses (explanations of if-then relationships).

For implementation science the Improved Clinical Effectiveness through Behavioural Research Group (2006) distinguishes among grand theories, mid-range theories, and micro-theories. These are summarized in Table 1.

Table 1. Levels of theory as defined by the Improved Clinical Effectiveness through Behavioural Research Group (2006).

Scope	Definition	Purpose
Grand or Macro Theory	A grand or macro theory is a very broad theory that encompasses a wide range of phenomena	A grand theory is a general construction about the nature and goals of a discipline
Mid-Range Theory	A mid-range theory is more limited in scope, less abstract, addresses specific phenomena, and reflects practice. It encompasses a limited number	Mid-range theory is designed to guide empirical inquiry. Mid-range theories are made up of relatively concrete concepts that are operationally defined and relatively

	of concepts and a limited aspect of the real world.	concrete propositions that can be empirically tested.
Micro Theory	A micro, practice or situation-specific theory (sometimes referred to as prescriptive theory) has the narrowest range of interest.	Focuses on specific phenomena that reflect clinical practice, and are limited to specific populations or to a particular field of practice.

The Active Implementation Frameworks (Blase & Fixsen, 2013; Fixsen, Blase, Metz, & Van Dyke, 2015; Fixsen et al., 2009) and the other implementation frameworks that have been developed can be considered mid-range theories of implementation.

Predictions and Hypotheses

If reliable predictions define science and if testing predictions is the work of scientists, then implementation science is a science to the extent that a) predictions are made and b) those predictions are tested in practice using the scientific method.

Predictions in the if-then logical format necessarily require a way to first produce the "if." In science, Wootton (2015) distinguishes between discovery and invention. If something already exists, a scientist or explorer can discover it. Gravity already existed before Galileo described the "law of fall" and Newton described it in a mathematical formula. In physics, chemistry, biology, and other "hard sciences" scientists can study natural phenomena that exist everywhere (e.g. every living thing has chromosomes that can be studied; chemical elements already exist and are waiting to be observed). The independent variable exists in nature and if-then relationships can be discovered by researchers.

On the other hand, inventions are new and do not exist in nature. In the so-called "soft sciences" the independent variable must be produced by the scientist. Unlike waiting for a solar eclipse to test Einstein's predictions of spacetime (Pasachoff, 2009), implementation scientists cannot wait for an expert implementation team (a postulated implementation independent variable) to form and begin to function and then assess the outcomes (Fixsen et al., 2005; Appendix D: Hypotheses for Advancing Implementation Science). This may never happen in any predictable and assessable way.

Implementation Independent Variables (if this …)

The independent variables in implementation science are inventions. To have an implementation science, implementation scientists must be able to produce the independent variable on demand so that predictions of its effects can be measured (if this, then that). At this point, soft-science increases in complexity. Consider the logic:

1. An implementation scientist must be able to produce the independent variable on demand (if this). For example, a prediction might be that if a high functioning implementation team supports practitioners' use of an evidence-based program, then the practitioners will use the evidence-based program with high fidelity and good outcomes.
2. Production of the independent variable itself requires implementation knowledge and skill. For example, who produces implementation teams? How can researchers produce implementation teams reliably and effectively so they can be studied within an experimental design?
   a. A contributor to the complexity of implementation science is that each independent variable required to test a prediction is, itself, a dependent variable in the context of a test of the prediction.

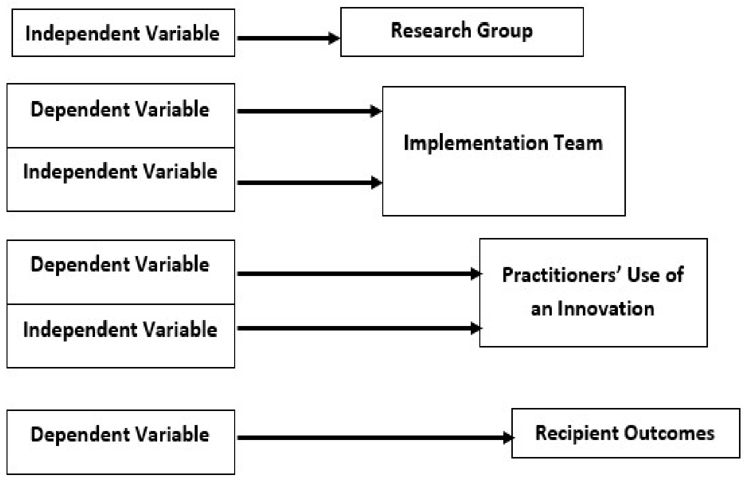

Figure 1. Each independent variable is a dependent variable in implementation research.

The overall logic is presented in Figure 1. A research group (an independent variable) must be sufficiently skilled in implementation practice to function as an Implementation Team (a dependent variable) so that the Implementation Team (an independent variable) can produce practitioners who use an innovation with fidelity (a dependent variable). At the next level, the practitioners' use of an innovation with fidelity (an independent variable) is assessed in terms of benefits to recipients (a dependent variable).

In this example, the fidelity of practitioners' use of an innovation is both a dependent variable and an independent variable. Perhaps this dual role helps to explain the number of studies of fidelity and its role in producing desired outcomes and in analyzing outcome data (e.g. Bartley, Bright, & DePanfilis, 2017; Durlak & DuPre, 2008; Hurley, Lambert, Gross, Thompson, & Farmer, 2017; Naleppa & Cagle, 2010; Sanetti & Kratochwill, 2014; Tommeraas & Ogden, 2016).

In addition, without an assessment of the fidelity of the implementation independent variable it is difficult to interpret the outcomes of intervention independent variables (Harvey, McCormack, Kitson, Lynch, & Titchen, 2018; Rycroft-Malone et al., 2018; Seers et al., 2018; Tiruneh et al., 2018).

How, then, can implementation independent variables be produced on purpose? Perhaps by using implementation best practices to produce them? The circularity of developing implementation science as a science becomes apparent – some implementation knowledge and skill are required to produce implementation independent variables suitable for study. The lack of assessments of Implementation Team performance is a problem when attempting to advance implementation as a science. Ultimately (see Figure 1), Implementation Team performance is a function of a research group's ability to use implementation best practices (Fixsen, Ward, Blase, et al., 2018) to establish effective Implementation Teams that support high fidelity use of an independent variable.

This calls for a different perspective on developing theory and testing hypotheses and predictions in implementation science. For example, Galileo had a powerful (at that time) telescope built to enhance his observations of the night sky. However, he did not have to produce the stars that were the object of the observations. The difference is implementation scientists need to "produce the stars." That is, they need to produce high functioning Implementation Teams if they want to study the potential effects of Implementation Teams. The independent variable is not already there waiting to be observed.

To establish implementation independent variables, implementation science must rely on implementation practice. Fortunately, implementation practice has been developing and improving since the middle of the 20th century. Best practices have been summarized and operationalized and lists of best practices have been organized into coherent frameworks to guide implementation practice in human services and other fields (for examples, see: Allanson et al., 2017; Blanchard et al., 2017; Blase, Fixsen, Naoom, & Wallace, 2005; Fixsen et al., 2005; Glennan Jr., Bodilly, Galegher, & Kerr, 2004; Greenhalgh, Robert, MacFarlane, Bate, & Kyriakidou, 2004; Meyers et al., 2012; Øvretveit, Mittman, Rubenstein, & Ganz, 2017; Schoenwald & Garland, 2013).

Implementation best practices are needed to establish implementation independent variables at the right time and place and with the quality required to test predictions and hypotheses.

Observation

Predictions in the if-then logical format require a way to observe the "then" to see if a prediction is disconfirmed in practice. When describing the invention of science, Wootton (2015, pp. 258-259) explains that "Before the Scientific Revolution facts were few and far between: they were handmade, bespoke rather than mass produced, they were poorly distributed, they were often unreliable."… "Establishing facts depends upon instruments ... which have to be standardized." … "Precision was pointless when units of measurement were local." Centuries ago, telescopes and microscopes provided new ways to observe with clarity and precision that far exceeded human sensory capacity. As the technologies improved, observation became reliable and repeatable from one group of scientists to another to inform science.

In implementation science, observation is a problem. Various measures exist and fit the "handmade" and "often unreliable" versions that pre-date the Scientific Revolution. Lewis et al.

(2015) catalogued 104 existing implementation-related measures and found them lacking in reliability, validity, and conceptual clarity. Proctor et al. (2011) described potential implementation measures derived from a review of concepts in the literature. Of the eight proposed measures, three (adoption, cost, fidelity) relate to implementation; three (acceptability, appropriateness/fit, feasibility) concern the innovation and two (sustainability, penetration). relate to scaling. Weiner et al. (2017) developed new implementation outcome measures where each item concerns the innovation ("This EBP meets my approval;" "This EBP seems applicable"), similar to acceptability, appropriateness/fit, and feasibility in the Proctor et al. list. Finally, reviews of the implementation research literature note that it is unusual for any measure to be used by more than one researcher (Allen et al., 2017; Fixsen et al., 2005).

The lack of useful and agreed-upon measures is a problem. If implementation independent variables (inventions) must be produced, then there must be some way to detect the presence and strength of the implementation independent variable in practice. Otherwise, the risk of Type III errors increases. Statistics calls attention to Type I errors (false positives) and Type II errors (false negatives). In implementation, a Type III error occurs when one is attempting to study the impact of a variable that does not exist in practice (Dobson & Cook, 1980; Harvey et al., 2018). In implementation studies, the fidelity with which implementation supports are provided is an important factor. If implementation dependent variables are to contribute to a science of implementation, then there must be commonly used measures of implementation outcomes. Pinnock et al. (2017) have proposed criteria for publishing research on implementation. The criteria include specific descriptions of intervention methods and outcomes, and specific descriptions of implementation methods and outcomes. These criteria hold promise for advancing the field.

The lack of repeated measures is a problem. Implementation is widely acknowledged as a complex process that may take several years to accomplish desired outcomes. Yet, few research studies examine implementation variables over time and use data to bring the process to light. Some examples are provided by Panzano and colleagues (Massatti, Sweeney, Panzano, & Roth, 2008; Panzano & Billings, 1994; Panzano & Roth, 2006; Panzano et al., 2004) who assessed 91 agencies every nine months for several years and identified patterns of adoption, use, deadoption, and readoption of selected evidence-based programs. McIntosh, Mercer, Nese, and Ghemraoui (2016) had repeated measures of intervention fidelity across five years for over 5,000 schools and found distinct patterns for achieving, sustaining, and losing fidelity. Independent studies have documented the progress of scaling for over a decade in Scandinavian countries using repeated measures of fidelity (Sigmarsdóttir et al., 2018; Tommeraas & Ogden, 2016). The studies document the consistent fidelity resulting from the use of consistent implementation supports (Ogden et al., 2012). In other studies, repeated measures of implementation capacity development have been conducted every six months for 18 months (Chaple & Sacks, 2016; McGovern, Matzkin, & Giard, 2007) and for five years (Fixsen, Ward, Ryan Jackson, et al., 2018; Ryan Jackson et al., 2018). These studies show the impact of implementation capacity development on attaining and sustaining criterion performance in organizations and systems. Repeated measures are used in global health environment to track the use of nationally sanctioned innovations and to document the improvements in innovations as they are used in practice (Adondiwo et al., 2013; Thomassen, Mann, Mbwana, & Brattebo, 2015).

These longitudinal studies are not typical, but they should be. After, before and after, one-time, or short-term assessments are interesting but may add little to the science of implementation. To do something once or even a few times is interesting. To be able to do something repeatedly with useful outcomes and documented improvements over decades will produce socially significant benefits for whole populations (Fixsen, Blase, & Fixsen, 2017). Data on the processes of implementation over time are badly needed.

Measures of implementation and capacity have been developed to assess factors in the Active Implementation Frameworks. A generalizable measure of Implementation Drivers (Fixsen, Blase, Naoom, & Wallace, 2006; Fixsen, Ward, Blase, et al., 2018) assesses Competency Drivers (staff selection, training, coaching, and fidelity), Organization Drivers (facilitative administration, decision support data system, system intervention), and Leadership Drivers (technical, adaptive). This measure has been used to assess implementation in large scale applications (Ogden et al., 2012; Tommeraas & Ogden, 2016). Another measure tracks the extent to which an organization is using an innovation with fidelity (Fixsen & Blase, 2009) so that overall fidelity and overall outcomes can be assessed. A set of measures to assess implementation capacity in organizations and systems have been developed based on the Active Implementation Frameworks (Fixsen, Ward, Duda, Horner, & Blase, 2015; Russell et al., 2016; St. Martin, Ward, Harms, Russell, & Fixsen, 2015; Ward et al., 2015). These measures assess leadership and executive management investment, system alignment, and commitment to Implementation Team development.

Linked Implementation Teams have been developed and evaluated in complex state education systems using these measures to guide action planning (Fixsen, Ward, Ryan Jackson, et al., 2018; Ryan Jackson et al., 2018; Ward, Ryan Jackson, Cusumano, & Fixsen, 2018). Measures for assessing Implementation Stages have been established (Saldana, Chamberlain, Wang, & Brown, 2012) and used in practice as summarized by Fixsen, Blase, and Van Dyke (2018). With the Active Implementation Frameworks as a mid-range theory of implementation (Improved Clinical Effectiveness through Behavioural Research Group, 2006), theory-based predictions are made, the Active Implementation Frameworks are used to establish the independent variables, and measures informed by the Active Implementation Frameworks are used to test the predictions in practice. In this way, the theory continues to improve and become more evidence-based.

Conclusion

At this point, implementation research is working toward meeting the criteria for becoming a science of implementation. It is not yet a “system of knowledge covering general truths” but mid-range theories are developing as a source of predictions (if-then) that can be tested in practice (then-what).

The next steps are to develop common concepts that are described using common language and assessed with common measures. If we have common concepts, common language, and common measures then implementation science can be crowdsourced globally where many researchers are working to generate concept-related knowledge that readily can be communicated and used by all.

Yet, the field is on the cusp on rapid and dramatic development. Potentially powerful implementation variables have been identified and operationalized. The efforts of so many over the past half century are about to bear fruit.

References

Adondiwo, A., Kubio, C., Kanyoke, E., Bagni, F., Dasoberi, I. N., Amenga-Etego, I. A., . . . Barker, P. M. (2013). Using quality improvement methods to test and scale up a new national policy on early post-natal care in Ghana. *Health Policy and Planning, 29*(5), 622-632. doi:10.1093/heapol/czt048

Allanson, E. R., Tunçalp, Ö., Vogel, J. P., Khan, D. N., Oladapo, O. T., Long, Q., & Gülmezoglu, A. M. (2017). Implementation of effective practices in health facilities: a systematic review of cluster randomised trials. *BMJ Global Health, 2*(2). doi:10.1136/bmjgh-2016-000266

Allen, J. D., Towne, S. D., Maxwell, A. E., DiMartino, L., Leyva, B., Bowen, D. J., . . . Weiner, B. J. (2017). Measures of organizational characteristics associated with adoption and/or implementation of innovations: A systematic review. *BMC Health Services Research, 17*(1), 591. doi:10.1186/s12913-017-2459-x

Bartley, L. H., Bright, C. L., & DePanfilis, D. (2017). Contributors to Fidelity of Child Welfare-Related Interventions: A Review. *Journal of Public Child Welfare, 11*(4-5), 433-463. doi:10.1080/15548732.2017.1340222

Blanchard, C., Livet, M., Ward, C., Sorge, L., Sorensen, T. D., & McClurg, M. R. (2017). The Active Implementation Frameworks: A roadmap for advancing implementation of Comprehensive Medication Management in Primary care. *Research in Social and Administrative Pharmacy, 13*(5), 922–929. doi:10.1016/j.sapharm.2017.05.006

Blase, K. A., & Fixsen, D. L. (2013). *Core intervention components: Identifying and operationalizing what makes programs work*. Retrieved from Washington, DC: http://aspe.hhs.gov/hsp/13/KeyIssuesforChildrenYouth/CoreIntervention/rb_CoreIntervention.cfm

Blase, K. A., Fixsen, D. L., Naoom, S. F., & Wallace, F. (2005). *Operationalizing implementation: Strategies and methods*. Retrieved from National Implementation Research Network, University of South Florida:

Bryk, A. S. (2016). Accelerating How We Learn to Improve. *Educational Researcher, 44*(9), 467-477. doi:10.3102/0013189X15621543

Chaple, M., & Sacks, S. (2016). The Impact of Technical Assistance and Implementation Support on Program Capacity to Deliver Integrated Services. *The Journal of Behavioral Health Services & Research, 43*(1), 3-17. doi:10.1007/s11414-014-9419-6

Dobson, L., & Cook, T. (1980). Avoiding Type III error in program evaluation: results from a field experiment. *Evaluation and Program Planning, 3*, 269 - 276.

Durlak, J. A., & DuPre, E. P. (2008). 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*, 327-350. doi:10.1007/s10464-008-9165-0

Fixsen, D. L., & Blase, K. (2009). *Implementation Tracker*. Retrieved from Chapel Hill: Active Implementation Research Network: https://www.activeimplementation.org/resources/implementation-tracker/

Fixsen, D. L., Blase, K., Naoom, S., & Wallace, F. (2006). *Measures of Core Implementation Components*. Retrieved from National Implementation Research Network, Florida Mental Health Institute, University of South Florida:

Fixsen, D. L., Blase, K. A., & Fixsen, A. A. M. (2017). Scaling effective innovations. *Criminology & Public Policy, 16*(2), 487-499. doi:10.1111/1745-9133.12288

Fixsen, D. L., Blase, K. A., Metz, A., & Van Dyke, M. K. (2015). Implementation science. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed., Vol. 11, pp. 695-702). Oxford: Elsevier, Ltd.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*(5), 531-540. doi:10.1177/1049731509335549

Fixsen, D. L., Blase, K. A., & Van Dyke, M. K. (2018). *Assessing implementation stages*. Retrieved from Chapel Hill, NC: Active Implementation Research Network: www.activeimplementation.org/resources

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*: National Implementation Research Network, University of South Florida, www.activeimplementation.org.

Fixsen, D. L., Ward, C., Blase, K., Naoom, S., Metz, A., & Louison, L. (2018). *Assessing Drivers Best Practices*. Retrieved from Chapel Hill, NC: Active Implementation Research Network https://www.activeimplementation.org/resources/assessing-drivers-best-practices/

Fixsen, D. L., Ward, C., Ryan Jackson, K., Blase, K., Green, J., Sims, B., . . . Preston, A. (2018). *Implementation and Scaling Evaluation Report: 2013-2017*. Retrieved from National Implementation Research Network, State Implementation and Scaling up of Evidence Based Practices Center, University of North Carolina at Chapel Hill:

Fixsen, D. L., Ward, C. S., Duda, M. A., Horner, R., & Blase, K. A. (2015). *State Capacity Assessment (SCA) for Scaling Up Evidence-based Practices (v. 25.2)*. Retrieved from National Implementation Research Network, State Implementation and Scaling up of Evidence Based Practices Center, University of North Carolina at Chapel Hill:

Glennan Jr., T. K., Bodilly, S. J., Galegher, J. R., & Kerr, K. A. (2004). *Expanding the reach of education reforms*. Santa Monica, CA: RAND Corporation.

Greenhalgh, T., Robert, G., MacFarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly, 82*(4), 581-629.

Harvey, G., McCormack, B., Kitson, A., Lynch, E., & Titchen, A. (2018). Designing and implementing two facilitation interventions within the 'Facilitating Implementation of

Research Evidence (FIRE)' study: a qualitative analysis from an external facilitators' perspective. *Implementation Science, 13*(1), 141. doi:10.1186/s13012-018-0812-z

Hurley, K. D., Lambert, M. C., Gross, T. J., Thompson, R. W., & Farmer, E. M. Z. (2017). The Role of Therapeutic Alliance and Fidelity in Predicting Youth Outcomes During Therapeutic Residential Care. *Journal of Emotional and Behavioral Disorders, 25*(1), 37-45. doi:10.1177/1063426616686756

Improved Clinical Effectiveness through Behavioural Research Group. (2006). Designing theoretically-informed implementation interventions. *Implementation Science, 1*(4). Retrieved from http://auraserv.abdn.ac.uk:9080/aura/bitstream/2164/114/1/Iceberg2952.pdf.

Kessler, R. C., & Glasgow, R. E. (2011). A proposal to speed translation of healthcare research into practice: Dramatic change is needed. *American Journal of Preventive Medicine, 40*(6), 637-644.

Lewis, C., Fischer, S., Weiner, B., Stanick, C., Kim, M., & Martinez, R. (2015). Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. *Implementation Science, 10*(1), 155. doi:10.1186/s13012-015-0342-x

Massatti, R., Sweeney, H., Panzano, P., & Roth, D. (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*(1-2), 50-65. doi:10.1007/s10488-007-0141-z

McGovern, M. P., Matzkin, A. L., & Giard, J. (2007). Assessing the Dual Diagnosis Capability of Addiction Treatment Services: The Dual Diagnosis Capability in Addiction Treatment (DDCAT) Index. *Journal of dual diagnosis, 3*(2), 111-123. doi:10.1300/J374v03n02_13

McIntosh, K., Mercer, S. H., Nese, R. N. T., & Ghemraoui, A. (2016). Identifying and Predicting Distinct Patterns of Implementation in a School-Wide Behavior Support Framework. *Prevention Science, 17*(8), 992-1001. doi:10.1007/s11121-016-0700-1

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*(3-4), 462-480. doi:10.1007/s10464-012-9522-x

Naleppa, M. J., & Cagle, J. G. (2010). Treatment fidelity in social work intervention research: A review of published studies. *Research on Social Work Practice, 20*(6), 674-681. doi:10.1177/1049731509352088

Ogden, T., Bjørnebekk, G., Kjøbli, J., Patras, J., Christiansen, T., Taraldsen, K., & Tollefsen, N. (2012). Measurement of implementation components ten years after a nationwide introduction of empirically supported programs – a pilot study. *Implementation Science, 7*, 49. Retrieved from http://www.implementationscience.com/content/pdf/1748-5908-7-49.pdf.

Øvretveit, J., Mittman, B., Rubenstein, L., & Ganz, D. A. (2017). Using Implementation Tools to Design and Conduct Quality Improvement Projects for Faster and More Effective Improvement. *International Journal of Health Care Quality Assurance, 30*(8), 755-768. doi:10.1108/IJHCQA-01-2017-0019

Panzano, P. C., & Billings, R. S. (1994). The influence of issue frame and organizational slack on risky decision making: A field study. *Academy of Management Proceedings, 1994*(1), 377-381. doi:10.5465/ambpp.1994.10345891

Panzano, P. C., & Roth, D. (2006). The decision to adopt evidence-based and other innovative mental health practices: Risky business? *Psychiatric Services, 57*(8), 1153-1161.

Panzano, P. C., Seffrin, B., Chaney-Jones, S., Roth, D., Crane-Ross, D., Massatti, R., & Carstens, C. (2004). The innovation diffusion and adoption research project (IDARP). In D. Roth & W. Lutz (Eds.), *New Research in Mental Health* (Vol. 16, pp. 78-89). Columbus, OH: Ohio Department of Mental Health Office of Program Evaluation and Research.

Pasachoff, J. M. (2009). Solar eclipses as an astrophysical laboratory. *Nature, 459*, 789. Retrieved from http://dx.doi.org/10.1038/nature07987. doi:10.1038/nature07987

Perl, H. I. (2011). Addicted to discovery: Does the quest for new knowledge hinder practice improvement? *Addictive Behaviors, 36*(6), 590-596. doi:10.1016/j.addbeh.2011.01.027

Pinnock, H., Barwick, M., Carpenter, C. R., Eldridge, S., Grandes, G., Griffiths, C. J., . . . for the StaRI Group. (2017). Standards for Reporting Implementation Studies: (StaRI) Statement. *British Medical Journal, 356*(16795). doi:http://dx.doi.org/10.1136/bmj.i6795

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. New York: Harper Torchbooks.

Popper, K. (2002). *The logic of scientific discovery*. New York: Routledge.

Proctor, E. K., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., . . . Hensley, M. (2011). Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Administration and Policy in Mental Health, 38*, 65-76. doi:10.1007/s10488-010-0319-7

Russell, C., Ward, C., Harms, A., St. Martin, K., Cusumano, D., Fixsen, D., . . . LeVesseur, C. (2016). *District Capacity Assessment Technical Manual*. Retrieved from National Implementation Research Network, University of North Carolina at Chapel Hill:

Ryan Jackson, K., Fixsen, D., Ward, C., Waldroup, A., Sullivan, V., Poquette, H., & Dodd, K. (2018). *Accomplishing effective and durable change to support improved student outcomes*. Retrieved from National Implementation Research Network, University of North Carolina at Chapel Hill:

Rycroft-Malone, J., Seers, K., Eldh, A. C., Cox, K., Crichton, N., Harvey, G., . . . Wallin, L. (2018). A realist process evaluation within the Facilitating Implementation of Research Evidence (FIRE) cluster randomised controlled international trial: an exemplar. *Implementation Science, 13*(1), 138. doi:10.1186/s13012-018-0811-0

Saetren, H. (2005). Facts and myths about research on public policy implementation: Out-of-fashion, allegedly dead, but still very much alive and relevant. *The Policy Studies Journal, 33*(4), 559-582.

Saldana, L., Chamberlain, P., Wang, W., & Brown, H. C. (2012). Predicting program start-up using the stages of implementation measure. *Administration and Policy in Mental Health, 39*(6), 419-425. doi:10.1007/s10488-011-0363-y

Sanetti, L. M. H., & Kratochwill, T. (Eds.). (2014). *Treatment Integrity: A foundation for evidence-based practice in applied psychology* Washington, DC: American Psychological Association Press (Division 16).

Schoenwald, S. K., & Garland, A. F. (2013). A review of treatment adherence measurement methods. *Psychological Assessment, 25*(1), 146-156. doi:10.1037/a0029715

Seers, K., Rycroft-Malone, J., Cox, K., Crichton, N., Edwards, R. T., Eldh, A. C., . . . Wallin, L. (2018). Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the

Promoting Action on Research Implementation in Health Services (PARIHS) framework. *Implementation Science, 13*(1), 137. doi:10.1186/s13012-018-0831-9
Sigmarsdóttir, M., Forgatch, M., Vikar Guðmundsdóttir, E., Thorlacius, Ö., Thorn Svendsen, G., Tjaden, J., & Gewirtz, A. (2018). Implementing an Evidence-Based Intervention for Children in Europe: Evaluating the Full-Transfer Approach. *Journal of Clinical Child & Adolescent Psychology*, 1-14. doi:10.1080/15374416.2018.1466305
St. Martin, K., Ward, C., Harms, A., Russell, C., & Fixsen, D. L. (2015). *Regional capacity assessment (RCA) for scaling up implementation capacity*. Retrieved from National Implementation Research Network, University of North Carolina at Chapel Hill:
Tabak, R. G., Khoong, E. C., Chambers, D. A., & Brownson, R. C. (2012). Bridging research and practice: Models for dissemination and implementation research. *American Journal of Preventive Medicine, 43*(3), 337-350.
Thomassen, O., Mann, C., Mbwana, J. S., & Brattebo, G. (2015). Emergency medicine in Zanzibar: the effect of system changes in the emergency department. *International Journal of Emergency Medicine, 8*(1), 22. doi:10.1186/s12245-015-0072-5
Tiruneh, G. T., Karim, A. M., Avan, B. I., Zemichael, N. F., Wereta, T. G., Wickremasinghe, D., . . . Betemariam, W. A. (2018). The effect of implementation strength of basic emergency obstetric and newborn care (BEmONC) on facility deliveries and the met need for BEmONC at the primary health care level in Ethiopia. *BMC Pregnancy and Childbirth, 18*(1), 123. doi:10.1186/s12884-018-1751-z
Tommeraas, T., & Ogden, T. (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, 44*, 203-216. doi:10.1007/s10488-015-0712-3
Ward, C., Ryan Jackson, K., Cusumano, D., & Fixsen, D. (2018). *State Capacity Development Plan*. Retrieved from National Implementation Research Network, University of North Carolina at Chapel Hill:
Ward, C., St. Martin, K., Horner, R., Duda, M., Ingram-West, K., Tedesco, M., . . . Chaparro, E. (2015). *District Capacity Assessment*. Retrieved from National Implementation Research Network, University of North Carolina at Chapel Hill:
Weiner, B., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., . . . Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implementation Science, 12*(1), 108. doi:10.1186/s13012-017-0635-3
Winter, S. G., & Szulanski, G. (2001). Replication as Strategy. *Organization Science, 12*(6), 730-743.
Wootton, D. (2015). *The invention of science: A new history of the scientific revolution*. New York: Harper Collins.