Measuring Implementation, Spread and Sustainability of Educational Innovations: Innovating for Coordinated Collaborative Research

Jeanne Century
Amy Cassata
Cassie Freeman
Mollie Rudnick

Researchers Without Borders Initiative
Center for Elementary Mathematics and Science Education (CEMSE)
University of Chicago

American Educational Research Association
Vancouver, BC, April 14, 2012

This work has been funded in part by the National Science Foundation.

INTRODUCTION

This paper is a proposal; or perhaps it's an invitation, or even best called a "plea" to the field to bring about in ourselves the kind of courageous innovation we seek to bring to schools every day. This is a request to researchers who *study* change to make themselves the "*changees.*" It is a suggestion that moving our individual interests to the background will enable us to prioritize a powerful, collective effort to develop, implement, spread, and continuously support improvements in education. In short, we propose the creation and continued development of shared concepts and tools, and common language for examining enactment of educational innovations.

Researchers who wish to understand how educational innovations are used, spread to new places and sustained over time have an unusual opportunity. We have the chance to capitalize on the good work conducted in other fields to create a learning infrastructure for ourselves that will benefit our work and most importantly, our communities, schools and students. Working collaboratively will require that we communicate in ways that historically have not been part of our practice. It will call for behaviors and interactions that have, to some extent, been outside of what is considered acceptable in academic circles. More specifically, we will need to compromise as we develop and refine our terms and definitions. We will have to surrender our established ways of developing and protecting our work. And we will need to share our work with one another as it progresses – in an open, collaborative, research environment.

Many of today's reforms call for teachers and school administrators to move away from their comfortable daily practice with the expectation that there will be a high return on their risk. The same expectation applies to researchers. Changing our practice will contribute to the development of coherent conceptual frameworks, clear terms and definitions, and shared measures and tools. Then, we will be able to develop findings that will not only highlight our personal efforts but also contribute to the field. As Backer and David state, "the academic research culture has a significant impact on technology transfer….the reward system in academia is not geared to promote technology transfer; articles in prestigious refereed journals and conference presentations are the coin of the realm, not evidence of field use" (p. 263). By shifting our efforts to understand innovation implementation collectively, rather than individually, we will, ourselves, be innovators and together we will be able to more readily impact practitioners' daily work.

Our call for self-innovation grows out of two National Science Foundation funded projects developed at the Center for Elementary Mathematics and Science Education (CEMSE) at the University of Chicago over five years ago. The first project focused the development of a conceptual framework and suite of instruments for measuring the fidelity of implementation of mathematics and science instructional materials. The second project focused on sustainability of innovations and the identification of factors that

contribute to and inhibit reform endurance. At the time we began these independently conceived projects, we didn't realize that they would eventually come together to put us on the path we are proposing here.

The development and theory underlying the first project, "Applied Research on Science Materials Implementation: Bringing Measurement of Fidelity of Implementation (FOI) to Scale" (2007-2010)" has been well documented (Century, Rudnick & Freeman, 2010; Century, Rudnick & Freeman, 2011; Century, et al., in press) so we will not review it here. We will, however, address how the framework we developed in that project has become part of our larger landscape of understanding "implementation" as it relates to the enactment of educational innovations.

The other project, "Accumulating Knowledge on Scaling and Sustaining Reform: A Foundation for Future Research" (2007-2011), set out to investigate a long-standing but often disregarded question in education research: How can we get our innovations to last? At the time the study began, the desire to "scale up" innovations was very visible in the field, but few explicitly examined the question of what happens next. Although perhaps an implicit concern, the issue of innovation sustainability, both over time and across contexts, had rarely been the focus of explicit study in education. Rather, educators appear to have relied on two implicit answers to this question: "Get more money"; and "If it works, people will continue doing it." While both of these seem reasonable, neither statement is the answer to reform endurance. Each is only one very small part of a constellation of factors that affect implementation and sustainability.

Many other questions unfolded as we progressed with this work: How does one introduce and support an evidence-based practice to its intended users so that it is effectively implemented? How and why do practices spread from place to place? Where is the boundary between the intervention and the factors that affect it? How much adaptation can there be? Which factors affect implementation the most? How do the factors that affect implementation change over time? In considering these questions in the context of our own work, we became increasingly convinced that the best and perhaps only way to find answers to these challenging questions would be to make a commitment to working, openly and collaboratively, with others.

The next part of this paper will describe our initial work in identifying the factors that contribute to and inhibit sustainability of innovations. Then, we will discuss how the findings of that study established a foundation for our current work that involves measuring the presence of these factors in the contexts of two different kinds of educational innovations. We will discuss how this work and the work of the fidelity of implementation study mentioned above came together to create a landscape for measuring the overall implementation process. And finally, we will conclude with recommendations for our field for creating a conceptual framework, common language, and an open collaborative dialogue that in time will help us to understand how to bring effective innovations to our schools.

THE SUSTAINABILITY STUDY

Rationale

Some researchers suggest that the factors affecting sustainability narrowly focus on the attributes of the innovation to be sustained (Akerlund, 2000; Han & Weiss, 2005). Others focus on the social and human capital of the individuals in an organization that is attempting to sustain an innovation (Atkins, Graczyk, Frazier, & Abdul-Adil, 2003; Elias, Zins, Graczyk, Weissberg, 2003). And still others identify organizational and extra-organizational factors that influence the sustainability of an innovation (Akerlund, 2000; Evashwick & Ory, 2003; Goodson, Smith, Evans, Meyer, & Gottleib, 2001; Harvey & Hurworth, 2006; Shediac-Rizkallah & Bone, 1998). It became clear to us that these and other researchers have implicit understandings of the concept of sustainability and that without clarifying the conceptual and operational definitions underlying their work, we would not make progress as a field. Further, in order to account for the wide range of factors identified by researchers, we saw that we would need to create a unified framework that would place those factors in a single landscape.

Before moving forward, it is necessary to establish some operational definitions for this paper. We, like many researchers who have published about implementation, use the terms *program*, *innovation*, and *intervention*, interchangeably (Durlak & Dupre, 2008; Proctor, et al., 2011). Further, because we are focused on educational *change*, we assume that the program, or innovation is new to the user(s) and is thus considered an innovation for them, even if the practice isn't new to the field. And finally, we will be discussing two kinds of implementation in this paper: *innovation implementation* (the extent to which the innovation, *itself* is implemented in whole or part) and *the implementation process* (the *factors* that contribute to and/or inhibit the implementation of the innovation).

The necessity to offer definitions in this paper is a testament to the challenges we face in the field that our sustainability project sought to address. We have no shared understanding of the terms "sustainability." It has become an umbrella term for a family of ideas that include "institutionalization," "institutional change," "longevity," "maintenance," "routinization," "stabilization," "incorporation," "continuation," "integration," "acceptance," and persistence" (Goodman et al, 1993). As we began to look at these terms in our work, they initially appeared to mean essentially the same thing; but a closer look at their use in context revealed that in fact, the concepts behind them were often very different. It had become clear to us that before any real progress could be made regarding understandings about the factors that affect innovation implementation, spread and sustainability, our field had to develop conceptual coordination and rigor and clarity in our language.

We began our study aware that although the lack of endurance of reforms had been well documented, it had not yet been rigorously studied (Tyack & Cuban, 1995; Datnow, 2001). Simply put, "there is a paucity of field-based, empirical research that focuses intensively on the issues revolving around the question of whether or not…changes can and will be institutionalized in schools" (Prestine, 2000, p. 126). Over twenty years ago, sustainability was identified as being, at best, a "'latent' concern" (Goodman & Steckler, 1987/88 as cited in Shediac-Rizkallah & Bone, 1998, p. 87). The problem persists. While some attention to sustainability has increased, with little progress made in the way of systematic research, conceptual development, or instrumentation for measuring sustainability, it (as well as "implementation" and "spread") remains a hidden albeit increasingly critical issue.

Although education was the focus of our work, the questions we were asking were also key challenges in health, medicine, mental health, and other fields, so we turned to their literature as well. Further, because there was little research specifically focused on "sustainability" in education, we had to broaden the scope of our search topics for our theoretical underpinnings. This opened the door for our investigation into the research on implementation and diffusion of innovations including "diffusion of knowledge," "dissemination," and "knowledge transfer." While the expansion of our search yielded more useful sources, the literature bases were simultaneously overlapping, confused and poorly defined which left us uncertain as to whether we were making progress, or taking unintentional steps backward.

Over time, however, we came to understand that the factors that contribute to and inhibit "sustainability" are of the set that contribute to "diffusion" and that these factors all influence the implementation process. This initial realization planted the seeds for a now more fully developed vision for identifying a single set of factors, each with clear and specific, measurable definitions. These factors all affect implementation of innovations to some degree, with their impacts varying from the very beginning of use (or what some refer to as "adoption") through a process of "spread" or what some would call "scale-up," all the way to what some would call the "phase" or "stage" of sustainability. We will discuss this further in the later sections of this paper.

So, what are the factors that influence implementation, spread and sustainability? Some cite structural concerns (e.g. workload, inadequate computerization, and access to materials); attitudinal issues (e.g. focus on other priorities, relationships with others); and educational issues (methods by which information is learned) as potentially affecting use of an innovation (Cranney et al, 2001; Freeman & Sweeney, 2001). Others suggest that knowledge may be used simply because "it is engaging or compelling, and because the person presented with it can imagine how it would apply to him or her" (Louis & Jones, 2001, p. 20). In sum, one must account for many variables when developing a model of or theory about the diffusion process in education (Robertson, 1967). Our project set out to identify and clearly define these variables.

Sustainability Study Methods

This section of the paper reports specifically about the study to identify the factors that contribute to and inhibit sustainability of innovations in education. This project, the "Accumulating Knowledge on Scaling and Sustaining Reform: A Foundation for Future Research" project, funded through NSF's REESE program sought to:

1. provide a foundation for accumulating knowledge about scaling and sustainability of innovations in education;
2. identify knowledge about scaling and sustainability of innovations from other disciplines that could inform researchers', reformers' and policy makers' improvement efforts in education; and
3. establish a forum for cross-discipline collaboration and sharing knowledge on scaling and sustainability of innovations.

In addition to the search and review process described below, we used a technique known as "concept development" as described in Rodgers and Knafl's (2000) book *Concept Development in Nursing*. This book includes what the authors describe as foundations, techniques and applications for concept development; and suggests this technique can address some of the most obvious conceptual problems in nursing including "vague terminology, ambiguity regarding the definitions of important concepts in nursing, and inconsistencies among theories" (Rodgers & Knafl, 2000, p. 4). Given that the barriers to accumulating knowledge about implementation, spread and sustainability in education were similar, we felt this approach would be applicable. Specifically, we explored and articulated conceptual and operational definitions using methods of "simultaneous concept analysis" (Haase et al, 2000). By assimilating and evaluating individual conceptual and operational definitions of sustainability and related terms, we were able to clarify individual terms such as "diffusion" and differentiate them from related terms such as "dissemination." As we engaged in the concept development process, we organized the concepts we identified into a conceptual framework that grouped the concepts into categories. A more specific description of the framework and its applicability to current studies is given in the latter part of this paper.

Search Procedure

Pilot Search—The project began with a pilot of our search strategy. The pilot was a test of our search procedures and informed our preliminary coding development. We began by using EndNoteX to search the SSCI Web of Science Index. The searches were based on three groups of terms: group 1 included terms that are roughly synonymous with sustainability or scale-up; group 2 included terms that indicated the object of sustainability (e.g. program, reform, innovation); and group 3 named the field or discipline (see Table 1). Each

search was designed to return results that included at least one word from each of the three groups. The search strings were TS = (sustain* OR diffuse* OR endur* OR scale* OR institutionalize) AND TS = (implementation OR program OR capacity OR practice OR policy OR innovation OR curriculum OR knowledge OR organizational change) AND TS = (science edu* OR math* edu*).

Table 1.
Pilot Process Search Terms

Search Term 1	Search Term 2	Search Term 3
Sustain*	Implementation	Science Edu*
Diffus*	Program	Math* Edu*
Endur*	Capacity	
Scale*	Practice	
Institutionalize	Policy	
	Innovation	
	Curriculum	
	Organizational Change	
	Knowledge	

This process returned 554 articles. We reviewed the abstract of every source and downloaded the full text of those articles that seemed relevant. We reviewed the selected sources in greater depth and rejected those that ultimately were not relevant. Once we identified sources that appeared to be useful and relevant, we retrieved the sources identified in their references, and put these through the process described above. Additionally, we ran Google searches using the same search terms used in the EndNote search. We then repeated the process by searching the ISI Web of Science Index and the ERIC at EBSCO index. The pilot search process yielded useful information about necessary adjustments we needed to make to the search process for the full search.

Full Search—As a first step in the full search, we finalized the selection of search terms. Building on the pilot process, we identified what we referred to as first-, second- and third-degree terms. First-degree terms were those that captured our primary phenomenon of interest – sustainability – and were essentially synonyms for the concept of sustainability in the many ways it is used. Second-degree terms were those that answered the question, "sustainability of *what*?" and included words such as *change* and *program*. Third-degree words described the context for the sustained change – ranging from science education to marketing (see Table 2). Every search had at least one word from the first-degree list *and* one word from the second- *or* third-degree list.

We began by performing Boolean searches on Google and of the Web of Science SSCI, Library of Congress, and Proquest Dissertation databases using EndNote X. Initial searches with these terms returned too many irrelevant results so we narrowed our searches by subtracting terms that had been the least fruitful

including *adher**, l*ongevity*, *utiliz**, *persist**, and *routinization* from the list of first-degree terms and *curricul**, *capacity, knowledge, and organizational change* from the second-degree terms. We felt confident that any of the relevant sources that would have been returned for the terms we eliminated would be identified through the use of the remaining search terms. Additionally, we decided to discontinue searching the Library of Congress

Table 2.

Full Search Process Search Terms

First-degree	Second-degree	Third-degree
Sustain*	Reform*	Educat*
Scal*	Change	Science Edu*
Diffus*	Innovat*	Math Edu*
Institution*	Program	Business
Adher*	Policy	Marketing
Maintain*	Curricul*	Health
Longevity	Implement*	Economic*
Utiliz*	Capacity	
Dissem*	Progress	
Persist*	Knowledge	
Routinization	Organization*	
Endur*	Organizational Change	

and only use books that were cited in sources identified earlier in the search process. Finally, we decided to eliminate searches with Google until we had more specific search terms and people to search for, based on the results of earlier searches.

At the end of the full search process, we had identified 69,801 sources for abstract review. Table 3 illustrates the number and type of sources identified at the end of the general search process.

Table 3.
Sources Identified by Type and Field

Field			Source Type		
	Total	% of total	Journal	Dissertation/Thesis	All Types
Science Edu*	2,750	3.9%	1,968	782	2,750
Math Edu*	1,238	1.8%	367	871	1,238
Edu	28,555	40.9%	19,806	8,749	28,555
Health	9,197	13.2%	5,353	3,844	9,197
Business	8,255	11.8%	3,352	4,903	8,255
Marketing	2,505	3.6%	1,001	1,504	2,505
Economics	17,301	24.8%	9,126	8,175	17,301
Total	69,801	100%	40,973	28,828	69,801

Abstract Collection and Review

We used EndNoteX to draw abstracts from the Social Science Citation Index and then compiled them into pdfs according to the third degree search terms (e.g. science education, business, health) so that team members could review them in subject matter groups. In order to be included in the full review, a source abstract had to provide evidence that the full source would meet at least one of the following criteria:

1) the source provides a conceptual or operational definition of one of the first degree search terms;
2) the source identifies a factor that influences one of the first degree terms; and/or
3) the source describes an approach, methodology, or instrument for studying or measuring sustainability.

As the process proceeded, the team decided to also include sources with abstracts that suggested the source had some information about dissemination, diffusion, and/or implementation that might be relevant to sustainability. There was also a provision for inclusion of abstracts that didn't meet the other criteria but appeared to still be of interest for an unanticipated reason.

Once the abstracts were compiled into pdfs by subject area search term (e.g. science education, mathematics education, marketing, etc.), each member of the team was assigned to one of the fields (or a portion of a field for the larger abstract sets) and coded their assigned abstracts using a pdf tagging process. The tags used for inclusion are in Table 4.

Table 4.
Abstract Tags for Review Inclusion

Tag	Definition
Operational definition	Abstract suggested the source contained an operational definition of one of our terms of interest;
Conceptual definition	Abstract suggested the source contained a conceptual definition of one of our terms of interest;
Influential factor	Abstract suggested the source identified a factor that contributed to or inhibited sustainability or another first degree search term;
Methodology	Abstract suggested the source contained a description of a methodology for studying "sustainability
Diffusion and scale-up review	Abstract suggested that the source contained relevant information about diffusion and scale up
Implementation review	Abstract suggested the source contained relevant information about implementation
Sustainability review	Abstract suggested the source contained other relevant information about sustainability
Alternative inclusion	Abstract suggested the source should be included for a reason other than the others described

A ninth tag, "not useful," was used to identify abstracts that had been reviewed and established to be of no relevance to our work. Initially, each abstract was tagged with all relevant tags. In order to streamline the process, however, we shifted to identifying sources with only a single tag since that was sufficient to warrant inclusion in the full text coding process.

Summary of Abstract Review

We reviewed over 69,000 abstracts and identified a total of 657 sources for the full text coding process. Of these, 572 were retrievable. Table 5 shows the total numbers of abstracts identified with each tag. There is not a 1:1 correspondence of tags to sources because as noted above, the team began the process by tagging with all relevant tags but later moved to a process identifying a single tag for each abstract.

Table 5.
Total Numbers of Tags

Operational Definition	9
Conceptual Definition	25
Methodology	44
Influential Factor	147
Sustainability Review	421
Implementation Review	219
Diffusion and Scale-Up Review	178
Alternative Inclusion	1

Report Coding

We collected the 572 sources and divided them into groups based on their subject area search term (business, economic, educat, health, marketing, math edu, science edu). Then, we organized the documents into batches for members of the team to read and code using NVivo 7. Table 6 summarizes the numbers of sources by subject area search term.

Table 6.
Total Sources by Subject Area Search Term

Business	71
Economic	142
Educat*	354
Health	269
Marketing	24
Math* Edu*	16
Science Edu*	142

Prior to independent study coding, the team completed a practice coding process with eight sources identified during the pilot collection process and determined to be of high relevance. In order to begin, the team created a codebook that eventually came to be called the "Nodebook" (titled as such because NVivo refers to text codes as "nodes"). The team created the first draft of the Nodebook with a list of possible nodes they anticipated using, with the expectation that the Nodebook would be revised iteratively with the coding process. At the outset, there were three groups of nodes: 1) words the team anticipated would appear in the sources that would be important to track (e.g. "sustainability"); 2) anticipated factors that would affect

sustainability (e.g. financial resources); and 3) other types of information worth tracking (e.g. operational and conceptual definitions of sustainability). This practice coding process allowed the team members to became familiar with the use of NVivo, the Nodebook, and the coding process. Once the eight sources were coded in NVivo, the team members discussed their coding decisions and made suggestions about the coding process in order to ensure shared understanding and consistency in the process.

Once the practice coding process was complete, the team members began independent coding with a minimum of ten studies each week. The coding was distributed in such a way that each team member became an "expert" in two subject area categories (e.g. business, economic, science). In addition, each team member coded a smaller number of sources from a different category. With this design, team members developed expertise in particular areas, grew to understand the nuances of subject-specific language, could more easily identify relationships or references across sources within a field, and could identify connections across fields.

As the coding proceeded, the team met twice weekly to share findings and emerging ideas and discuss issues that arose in the coding process. During these meetings, the team discussed the introduction of new nodes for the Nodebook, revisions to the definitions of existing nodes, and the potential elimination or adaptation of nodes. As the source review proceeded, the team refined the definitions of each node and incorporated illustrative examples from the sources themselves into the Nodebook.

It had been the team's intention to create a second "generation" of sources by reviewing the reference lists of sources collected during the search. However, as coding of the sources continued, it became clear that limited time and resources would not allow for a full review of all citations. Instead, the team created a node called "reference" and used it to identify citations in the source text that appeared to be of some potential interest. Then, the team looked at all of the tagged references to determine which had already been captured in the first search and which would need to be collected. Most of the citations tagged in this manner were books and conferences proceedings, but the process did uncover some additional journal articles sources that we included in the review.

During the coding process the team observed that some of the nodes appeared to be related to each other, and that some of the broader nodes encompassed some of the more specific ones. For example, the node "Characteristics of It" (characteristics of the innovation) was a broad term that encompassed more specific nodes such as "Adaptability", "Complexity", "Scope", and "Specificity." In order to capture these relationships, the team used the "tree" node function in NVivo. This meant that when coders identifies source text to be associated with one of the specific nodes (e.g. adaptability) and coded it as such, it would also be automatically coded with the broader tree node (characteristics of it).

Categorization of Sources

After coding each document, team members created an annotation (a mechanism in NVivo) to summarize it, and then assigned it to one of the following categories: seminal, significant, contributing, of interest, limited importance, and not useful. Table 7 shows the distribution of these sources across the categories. Given their potential importance, all team members were tasked with reading all of the articles that were deemed "seminal," and one team member was tasked with reading and doing additional coding on all of the "significant" and "contributing" works.

Table 7.
Sources by Usefulness

Seminal	16
Significant	18
Contributing	60
Of Interest	145
Limited Importance	133
Not Useful	198
Total	572

Analysis and Results

As the coding process neared completion, the team began analysis. Using the report function in NVivo, the team created "reports" on various nodes. A report is a compilation of all of the text from all sources coded with a particular node. In reports, text excerpts are organized by author and then reference number (some sources have multiple excerpts, or "references" for a single node). Reports varied in length from only a few pages to up to thirty or more, depending on the node.

During the coding process, the team began to review selected reports based on questions about particular nodes that arose during meeting discussions. These reports would not be complete until the coding process came to an end, but these initial report reviews yielded emerging ideas about the factors that affect sustainability and their relationships to one another. Once all sources had been coded, the team decided on a specific order for report review, starting with concepts and definitions pertaining to sustainability and then moving on to the factors affecting implementation and sustainability. As the team proceeded, new categories of nodes emerged that provided additional structure to the order of the report review process.

All team members read each report and met regularly to discuss the main ideas emerging from the reports and highlight particular references that supported those ideas. These discussions led to more refined understandings of the nodes, revision of node definitions, addition of nodes, and over time the articulation of relationships between the nodes and ways to categorize them. Periodically, this process led to what we called "thought memos." Upon discussing reports, team members documented their thinking about particular nodes and then used those documents as the focus for discussion in future meetings.

Through conversations with Rodgers as well as reading her book titled *Concept Development in Nursing: Foundations, Techniques, and Applications* the team was better able to plan for how the nodes could be used to create a conceptual framework. Rodgers suggests creating model cases for concepts of interest. In our case, these model cases would be the "operational definitions" for each node. Concurrent with reviewing the reports and revising the Nodebook, the team began the creation of a conceptual framework for organizing the emerging ideas. Rodgers touts a "low-tech" method using butcher paper and note cards. The team decided to use "sticky notes" and began to organize the different concepts on a wall, shifting them with each revision of our thinking about the relationships each node had with another and their sensible groupings.

As each node was discussed and each report reviewed, the team again revisited the node definition and refined it as necessary. The team also reviewed examples of the nodes that had been identified in the reports for accuracy and then discussed the relationships between those nodes and other nodes. This process led to questions about other not-yet-reviewed nodes and thus provided an order in which to review subsequent reports. Throughout the process, the sticky note associated with each node was moved on the wall as appropriate to newly identified relationships and categories. Ultimately, we developed a framework that included 28 factors, 11 processes or strategies, and 6 "concepts" or definitions (see Table 8). Since completion of this first iteration, we have revised the framework to include more conceptual clarity and specific definitions and to develop more consistency with the work of others. The evolution of the framework was driven by the need to bring what had until that point been theoretical ideas to practical use. That process and the revised framework are described below.

Table 8. Original List of Factors Contributing to and Inhibiting Sustainability of Innovation

Category	Node	Definition
Reasons for Decisions	Buy-In	User decision to adopt or use an innovation because the user wants to.
	Compliance	User decision to adopt or use an innovation because the user(s) feel they have to.
	Reward	User decision to adopt or use an innovation because they are compensated.
Characteristics of People in the Organization	Characteristics of Implementer	The attributes of the individuals who have responsibility for implementing an innovation. This includes demographics, attitudes, and sophistication of the user.
	Quality of Leaders	Evidence of an individual providing guidance, direction or support for an innovation. A leader is an implementer who formally or informally provides guidance, direction, or support for the innovation.
	Sophistication of the User	The extent to which user(s) (organization or individuals) have experience with and understanding of the innovation and/or other innovations.
Elements of the Environment	Locus of Decision-Making	The people involved in decisions relating to an innovation.
	Opportunities for Learning	Sources of growth and development for implementors (including leaders) of an innovation and include professional development, forums for collaboration, mentoring, meetings with consultants, etc. These are often referred to as opportunities for human capacity building.
	Incentives	Any compensation for adopting or using the innovation.
Elements of the Environment-Inside	Internal Organizational Structures	The formal rules and guidelines for operations and the organizational chart of an organization in which an innovation resides.
	Internal Social Climate	The extent and nature of human interaction (e.g., collaboration, communication) in an organization.
	Resource Allocation	The distribution of human and material resources in the organization in which an innovation resides.
Elements of the Environment-Outside	External Climate: Political	The policies (e.g., federal, state, local, institutional) outside of the organization within which an innovation resides that support of hinder the innovation.
	External Climate: Social	Beliefs and attitudes in the population surrounding the organization in which the innovation resides.
	Networks	Connections between the user and external organizations or individuals.
	Resource Allocation	The distribution of human and material resources outside the organization in which the innovation resides.
Fit	"Fit" – Current Practice	The extent of alignment between practices supported by an innovation and current practices.
	"Fit" – Needs	The extent to which the user's (or organization's) actual or perceived needs are met by an innovation.
	"Fit" – Values	The extent of alignment between the user's (or organization's) values and an innovation.
Properties of the innovation	Properties of the innovation	The attributes of an innovation that are uninfluenced by other factors at a given point in time.
	Adaptability	The ability of an innovation to be altered. This is comprised of complexity and specificity.
	Complexity	The number of an innovation's parts and the extent of their interdependence.
	Scope	The population, in number and category, of participants (these might be referred to as enactors or receivers of the innovation).
	Specificity	The degree to which the opertaionalization of an innovation is explicitly described.
	Effectiveness	Evidence (can be empirical or informal) that an innovation accomplishes desired outcomes.
	Visibility	The ability of an innovation to be seen/recognized.
Strategies	Collaborative Change Process	The engagement of the users in decisions related to an innovation.
	Feedback	Asking users for input, but not engaging them in decisions related to an innovation.
	Formative evaluation	Ongoing evaluation that provides users with data to inform decision making about an innovation.
	Mandate	Telling users to engage in adoption or use or other actions related to an innovation (without user input).
	Planning	Ongoing strategizing about the continuation of an innovation.
Time	Duration	A period of time that marks the presence of an innovation in any form.
	Phase/Stage	A discernable period in the implementation of an innovation that is distinguished from other phases or stages.
Trust	Trust	Willingness to engage in an innovation because the request comes from a credible source where a previous relationship exists.
Movement	Movement	The difference in number of users between one point in time and another. A decrease in number of users is reduction. An increase in numbers is expansion.
Locus of Movement	User to user	The transfer of an innovation in any form from one user (e.g., classroom, school, organization, individual) to another.
Type of Movement	Originator to user	The establishment of an innovation that originated outside of the user.
	Dissemination	Active distribution.
	Diffusion	Passive distribution.
	Adoption	User decision to implement.
Ways of lasting	Sustainability	Lasting of an innovation characterized by its evolution and adaptation.
	Maintenance	Lasting of an innovation characterized by a lack of change.
	Institutionalization	Incorporation of an innovation into everyday practice, such that it is no longer recognized as an innovation.
	Use	The operationalization of an innovation at a given point in time.
What lasts?	The innovation	A new or different belief or practice.
	Ongoing effectiveness	Achievement of desired outcomes. Desired outcomes may evolve over time.

BUILDING FROM THE CONCEPTUAL FRAMEWORK TOWARD CLEAR LANGUAGE AND EMPIRICAL STUDIES

Evidence of Confused Language

One of the expected, but still striking, early findings was the variability in the definitions of "sustainability" and the its contributing factors. For example, one source described sustainability of a program as the "extent to which a new program becomes embedded or integrated into the 'normal' operations of an organization" (O'Loughlin, Renaud, Richard, Gomez, & Paradis, 1998) while another described sustainability as "a process of developing local capacity to enable a program to be maintained at the stakeholder/community level" (Harvey & Hurworth, 2006). Even some definitions of "implementation" seemed to align with others' definitions of sustainability: "the transitional period during which targeted organization members ideally become increasingly skillful, consistent, and committed in their use of an innovation" Klein and Sorra (1996).

A stark example of the field's confusion about definitions and its effect on our ability to understand one another is in a game we created as a mixer for a meeting related to the sustainability project. In this game, researchers are asked to match the words "maintenance," "diffusion," "dissemination," "scale-up," "sustainability," "adoption," and "institutionalization" with definitions of those words used in published articles. In the game, the results are displayed as pie graphs. If players match the terms and definitions correctly, the pie charts showing the results will all be single colors. However, that has not been the case. See Figure 1 below for the results of the game as of this publication. The first pie graph shows that 24% of the definitions the researchers associated with "maintenance" were actually definitions associated with "institutionalization" in the literature. Likewise, 8% of the definitions associated with maintenance were actually definitions for "diffusion." This and the other graphs demonstrate that there is far to go before researchers can communicate with one another about implementation, spread, and sustainability with confidence that each understands the other. To do this activity yourself, go to: http://www.researcherswithoutborders.org/projects/sustaining-innovation.

Operationalizing the Framework in Studies of Innovations

As mentioned above, our need to develop more clear and specific definitions of the factors in our framework was driven by more than theory. After finishing the original framework development, we received funding for two more studies that would require us to develop each of the factors into measurable constructs. The first study is examining implementation of Science, Technology, Engineering and Mathematics (STEM) Schools in the Ohio STEM Learning Network and the spread of those schools' structures and practices to other schools in their communities, state, and beyond.

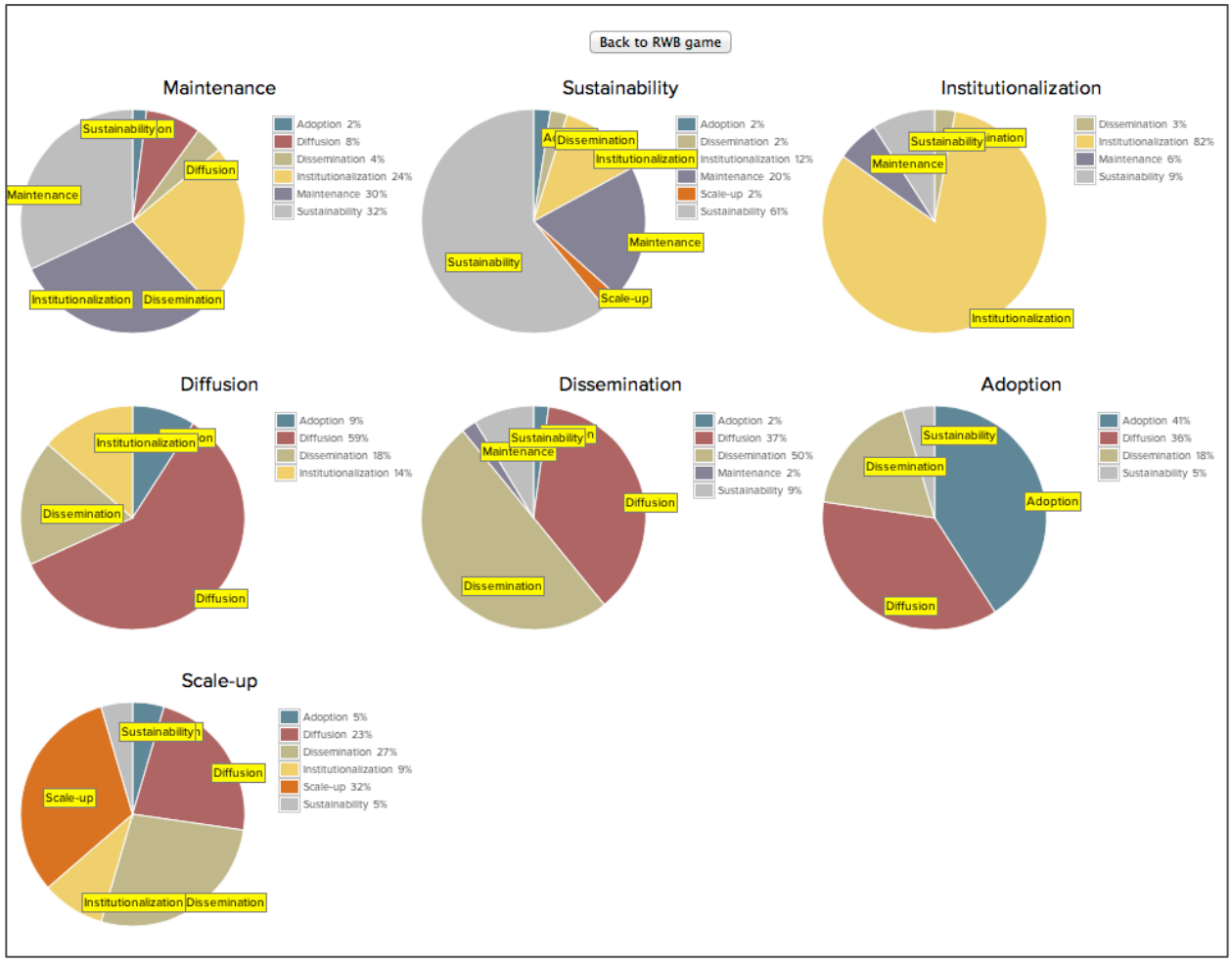

Figure 1. Results of the "Sustainability Game" http://researcherswithoutborders.org/projects/sustaining-innovation

In order to move from our conceptual framework to actual measurement for this new study, it was necessary for us to reconsider the definitions we had created for each factor. More specifically, we had to consider potential indicators and how to operationalize them in the educational context. As part of this process, we searched for and collected existing research and associated instruments with items pertaining to the framework factors. We identified 85 instruments (including 8 focused on leadership, 28 focused on organizational climate, and 12 focused on self-efficacy) and used them as a starting point for developing our factor measures. Table 9 illustrates our initial revision of the factor list as we operationalized the factors and their categories to use in our STEM school study.

We created first drafts of the instruments using a combination of items adapted from existing instruments and newly created items. The item discussions took place iteratively with the conversations focused on creating clearer definitions of the factors, as we considered more specifically the items we wanted

to use to measure each factor. This process also included considerations of the most appropriate data sources (e.g. teachers, school leaders, students, etc.) and instruments (e.g. questionnaires, interviews, observations) for each factor. We arrived at a combination of instruments including teacher, school leader, and student questionnaires, teacher and school leader interviews, partner interviews, and school observations.

Table 9. Initial Factor Revision

Category	Original Factors (previous REESE project)	STEM School Study factors
Reasons for Decisions	Buy-In	Subsumed into "Intrinsic Motivation"
	Compliance	Subsumed into "Intrinsic Motivation"
	Reward	Became "Benefits for Participation"
Characteristics of People in the Organization	Characteristics of Implementer	Broken down into "Innovativeness," "Descriptive Characteristics," and "Resourcefulness and Coping"
	Quality of Leaders	Became "Transformational Leader"
	Sophistication of the User	Became "Understanding of the Innovation" and "Self-Efficacy"
Elements of the Environment	Locus of Decision-Making	Kept
	Opportunities for Learning	Split into "Opportunities for Learning-Presence" and "Opportunities for Learning-Utility"
	Incentives	Became "Benefits for Participation"
	---	Resource Sufficiency (added)
Elements of the Environment-Internal	Internal Organizational Structures	Kept
	Internal Social Climate	Split into "Communication" and "Collaboration"
	Resource Allocation	Kept
	---	Clarity of Organizational Goals (added)
Elements of the Environment-External	External Political Climate	Kept
	External Social Climate	Kept
	---	Networks (added)
	Resource Allocation	Kept
Fit (organization or individual)	"Fit" – Current Practice	Kept
	"Fit" – Needs	Kept
	"Fit" – Values	Kept
Properties of the Innovation	Adaptability	Split into "Adaptability-Perceptions" and "Adaptability-Descriptive"
	Complexity	Kept
	Scope	Kept
	Specificity	Kept
	Effectiveness	Kept
	Visibility	Split into "Visibility-Descriptive," "Visibility-Implementation/Outreach Activities," and "Visibility"
	---	Ease of Use (added)
	---	Results Demonstrability (added)
Strategies	Collaborative Change Process	Subsumed into "Locus of Decision-Making"
	Feedback	Subsumed into "Locus of Decision-Making"
	Formative evaluation	Kept
	Mandate	Subsumed into "Locus of Decision-Making"
	Planning	Kept
	---	Leveraging (added)
Trust	Trust	Subsumed into "Collaboration"
Student Attitudes	---	Student Attitudes

The second current study is examining the implementation of the *Everyday Mathematics* program in several districts across the country with a particular focus on differences in the roles of the factors in communities where the duration of use is widely varied. This study complements the STEM school study because they share similar questions, but focus on innovations with very different characteristics. The *Everyday Mathematics* study has provided an opportunity to test our factors and framework with an innovation that is very specific, and has been in place, in at least two of the study sites, for years longer than the STEM schools have been in existence. As mentioned above, this framework is in a process of evolution. The current version, with current definitions is in Table 10.

Table 10. Factors, Processes and Mechanisms that influence the Implementation, Spread, and Sustainability of Educational Innovations

Characteristics of the Innovation	*The attributes of an innovation that are uninfluenced by other factors at a given point in time.*	
	Complexity	The number of an innovation's parts and the extent of their interdependence.
	Specificity	The level of detail in which the operationalization of an innovation is described.
	Adaptability	This is comprised of complexity and specificity.
	Scope	The innovation's target area(s) within the field of education.
	Empirical Effectiveness	Evidence that an innovation accomplishes desired outcomes.
	Results Demonstrability	The extent to which the impacts of an innovation can be communicated/shown to others.
Characteristics of Individual Users	*The attributes of users (or potential users) of the innovation. Users include teachers, school leaders, and students.*	
IN THE CONTEXT OF THE INNOVATION		
	Self-efficacy	Competency and confidence in the user's (or potential user's) own abilities to enact the innovation.
	Understanding of the Innovation	The extent to which a user (or potential user) understands the strategies, components, and goals of an innovation.
	Attitude toward the Innovation	The extent to which the user (or potential user) is in favor of (or not) use of the innovation.
	Attitudes towards things related to the innovation	Enjoyment of topics/areas related to the innovation. (Also related to intrinsic motivation.)
	Intrinsic Motivation	Influence on an individual's decision-making that comes from their level of enjoyment of, sense of commitment to, and sense of ownership toward an innovation.
	Extrinsic Motivation	Influence on an individual's decision-making that comes from external incentives such as recognition, money, and power, or to avoid negative consequences from an external source (e.g., to avoid being penalized).
NOT IN THE CONTEXT OF THE INNOVATION		
	Innovativeness	The extent to which a user (or potential user) seeks out, creates, and/or enacts new ways of doing things.
	Resourcefulness and Coping	The ability of a user (or potential user) to combat stress and persist with difficult goals/tasks.
	Networked-ness	The tendency for an individual to participate in a social network (inside or outside of the organization) Example for inside the organization: Level of social participation (e.g., attendance at meetings, density of network). Example for outside the organization: Number of links a person has outside the organization. (Reference: Greenhalgh et al (2004), *How to Spread Good Ideas*)
	Time Management and Organizational Skills	The act or process of planning and exercising conscious control over the amount of time spent on specific activities, especially to increase efficiency or productivity. The skills enable people to plan and carry on activities effectively. To put order to a situation, objects, or people.

Table 10. (cont.)

USER PERCEPTIONS OF THE INNOVATION		
	Perceived Adaptability	The user's (or potential user's) perceptions of an innovations permissible flexibility.
	Perceived Visibility	The extent to which a user or potential user has seen or heard about the innovation.
	Ease of Use	The user's (or potential user's) perceptions of an innovation's ease of implementation.
	Perceived Effectiveness	The user's (or potential user's) impression that an innovation accomplishes desired outcomes.
DESCRIPTIVE CHARACTERISTICS OF THE INDIVIDUAL USER		
	Demographic	Includes gender, age, SES
	Education	Includes formal education and training
	Experience	Includes number of years in position, experience with the innovation
CHARACTERISTICS OF THE LEADERSHIP (A SPECIAL CASE OF CHARACTERISTICS OF THE INDIVIDUAL USERS):		
	Instructional Leadership	The extent to which the leadership communicate knowledge, expertise, and understanding about the innovation content and pedagogy. a. Leader's Perception: The leadership's views of their own skills, knowledge and qualities. b. Users' Perceptions: The individuals in the organization's views of the leadership's skills, knowledge and qualities.
	Personal Support	The extent to which the leadership provides supports for individuals' personal and emotional needs. a. Leader's Perception b. Users' Perceptions
	Innovation Advocacy	The extent to which the leadership visibly and vocally communicates about the merits and benefits of the innovation. a. Leader's Perception b. Users' Perceptions
	Innovativeness	The extent to which the leaders seek out and enact new ways of doing things. a. Leader's Perception b. Users' Perceptions
Characteristics of the Organization		
CHARACTERISTICS RELATED TO PEOPLE IN THE ORGANIZATION		
	Organizational Innovativeness	The extent to which users perceive that their organization is a place that seeks out, creates, and/or enacts new ways of doing things.
	Organizational Efficacy	The extent to which people in the organization perceive it as being competent at and capable of implementing the innovation.
	Clarity of Organizational Goals	The extent to which people in the organization perceive the intended direction of the organization as it pertains to the innovation are clear.
	Collaboration	The extent to which people in the organization perceive that interactions between individuals in an organization are rooted in trust, mutually beneficial, and make progress toward a goal. *[Note: Communication and collaboration are tied to Organizational Climate and the Operational Network in the Organization]*
	Communication	The extent to which people in the organization perceive that the organization shares information about the innovation that is sufficient for individuals to enact their roles.
	Shared Beliefs and Values	The extent to which people in the organization perceive that the norms, values and beliefs about education are shared by individuals in an organization. (Domitrovich et al, 2008) *[Note: Some people refer to this as Organizational Culture]*
	Locus of Decision Making	The extent to which different individuals perceive they are involved in decision-making processes.
	Resource Sufficiency	The extent to which users feel they have enough resources (financial, material, human) to implement the innovation.
	Time Sufficiency	The extent to which users feel they have enough time to implement the innovation.
	Utility of Opportunities for Learning inside the Organization	The extent to which users feel the opportunities for learning inside the organization are useful.
	Utility of Opportunities for Learning in the Environment	The extent to which users feel the opportunities for learning outside the organization are useful.

Table 10. (cont.)		
Descriptive Characteristics of the Organization		
	Organizational Structures	The formal rules, policies and guidelines for operations of an organization in which an innovation resides. Includes decision-making structures, reporting structures, supervisory structures.
	Financial Resource Allocation	The distribution of financial resources to the innovation relative to the financial resources available.
	Physical Environment	The characteristics of the physical space in which the innovation is enacted.
	Population Characteristics	Demographic, education, and experience of the whole organizational population. This also includes organization size, student mobility, and staff turnover.
	Extraneous Events, Initiatives, and/or Incidents	These are events or initiatives that occur within the organization that can cause distraction from or support for the innovation. (Poor student behavior is included here).
	Stakeholder Community Support	The human, material, and other resources provided by the community of individuals and organizations who are invested in the success of the organization (e.g., volunteers, partners, families).
	Presence of Opportunities for Learning Inside the Organization	The extent to which sources of growth and development for users (including leaders) of an innovation present inside the organization.
Strategies	*Actions taken by the organization that pertain to implementation and spread of the innovation.*	
	Ongoing Improvement Structures	Includes Planning (Ongoing strategizing about the continuation of an innovation) and Formative Evaluation (Ongoing evaluation that provides users with data intended to inform decision-making about an innovation).
	Leveraging	One organization capitalizing on another organization's resources, contacts, or practices across participants in the network.
	Dissemination	Steps taken to actively share the innovation with others.
	Implementation Strategy	The population in number and category of participants in the innovation at any given point in time.
Elements of the Environment		
	Political Environment	The policies, guidelines and rules outside the organization that can affect implementation and spread of the innovation. (District mandates are included here).
	Community Beliefs and Values	The extent to which the population (e.g., community, parents) surrounding the organization in which the innovation resides has beliefs and values that pertain to education and/or the innovation.
	Descriptive Characteristics of the Community	This includes community size, SES, and geographic region.
	Presence of Opportunities for Learning in the Environment	The extent to which sources of growth and development for users (including leaders) of an innovation are present outside the organization.
	Network Structures	Intentionally designed opportunities for interactions between individuals in different organizations
	Descriptive Visibility	Evidence of the innovation in places it can be seen or heard.
	Extraneous Events or Initiatives	These are events that occur around the organization that can cause distraction from the innovation.
Networks		
	Enacted Networks	The nature and extent to which interactions occur between individuals in different organizations. This exists in the space between Elements of the Environment and Characteristics of the User.

Note about Fit: *We don't have a clear home for "fit" and are still in conversations about its place in the framework. For now, we recognize that* ***Fit with Needs (****the extent to which the user's actual or perceived professional needs are met by an innovation)****, Fit with Current Practice*** *(the extent of alignment between an innovation and the user's knowledge, skills and practice)* ***a****nd Fit with Values and Beliefs (the extent of alignment between the user's values and an innovation) are tied to motivation. So, we placed the fit factors here. Like Networks reside in the "space" between Characteristics of the User and Elements of the Environment; Fit might reside in the space between Characteristics of the User and Characteristics of the Innovation.*

In the section below, we follow the form found in other sources (Greenhalgh, 2004; Wejnert, 2002; Durlak & Dupre, 2008) and provide our explanation of the categories and the factors that reside in each. A difference, however, that exists between our work and many others is that we have moved beyond theory and are in the process of actually measuring these factors. Proctor et al. (2011) report on their efforts to develop definitions that reflected as much as possible the consistent use of the terms in the field as well as their intention to "sharpen" distinctions between similar constructs (Proctor et al., 2011). We sought to do the same but had to do so while considering the definition and construct more practically. We had to consider how the construct could actually be measured, within all of the constraints and limits of research conducted in the day-to-day world of schools and education.

In the following section, we do not repeat the definitions shown in Table 10. Rather, we further explain the reasoning behind the definition and distinctions that we made in the process.

Characteristics of the Innovation: During our factor definition process, we considered which of the factors represented *actual* or more "objective" contexts versus *perceived* or more "subjective" contexts. This effort grew from the fact a practical measurement task framed our work and we had to ensure that we were clear and specific about what was actually possible to measure and how each factor should be defined given the constraints of measurement. For example, as we considered "adaptability," we did so from two perspectives. First was the extent to which an innovation could be deemed adaptable based on an objective scale or rubric that could be applied to any innovation. The other perspective, however, was the extent to which the user *perceived* the innovation as being adaptable. This first category then – characteristics of the innovation – includes factors that are among the most objective. Others have identified this same category (Durlak & Dupree, 2008) sometimes referring to it by a different but similar name, "innovation characteristics." (Wejnert, 2002).

The definitions of "complexity" and "specificity" in Table 10 are self-explanatory. We are in the fortunate position of having two projects that focus on very different innovations with regard to their complexity and specificity. For example, the *Everyday Mathematics* program is a very specific innovation that indicates very specific steps for users (in this case, teachers). The STEM Schools, on the other hand, have been grounded in very broad design principles that provide a framing for the innovation, but little specific direction. Complexity and specificity in turn together contribute to a third factor, "adaptability." When there is greater latitude in the guidelines outlining the steps to take in an innovation, and many pieces and parts, there is potentially more opportunity for adaptation. However, the extent to which that adaptation happens is related to other influential factors.

"Scope" refers to the innovation's targeted area of change. More specifically, it refers to whether the change process focuses on a particular aspect of instruction (e.g. questioning – small scope), instruction overall (e.g. pedagogical strategies), one or more topics (e.g. mathematics, science), or multiple aspects of the system (e.g. classroom, whole school, district – large scope). The scope of an innovation can have both positive and negative relationships to other factors and ultimately to outcomes. "Empirical Effectiveness" refers to the evidence of program effectiveness that is outside of the user's direct perception. This includes empirical studies, recommendations, and other testimonials. *Perceived* effectiveness resides in another category (characteristics of the user). There are other factors that we haven't yet included in our framework (e.g. trialability, feasibility) that others have identified that we haven't yet incorporated as they may or may not be appropriate for the educational context.

Characteristics of Individual Users: Greenhalgh et al., (2004) suggest that people 'seek innovations, experiment with them, evaluate them, find (or fail to find) meaning in them, develop feelings (positive or negative) about them, challenge them, worry about them, complain about them, "work around" them, gain experience with them, modify them to fit particular tasks, and try to improve or re-design them – often through dialogue with other users." (p. 598). Their assessment of the role of these very personal and emotional user characteristics cannot be underestimated. Most often, user characteristics are reduced to what in our work we refer to as "descriptive" objective characteristics (e.g. years of teaching, content knowledge background, teacher preparation). While undoubtedly important, other characteristics – what others refer to as "provider characteristics" (Durlak and Dupre, 2008) or "characteristics of innovators" (Wejnert, 2002) are an essential part of bringing change to schools.

In educational settings, users can include administrators, informal educators, teachers, students and potentially a whole range of others. For the purposes of describing the place of user characteristics in the conceptual framework, we group them in the same category. However, it is important to recognize that operationalizing the measurement of these factors in an actual study requires making clear distinctions in the users and the different forms these factors can take.

Regardless of the particular user group, however, we determined that "User Characteristics" fall into four categories. The first category includes those characteristics that are *affected or shaped by the innovation itself.* These characteristics include self-efficacy, understanding of the innovation, attitude toward the innovation, attitude toward things related to the innovation, intrinsic motivation and extrinsic motivation. These user characteristics change as the content, pedagogy, style, and focus of the innovation shifts. See the definitions of these characteristics in Table 10.

The second category of user characteristics are those that are shaped less by the innovation, and pertain more directly to the individual and how that individual generally functions in the educational settings

and perhaps daily life. These are what Wejnert (2002) refers to as "personal qualities." The first in our list is innovativeness – the extent to which a user, in general, is inclined to seek out and try new things. Second, is "resourcefulness and coping." This refers to how naturally and easily a user manages stress and persists with difficult tasks. Networked-ness has to do with the extent to which an individual is, in general, a networked person in educational contexts that can support innovation implementation and spread.

A third set of user characteristics focuses on users' perceptions of the innovation. Although these factors pertain to the innovation itself (similar to "Characteristics of the Innovation"), they reside in the mind of the user and as such, we consider them user characteristics. Perceived adaptability, as mentioned earlier, refers to the user's *perception* of how adaptable an innovation is (as opposed to how adaptable it is structured and/or intended to be). Similarly, *perceived* effectiveness differs from empirical effectiveness (listed in Characteristics of the Innovation) in that it addresses the *users'* impression of the extent to which the innovation accomplishes its desired outcomes, from their direct experiences, as opposed to the users' processing of *others'* reporting of the innovation's effectiveness. Perceived visibility refers to the extent to which the user has awareness of the innovation prior to and during implementation. Ease of use is sufficiently described in Table 10.

The final general group of user characteristics are the "descriptive characteristics" of the user. These are the characteristics that most studies focus on including user demographics, education, and years and types of experience. Although these are often the only user characteristics documented, it is clear from our work and the work of others that they are only a small part of the many characteristics that affect implementation, spread and sustainability.

Continuing in the "Characteristics of the User" group are characteristics of the leader. We consider the leader a special case of the user with a specialized set of characteristics. We recognize that many have studied and identified the role of leaders and their characteristics in bringing about reform. Our list is not a comprehensive one, but rather includes those characteristics that we have felt were particularly important for the innovations that we are investigating. These leadership characteristics are: instructional leadership, personal support, innovation advocacy, and innovativeness. Table 10 provides descriptions of these leader characteristics.

Looking at the table, one can see that we have subdivided leadership characteristics into two groups. First, are the leader's perceptions of his or herself; and second, are the other *other* users' (those who are "following" that leader) perceptions of the leader. Once again, this refinement came from moving from a theoretical framework toward practical questions of measurement. Further, we considered that discrepancies between these two aspects of leader characteristics could greatly affect implementation. Imagine, for example, a situation where a principal perceives that he is providing personal support but the teachers feel he is not.

Such a situation can be even more challenging than one in which they both agree that he is not providing personal support.

Characteristics of the Organization: Characteristics of the organization, or what others have referred to as "organizational capacity," (Durlak and Dupre, 2008) "organizational climate" and "operational network" are commonly identified as factors that affect implementation, spread and sustainability. We have divided these characteristics into two groups: characteristics pertaining to the people in the organization; and characteristics pertaining to the non-human aspects of the organization (e.g. policies, schedules, organizational charts). The definitions of the first group of characteristics are all worded in terms of the organizational participants' *perceptions* or the extent to which they "feel" something is present. We worded the definitions this way because as we began to operationalize measures of these factors, we realized that the only way to measure these factors within a reasonable timeline and the scope of a study, would be to ask the organization members.

Some of the characteristic definitions have nuanced elements that warrant explanation. For example, "communication" doesn't refer to communication in general. Rather, it refers to communication *about the innovation* and the extent to which users feel they have sufficient information to enact their expected roles. Similarly, we define collaboration as more than simply working together. Collaboration includes interactions that are rooted in trust, mutually beneficial and make progress toward a goal. We have refined the definitions this way in order to get at the dimensions of these concepts that most affect implementation. Further, in addition to measuring the *presence* of opportunities for learning (see the "Elements of the Environment section below) in this group of factors we measure the extent to which the users perceive opportunities for learning as being helpful. Simply documenting the presence of professional development is not sufficient for shedding light on implementation. Just as important is establishing the extent to which the professional development participants *perceive* that experience as being helpful.

Elements of the Environment: The elements of the environment, or what others have referred to as "community factors" (Durlak and Dupre, 2008) or "environmental context" (Wejnert, 2002) include factors that are commonly referenced in theoretical pieces on implementation. Some of the factors in this group have counterparts in the other sections and are worth clarifying here. More specifically, "descriptive visibility" differs from "perceived visibility" (in "Characteristics of the User") in that descriptive visibility refers to the extent to which the innovation is *actually* visible or present in the media, organizational documents, and organizational and individual communications. This differs from perceived visibility, which refers to the extent to which a user is aware of the innovation. It seems likely that there would be a relationship between these two, but that relationship is not presumed and can be verified (or not) through empirical means.

Strategies: Strategies are the actions taken by the users that pertain to implementation and spread of the innovation. Ongoing improvement structures, for example, include the interactions that happen within the organization that focus on the use of data and informed strategic decision-making. Implementation strategy has to do with the scope and nature of the innovation "rollout" and ongoing implementation. And finally, we distinguish "dissemination" from other kinds of "diffusion" and spread by referring to it as *active* sharing of the innovation with others. As Greenhalgh et al., (2004) suggest, there is a spectrum of movement from "let it happen" to "help it happen" to "make it happen" that needs to be measured and accounted for.

Networks: Networks, our final category of factors, is undoubtedly the least clear for us. "*Network structures*" is identified in "elements of the environment" as formal networking structures designed to facilitate interactions between individuals in organizations. "Enacted Networks," on the other hand, refer to the nature and extent to which interactions occur between individuals in different organizations whether structured or not. Networks is its own category because it spans the space between individuals and organizations. Enacted networks only really exist at an individual level, but the extent to which those individuals are connected with their organizations is key.

A key point here is that this *single set* of factors represents a range of contexts and conditions that potentially influence *implementation*, *spread* and *sustainability* of innovations. Our factor framework does not categorize the factors into different functional groups (e.g., those that contribute to and inhibit implementation, those that influence spread, and those that support or deter sustainability). This decision is intentional and is based on the fact that all of these processes are happening simultaneously to varying degrees. Some suggest that change happens in phases (Yin, 1978; Aaron, Hurlburt & Horowitz, 2011). While this may help structure our thinking about the evolution of the innovation, it also may mislead us and keep us from understanding how contexts and conditions that are brought to bear at the outset of an innovation can have implications for its sustainability later on. Still, whether one thinks about change as a process of phases or one that is more evolutionary, the fact remains that in our studies, we must use a common set of factors; a common landscape.

The Relationship Between Measuring "Innovation Implementation" and the "Implementation Process"

Even as we were clarifying what sustainability meant, we found that we weren't going to make much progress until we answered another basic question: "Sustainability of *what*?" The sources that examined implementation, spread and sustainability of innovations tended not to tackle the methodological challenge of determining what the "it" was. They focused on the factors that affected the implementation of the

innovation, but did not give attention to documenting the degree to which the innovation (the "it") was actually present.

It was in this question that we recognized the fortunate confluence of our two projects. As mentioned at the beginning of this piece, at the same time that we were funded to identify factors affecting sustainability of innovations, we also received funding to develop a set of instruments for measuring fidelity of implementation of instructional materials. This project led us through a process of creating a theoretically grounded framework for measuring implementation of educational innovations in general. We found that that framework could be applied to many kinds of educational innovations and using a component analysis approach, could support very clear and specific description of implementation.

In addition to creating the framework and measurement tools, our project highlighted differences between focusing on "fidelity" and focusing on "use." In the educational research literature, there is an increasing emphasis on the importance of using innovations shown to be effective with fidelity. (add references) However, it's clear that practitioners in all fields seldom, if ever, use an innovation exactly as intended. Further, many assert that adaptations can in fact improve the enactment of an innovation as much as harm impact. (add references) We came to see the flexibility and fluidity with which even very complex and specific innovations were implemented, sometimes with potentially better results and sometimes not.

Because we had both projects underway simultaneously, we could see a gap that needed to be closed. We saw that studies of the enactment of the **innovation** have typically resided in the realm of those examining *fidelity of implementation* or, more loosely referred to *innovation implementation*. At the same time, work examining the factors that affect *implementation* or the *implementation process* have resided with organizational and implementation studies. It was clear to us, that one conceptual framework (e.g. the framework of factors affecting implementation (the implementation process and sustainability) would not be useful without another framework supporting careful, specific description of innovation enactment. This is where our two early projects came together to help us consider ways to establish a foundation for describing, measuring and ultimately understanding the whole implementation landscape.

The framework above outlines a set of factors that contribute to and inhibit implementation, spread and sustainability of innovations. In our terms, these are factors that affect the *Implementation Process*. The implementation landscape, however, must be completed with measures of the status of the innovation itself or what is often referred to as "fidelity of implementation." We prefer to describe this as "*Innovation Implementation*" because we don't want to suggest that complete fidelity is always the ideal – even with innovations that have been established as "best practices."

Our framework for measuring innovation implementation is described at length in other sources, however, some basic principles underlie this framework that are worth articulating here: 1) *Innovations are*

developed based on an intended program model that includes elements, or "critical components," that may be either explicit or implicit. Critical components are the measurable elements of the intended program model that are essential to its implementation. Each critical component may be present to a varying extent when a program is enacted. 2) *Elements of innovations can be organized according to a set of shared categories.* It is important to systematically capture the facets of program use, such as interactions, structures, resources provided to the implementer and what the innovation participants/recipients are doing. 3) *Fidelity of implementation of all critical components is not necessarily optimal; Implementation never happens with complete fidelity, and this may be appropriate depending on contexts and conditions*. Sanetti & Kratochwill[2] suggest that "high" fidelity or "more" fidelity does not always lead to improved outcomes, and that in certain circumstances adaptations or deviations from the original model may in fact have a positive effect if they make the intervention more contextually relevant. The education world has been focused on finding innovations that "work" or as Durlak and Dupre (2008) put it, a "better mousetrap." They assert, however, that over-emphasizing the creation of the better mousetrap without fully considering how it will translate to use will not necessarily make any progress. The phrase "build a better mousetrap and the world will beat a path to your door" (Durlak and Dupre, 2008) doesn't necessarily hold true in education.

While few dispute the importance of measuring enactment of innovations, researchers continue to note the variability that comes from implementation within the same study (Durlak and Dupre, 2008.) No single intervention is always successful for changing behavior (Grimshaw et al., 2001). Rather, the key is to start with an intervention shown to contribute to desired outcomes and consider ways to adapt it while still retaining the effective elements. CEMSE's framework accounts for this point of view by having the specificity that can capture the range of ways program elements may be enacted, and by having the flexibility to account for components that may be added or modified as a program evolves.

One may wonder why the shift in this narrative from measuring the factors that influence the implementation process to measuring the status of innovation implementation. The answer lies in the fact that this distinction is seldom made. Fidelity of implementation as an area of study has its own body of literature focused specifically on what we refer to as *innovation implementation* measures – those that measure the enactment of the program of interest. Measuring the *implementation process* also has a body of literature that includes other projects that have identified sets of factors much like ours. However, there is no clarification in these works that distinguish measures of innovation enactment from measures of the factors that affect implementation. And perhaps just as important, no discussions in the literature about where the lines between the "it" (the innovation) and the factors that affect "it" are drawn. There is no single answer to this question, but there is little discussion about steps taken to identify these boundaries so that we can measure each – the innovation and the factors affecting its implementation – with confidence.

Our Proposal

"Science cannot study what it cannot measure accurately and cannot measure what it does not define." (Durlak and Dupre, 2008). Now is the time to define, measure and study. Our proposal is inspired by a recent publication by Damschroeder, et al. (2009) entitled, "Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science." In this piece, the authors suggest creating a conceptual framework that acknowledges and incorporates the work of others and provides the "taxonomy, terminology and definitions" that can be applied in multiple contexts for multiple purposes. Others (Proctor et al., 2011) have also sought to create a common language or taxonomy for measuring and sharing knowledge about implementation in the health and medical fields. This is what we are proposing for the field of education.

Damschroeder's categories align well with ours (see Table 11). The exact names of the categories themselves vary, but it's clear that they have arrived at the same categories that our work has arrived at – quite independently. They, like us, comment about the varied efforts to tackle measurement of program implementation, but lament the fact that people use different words and definitions. (Damschroder et al., 2009). This is an opportunity for the field of education to benefit from the work already done in health and medicine field and "get it right" from the start.

Table 11.
Categories of Factors

Proposed Categories for Education	Damschroeder, et al.
Characteristics of the Innovation	Intervention Characteristics
Characteristics of the User	Characteristics of the Individuals Involved
Characteristics of the Organization	Inner Setting
Characteristics of the Environment	Outer Setting
Strategies	Process of Implementation

In the time since we completed our study, we have taken note of the fact that the list that we derived from our comprehensive literature review and coding process was quite consistent with frameworks that others had developed from other reviews, personal experience, and theory (Durlak and Dupre (2008). Some works, however, tend to posit relationships between the factors. Damschroder's framework, like ours, is "atheoretical" – in other words, it does not assert relationships between the factors suggesting that much more research is necessary to create an empirical base for those factors. They explain that their framework represents a '...professional consensus within a particular scientific community. It stands for the entire constellation of beliefs, values, and techniques shared by members of that community... [and] need not specify the direction of relationships or identify critical hypotheses". (p. 3). We strongly agree. If we

communicate and coordinate our language now, we will be able to accumulate knowledge across studies to develop an empirically based understanding of the relationships between factors and innovations in the future.

We've been on a counter-trajectory with work in the health sciences. They have long worked in a tradition of establishing evidence-based practices through efficacy and effectiveness studies but over recent decades have been striving to advance understandings about how to translate those "EBPs" into use. Education, on the other hand, has focused on identifying practices that "work" with an apparent assumption that once those practices are identified people will use them. This is the mistaken assumption. This paper has articulated over 30 factors that affect individual and organizational decisions to use, spread and sustain a program or practice in education. Economists, physicians, social workers, and others have long recognized that individuals' and organizations' decisions about whether to adopt, implement, and then maintain an innovation are more likely than not to be a "lengthy process rather than a single instantaneous act" (Herie & Martin, 2002, p. 5). We need to understand the factors that cause individuals to use of a "piece" of information (Cousins & Leithwood, 1993); it is more than proven effectiveness alone.

We find the overlap between Damschroder's and other health researchers' work (Durlak and Dupre, 2008; Greenhalgh, 2004; Fixsen, et al., 2005; Grimshaw, 2001) confirming, but at the same time, recognize that educational innovations and educational settings will require their own particular adaptations. Review upon review (Greenhalgh, 2004; Fixsen, et al., 2005) arrive at very similar categories; similar lists of factors and definitions. We don't need any more reviews. Rather, we need mechanisms for bringing clarity and coherence to the factors, processes and mechanisms that have already been identified so that we can systematically measure them and begin to accumulate knowledge. We must use this opportunity to arrive at our shared taxonomy and common conceptual organizer and language – to set up a foundation for learning. Proctor, et al., (2011) stated it well, "Our purpose is to advance the clarity of language, provoke debate, and stimulate more systematic work toward the aims of advancing the conceptual, linguistic, and methodological clarity in the field."

In theory, change can happen on its own. But there is no guarantee it will come and even if it does, no assurance that it will be the change we want. We cannot be passive (Grimshaw et al., 2006; Grimshaw, et al., 2001; Greenlagh et al., 2004). Nor can we expect that simply identifying and mandating of an evidence-based practice will result in widespread, let alone sustained, implementation (Ragahavan, 2008). *This* is what the literature on implementation says. If we are to hold ourselves to the same standards we hold up to others with regard to putting rigorous research into practice, we cannot deny the imperative to make progress in studying and measuring implementation.

Proctor (2011) calls for several areas to further the research agenda including: consistency of terms, consideration of the level of analysis for which outcomes and when, development of measures, theory development, modeling relationships among factors. Further, Greenhalgh, et al. (2004) has identified a set of guidelines for the next generation of research on diffusion of health service organizations. Nearly all of them apply here. First is Greenhalgh et al.'s suggestion that research should be "theory driven" – meaning, studies should examine particular relationships between factors and program enactment. Sufficient work has been done to identify potential factors and the time has come to begin to gather empirical data. Second, they suggest that research be "process-oriented" rather that "package-oriented" – indeed, implementation studies need to do more than view innovations as "wholes." Rather, they should examine the particular features or components of programs. Greenhalgh also advocates for several other aspects of research we have sought to further in our work. More specifically, that we use common definitions, measures and tools, and do our work with, as they put it "meticulous detail."

Finally, Greenhalgh et al. (2004) and others advocate for a participatory, collaborative approach to implementation research. And that is ultimately, the purpose of this paper and our on-line open collaborative environment called "Researchers Without Borders." In the creation of this space, we have attempt to begin a dialogue and working systems that will help *us* to innovate. We hope it will help us move toward common language, shared conceptual understanding and frameworks and measures for accumulating knowledge about implementing educational innovations. By taking these steps now, we will create a foundation for researchers in the future and a basis for learning about ways to improve our schools with effective innovations that will last.

REFERENCES

Aarons, G., Hurlburt, M., & Horowitz, S. (2011). Advancing a conceptual model of evidence-based practice Implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services.* 38:4-23.

Akerlund, K. M. (2000). Prevention program sustainability: The state's perspective. *Journal of Community Psychology, 28*(3), 353–362.

Atkins, M. S., Graczyk, P. A., Frazier, S. L., & Abdul-Adil, J. (2003). Toward a new model for promoting urban children's mental health: Accessible, effective, and sustainable school-based mental health services. *School Psychology Review, 32*(4), 503–514.

Backer, T. E. & David, S. L. 1995. Synthesis of Behavioral Science Learning about Technology Transfer. National Institute on Drug Abuse. NIH Publication 95-4035. Pp. 262-289.

Berman, P. & McLaughlin, M. W. (1978). *Federal programs supporting educational change: Vol. 3. Implementing and sustaining innovations* (R-1589/8-HEW). Santa Monica, CA: RAND.

Bush, George W. (2006). American competitiveness initiative. Retrieved March 19, 2007, from http://www.whitehouse.gov/stateoftheunion/2006/aci/

Century, J. R. & Levy A. J. (2002). *Researching the sustainability of reform: Factors that contribute to or inhibit program endurance.* Newton, MA: Education Development Center, Inc.

Century, J., Rudnick, M., Freeman, C., Cassata, A. (In press). *Measuring Enactment of Innovations and the Factors that Affect Implementation and Sustainability: Moving Toward Common Language and Shared Conceptual Understanding.* Journal of Behavioral Health Services and Research.

Century, J., Rudnick, M., Freeman, C. (2011). Measuring Enactment of the Everyday Mathematics Program in Heck, D. J., Chval, K. B., Weiss, I. R., and Ziebarth, S. W. (Eds.) *Approaches to Studying the Enacted Mathematics Curriculum.* Charlotte, NC: Information Age Publishing.

Century, J., Rudnick, M., Freeman, C. (2010). *A framework for measuring fidelity of implementation: a foundation for share language and accumulation of knowledge.* American Journal of Evaluation. 31(2) 199-218.

Committee on Prospering the Global Economy of the 21st Century: An agenda for American science and technology. (2006). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington, DC: National Academy of Sciences.

Cousins, J. B. & Leithwood, K. A. (1993). Enhancing knowledge utilization as a strategy for school improvement. *Knowledge: Creation, Diffusion, Utilization, 14*(3), 305–333.

Covaleski, J. (1994). The educational system and resistance to reform: The limits of policy. *Education Policy Analysis Archives, 2*(4).

Coyne, IT. (1997). "Sampling in qualitative research: Purposeful and theoretical sampling; merging or clear boundaries." *Journal of Advanced Nursing,* 26, 623-630.

Cranney, M., Warren, E., Barton, S., Gardner, K., & Walley, T. (2001). Why do GPs not implement evidence–based guidelines? A descriptive study. *Family Practice, 18* (4), 353–355.

Datnow, A. (2001, April). *The sustainability of externally developed school reforms in changing district and state contexts.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Damanpour, F. (1991). Organizational innovation: a meta-analysis of effects of determinants and moderators. *The Academy of Management Journal*, *34*(3), 555-590.

Damschroder, L., Aron, D., Keith, R., Kirsh, S., Alexander, J., Lowery, J. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science.* 4:50.

Durlak and Dupre, 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.

Elias, M. J., Zins, J. E., Graczyk, P. A., & Weissberg, R. P. (2003). Implementation, sustainability, and scaling up of social-emotional and academic innovations in public schools. *School Psychology Review, 32*(3), 303–319.

Emory, R. (1981). *Institutionalization: How can we continue good practices and functions when funding ends? Part I: A synthesis of findings.* Paper presented at a meeting of the Northwest Regional Exchange Advisory Board, OR. (ERIC Document Reproduction Service No. ED 453 583)

Evashwick, C., & Ory, M. (2003). Organizational characteristics of successful innovative health care programs sustained over time. *Family & Community Health, 26*(3), 177–193.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (No. Louis de la Parte Florida Mental Health Publication #231). Tampa: University of South Florida.

Freeman, A. C., & Sweeney, K. (2001). Why general practitioners do not implement evidence: qualitative study. *British Medical Journal, 323* (7321), 1100–1102.

Giacquinta, J. (1973). The process of organizational change in schools. *Review of Research in Education. 1*(1973), 178-208.

Goodman, R. M., McLeroy, K. R., Steckler, A. B., & Hoyle, R. H. (1993). Development of level of institutionalization scales for health promotion programs. *Health Education Quarterly, 20*(2), 161–178.

Goodson, P., Smith, M. M., Evans, A., Meyer, B., & Gottlieb, N. H. (2001). Maintaining prevention in practice - Survival of PPIP in primary care settings. *American Journal of Preventive Medicine, 20*(3), 184–189.

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.

Grimshaw, J., Shirran, L., Thomas, R., Mowatt, G., Fraser, C., Bero, L., Grilli, R., Harvey, E., Oxman, A., O'Brien, M. (2001). Changing provider behavior: An overview of systematic reviews of interventions. *Medical Care.* 39: 8. Pp. II2-II45.

Grimshaw, J., Eccles, M., Thomas, R., MacLennan, G., Ramsay, C., Fraser, C., Vale, L. (2006). Toward evidence-based quality improvement. *Journal of General Internal Medicine.* 21:S14-20.

Haase, J., Leidy, N., Coward, D., Britt, T., & Penn, P. (1993). Simultaneous concept analysis: A strategy for developing multiple interrelated concepts. In B.L. Rodgers & K.A. Knafl (Eds.), Concept development in nursing foundations, techniques and applications (pp. 175-192). Orlando, FL: Saunders.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33*(6), 665–679.

Harvey, G. & Hurworth, R. (2006). Exploring program sustainability: Identifying factors in two educational initiatives in Victoria. *Evaluation Journal of Australiasia 6*(1):36-44

Herie, M. & Martin, G. (2002). Knowledge diffusion in social work: A new approach to bridging the gap. *Social Work, 47*(1), 85—95.

Hutchinson, J. & Huberman, M. (1993). Knowledge dissemination and use in science and mathematics education: A literature review. Retrieved 3 April, 2003 from http://nsf.gov/pubs/stis1993/nsf9375/nsf9375.txt

Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *Academy of Management Review, 21*(4), 1055–1080.

Loucks, S. & Hall, G. (1977). Assessing and facilitating the implementation of innovations: A new approach. *Educational Technology, 17*(1), 18–21.

Louis, K. S. & Jones, L. M. (2001). *Dissemination with impact: What research suggests for practice in career and technical education.* St. Paul, MN: University of Minnesota, National Research Center for Career and Technical Education.

O'Loughlin, J., Renaud, L., Richard, L., Gomez, L. S., & Paradis, G. (1998). Correlates of the sustainability of community-based heart health promotion interventions. *Preventive Medicine, 27*(5), 702–712.

Prestine, N. (2000). Disposable reform? Assessing the durability of secondary school reform. *Planning and Changing, 31*(3 & 4), 124–147.

Proctor, E., Silmere, Hiie, Raghavan, R., Hovmand, P, Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges and research agenda. *Administration Policy in Mental Health and Mental Health Services Research.* 38: 65-76.

Ragahavn, R., Bright, C., Shadoin, A. (2008). Toward a policy ecology of implementation of evidence-based practices in public mental health settings. *Implementation Science.* 3:26 p. 1-9.

Robertson, T. (1967). The process of innovation and the diffusion of innovation. *Journal of Marketing, 31*(1), 14-19.

Rodgers, B. & Knafl, K. (2000). Introduction to concept development in nursing. In B. L. Rodgers & K. A. Knafl (Eds.), *Concept development in nursing: Foundations, Techniques, and Applications*. (pp. 1-6). Philadelphia: Saunders.

Sarason, S. (1996). *Revisiting "The culture of the school and the problem of change."* NewYork: Teachers College Press.

Schatzman, Leonard and Strauss, Anselm. (1973). *Field Research: Strategies for a Natural Sociology*. Englewood Cliffs, NJ: Prentice-Hall.

Shediac-Rizkallah, M.C., & Bone, L. R. (1998). Planning for the sustainability of community-based health programs: Conceptual frameworks and future directions for research, practice and policy. *Health Education Research Theory and Practice, 13*(1), 87–108.

Strauss, A., & Corbin, J. (1990). Basics of Qualitative Research: Grounded theory procedures and techniques. Newbury Park, CA: Sage Publications.

Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform.* Cambridge, MA: Harvard University Press.

Wejnert, B. (2002). Integrating models of diffusion of innovations: A conceptual framework. *Annual Review of Sociology*. 28: 297-326.

Yin, R., Quick, S., Bateman, P. & Marks, E. (1978, March). *Changing urban bureaucracies: How new practices become routinized.* Santa Monica, CA: RAND.