Clarity out of chaos: Use of theory in implementation research

Laura J. Damschroder

VA Center for Clinical Management Research, Ann Arbor, MI, USA

ARTICLE INFO

Keywords:
Implementation science
Theory
Frameworks
Models

ABSTRACT

Implementation science has been recognized as a potential catalyst for health system reform, in part, because of its contribution of well-grounded conceptual theories, often encapsulated in frameworks. Well-designed frameworks provide a semantic structure, a common language by which to guide systematic approaches to studying implementation and testing interventions. An overview of the types and roles of theory in advancing implementation science is offered in this article. Resources for selecting appropriate frameworks are described along with illustrative examples. The case is made that well-developed theory is what enables knowledge to emerge out of seeming chaos and for translation of that knowledge to be widely and reliably implemented into routine practice so that health and well-being of patients is maximized by delivery of interventions that are rooted in that knowledge.

[There is] nothing so practical as good theory

(Levin, 1951a)

All models are wrong...George Box1976

(Box, 1976)

1. Introduction

Implementation science has been recognized as a potential catalyst for health system reform, in part, because of its contribution of well-grounded theories including conceptual frameworks (Fisher et al., 2016). This scientific discipline was born out of recognition of the proverbial “valley of death” that characterizes the chasm between evidence-based scientific discoveries and the patients who need them (Butler, 2008). Surveys of organizational leaders reveal that most attempts to implement innovations targeted at improvements within organizations fail (Meaney and Pung, 2008; Rafferty et al., 2013); effective interventions take much too long to get integrated into routine clinical practice to benefit patients (Balas and Boren, 2000). Implementation science focuses on developing and testing methods to broadly spread successful sustained implementations across diverse settings. The interconnected black boxes in Fig. 1 represent a high-level schematic of key foci for implementation science and highlight that these foci are embedded, interact, and are influenced by multiple levels of contextual domains including individuals and inner and outer settings. Identifying an appropriate evidence-based innovation (EBI), implementing, and then sustaining it in clinical practice is a complex undertaking because of the dynamic interplay between the targeted EBI, the need to assess and understand diverse contexts, adapt EBIs to clinical context and processes, and select and execute implementation strategies tailored to context to get EBIs into routine use, all dimensions of which change over time. Implementation scientists seek to understand the role and impact of each of these dimensions. Developing and testing theory is an important means by which to achieve these challenging goals.

2. The role of theories, models, and frameworks

Theory development is essential for encapsulating and then advancing our knowledge about which EBIs work best in which contexts and to guide development of reliable approaches to ensure successful implementation of those EBIs into routine practice. But what is theory and how is it encapsulated and tested? Theory may be less formal prospective statements of “if I do a, then b will happen” or retrospective “reason-giving” explanations for observed outcomes (Davidoff et al., 2015). Building scientific knowledge demands formally stated, encapsulated theories that are conjectural (and testable) propositions that are explicit but have varying degrees of specificity based on the current state of knowledge.

Within implementation science, theories are encapsulated as generalized theories, models, or frameworks. Rogers’ Diffusion of Innovation may be regarded as generalized theory (Rogers, 2003) with broad applicability of oft-used principles that include e.g., recruiting opinion leaders and working through established professional networks to get new practices implemented and sustained in clinical practice (Davidoff et al., 2015).
Models, on the other hand, are narrower in scope; they provide simplified but useful representations of complex realities; they may be prospective representations that describe program components that are expected to yield desired outcomes or may describe theoretical constructs that are thought to mediate or moderate the level of observed outcomes. Importantly, models encapsulate “theories of change” or “theories of explanation.” For example, Klein and colleagues developed a detailed testable model of factors to explain variations in success of implementing new software in manufacturing settings. Their model had six components with specified relationships that encapsulated their theory of change: 1) management support for implementation and 2) financial resource availability; 3) high-quality implementation policies and practices, which leads to 4) positive implementation climate, which in turn leads to 5) effective implementation (i.e., high use of the new software), which leads to 6) positive impact of using the new software on key outcomes. The authors include detailed rationale for each construct and expected relationships between those constructs along with measures and an analytic strategy for assessing validity of their model. Since the time of their publication, other researchers adapted and refined the model for use in health care (Damschroder et al., 2011; Helfrich et al., 2010, 2007; Kirk et al., 2016).

Theory may also be encapsulated by frameworks, which describe more loosely structured constellations of theoretical constructs (often without specifying relationships between them) or prescriptive approaches for accomplishing implementation goals. Well-designed frameworks articulate theories about semantic structures of constructs (Larsen et al., 2013; Wacker, 1998) and provide a common language by which to guide systematic approaches for studying implementation context and testing implementation interventions. Such frameworks provide clarity in terms and definitions which allows clear articulation of complex dynamic phenomena that are so common in the realm of implementation science.

In summary, generalized theories are broad statements that provide a lens by which to ground approaches for studying implementation. These theories must have sufficient scientific consensus with supporting evidence though they may have pockets of dissent within the scientific community (Jaccard and Jacoby, 2009). Models are narrower in scope than frameworks or generalized theories but should be grounded within them. Frameworks provide a foundation (Larsen et al., 2013; Wacker, 1998) for developing measures (Arnulf et al., 2014) and explanatory models or models of change (Kazdin, 2007; Lewis et al., 2018; Sofaer et al., 2003). Frameworks and models are bridges for connecting findings across diverse studies because they provide common language, definitions, foundations for measurement and assessment, and building generalized theories.

3. Classification and function of theories

The science of implementation is relatively young, without the benefit of the long decades of research necessary to establish widely accepted, more highly specified models of change nor broadly established generalized theories. An exception to the latter is May’s formal development “Toward a General Theory of Implementation” (May, 2013) that builds on Normalization Process Theory (May and Finch, 2009). To date, implementation scientists rely heavily on frameworks because they provide flexibility in application. Nilsen classified published implementation frameworks into three categories (Nilsen, 2015); the white boxes in Fig. 1 show how each of these three categories of theories can contribute knowledge about each of the four key foci of implementation science. Though Nilsen focuses on frameworks and the next sections do so as well, his classification scheme also applies to models and generalized theories. Table 1 lists example frameworks by category.

3.1. Process theories

Nilsen’s first category includes process frameworks that specify steps or phases to execute for accomplishing implementation goals. The aim...
of these frameworks is to provide practical guidance for planning and executing implementation endeavors. For example, the Knowledge to Action framework describes high-level phases of knowledge transfer starting with monitoring knowledge use to identifying problems to adapting knowledge for local context to selecting, tailoring, and implementing interventions. The Getting to Outcomes framework describes 10 steps to implement EBIs starting with how to select appropriate EBIs to solve a documented problem to planning for sustaining implementation efforts. An honorable mention, though not characterized as a framework, is the Expert Recommendations for Implementing Change (ERIC), which provides a list of implementation strategies. ERIC provides language (terms and definitions) by which to identify and describe the range of implementation strategies included in the list (Powell et al., 2015).

### 3.2. Determinant theories

Processes do not always proceed as planned or guided by theory. Nilsen's second category describes determinant frameworks as those that specify constructs (often, but not always, conceptualized as potential independent variables) that may influence processes or predict implementation outcomes (dependent variables). Fig. 1 shows that determinant frameworks can help to identify moderators that may affect or confound the relationship between the targeted EBI, implementation processes used, and/or outcomes or may be used to describe underlying mechanisms of change. Determinant frameworks are foundational for advancing understanding of implementation endeavors by explaining variation in observed outcomes in retrospect or predicting outcomes a priori. Determinant frameworks provide structure for exploratory evaluations that can lead to more specific hypothesized models that include constructs derived from a determinant framework and along with hypothesized relationships between those constructs that can be tested in follow-on studies. Thus, determinant frameworks are foundational for building more specific theories of change.

Examples of determinant frameworks (see Table 1) include Wilsey-Stirman and colleagues’ framework (Stirman et al., 2019) that articulates ways EBIs (and implementation strategies) may be adapted to local clinical settings. The importance of adaptation in predicting and explaining implementation outcomes is illustrated by an example. The landmark randomized clinical trial of the Diabetes Prevention Program (DPP) demonstrated a 58% reduction in diabetes compared to “usual care” (Knowler et al., 2002). The trial was stopped early because outcomes were so clearly positive. Since then, many implementation studies have been conducted to replicate results of DPP in broader populations and diverse clinical settings. In the original trial, the intervention was delivered by an interventionist to individual patients. However, most clinical and community entities deliver DPP groups of patients, instead of individually. This was a necessary adaptation of the program in response to the realities of constrained resources. The adaptation framework describes dimensions by which to document and plan adaptations by specifying for example, what content is being modified, how it is fit to context, and level of delivery.

Many failures of implementation are rooted in local contextual factors. Another example determinant framework is the Consolidated Framework for Implementation Research (CFIR). The CFIR describes 39 theoretical constructs across five contextual domains (Damschroder et al., 2009). The CFIR provides language (terms and definitions) for constructs that allows researchers and practitioners to clearly and consistently articulate factors that potentially affect implementation outcomes. A case example using the CFIR is provided below.

### 3.3. Evaluation theories

The third category includes evaluation frameworks that specify multiple levels of outcomes and processes to assess. Proctor and colleagues published a framework comprising eight types of implementation outcomes; for example, penetration outcomes indicate the extent to which an EBI has been integrated into clinical practice (Proctor et al., 2011). The RE-AIM framework describes outcomes across five domains: Reach (engaging the targeted population); Efficacy/Effectiveness of the EBI; Adoption by targeted setting; Implementation consistency, costs, and adaptations; and Maintenance of EBI effects in individuals and settings over time.

---

**Table 1**

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Purpose</th>
<th>Online technical assistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>Knowledge to action (Graham et al., 2006)</td>
<td>Phases used to transfer scientific knowledge into practice</td>
<td>N/A</td>
</tr>
<tr>
<td>Hybrid: Determinant and</td>
<td>Getting to outcomes (Chinman et al., 2008)</td>
<td>To classify a broad range of adaptations that made be made to EBIs “...a broad range of modifications that may be made to evidence-based interventions...for more precise determination of the effects of such modifications on clinical or implementation outcomes of interest.”</td>
<td>N/A</td>
</tr>
<tr>
<td>process</td>
<td>Adaptation framework (Stirman et al., 2019, 2013)</td>
<td>Broad list of theoretical constructs across five domains “...to promote implementation theory development and verification about what works where and why across multiple contexts”</td>
<td>N/A</td>
</tr>
<tr>
<td>Hybrid: Determinant and</td>
<td>Consolidated Framework for Implementation Research (CFIR)</td>
<td>Four-phase framework of implementation processes derived from extant literature</td>
<td><a href="http://www.cfirguide.org">www.cfirguide.org</a></td>
</tr>
<tr>
<td>process</td>
<td>Exploration, Preparation, Implementation, and Sustainment (EPIS) (Aarons et al., 2011)</td>
<td>Developed based on “the proposition that for implementation of evidence to be successful, there needs to be clarity about the nature of evidence being used, the quality of context, and the type of facilitation needed to ensure a successful change process.”</td>
<td><a href="https://EPISframework.com">https://EPISframework.com</a></td>
</tr>
<tr>
<td>Evaluation</td>
<td>Implementation outcomes (Proctor et al., 2011)</td>
<td>“...to encourage program planners, evaluators, readers of journal articles, funders, and policy-makers to pay more attention to essential program elements including external validity that can improve the sustainable adoption and implementation of effective, generalizable, evidence-based interventions.” (from web)</td>
<td>N/A</td>
</tr>
<tr>
<td>Evaluation</td>
<td>RE-AIM (Reach, Effectiveness, Adoption, Implementation Maintenance) (Glasgow et al., 1999)</td>
<td>N/A</td>
<td><a href="http://www.re-aim.org">http://www.re-aim.org</a></td>
</tr>
</tbody>
</table>
Two “hybrid” frameworks are included in Table 1 to demonstrate that not all frameworks fit neatly into a single category. The EPIS framework, named for its four phases of implementation (exploration, preparation, implementation, and sustainment), also includes specific constructs that potentially influence ability to accomplish each phase (e.g., organizational characteristics). The PARIHS framework is widely used and includes two determinant domains (descriptions of constructs related to evidence and context domains) in addition to a third domain focused on facilitation as a process approach for implementation.

3.5. Selecting theory

The range of theories from which to select can be dizzying. Several researchers have created readily accessible online tools and guides to help researchers choose and apply appropriate theories, mostly focused on frameworks, for their research including VA’s Quality Enhancement Research Initiative’s (QUERI) Implementation Guide (https://www.queri.research.va.gov/implementation/), Birken and colleagues’ Theory, Model, and Framework Comparison and Selection Tool (T-CaST; https://imisci.tracs.unc.edu/tcast/) (Birken et al., 2018), an online guide based on Tabak and colleagues’ review of frameworks (Tabak et al., 2012) (http://www.dissemination-implementation.org/), and Colorado University’s online guide (http://crispebooks.org/). These resources are however, do not cover the full range of theories but rather provide a starting point.

4. Example application of a determinant framework

A case example serves to illustrate the key role of a determinant framework in implementation research. Damschroder and colleagues conducted an implementation evaluation of a Telephone Lifestyle Coaching (TLC) program in Veterans Health Administration (VHA) in 24 medical centers (Damschroder et al., 2017b). The TLC offered behavioral coaching for six lifestyle topics from which patients could choose to focus: Eat Wisely, Be Physically Active, Be Tobacco Free, Strive for a Healthy Weight, Manage Stress, and Limit Alcohol. Penetration (defined in Proctor and colleagues’ evaluation framework (Proctor et al., 2011)) was identified as an implementation outcome and measured as the rate of referrals (number of Veterans referred to TLC / number of Veterans enrolled in primary care) to TLC by clinicians within each of the medical centers: higher rates indicated more robust implementation of TLC referral processes. Over 9000 veterans were referred to TLC across the 24 pilot facilities. Rates within facilities varied widely: the highest rate of referral, 19 months after the program was launched, was seven times higher than the facility with the lowest rate of referral. The authors sought to explain differences in referral rates; specifically, what were barriers and facilitators to implementing TLC? They selected 12 sites to maximize variation based on a baseline measure of readiness for implementing change (high, medium, low) and complexity (high, low; an indicator of size, range of services available, and other factors) to qualitatively explore barriers and facilitators that might explain the wide variation in outcomes across the pilot facilities.

The authors used the CFIR determinant framework to guide data collection and analysis. Over 100 semi-structured interviews were conducted. The CFIR, with its inclusion of 39 constructs across five domains of context that may influence implementation outcomes, was used to guide the interviews. An online technical assistance website is available to help researchers create interview guides and approaches for qualitative data coding, analyses, and interpretation (see www.cfirguide.org). Qualitative interview data were transformed to quantitative ratings based on whether qualitatively coded constructs manifested as barriers or facilitators following guidance provided on the online technical website and published guidance (Damschroder and Lowery, 2013). Pearson correlations between ratings for each CFIR construct and outcomes, were used to assess the strength of association between the qualitatively derived ratings for each construct with referral rates across facilities. Correlations of 0.5 or higher were highlighted as possibly associated with implementation outcomes. Table 2 lists seven CFIR constructs that were associated with referral rates.

Increasingly, multiple frameworks are being used in studies to address multiple facets of implementation. For example, Damschroder and colleagues evaluated implementation of the DPP in three VHA medical centers (Damschroder et al., 2017a). In this study as well, they used the CFIR to guide identification of contextual barriers and facilitators potentially associated with outcome. They also used the RE-AIM evaluation framework to guide mixed methods assessments across five domains of outcomes (Damschroder et al., 2015). The use of both frameworks enabled the authors to more specifically link barriers and facilitators with multiple types of outcomes. For example, they found that the relatively low priority placed on referring patients to DPP, adversely impacted Reach of the program to patients who would benefit from participation while the failure of mid-level managers to help resolve hiring and space issues impacted teams’ ability to Implement the program; information that can be used to more concretely guide future implementations.

5. Value of theories

5.1. Building the knowledge-base

Deep, integrated use of well-grounded theory can help to easily connect findings across implementation studies. It is relatively easy to compare the evaluation of the TLC to other studies that use the same framework. For example, a study of DPP implemented in three medical centers in VA revealed that the Compatibility construct (part of the Inner Setting Domain in the CFIR) may be associated with implementation

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Seven CFIR constructs associated with implementation outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construct (correlation with outcomes)</strong></td>
<td><strong>Short construct definition</strong></td>
</tr>
<tr>
<td>Structural Characteristics (p = .73)</td>
<td>The social architecture, age, maturity, and size of an organization</td>
</tr>
<tr>
<td>Networks &amp; Communications (p = .58)</td>
<td>The nature and quality of webs of social networks and the nature and quality of formal and informal communications within an organization.</td>
</tr>
<tr>
<td>Compatibility (p = .55)</td>
<td>The degree of tangible fit between meaning and values attached to the intervention by involved individuals, how those align with individuals’ own norms, values, and perceived risks and needs, and how the intervention fits with existing workflows and systems.</td>
</tr>
<tr>
<td>Organizational Incentives &amp; Rewards (p = .98)</td>
<td>Extrinsic incentives such as goal-sharing awards, performance reviews, promotions, and raises in salary, and less tangible incentives such as increased stature or respect.</td>
</tr>
<tr>
<td>Engaging: Implementation Lead(s) (p = .64)</td>
<td>Commitment, involvement, and accountability of leaders and managers with the implementation.</td>
</tr>
<tr>
<td>Engaging: Stakeholders (p = .66)</td>
<td>Attracting and involving appropriate individuals in the implementation and use of the intervention through a combined strategy of social marketing, education, role modeling, training, and other similar activities.</td>
</tr>
<tr>
<td>Planning (p = −0.68)</td>
<td>The degree to which a scheme or method of behavior and tasks for implementing an intervention are developed in advance, and the quality of those schemes or methods.</td>
</tr>
</tbody>
</table>
outcomes (Damschroder et al., 2017a). This construct is also listed in Table 2 as being associated with outcomes related to implementing TLC. Thus, it would be important to include this construct in hypothesized models of change in future research studies. Note that other theories provide equally useful and important ways of articulating potential determinants of success. The key for researchers is to clearly describe how the framework informed planning and conduct of the research, its usefulness in bringing clarity to the work, and whether it leads to further theory development.

5.2. Advancing the science

The growing scientific field of implementation has leveraged developments in older scientific disciplines to construct the wide array of published frameworks (Nilsen and Bernhardsson, 2019). Greenhalgh and colleagues published among the earliest and broadest reviews of published literature from 13 scientific disciplines (Greenhalgh et al., 2004), which several more recent frameworks draw upon, including the CFIR (Damschroder et al., 2009). Systematic and narrative reviews continue to identify and classify many dozens of implementation frameworks (Nilsen and Bernhardsson, 2019; Tabak et al., 2012) and new ones continue to be published (e.g., Woodward et al., 2019). A smaller number of frameworks, however, are becoming more dominant and influential as the science of implementation matures (Norton et al., 2017; Skolarus et al., 2017).

Focused effort is needed to transparently apply and test existing frameworks (and models). Tabak and colleagues’ (Tabak et al., 2012) review of 61 frameworks highlighted the paucity of testing of frameworks and theory development; this finding was affirmed more recently by Holt and Chambers (2017). For example, separate systematic reviews of studies that relied on two widely used frameworks (CFIR and the PARIHS), each revealed that though each framework had been used across a diversity of settings, publications did not describe methods that reflect sufficiently deep integration or testing of their chosen framework (Helfrich et al., 2010; Kirk et al., 2016). Frameworks must be critiqued to help ensure they support further theory development including creating testable explanatory models or models of change. For example, researchers have critiqued the CFIR (Barwick et al., 2019; Ilott et al., 2012)), which will help to further strengthen that framework.

Well-grounded and validated measures, rooted in clearly conceptualized (defined, labeled, and described) constructs, are needed to develop and test more specific theoretical models. For example, Weiner and colleagues developed a theory of organizational readiness for change that started with a well-described framework of determinants and outcomes (Weiner, 2009). This framework was used to guide development of validated quantitative measures (Shea et al., 2014) that can be used in testable models of change. Aarons and colleagues have likewise identified and applied a series of measures (Aarons et al., 2014, 2016) rooted constructs that are labeled and described in their Exploration, Preparation, Implementation, and Sustainment (EPIS) framework (Aarons et al., 2011).

6. Conclusion

Theory is intimately woven into all we do including implementations of complex EBIs (Davidoff et al., 2015). Generalized theory, models, and frameworks all encapsulate some level of theory but at varying levels of specificity and applicability. None are perfect representations of the world within which complex implementation processes play out. The adage that “all models are wrong” is not the end of the story, however (Box, 1976, p792). Box goes on to acknowledge that “…some are useful; the practical question is how wrong do they have to be to not be useful? (Box and Draper, 1987, p74)” A city map is useful though it cannot capture the full truth of the city – its vibrancy, the look and feel of it, and yet, anyone would agree that a city map is essential for newcomers. The signature of a great scientist is the ability to create simple but illuminating theories, frameworks or models, which require that each be rooted in well-grounded theories of language, theories of constructs, theories of predictions, and/or theories of causality. Implementation researchers need clear, collective, consistent use of theory to build knowledge about what works, where, and why (Damschroder and Hagedorn, 2011). Theory, whether encapsulated in models, frameworks, or generalized theories, needs to be clearly described, applied, and critiqued to advance the science of implementation. Indeed, there is nothing so practical as a good theory because good theory is what enables knowledge to emerge out of seeming chaos and to be translated into effective use for the benefit of humankind (Lewin, 1951b).

Funding

This work was supported by the Veterans Affairs (VA) Personalizing Options through Veteran Engagement (PROVE) Quality Improvement Research Initiative (QUERI) Program in Ann Arbor, Michigan (QUE 15-286).

References
