An Integrated Embedded Research Agenda for Agile Implementation Frameworks

WORKING DRAFT, PLEASE DO NOT CITE

Kirsten Kainz, Ph.D.
School of Social Work
Frank Porter Graham Child Development Institute
University of North Carolina, Chapel Hill

Allison Metz, Ph.D.
National Implementation Research Network
Frank Porter Graham Child Development Institute
University of North Carolina, Chapel Hill

The primary obstacle to faster progress in U.S. education reform is hard to put your finger on, because it's an absence, not a presence. It is not an interest group or a manifest social problem. It is the infrastructure we never built for identifying what works. It is the organizational framework we've not yet constructed for building consensus among education leaders across the country to identify what's working.
(Thomas Kane, 2015)

The absence of infrastructure illuminated by Thomas Kane is powerfully relevant for those who do and promote implementation science. Implementation scientists support the uptake of research evidence to promote effective practice in local settings. To do so requires an infrastructure not only for identifying research-based practices but also for naming the promising structures and processes that build effective practice in local settings. In many cases, this infrastructure is incomplete or totally absent.

Implementation scientists can co-construct the knowledge infrastructure with stakeholders and practitioners in context by embedding research – strategic inquiry, feedback, and sense-making systems - into the implementation work. Such co-construction will begin with a clear articulation of effective practice and its temporal and contextual nature. Effective practice is a suite of human actions and interactions that:

- occur in the present time and context,

- yield compelling evidence that they produce expected and desired effects in the present time and context; yet
- in many cases are based on research evidence produced in the past and in other contexts.

By extension, effective practices are agile, indicating that practices initially based on research evidence might have to adapt to produce consistently desirable outcomes over time and across context. Such practices will be informed by research evidence, but not chained to it, as the ultimate goal is to achieve desired outcomes. Implementation frameworks must also be agile, then, originating with a defensible set of implementation components and then eliminating existing or incorporating new components as new evidence is generated in the implementation setting. The foundation for decisions about eliminating and incorporating for agile implementation frameworks will be an embedded research agenda deeply integrated with implementation practice that leads to the co-construction and use of evidence of when and how to adapt to spread effective practice for desirable outcomes.

An embedded and integrated implementation research agenda responds to growing concern from multiple disciplines that traditional and rigid implementation research agendas have not been consistently useful for translating research evidence into effective practice and desirable outcomes in complex settings (Chambers, Glasgow, & Stange, 2013; Ghate, 2015; Glasgow & Chambers, 2012). From this concern a provocative two-sided question emerges: are traditional notions of causality impediments to an integrated embedded research agenda for agile implementation frameworks, or can causal thinking be used to create shared vocabulary and concepts to support embedded research for agile implementation frameworks? Toward an answer to that question, we do

three things in this paper. First, we lay out the current challenge for implementation frameworks in education and health and human services: implementation frameworks must cultivate relevant and reliable knowledge for how to implement and adapt in complex systems in ways that optimize local outcomes and accumulate transferrable knowledge for the spread of effective practices. Then, we explore how ideas from the fields of causal thinking can support the co-construction of knowledge for agile implementation frameworks that promote effective practice. Finally, we follow the review of causal thinking with a list of recommendations for an embedded and integrated agile implementation research agenda that is sufficiently sensitive, relevant, and rigorous for promoting effective practice across time and context. We argue throughout that causal thinking can serve as a trans-disciplinary foundation from which implementation scientists, practitioners, and stakeholders can build a rigorous science and co-construct relevant and reliable evidence of when and how to adapt in context.

The Current Challenge in Implementation Science

The argument for evidence-based practice in education and health and human services is fairly straightforward. In a system of finite human and financial resources decision makers cannot afford to waste time and money on ineffective practices, and must instead allocate resources to evidence-based practices for desired education outcomes. Implementation science builds from the assumption that evidence-based practices will consistently yield desired outcomes over time and place when implemented effectively. Of course, with those assumptions in place, the simple argument for evidence-based practice now applies to implementation science. We cannot afford to

waste limited resources on implementation practices that do not yield desired outcomes over time and across contexts.

Implementation frameworks, like operating theories and models, serve to guide the implementation of evidence-based practices; predict and explain barriers and facilitators of implementation; and structure the evaluation of implementation practices (Nilsen, 2015). Researchers and practitioners interested in facilitating and understanding change processes note that "implementation strategies have unparalleled importance in implementation science, as they constitute the 'how to' component" (Proctor, Powell, & McMillen, 2013, p. 1) of changing practices and optimizing the use of evidence to improve outcomes for children, youth and families. Implementation strategies are often packaged as frameworks. When searching for a framework by which to examine implementation, different models compete for our attention. There are many frameworks available currently, but there is no consistent evidence on how the specific components of frameworks contribute to the use of evidence in practice.

So, why are implementation frameworks so frequently used? The promise of implementation frameworks has been demonstrated through their use in investigating challenges and guiding solutions. In that regard, implementation frameworks are often used as heuristic devices to aid in discovery, reflection, and iterative learning through small cyclical tests of change. Implementation scientists and practitioners rely on implementation frameworks to guide planning, problem-solving, and continuous improvement efforts in the implementation setting. However, the specific implementation strategies used to effectively support and assess change processes are often poorly understood by implementation agents because there is no embedded research

agenda to support collectively identifying and contrasting effective implementation components and there are few opportunities for sense-making around what is and is not working in the implementation space.

Much of the evidence in support of implementation framework components has been gathered retrospectively through systematic reviews (Meyers, Durlak & Wandersman, 2012; Moullin, Sabater-Hernandez, Fernandez-Llimos & Benrimoj, 2015). But, if our aim is to construct multifaceted, multilevel change in complex service systems to improve outcomes, then we will need to recognize the building blocks of change (Powell, et al., 2015) as flexible, integrated, and trans-disciplinary. In doing so, the ways we study the effectiveness of implementation strategies will require new methods and diverse ways of knowing.

There is a need for a body of knowledge gathered prospectively to identify the value of implementation frameworks in a variety of relevant contexts. Such a body of knowledge could enumerate and define the dynamic processes by which and multi-disciplinary roles through which interventions are implemented, adapted, and optimized. That body of knowledge would outline the salient components of agile implementation frameworks and guide their use in practice.

Alongside the mounting need for a strong knowledge base to guide agile implementation in education and human services is an emerging message from the field of healthcare that implementation science is currently "held back in part due to the types of questions we ask, the tools we employ, and the constraint of researchers to fit into a traditional paradigm when the world of dissemination and implementation is at its nature complex, dynamic, and uncontrollable. The problems we study are often at odds with the

view of good or robust science." (Glasgow & Chambers, 2012, p. 48). The corresponding message for the study of implementation frameworks in education and human services is that over-emphasizing fidelity to a known treatment may blind implementers and implementation researchers to the numerous points of adaptation in context that do and must occur to facilitate optimal implementation. Multiple designs and methods are needed to investigate the complexity of implementation adaptation and to provide timely feedback for practitioners who work to achieve optimal outcomes. And, while multiple designs and methods will serve to increase the relevance of agile implementation framework research, they must also serve to ensure the rigor of such research.

In the next section we review ideas from the fields of causal thinking. Our intent is to produce points of synthesis that may serve those who seek to build embedded research for agile implementation that is both relevant and rigorous, supporting effective practice. Specifically, we borrow from scholars from different disciplines whose distinct yet inter-related contributions to causal thinking over time allow us to expose unique causal questions facing implementation scientists By combining thought from multiple disciplines we further explicate our early assertion that truly effective practice is agile and implementation frameworks must also be agile to support agile practice. Such agility in the implementation setting can be accelerated by an embedded and integrated research agenda that yields action-oriented, reliable, and timely information for how to implement and when to adapt so as to optimize outcomes in context (See Figure 1).

Causal Thinking

Implementation frameworks are implicitly, but not explicitly, causal. The frameworks are implicitly causal in that they imply that implementation supports will improve intervention implementation. The subsequent strand of this implied causal chain is that effective intervention implementation will yield desired intervention effects. The frameworks are not explicitly causal, however, in that by no means do framework authors claim they have identified the full set of supports that are both necessary and sufficient to increase intervention implementation efficacy and subsequent intervention effects. Nor do framework authors indicate that implementation supports will work every time in all contexts. In fact, unexpected variability in implementation effects occurs regularly in practice. These ideas are borne out most clearly in the emerging body of implementation research that incorporates complexity theory (Chandler, Rycroft-Malone, Hawkes, Noyes, 2016; Leykum, Pugh, Lawrence, Parchman, Noël, Cornell, & McDaniel, 2007; Rycroft-Malone et al., 2016)

As the field of implementation science has grown significantly over the last decade with the proliferation of frameworks, models, and theories, there is mounting interest in the limited capacity of implementation practitioners and researchers to use these frameworks, especially given the complexity of the work. The need for major research centers (e.g., National Institutes of Health) to train researchers in dissemination and implementation research and the importance of identifying competencies for research training programs is an intensifying conversation in the field (Padek, et al., 2015). The identification of specific competencies for implementation specialists or technical assistance providers also has garnered increased attention. The shortage of individuals trained in the science and practice of knowledge translation and implementation science

has been cited as a reason for our failure to optimize the use of evidence to improve population outcomes (Straus, et al., 2011.) In response to this competency gap, more is being written about the specific competencies needed to facilitate implementation and broker knowledge (Bornbaum, Kornas, Peirson, & Rosella, 2015; Berta, et al., 2015).

While identification of competencies and training toward competencies is viewed as an important step forward for promoting evidence use for effective practice, it may be that the most pressing need for the field of implementation science is a reframing of the relationship between research and implementation. Research can be strategically embedded within implementation practice to provide ongoing feedback and identify "clearer parameters within which adaptive responses can be made." (Ghate, 2015, p. 13). Rigor in the embedded research can be served by building on existing causal thinking and expanding and enriching what can be learned from traditional causal thinking with evidence generated in the complex implementation setting.

In the following paragraphs we provide a brief presentation of important ideas about causal thinking form the fields of philosophy, psychology, statistics, and economics. We use the thinking from each of these fields to throw light on important opportunities for research on agile implementation frameworks.

John Stuart Mill. Mill's (1874) analysis of cause and causal relations within a system of logic reveals a set of conditions for inferring cause that have been succinctly codified into three claims:

1. the cause must precede the effect;
2. the cause must covary or co-occur with the effect; and

3. there must be no plausible alternative explanation for the co-occurrence of the cause and effect. (Shadish, Cook, & Campbell, 2002)

The first two claims map neatly onto the implied causality of implementation frameworks. Implementation supports will precede implementation effectiveness, and degrees of implementation supports will be proportional to degrees of implementation effectiveness. In a more concrete example, implementation specialists build the capacity of implementation teams to develop, install, and improve the infrastructure needed to support a new innovation. As implementation specialists support implementation teams to construct and optimize the infrastructure over time, it is expected that the degree to which infrastructure components are enacted by accountable teams is the degree to which they will observe effective implementation. Effective implementation has been equated in the literature with high-fidelity implementation of manualized programs or practices. (see Metz, A., et al., 2015 for a case example).

The theory of change linking implementation team capacity with the strength of the infrastructure, high-fidelity implementation of programs, and population outcomes is based on common factors related to successful implementation identified in most implementation frameworks (Metz, Naoom, Halle, & Bartley, 2015). However, without carefully designed evaluation of implementation efforts, we are unable to rule out alternative plausible explanations for the degree of implementation observed. Mill's work points us toward the need for rigorous evaluation designs that allow us to: establish temporal precedence by examining implementation status before and after supports are put in place; establish the statistical covariation or observed co-occurrence of

implementation support and status; and rule out other plausible explanations for that co-occurrence.

The complexity of implementation practice adds levels of challenge for identifying the conditions for causality. In many cases, temporal precedence will be fairly straightforward to assess in simple research designs but will grow more complicated in iterative designs, which are common in implementation studies that examine improvement over time. That is, as implementation practices iterate toward improvement, establishing evidence that supports preceded practices will become more challenging, and the potential for variables not specified by the framework to creep into practice and drive implementation outcomes will increase. Research in this context will require a priori and ongoing attention to the third point raised by Mill, ruling out plausible alternative explanations. Here, researchers must be constantly, critically vigilant, ensuring that alternative explanations for observed implementation practices are posited, examined, and either ruled out or incorporated into the agile framework as an active component.

Implementation evaluations grappling with and incorporating concepts of complexity (Rycroft-Malone, et al., 2016) serve the field well by theorizing models of the active components of learning within an implementation endeavor and the hypothesized effects of implementation practices. Such careful articulation of components and hypothesized effects allows for equally careful study of consequences, intended and unintended, in complex adaptive systems. After all, we cannot name an unintended consequence without first naming the intended consequences. We cannot unearth an unknown without first claiming what is known. Detailed theorizing based on input from implementation

scientists, practitioners, and stakeholders moves the field closer to understanding the building blocks for effective practice over time and across contexts.

Donald Campbell. Campbell (1957) and colleagues (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002) have contributed much to the fields of causal thinking, and we focus on two contributions here – the definition of internal validity and enumeration of threats to internal validity. Internal validity is the degree to which an observed relation between two variables (e.g., implementation support and implementation practice) can be inferred as causal. We acknowledge upfront that a research agenda for identifying simple causal relations in complex adaptive systems is potentially misleading. And yet, implementation scientists seek to produce an effect through their practice. Therefore, a return to conditions for identifying cause and effect can be informative.

Threats to internal validity are those forces that lead us to attribute causation falsely. More specific, threats to internal validity are specific instances where forces other than the implementation practice under study are the true and unidentified cause for the observed effects. Known threats to internal validity in the implementation context include: failure to put supports in place before implementation begins; agent selection into or out of implementation supports; events that co-occur with implementation support, but are not part of the implementation framework; and increases in implementation efficacy over time due to experience, not implementation supports. In the implementation endeavor none, some, or all of these threats may be encountered.

The full set of potential threats to internal validity in implementation practice cannot be known by researchers ahead of time. Instead, collaboration with practitioners and

stakeholders can lead to a more thorough identification of potential threats and their subsequent elimination or incorporation defended by evidence from the embedded research. Note that agile implementation frameworks will be agile specifically because they allow for the elimination of components deemed ineffective and incorporation of components deemed effective via embedded research. The goal is for effective practice to produce desired outcomes.

Donald Rubin. Donald Rubin improved researchers' capacities to identify causal relations by articulating alternative design and analysis techniques that build on the logic of random assignment. What's desirable about random assignment, more generally referred to as experimentation, is the degree of certainty the method yields for inferences about the effect caused by an intervention, be it a policy, program, or treatment. Given effective randomization - and assuming random sampling from a known population, thorough implementation of a well operationalized intervention, and valid and reliable measurement - we can be certain that the calculated effect of the intervention is an unbiased estimate of the effect in the population. And, as the sample size approaches the population, the sample estimated treated effect is also consistent. Of course, random assignment is not always feasible or desirable, and random sampling, thorough implementation, and valid and reliable measurement are not guaranteed. Moreover, the control needed meet the assumptions of random assignment studies is often inconsistent with the needs to adapt and remain flexible in the implementation setting.

Rubin's Causal Model (1974, 2007, 2008) translates the mathematical logic underlying the estimation of treatment effects in random assignment studies to those estimated in non-random assignment conditions known as observational studies. In well-

conducted random assignment studies, agents in the control group are essentially exchangeable with those in the treated group. That is, we can assume based on randomization that control and treatment group participants are balanced on background characteristics and more importantly balanced on their potential response to any intervention. This initial balance before the introduction of an intervention allows us to say with a high degree of certainty that, but for the intervention, any outcome measured after intervention would be identical on average in the two groups. Rubin's work demonstrates that in observational studies – those that do not use random assignment – analytic steps can be taken to create exchangeable groups that are balanced on background characteristics, thus increasing our certainty that intervention is driving differences in outcomes across groups**.** There are, at a minimum, two implications of Rubin's work for embedded research on implementation frameworks.

First, we should put careful thought into the comparisons we construct when evaluating implementation supports. In some cases the comparisons will be across groups randomly or non-randomly assigned to receive implementation support or its absence. In those cases the exchangeability of the groups should be established before evaluating the effects of implementation supports. If exchangeable groups cannot be formed then we should carefully consider and explore the threats introduced by comparing groups that are different from one another. In other cases the comparisons will be across time periods within a single group that receives adapting forms of implementation supports. Again, exchangeability across time points will need to be established to make valid inferences about the effects of implementation supports. From this constraint emerges an opportunity to study changes in groups over time, focusing on

the relations between implementation and change while either ruling out or incorporating alternative causes.

Second, we should design our data collection protocols in the embedded research to include important background and context characteristics so that we can evaluate the exchangeability of our groups. Access to a rich and meaningful set of individual, group, and system characteristics will allow researchers to conduct analysis that: 1) rules out plausible alternative explanations for any changes in implementation efficacy due to pre-existing or evolving characteristics of the groups; or 2) identifies new supports for effective practice in the implementation setting.

James Heckman. The econometric approach to causality, as outlined by Heckman (2005, 2008), shifts focus away from simple identification a treatment effect in relation to a known cause (i.e., the intervention) to a broader understanding of the constellation of factors arrayed in the causal explanation of observed group differences. The econometric perspective of causal explanation highlights individuals as agents who actively choose conditions rather than passively accept assignment to conditions. Consequently, to understand cause and effect requires careful investigation of individuals' motivations, options, constraints, and rationale for choosing conditions. Moreover, the economic approach to causality posits that only with useful models of active agents' motivation, options, constraints, and rationales can researchers predict the future success of policies intended to influence individuals' outcomes.

Implementation specialists, and anyone whose work involves persuading adults to change their behavior, are likely to resonate with Heckman's perspective on individuals as active rather than passive agents. Individuals arrive in the change space accompanied

by experiences, beliefs, practical limits, and aspirations that affect their willingness and capacity to implement new practices. These factors, whether known or unknown, act as moderators and mediators of implementation supports. The implications of the econometric approach to causality on implementation evaluation indicate a pressing need to account for not only contextual barriers and enablers of implementation, but also the individual characteristics and thought processes that moderate and mediate implementation, and how those factors do or do not change over time in the implementation space. Knowledge of moderators and mediators and their potentially changing effects over time will bring us closer to modeling a full array of implementation supports that drive effective practice in various settings. This knowledge will be better when co-constructed by diverse agents in the implementation setting, including researchers who bring with them knowledge from the research evidence base.

Phyllis Illari and Frederica Russo. In their effort to connect philosophical theory with scientific practice Illari and Russo (2014) outline five scientific problems of causality. They are: inference, prediction, explanation, control, and reasoning. In our review of causal thinking thus far, the acts of inference making are circumscribed by Mill, Campbell, and Rubin. Heckman has addressed the inter-relatedness of explanation and prediction. What remains to be explored are issues of control and reasoning.

Illari & Russo's definition of control contains two parts modified slightly here to unmask their relevance for embedded research for agile implementation frameworks: 1) given our original implementation framework, how do we control factors within the implementation setting so that we can achieve greater certainty that our proposed implementation supports are driving effective practice while remaining open to emerging

evidence of alternative explanations; and 2) how do we control (read: design) implementation policies and frameworks so that they produce effective practice over time and across settings? Note that issues of control as defined by Illari and Russo are issues about how to act based on our best scientific knowledge.

Endeavors to control environments, policies, and practices will undoubtedly lead to questions about the limits of our capacity for valid causal inference making, for consistent causal prediction, and for thorough causal explanation. That is, efforts to exert control will ask us to unearth our reasoning about and justifications for causal thinking in implementation science. Do we really know enough with sufficient certainty to organize people's behavior productively over time and across setting? Recent research indicates that the co-production of evidence in implementation settings is challenged by the multiplicity of roles, viewpoints, and epistemologies invoked in the work (Rycroft-Malone et al, 2016). In order to produce a timely, relevant, and rigorous research base for agile implementation we need a research agenda that flexibly addresses the complexity of the workspace with new worldviews, goals, and methods (Glasgow & Chambers, 2015).

Toward a Research Agenda

Agile implementation frameworks are needed to shift focus away from the rigid uptake of isolated programs and instead promote scientific practices that integrate the use of evidence into day-to-day services and decision-making for optimal outcomes in context (Ghate, 2015). This shift toward agile implementation frameworks will require an ongoing and dynamic research agenda grounded in real practice and sufficiently rigorous to build a formal science for promoting effective practice over time and across settings.

Incorporating major ideas from causal thinking into a research agenda for agile implementation frameworks could move us toward a more formal science.

At the outset we acknowledge that examining cause and effect is not relegated to random assignment studies only. Causal thinking entails careful and accurate understandings of the critical relations among agents, actions, and aftershocks, which we define as those expected and unexpected events that result and reverberate from implementation activities. Causal thinkers have pointed to the misconceptions that might obscure such understanding in implementation science. These misconceptions include misunderstandings about:

- timing of events;
- alternate explanations for observed effects;
- the full system of factors that produces observed effects; and
- elements of systems that can and can't be controlled or manipulated to produce effects.

Causal thinking is a process of positing expectations, and then identifying the justifications for believing that our expectations have been met, justifications constructed by naming and ruling out threats to those beliefs. More precisely, "the path to causal inference is a constant process of elimination", (Kaufman & Poole, 2000, p. 107), and given the complicated and complex nature of effective practice that path is best traveled with practitioners and stakeholders. The promise of causal thinking for agile implementation frameworks is a promise for developing collective processes within an embedded research agenda to explore and rule out potential misunderstandings. Such embedded research will integrate inquiry, feedback, sense-making, and practice around emerging evidence of effective practice.

Multiple designs and multiple methods are needed across multiple research endeavors to carefully eliminate confounds, explicate causal relations, and generate reliable knowledge about how to support effective practice over time and across setting. And, a point of clarification is advantageous. Given the uniqueness of each implementation context we are no longer designing and evaluating frameworks to implement and bring evidence-based products to scale. Rather, agile implementation frameworks will serve to identify, adapt, and spread processes within and across complex systems that facilitate the translation of effective practice into common practice within local settings (Massoud et. al, 2006). Therefore, embedded research systems that support adaptation will be critical implementation supports within the implementation setting. Ultimately, new worldviews, goals, methods, and practical flexibility are needed to guide novel research endeavors (Glasgow & Chambers, 2012).

Worldview. We propose that the emerging specialization of Integration and Implementation Science (I2S: 2005; 2013) outlines a worldview that is especially helpful for elucidating the dynamics of complex systems such that agile implementation framework research can successfully eliminate confounding factors and identify relations among agents and actions in the implementation space so that effective process spread can happen. As articulated by Bammer, I2S combines theoretical and methodological pillars from multiple disciplines to provide more effective conceptual and research solutions to pressing social problems. The theoretical and methodological pillars are guided by systems theories (VonBertalanffy, 1967) and complexity science (Lissack, 1999), but do not require abandoning traditional research methods such as experimentation, observation, and quantitative and qualitative analysis. I2S does imply,

however, that these methods be enhanced to: 1) incorporate perspectives and participation from a broad set of stakeholders; and 2) reflect the complex systems that form implementation settings.

What distinguishes complex systems from complicated settings is that they embody self-organizing and emergent dynamics among agents and actions. Complexity does not render these systems incomprehensible or intractable; rather, "complex adaptive systems can form patterns and follow predictable paths of development. The identification of attractors or states, to which a system finally settles, is one clue as to why certain patterns (order) and not others are created" (Higginbotham, Albrecht & Connor, 2001 as cited in Bammer, 2005). Goals for agile implementation frameworks should include as part of the implementation process the identification of self-organizing and emergent phenomena in the implementation space, such that these phenomena may be explored, explained, and adapted to produce optimal practice and outcomes.

Goals. Agile implementation frameworks will embed research activities to map and engage with the implementation system so that reliable knowledge of how to adapt and what to adapt will lead to desired outcomes (See Table 1). This requires defining four distinct, yet interrelated, sets of goals for embedded and integrated research: convening relevant stakeholders and practitioners; articulating models and expectations for action and outcomes; evaluating frequently with attention to unexpected aftershocks in complex systems; and committing to adapting models and practices to achieve more desirable outcomes. These goals reflect a shift in purpose *away from* using frameworks to facilitate faithful uptake of interventions *toward* using agile frameworks to facilitate co-construction of knowledge for building effective and sustainable practice in context.

Convening relevant stakeholders and practitioners. Research demonstrates that successful uptake of evidence requires genuine and meaningful interaction among researchers, service providers, policy makers, consumers, and other key stakeholders (Flaspohler, Meehan, Maras, & Keller, 2012; Palinkas et al., 2011; Wandersman et al., 2008). Implementation efforts must address the various needs of these stakeholders (Palinkas et al., 2014). However, we know that in many instances, collaborations among stakeholders, including researchers and community members, are strained by a lack of mutual understanding of each other's goals and expectations (Stokols, 2006; Rycroft-Malone, 2016). The use of evidence is often a result of "iterative, messy, and dynamic" interactions (Nutley, Walter, & Davies, 2007, p. 39) among public agencies, policy makers, researchers, intervention developers, practitioners, communities, and families. Successful interactions take the shape of iterative "mutual consultations" that mediate the use of research evidence in complex service systems and political contexts. Still, the process of identifying effective practices and eliminating misunderstanding requires knowledge and input from practitioners, stakeholders, and implementation scientists, and so effective research for agile implementation will prioritize engagement.

Articulating models and expectations for action and outcomes. As evidence-based programs are "inserted" into existing systems, the status quo may be inflexible to the systems changes needed to accommodate for these programs. Metz & Albers (2014) observe that the infrastructure required for the use of research evidence in mainstream services is often invisible to policy-makers and funders, researchers, and service providers that seek to integrate, adapt, and optimize evidence to improve outcomes. The

'invisible infrastructure' reflects and maintains the status quo (Koerth-Baker, 2012), which "fights back" and jeopardizes effective implementation of innovations.

Efforts can be made to clarify models for action and expectations for results in the collaborative implementation setting, allowing time for diverse groups of practitioners, stakeholders, and implementation scientists to make sense of where there is overlap and where there is disagreement in models and expectations. At a minimum, implementation methods should operationalize a well defined strategy so that it is clear at the outset what practitioners will do as they carry out the innovation. Once an innovation is described in sufficient detail, researchers can begin to evaluate the congruence of the implementation model with observed practice, and ultimately with desired practice. From this research points for adaptation can be proposed, interpreted, and enacted. Enabling contexts also must be built that leverage and establish hospitable funding, regulatory and policy environments, engage key stakeholders, and promote ongoing learning (Metz, 2015).

Evaluating frequently with attention to unexpected aftershocks in complex systems. There is increasing emphasis on the role of continuous quality improvement in optimizing evidence-based programs and practices in a range of contexts and enabling ongoing learning among developers, interventionists, researchers and consumers to improve the sustainability of evidence in practice settings. Operational learning should be a core value of the implementation setting (Chambers, Glasgow, & Stange, 2013). A review of implementation frameworks by Damschroder and colleagues concluded "dedicating time for reflecting or debriefing before, during, and after implementation is one way to promote shared learning and improvements along the way" (Damschroder et al., 2009, pg. 11).

While the focus on evaluation for iterative improvement will support agile implementation, researchers and implementation teams must also attend to the unintended aftershocks of their work. Careful study of aftershocks, expected and unexpected, can lead to improved models of the implementation agenda and locations for adaptation to promote effective practice.

Committing to adapting models and practices to achieve more desirable outcomes. Aarons, Hurlburt, and Horwitz (2011) note that "implementation of an innovation will be successful to the degree that the innovation matches the mission, values, and service provider tasks and duties of the organization (p. 14)." Each of these dimensions – mission, values, tasks, duties – reflect implied models that silently direct behavior in the implementation space. The task for agile implementation research is to embed research activities and sense-making routines in the implementation space that allow emerging evidence of effective practice to direct collective action. Emerging evidence of what constitutes effective practice in the local setting can lead to adaptations in systems, organizations, and programs to meet the needs of local communities and target populations (Aarons, Green, Palinkas, Self-Brown, Whitaker, Lutzker, & Chapin, 2012). However, the work can be slow and challenging (Rycroft-Malone, 2016). Research methods that lead to timely identification of promising recommendations for action and effective sense-making routines that engage the knowledge and input of stakeholders, practitioners, and implementation scientists will be needed to accelerate and deepen the work.

Methods. An agile implementation research agenda will require strategic methods for convening, articulating, evaluating, and adapting in context. We highlight four

promising methodological techniques from different disciplines that when combined are especially well suited for the full suite of agile implementation research goals.

First, Midgley's (2006) depiction of systemic intervention includes a set of methodological principles and tools well aligned with the four goals of agile implementation, and especially informative for the convening and articulating goals. In specific, Midgely highlights the need for evolving and dynamic methods that gather multiple and diverse agents to learn together and from each other on an ongoing basis. Gathering and engaging diverse agents in the implementation agenda is critical for understanding the boundaries and limitations of our current knowledge. When not vetted and informed by diverse agents, boundaries of the original implementation framework may be introducing threats to understanding. By identifying and expanding boundaries diverse implementation learning groups can co-construct better models of implementation for effective practice, thus promoting agile implementation.

Second, work on participatory modeling in the systems sciences – strategic methods for co-creating conceptual models of complex systems with diverse agents – expands the set of tools proposed by Midgley and assists in the identification of boundaries and correction of misunderstandings about the inter-relations of agents, actions, and aftershocks in context (Johnsson, Andersson, Alken-Olsson, & Armheimer, 2007). Participatory modeling methods are part of a set of strategic efforts to use systems theories and complexity science to generate reliable practice-based evidence (Green, 2006). When diverse agents explicitly articulate and assess models of complex phenomena opportunities arise to identify and correct misunderstandings about the expected relations between implementation, effective practice, and desirable outcomes.

Positing, testing, and refining models leads to what Sterman (2006) has called double loop learning, the type of learning that is required for groups to achieve collective knowledge of effective practice in complex systems.

Third, Mayne's (2004, 2008) work on performance stories and contribution analysis merges traditional evaluation goals (e.g., focus on efficacy, cause and effect relations) with participatory techniques that foster shared understanding of performance expectations in public administration and government. Stakeholders work together to co-create models of change, proposing the agents and actions that will lead to desired aftershocks and ultimately to optimal outcomes. Expectations for performance are fixed initially, but it is anticipated that new knowledge will emerge in context and guide adaptations to performance. Implementation communities' collaborative assessments of performance create collective knowledge of effective practices that can be spread to other contexts.

Fourth, the work of Langley, Moen, Nolan, Norman, and Provost (2009), is a model for participatory improvement efforts, where subject matter knowledge of the intervention practice combines with profound knowledge to promote improvement. Profound knowledge refers to diverse agents':

- appreciation for the system, its parts, and their relation to the whole;
- understanding of optimal and non-optimal variation within the system;
- capacity for meaning-making, learning, and change within the system; and
- psychological and social factors that affect change capacity.

The identification of profound knowledge opens the door for unearthing and exploring known and unknown capacities in the context that otherwise would remain part of the invisible infrastructure sustaining the status quo. Simple methodological techniques can be used to build collective understanding of root causes, potential improvements, and

improvement progress. These techniques have proven useful in large scale health care improvement agendas (see Institute for Healthcare Improvement: www.ihi.org).

In summary, relatively new methodological practices emerging in multiple disciplines are exceptionally well suited for an embedded research agenda to support agile implementation. The emergence of these new methodological practices indicates a growing awareness across disciplines that complex systems require novel research responses for the purpose of building reliable and collective knowledge in context to serve effective practice and create desirable outcomes. Novel research responses may incorporate design features from experimentation and observational research as well as quantitative and qualitative and quantitative analysis. It is not a full departure from traditional design and analysis that distinguishes the novel research practices reviewed above. Rather, the factors that characterize these novel research practices are their simultaneous focus on methods for convening, articulating, evaluating, and adapting in complex systems.

Conclusion

We began this paper by articulating the challenge for implementation science: learn how to implement and adapt in complex systems in ways that optimize local outcomes and accumulate transferrable knowledge for the spread of effective practices. We proposed that integrated and embedded research to support agile implementation frameworks was needed. Embedded research that engages diverse agents to co-construct knowledge using tools from causal thinking, despite the limitations for complete explanation and prediction in complex settings, will bring us closer to optimizing practice and outcomes in context than current frameworks do.

We explored causal thinking as a resource for addressing the challenge of learning how to adapt in context, especially to the degree that causal thinking can be used to frame and develop relevant and rigorous embedded research for agile implementation. In the review of causal thinking from multiple disciplines we observed that although experimentation is a powerful technique for estimating unbiased effects of known causes, it is only one tool in the diversely populated tool-kit aligned with causal thinking. Causal thinking as described by multiple scholars requires more and deeper knowledge of how, why, and under what conditions effects happen. As such, causal thinking can be an especially useful for guiding a novel research agenda for agile implementation, where timely and reliable knowledge will support adaptation toward effective practice in unique contexts defined by place and time.

While causal thinking can support rigor in implementation research, grounding the research in the practice space will support relevance. Implementation frameworks are beginning to emphasize the need for continuous quality improvement through the systematic assessment and feedback of information and data related to planning, implementation, and outcomes (Chinman, Imm, & Wandersman, 2004). We propose that to achieve effective practice and desired outcomes, improvement processes must be viewed as more than a bounded component of implementation frameworks. Rather, integrated and embedded research activities with opportunities for sense-making are critical components of agile implementation frameworks. And, traditional improvement methods can be expanded to include a simultaneous focus on convening diverse agents; articulating models for action and expected outcomes; evaluating frequently with

attention to expected and unexpected aftershocks; and adapting models and action based on emerging knowledge.

Figure 1. Agility is defined as a system of multiple stakeholder practices that generate action-oriented, reliable, and timely information for when and how to adapt so as to optimize outcomes.

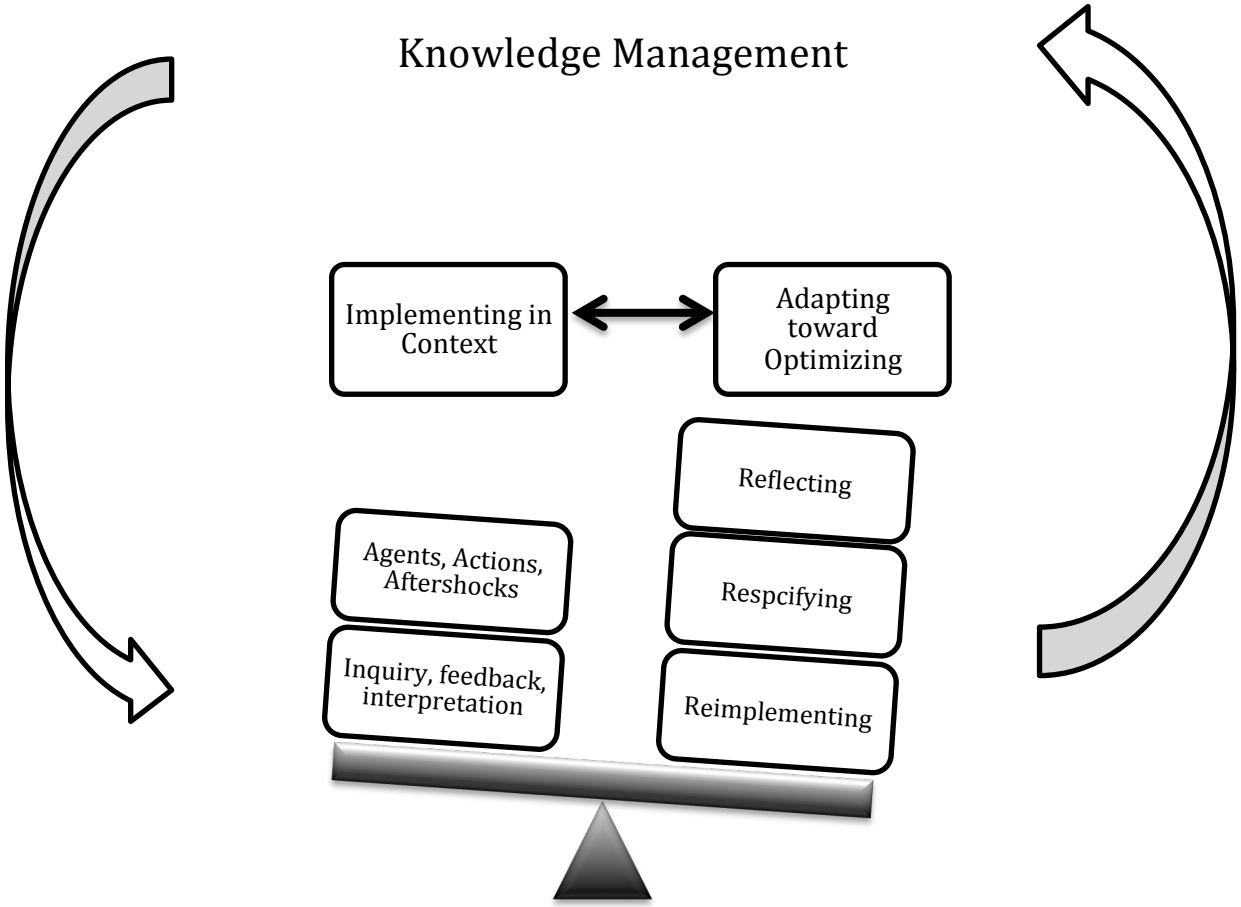

Table 1. Goals and Strands for Embedded Research for Agile Implementation Frameworks

Goal	Optimization Strand	How	What
Convening	Gathering stakeholders and practitioners	Invite members from many if not all levels of the system	Implementation, evaluation, and knowledge management plans that represent diverse perspectives within the system
Articulating	Modeling the system	Co-construct understanding and models of the system	A visual and descriptive model of the agents and actions in the system, as understood initially
	Naming the expectations	Co-design the expectations for action and results	A shared agenda for implementation, performance, and improvement
	Identifying boundaries	Engage stakeholders in identifying boundaries of what we can know using which methods with which samples	Insights into the limits of the initial system model, stakeholders represented, and implementation and evaluation plans
	Expanding boundaries	With collective input, expand/revise the research design, data collection protocol, methods, and sample based on boundary identification	Improved system model, stakeholder representation, and implementation evaluation plans
Evaluating	Evaluating performance	Create feedback systems so that stakeholders have access to timely performance information and opportunities to reflect on potential adaptations	Performance reports with timely and action-oriented feedback coupled with inquiry and decision making routines to guide interpretation
	Naming emergence	Use observation methods and	Insights into aspects of the system not accounted

		stakeholder feedback to identify unexpected desirable and undesirable aftershocks that emerge from implementation in location	for in the current system model and implementation and evaluation plans
	Naming self-organizing processes	Use observation methods and stakeholder feedback to identify local dispositions, resources, and actions that propel unexpected desirable and undesirable aftershocks	Insights into aspects of the system not accounted for in the current system model and implementation and evaluation plans
Adapting	Improving system models	Co-create improved system model based on synthesis of evidence from the performance and recognition feedback	A revised visual and descriptive model of the system that incorporates learning to date
	Improving performance expectations	Co-create recommendations for adaptation	An adaptation plan based on learning to date
	Continuous improvement	Foster inquiry within and across subsystems to support improved implementation and evaluation based on evidence of incremental outcomes	Better practice and outcomes over time and place

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (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.

Aarons, G. A., Green, A. E., Palinkas, L. A., Self-Brown, S., Whitaker, D. J., Lutzker, J. R., ... & Chaffin, M. J. (2012). Dynamic adaptation process to implement an evidence-based child maltreatment intervention. *Implementation Science*, *7*(32), 1-9.

Bammer, G. 2005. Integration and implementation sciences: Building a new specialization. Ecology and Society 10(2): 6. [online] URL: http://www.ecologyandsociety.org/vol10/iss2/art6

Bammer, G. 2013 *Disciplining Interdisciplinarity: Integration and Implementation Sciences for Researching Complex Real-World Problems*. ANU Press. http://press.anu.edu.au?p=222171

Berta, W., Cranley, L., Dearing, J. W., Dogherty, E. J., Squires, J. E., & Estabrooks, C. A. (2015). Why (we think) facilitation works: insights from organizational learning theory. *Implementation Science*, *10*(1), 141.

Bornbaum, C. C., Kornas, K., Peirson, L., & Rosella, L. C. (2015). Exploring the function and effectiveness of knowledge brokers as facilitators of knowledge translation in health-related settings: a systematic review and thematic analysis. *Implementation Science*, *10*(1), 162.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological bulletin*, *54*(4), 297.

Chambers, D. A., Glasgow, R. E., & Stange, K. C. (2013). The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. *Implement Sci*, *8*(1), 117.

Chandler, J., Rycroft-Malone, J., Hawkes, C., & Noyes, J. (2016). Application of simplified Complexity Theory concepts for healthcare social systems to explain the implementation of evidence into practice. *Journal of Advanced Nursing*, *72*(2), 461-480.

Chinman, M., Imm, P., & Wandersman, A. (2004). *Getting To Outcomes™ 2004*. Rand Corporation.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings*. Boston, MA: Houghton Mifflin.

Cook, J. M., Dinnen, S., Thompson, R., Ruzek, J., Coyne, J. C., & Schnurr, P. P. (2014). A Quantitative Test of an Implementation Framework in 38 VA Residential PTSD Programs. *Administration and Policy in Mental Health and Mental Health Services Research*, 1-12.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*, *4*(1), 50.

Ghate, D. (2015). From programs to systems: Deploying implementation science and practice for sustained real world effectiveness in services for children and

Families. *Journal of Clinical Child & Adolescent Psychology*, DOI: 10.1080/15374416.2015.1077449

Glasgow, R. E., & Chambers, D. (2012). Developing robust, sustainable, implementation systems using rigorous, rapid and relevant science. *Clinical and Translational Science*, *5*(1), 48-55.

Green, L. W. (2006). Public health asks of systems science: to advance our evidence-based practice, can you help us get more practice-based evidence?. *American Journal of Public Health*, *96*(3), 406.

Heckman, J. J. (2005). The scientific model of causality. *Sociological methodology*, *35*(1), 1-97.

Heckman, J. J. (2008). Econometric causality. *International Statistical Review*, *76*(1), 1-27.

Higginbotham, N., Albrecht, G., & Connor, L. (2001). Health social science: a transdisciplinary and complexity perspective| NOVA. The University of Newcastle's Digital Repository.

Illari, P., & Russo, F. (2014). *Causality: Philosophical theory meets scientific practice*. Oxford University Press.

Jonsson, A., Andersson, L., Alkan-Olsson, J., & and Arheimer, B. (2007). How participatory can participatory modeling be? A discussion of the degree of influence and stakeholder and expert perspectives in six dimensions of participatory modeling, *Water Science and Technology, (56*), 1, 207-214. http://dx.doi.org/10.2166/wst.2007.453

Kane, T. J. (2015). Frustrated with the pace of progress in education: Invest in evidence. http://www.brookings.edu/research/papers/2015/03/05-education-evidence-kane

Kaufman, J. S., & Poole, C. (2000). Looking back on "causal thinking in the health sciences". *Annual Review of Public Health*, *21*(1), 101-119.

Langley, G. J., Moen, R., Nolan, K. M., Nolan, T. W., Norman, C. L., & Provost, L. P. (2009). *The improvement guide: a practical approach to enhancing organizational performance*. John Wiley & Sons.

Leykum, L. K., Pugh, J., Lawrence, V., Parchman, M., Noël, P. H., Cornell, J., & McDaniel, R. R. (2007). Organizational interventions employing principles of complexity science have improved outcomes for patients with Type II diabetes. *Implementation Science*, *2*(1), 28.

Lissack, M. R. (1999). Complexity: the science, its vocabulary, and its relation to organizations. *Emergence*, *1*(1), 110-126.

Mayne, J. (2004). Reporting on outcomes: Setting performance expectations and telling performance stories. *The Canadian Journal of Program Evaluation, 19,* 31-60.

Mayne, J. (2008, May). Contribution analysis: An approach to exploring cause and effect. ILAC Brief, 16. Institutional Learning and Change Initiative: http://www.cgiar-ilac.org/files/ILAC_Brief16_Contribution_Analysis_0.pdf

Massoud M. R., Nielsen G. A. , Nolan, K., Schall, M. W. , Sevin, C. (2006). *A*

framework for spread: From local improvements to system-wide change. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement. Retreived from: www.IHI.org

Metz, A., Naoom, S., Halle, T. & Bartley, L. (2015). An Integrated Stage-Based Framework for Implementation of Early Childhood Programs and Systems. Research Brief for the Office of Planning, Research and Evaluation, Washington, DC.

Metz, A., Bartley, L., Ball, H., Wilson, D., Naoom, S., & Redmond, P. (2014). Active Implementation Frameworks for Successful Service Delivery. Research on Social Work Practice.

Metz, A. & Albers, B. (2014). What Does It Take? How Federal Initiatives Can Support the Implementation of Evidence-Based Programs to Improve Outcomes for Adolescents. Journal of Adolescent Health, Vol. 54. No. 3.

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.

Midgley, G. (2006). Systemic intervention for public health. *American Journal of Public Health*, *96*(3), 466.

Mill, J. S., (1874). *A System of Logic.* Eighth edition. New York: Harper and Brothers.

Moullin, J. C., Sabater-Hernández, D., Fernandez-Llimos, F., & Benrimoj, S. I. (2015). A

systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework. *Health Research Policy and Systems*, *13*(1), 16.

Nilsen, P. (2015). Making sense of implementation theories, models, and XXXXXX

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

Padek, M., Colditz, G., Dobbins, M., Koscielniak, N., Proctor, E. K., Sales, A. E., & Brownson, R. C. (2015). Developing educational competencies for dissemination and implementation research training programs: an exploratory analysis using card sorts. *Implementation Science*, *10*(1), 114.frameworks. *Implementation Science, 10,* XXXX

Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlain, P., Hurlburt, M., & Landsverk, J. (2011). Mixed method designs in implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(1), 44-53.

Palinkas, L. A., Fuentes, D., Finno, M., Garcia, A. R., Holloway, I. W., & Chamberlain, P. (2014). Inter-organizational collaboration in the implementation of evidence-based practices among public agencies serving abused and neglected youth. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*(1), 74-85.

Powell, B. J., Beidas, R. S., Lewis, C. C., Aarons, G. A., McMillen, J. C., Proctor, E. K., & Mandell, D. S. (2015). Methods to improve the selection and tailoring of implementation strategies. *The journal of behavioral health services & research*, 1-18.

Proctor, E. K., Powell, B. J., & McMillen, J. C. (2013). Implementation strategies: recommendations for specifying and reporting. *Implement Sci*, *8*, 139.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Statistics in medicine*, *26*(1), 20-36.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, 808-840.

Rycroft-Malone, J., Burton, C. R., Wilkinson, J., Harvey, G., McCormack, B., Baker, R., Dopson, S., Graham, I. D., Staniszewska, S., Thompson, C., Ariss, S., Melville-Richards, L., & and Williams, L., (2016). Collective action for implementation: a realist evaluation of organisational collaboration in healthcare. *Implementation Science*, published online on February 9, 2016 at: http://implementationscience.biomedcentral.com

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Shadish, W. R. (2010). Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological methods*, *15*(1), 3-17.

Sterman, J. D. (2006). Learning from evidence in a complex world. *American Journal of*

Public Health, *96*(3), 505-514.

Supplee, L & Metz, A. (2015). Opportunties and Challenges in Evidence-Based Social Policy. Social Policy Report**,** Volume 28, Number 4.

Stokols, D. (2006). Toward a science of transdisciplinary action research. *American journal of community psychology*, *38*(1-2), 63-77.

Von Bertalanffy, L. (1967). General theory of systems: Application to psychology. *Social Science Information*, *6*(6), 125-136.