CENTER FOR ELEMENTARY MATHEMATICS AND SCIENCE EDUCATION

ACCUMULATING KNOWLEDGE ON SCALING AND SUSTAINING REFORM: A FOUNDATION FOR FUTURE RESEARCH

TECHNICAL REPORT 3:
REPORT COLLECTION, CODING, AND NODEBOOK DEVELOPMENT

PROJECT OVERVIEW

The Center for Elementary Mathematics and Science Education (CEMSE) in the Physical Sciences Division of the University of Chicago is engaged in a project funded by the National Science Foundation's Research and Evaluation on Engineering and Science Education (REESE) Program. This project, "Accumulating Knowledge on Scaling and Sustaining Reform: A Foundation for Future Research," focuses on three goals:

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

The project has two strands of work: (1) literature review, analysis, and conceptual framework development; and (2) communication and dissemination.

The first strand of work has three iterative phases: Phase 1—Report Collection; Phase 2—Report Coding; and Phase 3—Report Analysis. Phase 1 entailed conducting comprehensive literature reviews on scaling and sustainability of innovations beginning with science education and then expanding to other areas of education and in turn, other fields that have developed knowledge on these topics, including business, marketing, health, and economics. Upon completion of the searches, the CEMSE team began a two-part analysis focused on identifying common themes and findings in the literature and on using a concept development process to identify the similarities and differences in the constructs that underlie the overlapping vocabulary used to describe scaling and sustainability in education and other fields. Building on this analysis, the team is creating a conceptual framework that can support the development of shared language, yet is rich enough to capture the range of ways scaling and sustainability of innovations in education have been described and understood. This conceptual framework with its clear, shared language will thus provide a foundation for accumulation of knowledge.

In Phase 1, the team conducted comprehensive searches for sources. That process is outlined in Technical Report 1. In Phase 2, the team first coded the abstracts of the sources identified in Phase 1 to determine their suitability for full coding. That is described in Technical Report 2. Then, they coded the full text of the sources that were identified for inclusion during the abstract review. That is the subject of this technical report. During Phase 3, which was concurrent and iterative with Phase 2 until Phase 2 was completed, the team has been compiling and analyzing the results of the coding to develop a summary of the literature and a conceptual framework for clearly describing scaling and sustainability of reform.

The second strand of work consists of the communication of the ongoing project work (including this technical report) and the creation of a collaborative research environment to share the literature summary and further develop the conceptual framework and other findings.

TECHNICAL REPORT INTRODUCTION

This technical report, *Report Collection and Coding*, describes Phase 2 that began in May 2008 and ended in September 2008. It includes source collection and the process of coding the sources using NVivo.

SOURCE COLLECTION PROCESS AND LIMITATIONS

Although the search procedures and criteria outlined in earlier technical reports allowed for a rigorous and thorough search of available resources, source collection was met with some technical and logistical limitations. First, our source coding process called for electronic versions of sources and some were not available electronically. Others were available electronically, but would not convert into the rich text format (.rtf) format required by NVivo, the coding softward. Of the 659 documents selected for a full read, 87 could not be downloaded or imported in the necessary formats and were not coded, leaving a total of 572 sources. The non-electronic sources can be reviewed at a later date.

REPORTS CODED

Once the sources were assembled, they were divided into groups based on their subject area search term (business, economic, educat, health, marketing, math edu, science edu). Then, the documents were organized into batches for members of the team to read and code using NVivo 7. Table 1 summarizes the numbers of sources by subject area search term.

Table 1.
Total Sources by Subject Area Search Term

Business	71
Economic	142
Educat*	354
Health	269
Marketing	24
Math* Edu*	16
Science Edu*	142
Total	572

Team members began by coding 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 edu). In addition, each team member coded a smaller number of sources from a different category. This allowed the team members to develop expertise in a particular area with regard to understanding nuances of the language and identifying relationships or references across sources. It also allowed each team member to gain exposure to other subject areas in order to better work together to identify connections across fields.

INITIAL PRACTICE CODING AND INITIAL NODEBOOK DEVELOPMENT

Prior to beginning the process of coding sets of studies, the team practiced coding with eight sources identified during the pilot collection process and determined to be of high relevance.

In order to begin the coding process, the team created a codebook that eventually came to be called the "Nodebook." The Nodebook was titled as such because NVivo refers to text codes as "nodes." The first draft of the Nodebook was created with a list of possible nodes the team anticipated using with the expectation that it 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 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 thoughts about the coding process in order to ensure shared understanding and consistency in the process.

ITERATIVE PROCESS OF SOURCE CODING AND NODEBOOK DEVELOPMENT/ADJUSTMENT

With the practice coding and initial Nodebook development complete, the team began the coding process. As the coding proceeded, the team met two times each week to share findings and emerging ideas and discuss issues that arose in the coding process.

One area of discussion focused on new nodes for the Nodebook. If a team member came across relevant text that did not fit with one of the existing nodes (e.g. a factor not previously identified) he/she would code it with an"NVivo code" that would allow it to be identified for discussion at the next meeting. At that time, the team would look at the source text to determine if it fit with an existing node, was not important (and therefore should not be coded at all), or warranted a new node. During these meetings, the team also discussed suggested revisions to the definitions of existing nodes and the potential elimination or adaptation of nodes. As the source review and Nodebook creation 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 in the first generation of collected sources. 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 references identified during the review of the 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 references turned out to be books and conferences proceedings, but the process did uncover some additional journal articles.

The nodes and their definitions are described more fully in Technical Report 5.

USING NVIVO TO CODE TREE NODES

During the coding and Nodebook development process the team noticed that some of the nodes were related to each other, and that some broader nodes encompassed more specific nodes. For example, the broad node "Characteristics of It"(characteristics of the intervention) was a broad term that encompassed some smaller, 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, which meant that any reference to adaptability of an innovation would be coded as both "adaptability" and "characteristics of it".

MEMO WRITING

During the coding process, team members occasionally identified issues or phenomena that appeared to be significant for understanding sustainability itself, or for understanding the state of the literature. When this occurred, the team member created an NVivo memo and shared it with the group. To create a memo, team members used the tree node function to create a new sub-node (titled with the name of the new memo) under the tree node "Memo" and then coded any relevant text that applied to that memo. Examples of memo topics include "use of metaphors" and "poor language use".

CATEGORIZATION OF SOURCES

After coding a document, team members created an annotation (a mechanism in NVivo) to summarize the document and then assigned it to one of the following categories: seminal, significant, contributing, of interest, limited importance, and not useful. Table 2 summarizes the distribution of these sources across the categories.

Table 2.
Sources by Usefulness

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

All team members were tasked with reading all of the articles that were deemed "seminal". In addition, one team member was tasked with reading and doing additional coding on all of the "significant" and "contributing" works.

TRACKING CODED SOURCES

The team kept track of the coding process in an excel spreadsheet. Each source was listed along with the inclusion criteria from the initial search (i.e. what the abstract was tagged with) and the key words/search terms that led to the source's initial identification. As each team member coded the sources, they updated the spreadsheet with the name of the coder, the date completed, and the source's category (e.g. seminal, significant, etc.). This proceeded until all sources were coded.