Fall 2014

The impact of National Science Foundation investments in undergraduate engineering education research: A comparative, mixed methods study

Jeremi Shavonda London
Purdue University

Recommended Citation

London, Jeremi Shavonda, "The impact of National Science Foundation investments in undergraduate engineering education research: A comparative, mixed methods study" (2014). *Open Access Dissertations*. 327.
https://docs.lib.purdue.edu/open_access_dissertations/327

Graduate School ETD Form 9
(Revised 01/14)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Jeremi Shavonda London

Entitled
The Impact of National Science Foundation Investments in Undergraduate Engineering Education Research: A Comparative, Mixed Methods Study

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Monica F Cox, Chair	Ann F. McKenna
Karl A. Smith	
Krishna Madhavan	
Robin S. Adams	

To the best of my knowledge and as understood by the student in the *Thesis/Dissertation Agreement. Publication Delay, and Certification/Disclaimer (Graduate School Form 32)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Monica Cox

Approved by: Ruth Streveler 08/29/2014

Head of the Department Graduate Program Date

THE IMPACT OF NATIONAL SCIENCE FOUNDATION INVESTMENTS IN UNDERGRADUATE ENGINEERING EDUCATION RESEARCH: A COMPARATIVE, MIXED METHODS STUDY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Jeremi Shavonda London

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2014

Purdue University

West Lafayette, Indiana

ACKNOWLEDGEMENTS

I am incredibly grateful for my faith. It is the overarching framework of my identity, and my impetus to live by principles of excellence. It is my source of encouragement when times are tough, and the affirmation that matters most when things go well.

Soli Deo Gloria. To You alone.

To the people I share my most cherished moments of life with— my parents (Ernest & Rosalind Hayes; Samuel & Mary London), husband (Ingram London), siblings (Jermaine, Erica, Ariel, Arissa, Tina, Lajuana, Shanice, and Samuel, Jr.) and extended family: Words fails to express the extent of my gratitude for your relentless love, and implicit reminders that, regardless of where my passion for research takes me, "home" is where my heart is. Similarly, my best friends are among the few who can get me to pause from "dissertating" to go get ice cream, and who effortlessly do things that cause me to laugh until it hurts. Chanee, Mia, Saade, Tiara, and Veronica: thanks for the little distractions that have helped smooth over the rough patches in my doctoral experience.

At each stage of my academic and professional journey, there has always been at least one notable person to serve as a role model and mentor; and who, over time, inevitably begin to feel like extended family. These are people who invest their most valuable asset in me to exchange experiences, offer guidance, and encourage me to "Marshal On". So, here's to you: Dr. Sharonica Hardin-Bartley (middle school English teacher), Dr. Pamela Randall-Garner (high school principal), Mr. Darryl Bennett (INROADS internship mentor), Dr.

Barrett Caldwell (I.E. undergraduate professor and Masters thesis advisor), and Dr. Don Millard (NSF internship mentor).

Many equate my productivity to that of a “well-oiled machine”; but few know that my accountability partners and colleagues are critical components of this “engine”. Special thanks to Dr. Ann-Marie Simmonds, Dr. Mary Pilotte, research team members in Purdue’s ENE and IE program, the ENE “Golden Girls”, and peer mentors in GEECS.

Three of my graduate school professors made distinct contributions to my development as a scholar. Dr. Demetra Evangelou: I appreciate you for emphasizing that reflection –although it may not feel like work— is a vital part of conducting research. Because of your instruction and your example, I set aside time regularly to think deeply about my research and to ponder other big ideas. Dr. Ruth Streveler: In the first class of my graduate education, you succinctly explained, “research is about claims and evidence”. This framework continues to help me make sense of what it means to be a consumer and now, a producer of knowledge. Dr. Hong Wan: Thank you for making the peskiest engineering techniques fun to learn, and for stressing the importance of not only knowing how to generate statistics, but also how to interpret and present the results in meaningful ways.

To my dissertation committee, Drs. Monica Cox, Ann McKenna, Karl Smith, Krishna Madhavan, and Robin Adams: You are the leading scholars in my field who have given this study countless hours of undivided attention, and have helped me navigate some of the murkiest waters associated with conducting research and interpersonal relationships. Finally, special thanks to Dr. Edith Gummer who offered sound research advice at a critical time in the conceptualization of this study. I appreciate each of you.

This study was partially supported by NSF Grant# 1123340.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

London, Jeremi S. Ph.D., Purdue University, December 2014. The Impact of National Science Foundation Investments in Undergraduate Engineering Education Research: A Comparative, Mixed Methods Study. Major Professor: Monica F. Cox.

The U.S. invests billions of taxpayers' dollars in research tied to the national priorities that contribute to its competitiveness in a global economy. As the federal funding agency with an explicit focus on engineering education, the National Science Foundation (NSF) contains a portfolio of projects focused on improving the quantity of engineering graduates and the quality of engineering programs. Within the agency, the Division of Undergraduate Education invests approximately $190 million (FY 2012) annually on science, technology, engineering and mathematics (STEM) education projects. Although the DUE portfolio includes a suite a projects with different foci supporting national initiatives and Principal Investigators (PIs) report their results in annual reports and conferences, there is little consistency on how impact is defined, evaluated, and measured.

While many agree on the importance of investing in research, the stiff economic climate necessitates that the research that demonstrates impact is what will continue to be supported. However, the dearth of scholarship on impact contributes to the lack of understanding around this topic. This study links the fragmented literature on impact to form a unified starting point for continuing the conversation. While existing literature includes three dimensions of research impact (i.e., scientific, societal, and domain-specific impact), this study focuses on the domain-specific impacts of engineering education research using two guiding frameworks, namely, Toulmin's Model (1958) and the Common Guidelines for Education Research and Development (Earle et al., 2013), and a multiphase mixed methods research design (Creswell & Plano Clark, 2011).

The qualitative phase of this study explores how researchers on NSF-funded STEM education R&D projects talk about the impact of their work; the findings reveal eight themes that are commonly discussed when PIs articulate the impact of their research, and two themes related to how they support their claims. The findings also indicate that the STEM discipline associated with the study and the project focus have more to do with the types of impact PIs claim than the amount of funding awarded to the project. As a result of identifying the points of alignment between PIs' perspectives on impact and existing literature, a preliminary description of what impact looks like in this context is proposed—using the three dimensions of research impact as an organizing framework. Although this study puts forth a preliminary description of the impact of STEM education research, extensions of this work are necessary before providing practitioners and policymakers with a valid, comprehensive framework characterizing what impact means in this context.

Ideas supporting the types of claims PIs make when discussing the impact of their work were used to develop a survey that was distributed to a small sample of current and former NSF Program Officers (POs) in the second phase of this study. The survey results provide preliminary evidence on how PIs and NSF PO' perspectives on research impact compare, and affirm that additional studies are needed. Consequently, implications for policy and practice and potential research directions are also discussed.

CHAPTER 1: INTRODUCTION

In 2013, the U.S. government invested $140 billion in basic and applied research (Sargent Jr, 2013). Federal investments in research support national security, public health, technological innovation, global economic competitiveness, and workforce development (National Academies Press, 2011). The most targeted investments in workforce development are in federal programs designed to increase knowledge in the science, technology, engineering and mathematics (STEM) fields and the attainment of STEM undergraduate and graduate degrees (Government Accountability Office, 2012). In 2010, thirteen government agencies administered over $3 billion to support more than 200 STEM education programs across educational levels and in informal learning environments (GAO, 2010). Of the thirteen agencies, the National Science Foundation (NSF) received the most funding, $1.1 billion, and administered 37 programs, the second largest number of STEM education programs within a single agency (GAO, 2010, p. 10).

NSF is an independent federal agency created by Congress in 1950 whose overarching mission is “to promote the progress of science” in the U.S. (NSF). Apart from the Department of Education, NSF is the only federal agency which --as part of its mission--supports science and engineering education projects from kindergarten to graduate education and beyond (GAO, 2012; NSF). Although NSF has a relatively small budget in comparison to other federal funding agencies, it is the funding source for approximately

24% of all federally supported research projects conducted in America's colleges and universities (NSF).

To understand the significance of this study, it is necessary to describe the general process associated with supporting a NSF-funded research project. See Figure 1.

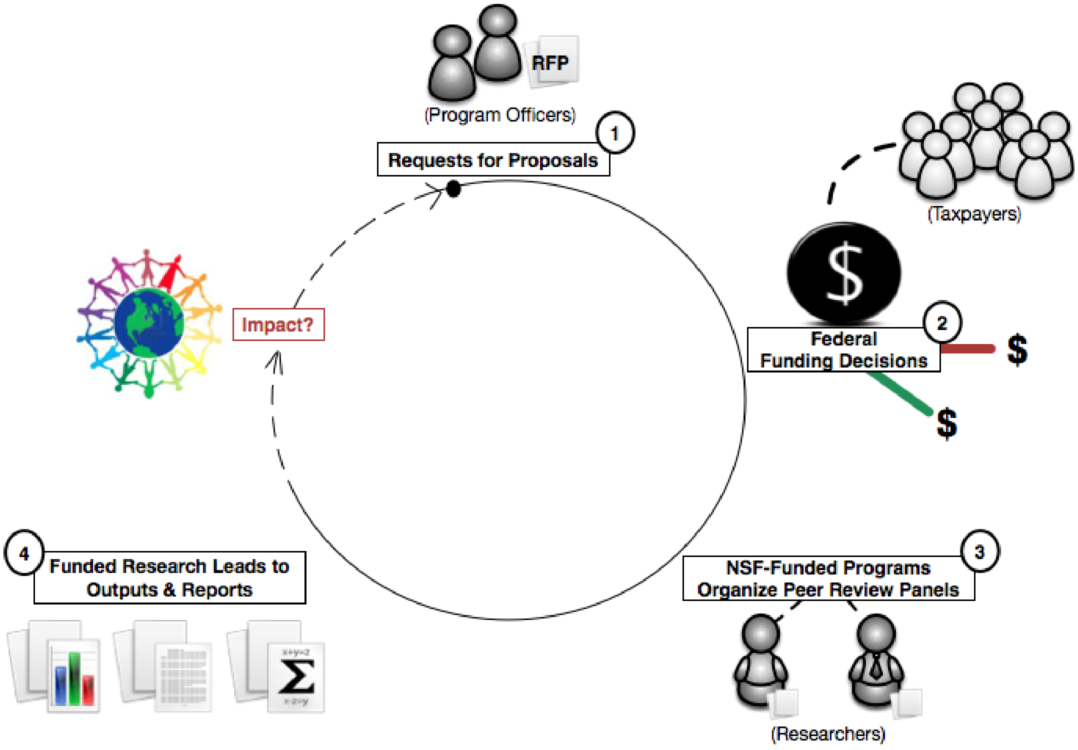


Figure 1: General Process for Supporting a NSF-Funded Research Project

This process formally begins when NSF program officers (POs) draft a Request for Proposals (RFP), which is a description of a program that could potentially address an issue of national importance. The RFP includes program goals and objectives, a logic model (indicating inputs, activities, outputs and expected outcomes) and a budget request. The program description is sent to government officials who review all requests for federal funds from all federal agencies. Based on national-level priorities, government officials make decisions about what should or should not be funded. This is an important step in the process since these programs are supported by taxpayers' contributions.

Assuming the proposed NSF program is funded, NSF publicizes the RFP as an invitation for researchers across the country to submit grant proposals in response to the elements of the program description. Once proposals have been received, NSF organizes panels of reviewers who evaluate the content and quality of the research proposal based on two general NSF criteria (i.e., intellectual merit and broader impacts) and alignment with programmatic goals. POs then use the panel evaluations as apart of the decision making process to determine whether or not to fund the proposed research. If a researcher is granted funding, the proposed study may begin on the date approved by the PO.

Each year, researchers conducting NSF-funded projects (also known as Principal Investigators or PIs) submit an annual report summarizing the progress and the outcomes of their study. At the end of the grant cycle, PIs submit a final report; this is the final step in this process. Under ideal circumstances, the next step would include an evaluation to determine if the aggregate project outcomes align with the expected outputs and outcomes in the program description. This is a vital step for determining the impact of federal investments in research; however, this is missing from the current process. This gap is the motivation for this study.

The NSF-funded projects of interest in this study are undergraduate engineering education projects, which are primarily funded by NSF's Division of Undergraduate Education (DUE). DUE has an annual budget of approximately $190 million (FY 2012) (Education and Human Resources Directorate, 2013) to fund projects that advance its mission: "to promote excellence in undergraduate STEM education for all students" (NSF).

DUE's programmatic activities represent a comprehensive approach to "strengthening STEM education at two- and four-year colleges and universities by improving curricula, instruction, laboratories, infrastructure, assessment, diversity of students and faculty, and collaborations" (NSF). DUE funds nine general areas of STEM research: biological sciences, chemistry, computer science, engineering, geological sciences, interdisciplinary, research and assessment, and social/behavioral sciences. Based on FY2011 – FY2013 data for one of DUE's largest core research and development (R&D) programs (i.e. "Transforming Undergraduate Education in STEM" or "TUES"), engineering received 25% of the awards and 18.4% of the funding (NSF, 2013b). In both cases, this represents the largest allocation across all disciplinary areas. This is one reason for focusing on engineering education research in this study; additional reasons are provided in the rest of this chapter. The remaining sections of this chapter elaborate on each step of the process depicted in Figure 1 in light of the focus on federal investments in undergraduate engineering education.

Engineering Education: A National Priority

With the increase in global challenges surrounding renewable energy, resilient infrastructure, and reliable healthcare, an engineer's role in society cannot be overstated (National Academy of Engineering, 2008b). Given that these problems are not confined to any single geographic region or population, finding solutions to them will have local, national, and global implications. The United Nation's UNESCO Engineering Initiative (2012) is one indicator of the interest among leaders worldwide in the training and success of the next generation of engineers who can tackle these issues. U.S. leaders concerned about the nation's global competitiveness are particularly interested in the education of engineers in the U.S. since "engineering impacts the health and vitality of a nation as no other profession does" (NAE, 2004, p. 37).

The surge of recent publications produced by the National Academy of Engineering (2004, 2005a, 2008a, 2010), the National Research Council (1999b, 2003, 2012), and the President's Council of Advisors on Science and Technology (2010, 2012) speak to the increasing interest in engineering education among U.S. policy makers at the highest levels of government. The bachelor's degree in engineering is the first professional degree for practicing engineers (NAE, 2004) unlike in other professions –such as medicine and law– where the first professional degree is oftentimes earned after completing a non-specialist (liberal arts or science) degree. According to the NSF/National Center for Science and Engineering Statistics (2008), nearly half of all scientists and engineers (49%) employed in industry, government, nonprofit sectors or are self-employed have a bachelor's degree as their highest degree. Therefore, for most practicing engineers the undergraduate engineering degree is the only one they will pursue. In light of this, it is important to ensure that undergraduate engineering programs are equipping graduates with the competencies they will need to effectively respond to and influence society's greatest technological and social needs.

The quality of engineering education in the U.S. is a national priority. It is difficult to articulate all of the quality issues, but Froyd, Wankat, and Smith (2012) discuss five major shifts in engineering education that have occurred over the last 100 years. These shifts not only highlight the major trends in this field, but also underscore many of the most pressing issues facing engineering educators and engineering education researchers. The five shifts relate to: 1) what and how engineering content should be taught; 2) changes in accreditation processes that emphasize outcomes-based assessments; 3) an emphasis on the engineering design process as a critical part of engineering education; 4) applications of learning science and social-behavioral research to educate engineers; and 5) use of the latest technology in engineering education. Collectively, important issues such as these affect the overall quality of undergraduate engineering education.

For the past decade, however, the focus has not only been on improving the quality of engineering education, but increasing the number of engineering graduates as well (Jackson, 2002; NAE, 2002; 2004, 2005a, 2010; NRC, 2007; NSTC, 2013; PCAST, 2010; 2012). NSF (2013c) national data reports that less than 5% of all bachelor's degrees awarded in the U.S. between 2001 and 2010 were in engineering disciplines. Some concerns about a shortage of engineers are motivated by the aging engineering workforce and fear that the statistics depicting the engineering pipeline suggest that there will not be a sufficient number of engineers to satisfy impending workforce demands (IBM, 2003; NRC, 2007). For many, the shortage of engineers poses a threat to the U.S.' global competiveness, especially as the number of engineers in competing countries –particularly China and India– continues to rapidly increase (Wadhwa, Gereffi, Rissing, & Ong, 2007). Many reports conclude that this shortage of engineers could have serious ramifications for the U.S. if the trend is not reversed (NAE, 2010; NRC, 2007; The National Commission on Mathematics and Science Teaching for the 21st Century, 2000).

On the other hand, there is an opposing view surrounding the "shortage" of engineers in the nation's workforce (Butz et al., 2003; Charette, 2013; Guess, 2008; Kennedy, Taylor, Urquhart, & Austin, 2004; Lowell & Salzman, 2008; Teitelbaum, 2004, 2007). Researchers that oppose the argument of a shortage of engineers in the U.S. recommend a shift from a focus on the perceived shortage to either: (1) improving our understanding of engineering workforce demands (Butz et al., 2003; Teitelbaum, 2007; Wadhwa, 2007), (2) increasing engagement in global collaborations to address technological needs (Lynn & Salzman, 2006, 2009); (3) conducting deeper analysis of the data to evaluate shortages in the context of discipline-specific needs or underrepresented groups (Lowell & Salzman, 2008; Salzman & Lowell, 2008; Wadhwa, 2011); or (4) improving the *type* of engineering graduate (Salzman & Lynn, 2010). More recently, however, the cry for more engineers and conversations about "shortages" has begun to resound beyond U.S. borders (Blau, 2011; Fackler, 2008; Richardson, 2012; Silverman, 2010; Srinivas, 2013; UK Parliament, 2012). Thus, even in other nations that are perceived as having engineering prominence, policy makers have come to a similar conclusion that there is a need for more engineers.

Research, the Path to Achieving National Priorities

There have been calls for more and better engineers for decades. However, "[b]usiness, academic, and government leaders across the engineering enterprise have repeatedly remarked that *systematic research* of how we educate engineers *must be the path* by which we transition from episodic cycles of educational reforms and move to continuous, long-lasting improvements in our education system" ("The Research Agenda for the New Discipline of Engineering Education," 2006, pp. 259, ephasis added).

Two articles published in consecutive years of the premier scholarly journal for the field of engineering education (Haghighi, 2005; "The Research Agenda for the New Discipline of Engineering Education," 2006) indicate that the turn of the 21st century was a pivotal time for this area of research. In the first of the two publications, "Quiet No Longer:

Birth of a New Discipline", Haghighi (2005) writes a short, galvanizing piece about what the academic discipline of engineering education needs in light of the national calls for more and better engineers. A coherent research agenda is first and foremost among the short list of needs. The second of the two publications builds directly on the first by providing that research agenda; it includes five focal areas for engineering education research ("The Research Agenda for the New Discipline of Engineering Education," 2006, pp. 259-261):

> **Area 1–Engineering Epistemologies:** *Research on what constitutes engineering thinking and knowledge within social contexts now and into the future.*
>
> **Area 2—Engineering Learning Mechanisms:** *Research on engineering learners' developing knowledge and competencies in contexts.*
>
> **Area 3—Engineering Learning Systems:** *Research on the institutional culture, institutional infrastructure, and epistemology of engineering educators.*
>
> **Areas 4—Engineering Diversity and Inclusiveness:** *Research on how diverse human talents contribute to solutions to the social and global challenges and relevance of our profession.*
>
> **Areas 5—Engineering Assessment:** *Research on, and the development of, assessment methods, instruments, and metrics to inform engineering education practice and learning.*

Because there are many points of overlap between the five research areas and the two overarching national priorities surrounding engineering education (i.e., to increase the number of engineering graduates and improve the quality of their education), this research agenda represents how the engineering education research community conceptualizes one path from the status quo to addressing these national priorities (Haghighi, 2005). Thus, the research agenda also hints at how engineering education researchers might utilize federal investments in this area of study. Because of the size of federal investments in engineering education research and what is at stake, taxpayers and stakeholders at the various levels want to know the extent to which federal investments are leading to desired outcomes; they are interested in understanding the impact of these investments (National Science and Technology Council, 2008a, 2008b).

National Interest in Understanding Research Impact

Recently, it has become imperative for federal agencies to prioritize initiatives that lead to insights about the promise of proposed research and the impact of research that has been conducted. This has not always been the case, however. *Science, the Endless Frontier* (Bush, 1945), a report given to President Franklin D. Roosevelt by his science advisor, put forth the rationale that research should be viewed as inherently valuable because of the importance of the knowledge that results from it-- even if the advances in knowledge could not be applied immediately. But this perspective is not widely shared among policymakers and the public today (Bornmann, 2012; Doz, Santos, & Williamson, 2001; NSTC, 2008a; 2008b).

There are two main reasons for this new focus. One reason is the push around the world for better infrastructure that supports better practices in: decision-making, management of R&D portfolios, documentation of the impact of investments in research, and the grounding of future research programming plans in compelling evidence (Bornmann & Marx, 2014; Fealing, Lane, Marburger III, & Shipp, 2011; NSTC, 2008b). In the U.S., this push starts at the White House. In correspondence sent to leaders of federal agencies preparing FY2011 budgets, they were asked to: "develop 'science of science policy' tools that can improve management of their research and development portfolios and better assess the impact of their science and technology investments" ("Memorandum for the Heads and Executive Departments and Agencies" as cited by Fealing, Lane, Marbuger III, & Shipp, 2011, p. 2). This further reinforces the shift in perspectives that prioritize activities that lead to insights on impact.

There is another reason why there is a shift in focus to the impact of research. Given the current economic climate, there is an even greater need to allocate federal resources more wisely. Moreover, there are conflicting funding priorities among those who have a stake in publicly supported research: on one hand, there are a number of social and technological problems that require the expertise of researchers (Bush, 1945; *FIRST Act of 2014* 2014; National Academy of Engineering, 2005b), but as Bornmann (2012) says, "…the growth of scientific research during the past decade has outpaced the public resources available to fund it" (p. 637). This makes for a very competitive environment for agencies and researchers requesting funds. As a result, those making request for funds will need to make strong cases that defend the link between federal investments in R&D and national priorities. The research that will continue to be federally supported is research that demonstrates impact.

Although all researchers are being asked to demonstrate impact, education researchers face unique pressures to demonstrate the impact of federal investments in research. The National Research Council report (2002), *Scientific Research in Education*, cites four key problems that underlie the lack of public support for education research: problems associated with research quality, fragmentation of research efforts, oversimplification of the role of research in education reform, and the longstanding disconnect between education scholarship and practice (Lagerman, 2000; Kaestle, 1993; Sroufe, 1997; Levin and O'Donnell, 1999; NRC, 1992; NRC, 2001d as cited by NRC, 2002). (The latter stems from a long history of education researchers and practitioners operating in different spheres, and rarely exchanging insights to inform one another's work (Lagerman, 2000, Mitchell and Haro, 1999, as cited by NRC, 2002)). In addition to these external pressures, STEM education researchers expect their colleagues to demonstrate the impact of their work so that the findings can inform decisions to pursue activities that promote undergraduate STEM education and abstain from those that do not (Miller & Pasley, 2012; National Science and Technology Council, 2008a).

Overseeing NSF's Research Portfolio

Because NSF is the federal funding agency with an explicit focus on supporting engineering education and makes the largest allocations in their area, the agency's practices for overseeing its research portfolio will be discussed in this section. NSF programs are held accountable for the program-level outcomes of their investments and hold PIs accountable for the outcomes of their funded projects. DUE (and NSF, in general) uses various processes to facilitate this- (1) review criteria, (2) annual reports, (3) PIs conferences, and (4) program evaluations. The basis of proposal evaluations primarily rests on two general NSF criteria (i.e., intellectual merit and broader impact) (National Science Board, 2011) and the alignment of the proposed study with programmatic goals. If awarded the funding, PIs must submit annual and final reports that provide updates on the status of completion and outcomes of the proposed activities. The reports are submitted to a repository and are only viewable by NSF employees, and thus not available to the public for review or critique. Grantees on active grants are expected to participate in periodic (e.g., bi-annual) PI conferences where they interact with other PIs and share updates on their work. Abstracts of the funded works are later compiled into a report and sometimes made publicly available online (e.g., NSF (2013a)). These three reporting mechanisms are useful for accountability and for allowing researchers to document individual project outcomes.

Although NSF has well-established criteria for selecting which proposals to fund and standard processes for sharing project-related updates during the grant lifecycle, a consistent way to determine the extent to which federally funded projects are making a difference is lacking (Allen et al., 2008; National Science and Technology Council, 2008a, 2008b; U.S. Department of Education, 2007). The current reporting mechanisms capture individual project outcomes but do not facilitate comparisons of impact across projects or over time. Given the vast amount of information contained in the project reports, it is also difficult to quickly aggregate projects results, determine program-level impacts, and make funding decisions based on past results. Oftentimes, these insights are garnered via the fourth accountability measure mentioned above, program evaluations (e.g., (Eiseman, Fairweather, Rosenblum, & Britton, 1996). However, such program evaluations necessitate extensive time, resources, and expertise (Rossi, Lipsey, & Freeman, 2004) and lead to reports that are oftentimes not circulated broadly enough for the wider research community to make use of the findings.

Lane and Bertuzzi (2011) sum up the problem this way: “The current scientific data infrastructure is based on identifying, funding, and managing high-quality science, not on understanding its impact” (p. 678). Without a way to characterize and ultimately evaluate the impact of NSF-funded research and education projects, it will continue to be difficult to determine the extent to which NSF's investments in undergraduate engineering education are affecting the quality of engineering education or the quantity of engineering graduates in the U.S. This problem is the motivation for this study.

Statement of Research Motivation & Significance

In short, the U.S. invests millions of taxpayers' funds in undergraduate engineering education research via the National Science Foundation, hoping that these investments will lead to improvements in the quality of engineering education and an increase in the number of engineers. NSF uses various mechanisms to oversee its research portfolio, but the current documentation processes are not sufficient for providing the meaningfully insights about the impact of NSF investments in research, particularly to those outside of the agency. The need for better decision-making in this stiff economic climate necessitates that the research that demonstrates impact is what will continue to be federally supported. Thus, exploring what impact looks like in this context would be valuable to policymakers, researchers, and practitioners alike.

In a paper on frameworks and review articles Schwarz, Mehta, Johnson, and Chin (2007) define a framework as the "exposition of a set of assumptions, concepts, values, and practices that constitutes a way of understanding the research within a body of knowledge" (p. 41). Over the last decade, frameworks have been developed to characterize the impact of research in domains such as health science research (Donovan & Hanney, 2011; Kuruvilla, Mays, Pleasant, & Walt, 2006) and arts & humanities research (Levitt et al., 2010). These frameworks help provide a shared language and understanding of impact as researchers communicate among themselves and share impact insights with those outside the community. However, within the context of engineering education, there is no shared vocabulary for discussing the impact of research, or a framework that characterizes the impact of federal investments in undergraduate engineering education research. This study seeks to begin to fill this gap in the literature. More specifically, this study begins to add to the body of knowledge by exploring how researchers on NSF-funded undergraduate STEM education R&D projects talk about the impact of their work, and how this compares with Program Officers' perspectives on impact. An understanding of how these two stakeholders think about impact is an appropriate first step toward the development of a conceptual framework that characterizes the impact of NSF investments in undergraduate engineering education R&D projects.

Organization of this Dissertation

Chapter 1 of this dissertation included the background of this study. Chapter 2 provides a synthesis of the bodies of literature that support this work. Chapter 3 includes the description of the frameworks guiding the data collection and analysis. Chapter 4 describes an overview of the methodology associated with this two-phased mixed methods study. Chapters 5 and 6 include the methods and results corresponding to the first phase of this study, respectively. The next two chapters include the methods and results corresponding to the second phase of this study, respectively. The last two chapters, Chapters 9 and 10, include a discussion about the results and their implications, and the conclusion.

CHAPTER 2: REVIEW OF RELEVANT LITERATURE

This chapter includes a review of the literature that supports this study on the impact of federal investments in undergraduate engineering education research. It is important to acknowledge that there have been other reviews of the literature on the topic of research impact (Bornmann, 2013; Jonathan Grant, Brutscher, Kirk, Butler, & Wooding, 2010; Salter & Martin, 2001; Walter, Nutley, & Davies, 2003), and there are areas of overlap in this chapter. Although good contributions to an area that is sorely lacking, these reviews have some deficiencies. For example, the definitions of research impact mentioned in existing literature do not capture the emerging dimension of impact that corresponds to characterizing research in a particular domain; and the recent U.S.-based initiatives to study this topic more aggressively are absent from the existing reviews. Additionally, the series of difficulties associated with studying the impact of research and the limited number of existing research impact frameworks are sprinkled across co-existing, disconnected bodies of literature, and the connections between them have not been articulated. The most significant advances in this field of study on impact will result from weaving the disconnected bodies of work together into a more comprehensive review of the literature on this topic.

This chapter has four main sections. The first section, "Describing Research Impact", includes a brief discussion of research and impact, in general, and concludes with a definition of research impact. The second, "Difficulties with Studying Research Impact", summarizes all of the difficulties in the literature into three major categories. The third section, “Existing Studies on Research Impact”, is a synthesis of the collections of literature that are most relevant to studying the impact of federal investments in undergraduate engineering education research. The last section is a brief acknowledgement of some of the STEM education research that is tangentially related to the topic of impact.

Defining “Research Impact”

“At its core, scientific inquiry is the same in all fields. Scientific research, whether in education, physics, anthropology, molecular biology, or economics, is a continual process of rigorous reasoning supported by a dynamic interplay among methods, theories, and findings” (National Research Council, 2002, p. 2). Six guiding principles that undergird all scientific inquiry are: 1) pose significant questions that can be investigated empirically; 2) link research to relevant theory; 3) use methods that permit direct investigation of the question; 4) provide a coherent and explicit chain of reasoning; 5) replicate and generalize across studies; and 6) disclose research to encourage professional scrutiny and critique (National Research Council, 2002). The disclosure of research is what allows researchers to formally join a conversation with other scholars who care about related topics, and is a prerequisite for “improv[ing] if not the whole world, at least [their] corner of it” (Booth, Colomb, & Williams, 2008, p. 11). An improvement in a corner of the world as a result of research is one way of conceptualizing research impact.

One problem with studying impact –in a research context or otherwise— is that there is no definitive meaning of the term and is oftentimes used interchangeably with other terms (e.g., outputs, third steam activities) (Brewer, 2011; Martin, 2007). One of the broadest definitions of impact in the literature is proposed by Halse and Mowbray (2011) who intended to capture a more "conventional meaning" of impact by defining it as "an affect that is a consequence or result of a particular process, event, action or phenomenon" (p. 51). However, Brewer (2011) describes impact as a proverbial terrain traversed by three groups of people: (1) those in the policy evaluation tradition (who often use impact to denote the involvement of 'users' as a critical part of the evaluation process); (2) those who are part of the audit culture in higher education (who use it as a mechanism for responding to questions surrounding accountability to the public); and (3) those who are part of the philosophy and sociology of knowledge tradition (who are concerned with the social production of knowledge). Each group's definitions of impact are often influenced by the motivations and priorities of key factions, which is why impact varies according to these three principal groups.

Although Brewer (2011) states that "there is no common ground between these extremes and no shared vocabulary to facilitate a universal conversation" (p. 255), STEM education researchers Lande, Adams, Chen, Currano, and Leifer (2007) coined the term "Scholarship of Impact" and developed a framework that may begin to constitute a connection across the impact terrain. In it, they define impact as "the measurement or evidence of change" (Lande et al., 2007, p. 2) while elaborating on the need for "impact studies", research which seeks to have greater synergy by essentially closing the gap between research and practical problems" (Lande et al., 2007, p. 2). According to this framework, one of the features of a good impact study is that researchers (in their case, engineering education researchers) are intentional about meeting the needs of the intended 'users' (e.g., educators) by including them in the early stages of the study, and not simply at the end when it is time to share the research findings. The idea of involving users in the process of conducting research not only aligns with those in the policy evaluation tradition, but also those interested in the social production of knowledge. While the motivation for this study (See Chapter 1) is related to perspectives of those in the policy evaluation tradition and the audit culture of higher education, the focus is on characterizing research impact in a particular domain and, as a result, most closely aligns with the philosophy and sociology of knowledge tradition.

Within this community of scholars, there are two competing definitions of what constitutes research impact (Donovan, 2011). For the purposes of this study, these two facets of research impact will be referred to as "scientific impact" and "societal impact". *Scientific impact* refers to advances in reliable knowledge (theories, methodologies, models, and facts) that primarily influence academic communities (Bornmann, 2013; Bornmann & Marx, 2014; Donovan, 2011; Godin & Doré, 2005). *Societal impact,* on the other hand, is broadly conceived as research results that influence social, cultural, environmental/natural, or economic capital of a nation (Bornmann, 2013; Bornmann & Marx, 2014; Donovan, 2011). Examples of this might include: stimulating new approaches to social issues; informing policy; improving our understanding of how we relate to one another's society and culture; reducing waste and pollution; and increasing productivity (Bornmann, 2013).

There is a general consensus around various aspects of scientific impact (i.e., what it is, how it is measured, who benefits, etc.) (Spaapen & van Drooge, 2011), and this understanding is fairly consistent across research domains. The ways in which it is measured is oftentimes related to the publications resulting from a study; such measurements are general practice across disciplines, and relate to standards of scientific rigor. Existing literature on the scientific impact of research is advanced—as indicated by scholarly journals, annual research conferences, and even some specialists (e.g., bibliometricians) (Bornmann, 2013). This is not the case for the societal impact dimension of research impact. Societal impact is less clear, harder to measure, and looks different across domains (Bornmann, 2013; Holbrook & Frodeman, 2011; Lane, 2009). For example, the way a set of engineering research results ultimately influence the environment will differ from how the results of a biological sciences study may– but both types of research results align with the environmental-facet of societal impact. There are no scholarly journals or research conferences dedicated to societal impact, and insights on this topic are in disconnected bodies of literature associated with various disciplines.

Bornmann states that there is "no direct link between the scientific quality of a research project and its societal value" (Bornmann, 2012, p. 673); said differently, there is not direct connection between the scientific and societal impact of a research project. Part of the reason for this may be because these two facets of research impact, though relatively comprehensive, do not capture the type of impact that is emerging in the literature. More recently, there has been an increase in the number of frameworks developed to characterize what this study calls *domain-specific impact*. Unlike scientific impact, which has little regard for the content of the research itself, and societal impact, which is too abstract to give the finest resolution to impact that occurs before it is realized at the societal level, the domain-specific dimension of research impact primarily focuses on what impact looks like in context.

Domain specific-impact is the influence of the methods or results of a R&D project on the people, priorities, and/or processes in the context of interest. Research impact frameworks have been developed for characterizing the impact of research in fields such as medicine (Donovan & Hanney, 2011; Kuruvilla et al., 2006) and arts & humanities research (Levitt et al., 2010). Some of the impact descriptors in the frameworks are not unique to the context of interest and align with scientific and societal impact (e.g., research impacts: publications and papers in Kuruvilla et al. (2006) and "impacts on policy" in Levitt et al. (2010)). However, impact descriptors like "health literacy" and "quality of care" are unique to health science research (Kuruvilla et al., 2006), just as "preservation of heritage" and "leisure and entertainment" (Levitt et al., 2010) are forms of impact that are more likely to be found among arts & humanities research and not in research, in general. (The literature review section entitled *Disciplinary Perspectives on Research Impact* provides the constructs of the frameworks that have been developed to explore research impact in specific disciplines.)

In short, research impact has three dimensions: scientific impact, societal, and domain-specific; Figure 2 presents the definition corresponding to each dimension for ease of reference.

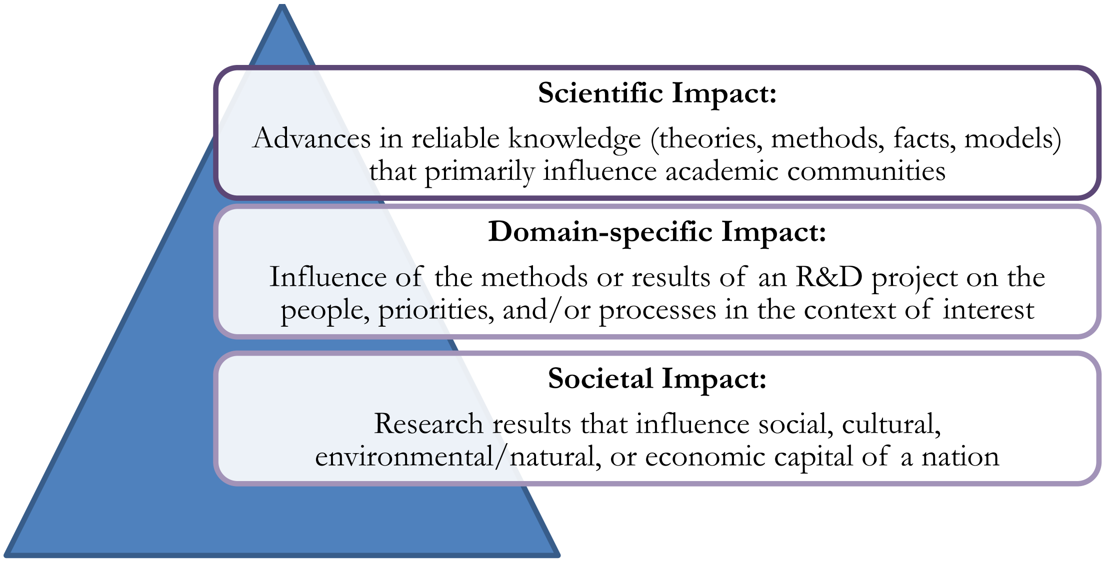


Figure 2. Defining the Three Dimensions of Research Impact

Although defined as three distinct dimensions, they do overlap: in many ways, one dimension enables another. Furthermore, the three dimensions can be understood together in terms of their order of impact. Societal impact –advances in reliable knowledge— is an example of a first order impact because the extent of the results is limited to conversations taking place among scholars with similar interests. Once the execution of the methods or the outcomes of a study begin to influence the context of interest –domain-specific impact—, the impact of the study is extending beyond conversations among scholars; this is an example of a second order impact. Finally, the aggregate influence of a particular line of research on national priorities –societal impact— is a third-order impact.

All three dimensions of research impact are important for studying federally funded studies, and are included in what is meant by the use of the term "impact" and "research impact" in this study. While the results of this study may include some insights on the scientific aspect of research impact in the context of interests (the area that has been studied the most extensively), its primary focus is on the societal and domain-specific impact of undergraduate engineering education research.

Difficulties with Studying Research Impact

Existing literature on research impact includes a myriad of difficulties associated with studying the impact of research. However, a synthesis of this literature has resulted in three categorizations across three headings: difficulties associated with connecting impact with research or the researcher; difficulties associated with assessment and evaluation; and difficulties associated with interpretations of impact. The following sections explore each of these headings in more detail.

Difficulties Associated with Connecting Impact with Research or the Researcher

The *attribution problem* is one of the most commonly cited reasons why studying impact is so difficult (Bornmann, 2013; Godin & Doré, 2005; Jonathan Grant et al., 2010; Martin, 2007; Rymer, 2011; J. E. Scott, Blasinsky, Dufour, Mandai, & Philogene, 2011; Spaapen & van Drooge, 2011). This is the difficulty with attributing impact to particular research projects or other inputs; this problem is also referred to as *impact accretion*. The reasons it is so difficult to make attributions is because impact diffuses through time and space, and all research builds on earlier research. Moreover, as research and development becomes more global, it is nearly impossible to make attributions to a particular research project or researcher; this is called the *internationality problem* (Bornmann, 2013; Martin, 2007). A similar challenge is referred to as the *causality problem*: the difficulty with tying impacts to causes (Bornmann, 2013; Martin, 2007). Additionally, the impact of research oftentimes depends on people outside of the research system (e.g., others who make intellectual and financial investments) (Rymer, 2011). Together, these issues make it difficult to connect the impact of research with a particular research project or researcher.

Difficulties Associated with Assessment and Evaluation of Impact

The difficulties associated with the assessment and evaluation of the societal and domain-specific impact of research relates to what should be assessed and how; when the evaluation should take place and who is qualified to conduct it; and unintended consequences of assessment and evaluation. One of the major issues with assessing impact starts with data. Unlike the data available for measuring the scientific impact of research, there is a lack of data on the societal impact of research (Spaapen & van Drooge, 2011); this is also true for domain-specific impact. The place for collecting data on impact is somewhat illusive—where one looks to observe it is not always apparent. Furthermore, the data this is available in dispersed across federal agencies and research institutions, and is not formatted consistently (Lane & Bertuzzi, 2011). Additionally, the current data infrastructure does not allow one to easily track connections between research and societal outcomes and are inadequate for decision-making (Fealing, Lane, Marburger III, et al., 2011; Lane & Bertuzzi, 2011; NSTC, 2008b). As it relates to how impact should be assessed, there are limits on the extent to which the impact of research can be quantified, and quantifying the research outcomes is not easy (Lane, 2009). Linear assessment models assume that the outputs of research are always a codified form of new scientific knowledge; however, this approach ignores knowledge that cannot be codified –for example, tacit knowledge that exists among trained people— but is just as important (Martin, 2007). Martin (2007) justifiably argues that there are "no perfect measures [of impact], only partial and imperfect indicators" (p. 10).

There are two difficulties associated with the timing of the assessment of impact. The *evaluation timescale problem* states that the timing of the evaluation will affect the impacts that are observed (Bornmann, 2013; Martin, 2007). This issue is particularly important in this context given that stakeholders plan to use the insights from impact research to inform decision-making— decisions will be made based on the information available at the time, not on what may happen in the future. Another time-related issue is the *temporality problem.* This is the time span between research and its embodiment in products, processes or social practices (Lane & Bertuzzi, 2011; J. E. Scott et al., 2011; Spaapen & van Drooge, 2011). "The time between the performance of research and when its benefits become apparent can be significant, unpredictable, and differ for different kinds of research" (Rymer, 2011, p. 3).

Some postulate that “it may take years, or even decades, until a particular body of knowledge yields new products or services that affect society” (Bornmann, 2012, p. 673). Rymer (2011) recommends assessing the impact of research in terms of what it aimed to achieve and capable of producing, not based on all the impacts that are possible.

Yet another problem associated with assessing societal impact of research is determining who should conduct the assessments. One logical recommendation is for researchers to conduct assessments of research impact. Researchers, however, tend to have one of two responses to such requests: feelings of disinterest or feelings of inadequacy (Bornmann, 2013; Holbrook & Frodeman, 2011; Spaapen, Dijstelbloem, & Wamelink, 2007). “Scientists generally dislike impact considerations, which they often see as challenging their authority and undermining the autonomy of the scientific enterprise” (Holbrook & Frodeman, 2011, p. 244). On the other hand, some researchers feel they do not have adequate expertise to evaluate the societal impact of research since such requests are beyond their disciplinary expertise (Holbrook & Frodeman, 2011). Because identifying the appropriate people to conduct assessments of research impact is an important part of studying it, researchers’ feelings of disinterest or inadequacy add to the challenges associated with studying this topic.

While there is value in generating ways to assess and measure impact in ways that take the aforementioned difficulties into consideration, there could also be a danger associated with conducting such assessments and evaluations. One potentially negative consequence of measuring the impact of research is that it can distort behavior (Rymer, 2011). Instead of using an improved understanding of impact *to inform* research decisions, researchers may begin to use it *to drive* their research. Researchers may begin to strive for the impact that gets measured, as opposed to conducting research based on guidelines of scientific inquiry (National Research Council, 2002).

Difficulties Associated with Interpretations of Impact

If all stakeholders viewed impact the same way, it would be easier to study the dimensions of research impact. This is not the case, however. There are three difficulties associated with interpretations of impact. The societal impact and domain-specific impact of research will vary based on the scientific work, since the research results will affect different aspects of society and the contexts of interest. As a result, there is no one model for assessing research impacts that will fit all research types, disciplines and institutions around the world (Bornmann, 2013; Martin, 2011; Molas-Gallart, Salter, Patel, Scott, & Duran, 2002; Rymer, 2011). Thus, any existing research impact assessments developed for one purpose will need to be modified to be relevant and applicable to another context of interest. In addition to the fact that impact looks different in different contexts, impact can come in different magnitudes: sometimes impact is very large but oftentimes it is very modest (Rymer, 2011). Rarely will all stakeholders agree on the worth of the impact (Rymer, 2011).

There is one final point related to the difficulties associated with studying the societal impact of research. It is easy to assume that impact implies a benefit or advancements. However, it is important to remember that impact may not always be desirable or positive (Bornmann, 2013; Martin, 2011). Moreover, there may be instances where the same research impact can be interpreted as positive, negative, or neutral—depending on the stakeholder's perspective (Bornmann, 2013; Martin, 2011; Rymer, 2011). Despite all the difficulties associated with studying societal and domain-specific impact, there are studies that have begun to address this topic.

Existing Studies on Research Impact

The literature on characterizing and evaluating the societal and discipline-specific impact of research is fragmented. There are four collections of studies that provide a foundation for this current study; linking research impact to hierarchies based on organizational and disciplinary structures is one way to weave the four collections of literature together. This proverbial hierarchical structure of these four bodies of literature would include four levels: a global, national, organizational, and disciplinary level. At the top of this hierarchy of literature are studies related to global initiatives focused on research impact. Because the current literature on research evaluation initiatives taking place outside of the U.S. do not include U.S.-based efforts, the second level includes studies centered on U.S.-specific attempts to describe and measure research impact. Level three includes the activities of federal funding agencies, specifically NSF, to understand the impact of their investments. The lowest level of studies on research impact includes discipline-specific studies focused on impact.

Governments around the world, federal organizations, and disciplines are institutions –whether physical or conceptual— that have something in common: they exist to engage in activities that promote a mission and set of priorities. Conducting evaluations is the way such institutions determine the extent to which the mission and priorities are achieved. Impact, to a large extent, is organizationally driven because the impact that is observed is usually with reference to an institution's mission and priorities and how impact is defined and evaluated in that context. As a result of this, any studies on impact that occurs at the highest levels of the hierarchy cannot give the highest level of resolution about the impact that will be realized at lower levels. The remainder of this section summarizes the work on research impact from a global perspective, national perspective, NSF perspective, and disciplinary perspective. The consistencies between these bodies of work will also be discussed. This chapter concludes by articulating the gap in the literature that this study seeks to fill.

Level 1: Global Perspectives on Research Impact

Research evaluation systems at the highest levels of government around the world are the first collection of work that provides a basis for this study on the use of public funds to support undergraduate engineering education research. Bozeman and Sarewitz (2011) define research evaluation as "any systematic, data based (including qualitative) analysis that seeks as its objective to determine or forecast the social or economic impacts of research and attendant technical activity" (p. 8). Thus, societal impact is the dimension of research impact that is of most interest in the research evaluations conducted at the national level around the world. A 2010 study conducted by the RAND corporation, a global policy think tank, provides an international review of how research agencies around the world are assessing research impact (Jonathan Grant et al., 2010). (The literature review section entitled *Linking the Global and U.S.-based Perspectives on Research Impact* provides more specific information on how impact is conceptualized and measured at this level.)

Level 2: U.S.-based Perspectives on Research Impact

One set of activities that were not captured in the RAND report and other literature on national evaluation systems around the world, but align with the other assessments therein, is the recent U.S.-based efforts to understand and measure research impact. What links the *Global Perspectives on Research Impact* and the *U.S.-based Perspectives on Research Impact* is the national-level focus on societal impact. In 2008, the National Science and Technology Council (NSTC) produced a document entitled, *The Science of Science Policy: A Federal Research Roadmap.* "The science of science policy (SoSP) is an emerging field of interdisciplinary research, the goal of which is to provide a scientifically rigorous, quantitative basis from which policy makers and researchers can assess the impacts of the Nation's scientific and engineering enterprise, improving their understanding of its dynamics, and assess the likely outcomes" (NSTC, 2008b, p. 1). NSF's Science of Science and Innovation Policy (SciSP) program is responsible for supporting and managing this portfolio of research awards (OSTP, 2010). The first SciSP awards were funded in 2007 and by 2013, 145 awards had been made (NRC, 2014a). PIs conducting this research include economists, sociologists, political scientists, psychologists and domain scientists. While the SoSP research is still a developing field, the NRC report of the PIs' conference organized the most recent SciSP awards by the following categories: adoption and diffusion of knowledge, understanding the impact of structures/process on science, advancing understanding of entrepreneurship and innovation, new approaches to studying science and innovation, implementing science policy, measuring and tracking science and innovation (NRC, 2014a). In light of this synopsis, it is reasonable to suggest that although SoSP work is focused on societal impact, the collection of studies in this emerging field of research may sometimes include insights on scientific impact and domain-specific impact as well. Although the SoSP work is still unfolding, a SoSP handbook (Fealing, Lane, Marburger III, et al., 2011) has been written to capture highlights of the U.S.-based work thus far.

In the U.S., the development of the STAR METRICS (Science and Technology for America's Reinvestment: Measuring the EffecT of Research on Innovation, Competitiveness, and Science) is arguably the most notable of the SoSP activities focused on measuring the impact of federally-supported research in the U.S. (Lane, 2009, 2010; Lane & Bertuzzi, 2011). STAR METRICS is a recently established partnership between a set of federal agencies and universities established to identify and/or develop mechanisms to document and/or measure the outcomes of federal investments in research. When the project was envisioned, it included two phases. Phase I focused on the "development of uniform, auditable, and standardized measures of the initial impact of ARRA [American Recovery and Reinvestment Act] and science spending on job creation" (National Academy of Sciences, 2014). On the other hand, Phase II focuses on the "collaborative development of measures of the impact of federal science investments in four broad categories: economic growth (through patents, firms start ups and other measures), workforce outcomes (through student mobility and employment), scientific knowledge (such as publications and citations), and later, social outcomes (such as health and environment)" (National Academy of Sciences, 2014). The next section, entitled *Linking Global and U.S.-based Perspectives on Research Impact,* provides more specific information on how impact is conceptualized and measured at the global and national level.

Linking Global & U.S.-based Perspectives on Research Impact

Although there are fragments in the literature on research impact, there are consistencies – in the dimensions of impact that are of interest and approaches used to explore it. What national research evaluation systems have in common is their primary focus on the scientific impact (i.e., advances in knowledge) and social impact (i.e., economic, social, cultural, and environmental effects) of research. A myriad of approaches are used to understand research impact; and they can operate at the level of projects, programs, organizations, or nations. Each method has its own characteristics, along with advantages and disadvantages. The approaches used to study impact are indicators of the mission and priorities of the institution conducting the assessment. The differences in methods, data and tools used in the approach will be a driving influence for the differences in impacts that are observed across assessments.

While the RAND study on the international review of research assessment systems around the world (Jonathan Grant et al., 2010) did not include the U.S.-based SoSP studies (NRC, 2014b), there is a connection between the approaches researchers around the world use to evaluate research impact. Table 1 includes a list of most commonly cited approaches for evaluating research impact and brief descriptions about them. It is a compilation of the methods mentioned in the RAND study (Jonathan Grant et al., 2010), SoSP handbook and PIs conference report (Fealing, Lane, Marburger III, et al., 2011; NRC, 2014a), and other literature on methods of studying research impact (Rymer, 2011; Salter & Martin, 2001; Wooding, Hanney, Buxton, & Grant, 2004).

As Table 1 indicates, approaches to studying research impact range from qualitative, to quantitative, and may include a mixture of both types of methods. Computational methods are also used to study impact.

Table 1. Approaches to Studying Impact

	Approach to Studying Impact	Description
Quantitative <-----------------------------------> Qualitative	Logic Modeling	A series of "if… then" statements which provides a picture of how a research program works; it includes a graphical representation of connections between inputs, outputs, and outcomes.
	Interview	A meeting of people to discuss the research project of interest
	Case Studies	An in-depth qualitative analysis of research projects that usually results in narratives of the research process and outcomes.
	Anecdotes	A brief qualitative description of an incident related to the research.
	Peer or Expert Review	An evaluation of a research project conducted by peers/other scholars in the field; similar to process used for reviewing scholarly publications.
	Decision Analysis	A problem-solving approach to making informed and objective decisions when encountering complex situations.
	Portfolio	A purposeful collection of work that reflects various aspects of research project.
	Program Evaluation	A use of qualitative and quantitative methods to assess the process and outcomes of a research program.
	Survey or Questionnaire	A series of closed- and/or open-ended questions that lends itself to statistical and/or qualitative analysis.
	Retrospective Analysis	A historical assessments of research processes and outcomes.
	Self-assessment	An analysis of one's actions or performance in relation to a standard.
	Benchmarking or Input Measures	A comparisons across programs, organizations, and countries.
	Cost-benefit Analysis	A quantitative assessment for determining the net gain (or loss) that will result from pursuing a set of actions.
	Bibliometric Analysis	A quantitative assessment of research publications, including its quantity, quality, and collaboration.
	Economic Modeling or Analysis	A quantitative assessment of the rate of the return on research investments.
	Productivity Analysis	A statistical analysis of the extent to which a project has produced outcomes of interest (e.g., publications, research collaborations).
	Risk Models	A quantitative description of the relationships between the risks and net gain associated with pursuing a set of actions.
Computational	Agent-based Models	A computational approach to simulate the actions and interactions of the key agents in a system.
	Visual Analytics	A computational approach to developing tools that aid in understanding and decision-making as a result of visually representing complex information in meaningful ways.

Various types of data are used to conduct these analyses, and the timing of when the evaluations occur with respect to when the research itself occurs varies as well. Moreover, just as the approaches to studying impact vary, so do the types of results that come from these approaches. This difference in methods and corresponding results is what leads to the differences in how institutions perceive their impact, and is oftentimes the lens through which they view the impact of others.

Table 2 presents which evaluation systems around the world use each of these approaches. As the table indicates, research evaluation systems around the world employ a breadth of approaches to studying impact—quantitative, qualitative, and computational approaches. The most commonly used methods are peer/expert review, and survey/ questionnaire. On the other hand, computational tools (i.e., agent-based models, visual analytics) are used the least often. The names of the assessment frameworks hint at another consistency in research evaluations systems used around the world. While it is reasonable to assume that there is interest in all four of the main aspects of societal impact (i.e., social, cultural, environmental, and cultural), social seems to be prioritized. More specifically, the impact of health science research (e.g., see Canada, the Netherlands, and U.S.) and evaluations of universities (e.g., see Japan and the UK) are the two specific types of impact that seem to be of special interest to policymakers at the highest levels of government around the world.

Table 2. Approaches Used to Study Impact in Research Evaluation Systems Around the World

	Research Evaluation Systems Around the World																				
	CONEAU: Comisión Nacional de Evaluación y Acreditación Universitaria	RFQ: Research Quality Framework	ERA: Excellence in Research for Australia	MORIA: Measurement of Research Impact and Achievement	Payback Framework	Framework Programme for Research and Technology Development	CNRS: Centre National de la Recherche Scientifique	NIAD-UE: National Institution for Academic Degrees and University Evaluation	Societal Impacts of Health Research	SEP: Standard Evaluation Protocol	ERiC: Evaluating Research in Context	PBRF: Performance-Based Research Fund	PNECU: Plan Nacional de Evaluación de las Universidades	Various Evaluations of Swedish Research	Economic Impacts of Investments in Research and Innovation Framework	RAE: Research Assessment Exercise	Higher Education Impact Model	STAR Metrics: Science & Technology in America's Reinvestment - Measuring the EffecT of Research on Innovation, Competitveness and Science (in development)	Science of Science Policy Research (other)	CDMRP: Congressional Directed Medical Research Programs	PART: Program Assessment Rating Tool
Approaches to Studying Impact	Argentina	Australia			Canada	EU	France	Japan	Netherlands			New Zealand	Spain	Sweden	UK			U.S.			
Logic Modeling					✓																
Interview						✓					✓		✓	✓	✓					✓	
Case Studies					✓	✓								✓	✓					✓	
Peer or Expert Review	✓	✓				✓	✓	✓		✓		✓	✓	✓		✓	✓				✓
Decision Analysis																			✓		
Portfolio	✓	✓	✓				✓					✓				✓					
Survey or Questionnaire	✓			✓	✓	✓	✓	✓	✓		✓		✓		✓		✓				✓
Self-assessment										✓											
Benchmarking or Input Measures		✓	✓	✓					✓			✓			✓	✓	✓				✓
Cost-benefit Analysis						✓															
Bibliometric Analysis			✓	✓	✓			✓			✓			✓							
Economic Modeling or Analysis						✓								✓	✓				✓		
Productivity Analysis														✓	✓				✓		
Risk Models																			✓		
Agent-based Models																			✓		
Visual Analytics																			✓		

Level 3: NSF's Efforts Focused on Research Impact

The third collection of work that informs this study on the impact of NSF's investments in undergraduate education research is NSF's efforts to understand impact. The Foundation's interest in the impact of a project begins with its two primary criteria used to review all research proposals: Intellectual Merit and Broader Impacts.

> *"The Intellectual Merit criterion encompasses the potential to advance knowledge; and [t]he Broader Impacts criterion encompasses the potential to benefit society and contribute to the achievement of specific, desired social outcomes" (NSB, 2011, p. 2).*

These two NSF review criteria also align with these two facets of research impact most commonly found in existing literature (scientific and societal impact); and to some extent, both concurrently address forms of impact that may be unique to a STEM domain (i.e., domain-specific impact). Consequently, one distinguishing characteristic of this body of work is that it has the potential to effortlessly link to all three dimensions research impact (scientific, domain-specific, and societal) if both Intellectual Merit & Broader Impacts are adequately addressed in the proposal and realized in the results of the proposed study. NSF was one of the first federal funding agencies around the world to emphasize both scientific and societal impact in their review process (Bornmann, 2013). Over the years, NSF has put a lot of attention on the Broader Impact criterion and the following quote tries to capture some evidence of progress: "NSF's attention to Criterion 2 has produced improvements in terms of the *quantity* of proposers and reviewers who address Criterion 2; yet the *quality* of the responses to Criterion 2 remains a persistent problem" (Holbrook, 2005, p. 445). Despite the agency's longstanding emphasis on broader impact, only a few recent scholarly studies have been focused on the NSF review criteria (NSB, Holbrook, 2005, 2012; Holbrook &

Frodeman, 2011; Kamenetzky, 2013; 2011; Roberts, 2009), with particular emphasis on the "Broader Impact" criteria. Some focus on the philosophical issues associated with it and the history of its changes in an attempt to add clarity (Holbrook, 2005). Others are related to peer reviewers' largely negative reactions to the second criteria (NAPA, 2001 as cited by Holbrook, 2005) and their difficulty with applying it to review proposals (Holbrook & Frodeman, 2011).

To date, only two studies have explored the contents of broader impacts narratives in NSF awards (Kamenetzky, 2013; Roberts, 2009). Key findings from Kamenetzky (2013) indicate that engineering proposals show a statistically significant difference in their likelihood to propose potential societal benefits and partner with potential users unlike proposers in biological or mathematical/physical sciences. Roberts (2009) states that "it appears that the potential societal benefits are probably overstated" (p. 212). The study goes on to explain that although some researchers may propose Broader Impacts in their grant proposals, the absence of dissemination plans suggests that they were no more likely to disseminate findings beyond the scientific community than those who only address the Intellectual Merit criterion in their proposal. This finding suggests an embedded assumption that Broader Impacts is accomplished by having a dissemination plan. Plans to disseminate findings may be an important step in realizing the "potential to benefit society and contribute to the achievement of specific, desired social outcomes" (NSB, 2011, p. 2).; however, outcomes of the dissemination plans in terms of its realized impact were not included in the purpose of this study, and a result, were not discussed.

Another set of work related to the impact of NSF investments in research relates to the development of data mining and visualization tools that allow policymakers to quickly understand their portfolio of awards and to use this understand to make informed decisions (Raghavan et al., 2011). One computational tool that was developed for this purpose is DIA2 (Deep Insights Anytime, Anywhere). It is a portfolio mining tool focused on NSF investments and is being designed for the STEM education community (Madhavan et al., 2012). In its mature state, DIA2 will enable "four broad classes of analytics capabilities: search and visualization, community and collaboration analysis, temporal modeling, and investment analysis" (Madhavan et al., 2012, p. 74). Using interactive computational tools such as DIA2 (www.dia2.org), users can quickly synthesize large amounts of text-based and quantitative data to understand how funding decisions have led to the current STEM education research landscape. Provided are two screenshots of DIA2 when using the "NSF Org Structure" and "Topic Explorer" widget (screenshots taken on September 19, 2014).

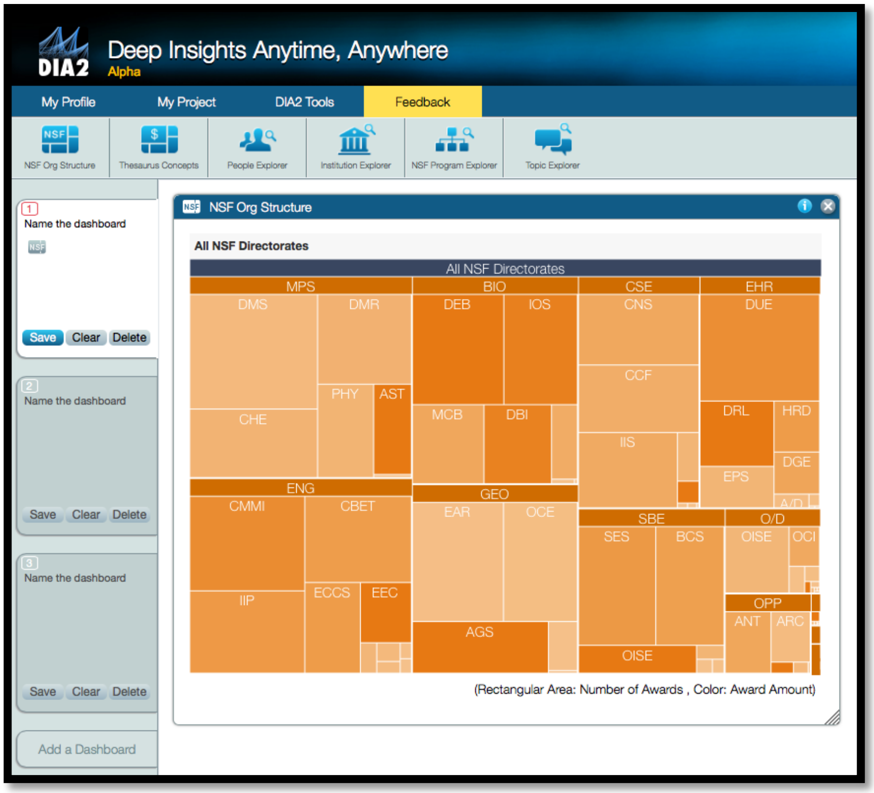


Figure 3. DIA2 Screenshot Using "NSF Org Structure" Widget

Which topic would you like to explore? engineering education

engineering education

No. of awards	1945
Award amount	$1.4bn
No. of PI/coPI	3475

3,475 entries

Name	Num. of awards
Norman Fortenberry	25
Matthew Ohland	18
Frank Huband	16
Lisa McNair	11
Marie Paretti	10

Figure 4. DIA2 Screenshot Using "Topic Explorer" Widget

Level 4: Disciplinary Studies on Research Impact

The fourth and final collection of studies that informs this current work were alluded to during the discussion on defining the *domain-specific* facet of research impact. Over the last decade, a number of studies focused on developing frameworks that specify the influence of methods or results of an R&D project on the people, priorities, and/or processes in the context of interest. *Disciplinary Studies on Research Impact* includes the most insights about scientific & domain-specific impact, although findings about a collection of research's societal impact is not uncommon.

In a paper on frameworks and review articles Schwarz et al. (2007) define a framework as the "exposition of a set of assumptions, concepts, values, and practices that constitutes a way of understanding the research within a body of knowledge" (p. 41). This is what is meant by the use of the term "framework" is this study. As part of their work, Schwarz, Mehta, Johnson, and Chin (2007) identify ten purposes for a frameowork. The two most relevant in this context are that frameworks "provide a new focus within a research stream", and "synthesize previous research in an actionable way for practioners" (Schwarz et al., 2007, pp. 32-33).

One framework was developed with the focus on NSF-funded informal STEM education and outreach research projects (Allen et al., 2008). Others focus on arts and humanities research at the University of Cambridge (Levitt et al., 2010), health sciences research (Donovan & Hanney, 2011; Kuruvilla et al., 2006), and science, in general (Godin & Doré, 2005). Publications, especially literature reviews on the topic of research impact, sometimes include a short list of the ways in which research and/or science, in general, may make a societal impact (Rymer, 2011; Salter & Martin, 2001; Walter et al., 2003). Apart from the common goal to characterize what impact looks from the perspective of a particular discipline, what links these frameworks together are the parallels between the approaches used to develop them, and the areas of overlap between the resulting constructs that make up the frameworks. See Table 3 for the purposes, methods overview, and summary of outcomes of the studies that led to these frameworks.

Table 3. Research Overview of Studies Resulting in a Research Impact Framework for a Specific Research Domain

	Research Overview			
Impact Framework	**Purpose of the Study**	**Data**	**Analysis and Validation**	**Result**
Informal Education and Outreach Framework (Allen et al,. 2008)	To develop a framework that identifies the broad categories of potential project impact	**Three qualitative data sources**: sample of proposals, final reports, and summative evaluations associated with NSF's Informal Science Education (ISE) program	Synthesis of project impacts included in project documents and evaluations.	5 categories of the impact of a informal science education and/or outreach project
Health Sciences Research Impact Framework (Kuruvilla et al., 2006)	To develop a conceptual framework that identifies potential areas of impact of health sciences research	**Three qualitative data sources**: Research impact assessment literature; research assessment criteria, for example, as set by the UK Research Assessment Exercise panels; case studies of research conducted by London School of Hygiene and Tropical Medicine; content experts' opinions	Synthesis of insights from the literature, research assessment criteria, and case study; content expert validation	4 broad categories of the impact of health sciences research, with additional 27 descriptive sub-categories
Payback Framework (Donovan & Hanney, 2011)	To examine the impact or 'payback' of health services research, and basic and clinical biomedical research.	**Three qualitative data sources**: Models of research utilization; existing assessments of the payback of research; case studies of existing health sciences research	Synthesis of insights from models and existing assessments. Case studies were used to assess feasibility and applicability	5 categories of the impact of health sciences research; logic model of the complete research process
Arts & Humanities Research at the University of Cambridge (Levitt et. al, 2010)	To provide an assessment of the reported impact of arts and humanities research	**Three qualitative data sources**: interviews with senior University of Cambridge arts and humanities researchers (n=22); interviews with external users of research outside the University of Cambridge (n=17); case studies of arts and humanities research at the University of Cambridge. **One quantitative data source**: survey of all arts & humanities researchers at the University of Cambridge (n=737).	Analytic Framework: Payback Framework (Donovan & Hanney, 2011). Interview informed the initial framework development. Surveys used to validate the interview data. Case study analysis to reveal the research process and impact of specific research projects or stream of research.	9 categories of the impact of the arts and humanities research at the University of Cambridge, adapted from the Payback Framework (Donovan & Hanney, 2011)
Typology of the Impact of Science (Godin & Dore, 2005)	To develop a typology that categorizes the contributions of science to society	**Two qualitative data sources:** Interviews with researchers at 17 publicly funded research centers; interviews with users of researcher results at 11 social and economic organizations	Synthesis of interview insights	11-dimension typology of the impact of science on society

There are many consistencies in the approaches used to develop the frameworks for characterizing impact in a specific research domain. As Table 3 indicate, qualitative data collection and analysis methods –namely document analysis, interviews, and case study analysis— were used as part of the framework development (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010). Three different approaches have been used to validate the resulting framework: integrating feedback from content experts (Kuruvilla et al., 2006), administering a survey to a large sample of researchers within the domain (Levitt et al., 2010), and applying the framework to a set of research projects (case study analysis) to evaluate its feasibility and applicability (Donovan & Hanney, 2011; Levitt et al., 2010).

As it was previously stated, the consistencies in the impact dimensions mentioned in the current literature reviews on this topic, and the disparate research impact frameworks are what link them. Table 4 lists the impact categories mentioned in three publications on research impact for research, in general (Molas-Gallart et al., 2002; Rymer, 2011; Walter et al., 2003). Tables 5 and 6 include the impact dimensions associated with the five frameworks for looking at research in a specific research domain (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010). Every impact category in the frameworks maps to at least one of the three dimensions of research impact: scientific impact, societal impact, and domain-specific impact. The mapping is based on the alignment of definition of research impact (See Figure 2) and the description of the impact dimension in the article describing the framework. Where applicable, only the main categories were mapped; sub-categories were not individually mapped.

In the frameworks on research in general (Molas-Gallart et al., 2002; Rymer, 2011; Walter et al., 2003) scientific impact is the most prominent. Domain-specific impact is well reflected in all of the frameworks as well. Societal impact rarely maps to the dimensions. In the frameworks on the impact of research in general (Molas-Gallart et al., 2002; Rymer, 2011; Walter et al., 2003), or in specific research domains (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010), all three facets of research impact were widely represented in each framework except Allen et al. (2008).

Table 4. Research Impact Dimensions Listed in Related Literature Reviews

Impact Framework & Corresponding Impact Dimensions	**Research Impact Dimensions**		
	Scientific Impact	**Domain-specific Impact**	**Societal Impact**
Research, in general (Salter & Martin, 2001)			
Increasing the stock of useful knowledge	✓		
Creating new scientific instrumentation and methodologies	✓		
Creating new firms			✓
Training skilled graduates		✓	
Forming networks and stimulating social interactions		✓	
Increasing the capacity for scientific and technological problem solving	✓	✓	
Research, in general (Walter, Nutley, and Davies, 2003)			
Changes in knowledge and understanding	✓	✓	
Changes in access to research	✓		
Changes in the extent to which research is considered, referred to, or read	✓	✓	✓
Citations in documents	✓		
Changes in attitudes and beliefs		✓	✓
Changes in behavior		✓	
Research, in general (Rymer, 2011)			
Advances in knowledge	✓		
Additional investment		✓	
Financial return		✓	
Economic impact			✓
Social impact		✓	✓
Environmental impact			✓
More effective teaching		✓	
Intangible impacts	✓	✓	✓

Table 5. Impact Dimensions in Research Impacts Frameworks in Specific Domains

Impact Framework & Corresponding Impact Dimensions	Research Impact Dimensions		
	Scientific Impact	Domain-specific Impact	Societal Impact
Informal Education and Outreach Framework (Allen et al,. 2008)			
Awareness, knowledge or understanding (of) STEM concepts, processes, or careers		✓	
Engagement or interest (in) STEM concepts, processes, or careers		✓	
Attitudes (towards) STEM-related topics or capabilities		✓	
Behavior (related to) STEM concepts, processes, or careers		✓	
Skills (based on) STEM concepts, processes or careers		✓	✓
Typology of the Impact of Science (Godin & Dore, 2005)			
Science: knowledge, research activities, training	✓	✓	
Technology: products and processes, services, know-how		✓	✓
Economy: production, financing, investments, commercialization, budget			✓
Culture: Knowledge, know-how, attitudes, values		✓	✓
Society: welfare, discourses and actions of groups			✓
Policy: policy-makers, citizens, public programs, national security		✓	✓
Organization: planning, work organization, administration, human resources		✓	
Health: public health, health systems		✓	✓
Environment: management of natural resources and the environment, climate & meteorology			✓
Training: curricula, pedagogical tools, qualifications, graduates, insertion into the job market, fitness of training/work, career, use of acquired knowledge		✓	✓
Symbolic: legitimacy/credibility/visibility, notoriety	✓	✓	✓
Arts & Humanities Research at the University of Cambridge (Levitt et. al, 2010)			
Academic impact	✓		
Impact on public knowledge creation			✓
Impact on policy		✓	✓
Impact on the preservation of heritage		✓	✓
Economic impact on the wider society		✓	✓
Impact on leisure and entertainment		✓	✓
Direct economic impact			✓
Impact through teaching		✓	
Cross-cutting findings	✓	✓	✓
Payback Framework (Donovan & Hanney, 2011)			
Knowledge	✓		
Benefits of future research and research use	✓	✓	
Benefits from informing policy and product development		✓	
Broader economic benefits			✓
Health and health sector benefits		✓	✓

Table 6. Impact Dimensions in Research Impact Frameworks for Specific Domains

Impact Framework & Corresponding Impact Dimensions	Research Impact Dimensions		
	Scientific Impact	Domain-specific Impact	Societal Impact
Health Sciences Research (Kuruvilla et al., 2006)			
Research-related impacts	✓	✓	✓
Type of problem/knowledge			
Research methods			
Publications and papers			
Products, patents, and translatability potential			
Research networks			
Leadership and Awards			
Communication			
Policy impacts		✓	✓
Level of policy-making			
Type of policy			
Nature of policy impact			
Policy networks			
Political capital			
Societal Impacts		✓	✓
Types of service: health/intersectoral			
Evidence-based practice			
Quality of care			
Information systems			
Services management			
Cost-containment and cost-effectiveness			
Service Impacts: health and intersectoral		✓	✓
Knowledge, attitudes, and behavior			
Health literacy			
Health status			
Equity and human rights			
Macroeconomic/related to the economy			
Social capital and empowerment			
Culture and art			
Sustainable development outcomes			

Acknowledging Related Literature

Before concluding this review of the literature, it is important to acknowledge another body of tangentially related research. Among STEM education research, there is a fast growing collection of studies on how transformational change happens in undergraduate STEM education – the need for it, what it looks like, what enables it, what impedes it, what processes and factors might lead to a particular type of change, etc. (Beach, Henderson, & Finkelstein, 2012; Besterfield-Sacre, Cox, Borrego, Beddoes, & Zhu, 2014; Boyer Commission on Educating Undergraduates in the Research University, 1998; Boyer, 1990; Burkhardt & Schoenfeld, 2003; Dancy & Henderson, 2008; Fairweather, 2010; Finelli, Daly, & Richardson, 2014; Gillespie, McKenna, & Pimmel, 2011; Henderson & Dancy, 2010; McKenna, Froyd, King, Litzinger, & Seymour, 2011; McKenna, Froyd, & Litzinger, 2014; NRC, 1999a; 2003; Seymore; Seymour, 2001; Seymour & DeWelde; Siddiqui & Adams, 2013). The April 2014 Special Issue of the *Journal of Engineering Education,* the premiere journal for engineering education research, represents a critical junction in the evolution of research on this topic because it provides a "constructive starting point for future conversations" on systemic transformation in engineering education (McKenna et al., 2014, p. 189). The authors present "perspectives on the breadth and complexity of systemic transformation in engineering education" and examples of "mechanisms that include the capacity for transformation" (McKenna et al., 2014, p. 189) Much of this work is focused on the need to change STEM education policies at various levels and employ more evidenced-based pedagogical practices. While it is anticipated that research impact will influence policy and pedagogy, existing research impact frameworks hint at the possibility that impact may extend beyond these facets of undergraduate education. The use of the *means vs. ends* idiom is useful for depicting the connection between this literature and the existing research that was reviewed in this chapter. While this literature on change and transformation is focused on the means to a particular end, this study on impact is focused on the ends itself. Together, the two bodies of work complement and reinforce one another.

Summary of Literature Review & Gap in the Literature

Impact is a term that has different meanings among different groups. Some of these differences stem from differences in philosophical perspectives on impact. Those in the group interested in the philosophy and sociology of knowledge tradition are interested in the impact of research. Although most literature on research impact usually talks about this in terms of two dimensions –scientific impact and societal impact—the recent development of frameworks that capture impacts in particular discipline add evidence to support that research impact has another dimension: domain-specific impacts. There are three major difficulties associated with studying the non-scientific facets of research impact: difficulties associated with connecting impact and research or researcher, difficulties associated with assessment and evaluation of research impact, and difficulties associated with interpretations of impact. Despite these difficulties, however, there have been a number of studies conducted to explore impact at national, institutional, and disciplinary levels. The few NSF-specific studies have been centered on Broader Impacts review criteria, and insights on the proposed impact of a study. Because the impact statements analyzed in these two studies were included in the proposal, these narratives capture the intended or anticipated impact of the study— at most. Until now, no study has looked at the research impact that is actually realized after a NSF-funded undergraduate STEM education study commences or is completed. It is reasonable to suggest that the research impact frameworks developed to characterize the impact in research in specific domains are an exposition on how a research community is beginning to conceptualize impact. While researchers in different communities that rarely connect conducted the existing studies on research impact, there are consistencies in the approaches used to characterize and/or measure impact, and in the dimensions of impact observed. In multiple instances, the process for developing a framework of research impact in a specific domain begins with an exploratory approach, then is validated using quantitative and/or qualitative methods. But in spite of these elements of continuity, there is a gap that still exists. Within the context of engineering education, there is no shared vocabulary for discussing the impact of research, or a framework that characterizes the impact of federal investments in undergraduate engineering education research. This study seeks to begin to fill this gap in the literature by exploring the perspectives of two key stakeholders (i.e., Principal Investigators and NSF Program Officers) on what it means for a federally funded STEM education project to have impact. An understanding of how these key stakeholders talk about impact is an important step toward the development of a conceptual framework that characterizes the impact of NSF investments in undergraduate engineering education R&D projects.

CHAPTER 3: GUIDING FRAMEWORKS

Toulmin's Model (1958) and the *Common Guidelines for Education Research and Development* (Earle et al., 2013) are the two frameworks that will guide the data collection and analysis in this study. The rationale for using both frameworks is that the strengths of one offset the limitations of the other, and vice versa. While Toulmin's Model provides a strong theoretical framing for this study, the *Common Guidelines for Education Research and Development* serve as a conceptual framework. In the next two sections of this chapter, both frameworks will be described independently. This chapter will conclude with a description of how *Toulmin's* model is being extended to include the *Common Guidelines for Education Research and Development*; together they provide a theoretical and conceptual lens for this study.

Theoretical Framework: Toulmin's Model

Toulmin developed a model (1958) in response to a need for a method that "would blend logic and epistemology into 'applied logic'" (Hitchcock & Verheij, 2006, p. 1) for the purpose of being more applicable for assessing the arguments made in everyday life. Toulmin focused on a particular use of arguments: to defend a claim made by asserting something (Hitchcock & Verheij, 2006b). Toulmin's model is normally used in the context of formal logic, debate, and argumentation. However, this is one of the first studies that use it to understand the arguments made by STEM education researchers about their work.

The *Toulmin Model* (also referred to as the *Toulmin Scheme*) (1958) describes the process of defending a claim against a challenger and is a pattern that is largely independent of a particular field. Hitchcock and Verheij (2006) succinctly describe *Toulmin's Mo*del in the introductory chapter of *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation*. The first step in the process is the assertion of a claim (C). In response to this claim, a questioner may ask, "What do you have to go on to make this claim?" The defender then appeals to data (D), relevant facts that are available to the defender. The challenger may then ask about how the data relates to the claim that is being made. To this, the defender may respond with a proposition, or warrant (W), that may take the form of: "Data such as D entitle one to draw conclusions, or make claims, such as C" (Toulmin, 1958, p. 98).

Booth et al. (2008), authors of the *The Craft of Research,* concur with Toulmin that warrants are "statements that connect reasons to claims" (p. 152), but clarify "academic warrants" as "specific principles of reasoning that particular communities of researchers develop over centuries of thinking and writing, and they are countless" (p. 154). Warrants are rarely stated because they are deeply embedded in tacit knowledge; however, three instances in which writers should explicitly state academic warrants are when: 1) readers are outside of the writer's field; 2) the principle of reasoning is new or controversial; and 3) readers will likely resist the claim because they simply do not want it to be true (Booth et al., 2008). These additional insights on warrants will be useful during data analysis.

To indicate the degree of force that a warrant confers on the conclusion being justified, a defender may qualify the conclusions with a qualifier (Q). Additionally, there may be a need for the defender to present conditions of rebuttal (R) "indicating the circumstances in which the authority of the warrant would have to be set aside" (Toulmin, 1958, p. 101). Lastly, the questioner may ask about the general applicability of the warrant; the defender's response to this question is called backing (B). The differences in the backing of an argument vary from one field to another. Toulmin proposed the following diagram (See Figure 5) to illustrate the contributions of each component:

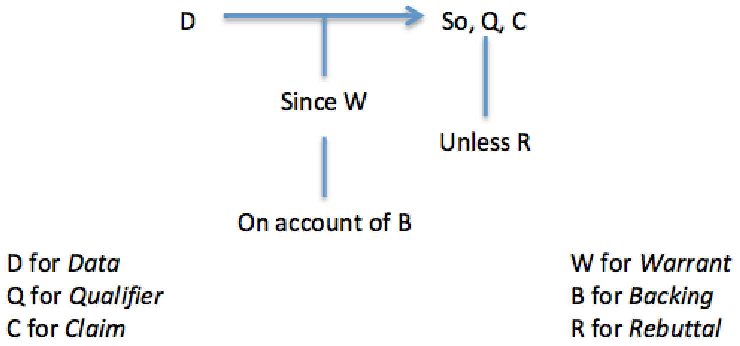


Figure 5: Diagram of Toulmin's Model (Toulmin, 1958, p. 104)

Toulmin's model, first published in his *Uses of Argument* book (1958), is a commonly-used reference in the context of rhetoric, logic, debate, and argumentation. However, Toulmin's Model not only applies to standards of arguments, but to verbal reasoning in general (Hitchcock, 2006). It provides a theoretical basis for this study because it offers an explanation on the construction of arguments and generalizes beyond its original context to one in which the goal is to understand arguments researchers make surrounding aspects of their research. This model is applicable to this study since the primary data source that will be used in this study includes research project abstracts, narratives that include claims researchers are making about the impact of their studies.

Although Toulmin's Model makes a strong theoretical contribution to the body of knowledge, rhetoric scholars have noted the difficulty with applying it to the evaluation of an actual argument (Klumpp, 2006; Tans, 2006; Voss, 2006). One of the reasons why Toulmin's model is difficult to apply is because not all arguments are structured the same. For example, some arguments may include multiple data or multiple warrants. Another challenge comes with appropriately identifying the constituents of the model in actual argument. Distinguishing between warrants and data, and data and backing are among the most commonly cited difficulties. Being able to apply Toulmin's model to identify claims about impact is a critical part of this study. The second framework used in this study makes up for this limitation.

Conceptual Framework: Common Guidelines for Education Research and Development

The second framework used in this study is the *Common Guidelines for Education Research and Development* (Earle et al., 2013) (referred to hereafter as the "Common Guidelines"). A Joint Committee of representatives from the U.S. Department of Education (DOE) and NSF first met in January 2011 to start the development. The guidelines were a response to a need to "establish cross-agency guidelines for improving the quality, coherence, and pace of knowledge development in science, technology, engineering, and mathematics (STEM) education" (p. 4). These guidelines articulate the "role of various types or 'genres' of research in generating evidence about strategies and interventions for increasing student learning" (p. 7). Figure 6 shows the "pipeline" of STEM education research types, while Table 7 lists each research type and a brief description of them.

Figure 6: Pipeline of STEM Education Research Types, from the Common Guidelines

Table 4: Description of Research Genres & Types Adopted from the Common Guidelines (Earle et al., 2013, p. 9)

Foundational Research and Early Stage or Exploratory Research (Research Types 1 & 2)
Foundational Research and Early Stage or Exploratory Research contributes to basic understandings of teaching and learning. Examples of this may include research focused on cognition, or processes involved in learning and instruction.
Design and Development Research (Research Type 3)
Design and Development Research builds on existing theory and evidence to develop solutions to achieve a goal related to education or learning. Examples of goals may include improved student engagement or skill mastery.
Impact Studies (Research Types 4-6)
Efficacy, Effectiveness, and Scale-Up Research contributes to evidence of impact by generating reliable estimates of the ability of a fully developed solution to achieve its goals.

Earle et al. (2013) provide guidelines for each of the six types of research. Each set of guidelines provides a basic description of the research purpose, justification (i.e. empirical, theoretical, practical, or policy) guidelines, and the type of evidence produced by a particular research type.

Although the six study types follow a logical sequence from the development of basic knowledge, design, testing, and scaling up, the developers of these guidelines acknowledge that the reality of research is much more complex (Earle et al., 2013). More specifically, Earle et al. (2013) explicitly state the following assumptions: knowledge development is not linear; investigation can sometimes move directly from development of core knowledge to scale-up research; and individual studies may incorporate elements that cut across research types (p. 10). Nevertheless, these guidelines describe –in broad strokes– the major types of federally funded STEM education research and development, and are fitting for the context of this study.

Lastly, the *Common Guidelines* refers to the last three of the six research types as "Impact Studies". Since the focus of this study is on the "research impact" of NSF's investments in undergraduate STEM education, it is necessary to highlight the differences in the use of the term "impact". For the *Common Guidelines*, the term "impact" implies a particular methodology that will lead to "reliable estimates" on the extent to which an intervention or strategy can achieve its intended outcome. In this study, the term "research impact" is used to denote the scientific, societal, and domain-specific impacts of research and development projects. Scientific impact of research includes advances in reliable knowledge, while societal impact is research outputs or products that benefit a dimension of society (e.g., social, cultural, environmental, economic, business or education practice) (Bornmann, 2013; Donovan, 2011). Domain-specific impacts are those that are unique to a particular research field, and result from the influence of methods or results of an R&D project on the people, priorities, and/or processes in the context of interest. All three types of research impact are deemed important for federally funded research, and are what is meant when the term "research impact" is used in this document.

The *Common Guidelines* (Earle et al., 2013) were developed to be used for organizational and clarification purposes by decision makers in federal agencies and grantees seeking federal funding (Earle et al., 2013). Again, the *Common Guidelines* describe the methodologies STEM education researchers typically use, the types of evidence that comes from particular methodologies, and justifications (i.e. empirical, theoretical, practical, or policy) associated with each type of research. Thus, the *Common Guidelines* are very practical (and easy to apply), but the authors do not claim these guidelines were developed from theory. As a result, they do not have the guiding or explanatory power that theory typically provides in research. Again, the strengths and limitations of both frameworks work together to provide a basis for this study that is grounded in theory, yet practical. In short, Toulmin's model explains the structure of arguments, in general, and the *Common Guidelines* inform how *Toulmin's Model* is applied to look for arguments about research impact.

Extending Toulmin's Model to Include the Common Guidelines

When integrating the elements of *Toulmin's Model* (1958) and the *Common Guidelines* (Earle et al., 2013), what results is an adapted model that fits the context of studying the arguments surrounding federally-funded STEM education R&D projects. More specifically, the *Common Guidelines* state that a particular *Research Type (RT)* affords certain types of *Evidence (E).* Simply put, evidence is some form of information presented in response to an inquiry. In light of this, *Evidence (E)* in the *Common Guidelines* is comparable to the *Data (D)* in *Toulmin's Model.* The *Warrant (W)* in Toulmin's Model that allows someone to make a claim based on *D* is comparable to the *Justification (J)* in the *Common Guidelines*, since the justification is a researchers' rationale (or basis) for various aspects of the research (e.g., purpose, methodology, outcomes). All of the other elements of Toulmin's model (i.e. Claims, Qualifiers, Backing, and Rebuttal) remain as they were originally. Figure 7 illustrates a way to extend Toulmin's Model to include the *Common Guidelines;* the subscripts (TM & CG) denote the original framework. The use of this integrated model as an analytical tool is described in the *Data Analysis* section of the Phase One Methods chapter.

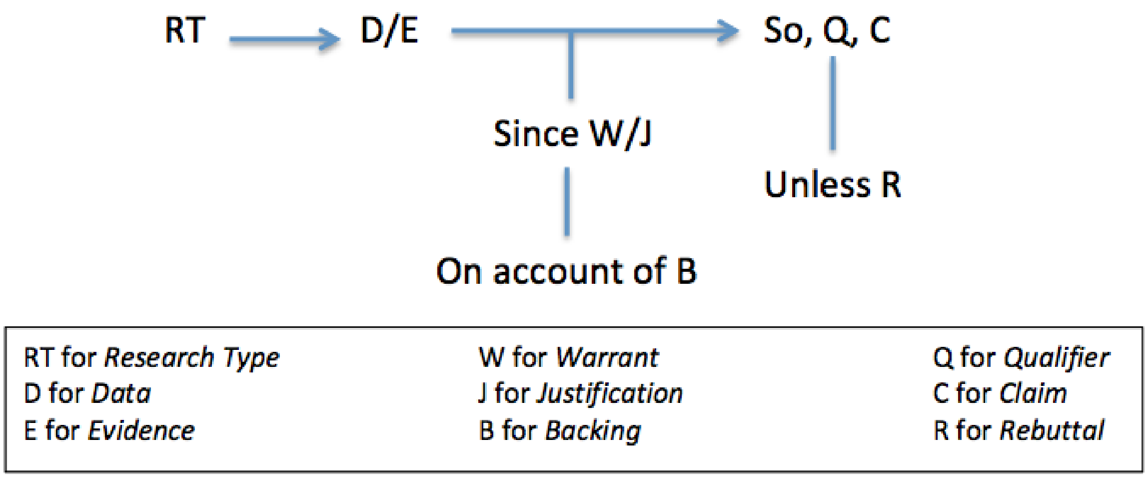


Figure 7: Extending Toulmin's Model to Include the Common Guidelines

CHAPTER 4: PURPOSE, RESEARCH QUESTIONS, AND RESEARCH DESIGN

This chapter begins by presenting the purpose of this study, and the research questions corresponding to it. A *Multiphase Mixed Methods* research design was used to answer the proposed research questions. The research design and epistemological assumptions associated with this mixed methods study will be presented after the research questions. The proceeding chapters will include full details of the data collection and analysis corresponding to the research questions proposed in each phase of the study.

Purpose

The purpose of this two-phased study is to: 1) explore how Principal Investigators on NSF-funded undergraduate STEM education R&D projects talk about impact; and 2) compare PIs' perspectives on impact with the perspectives of Program Officers who oversee NSF's STEM education R&D programs.

Research Questions

Research Question Guiding Phase I: What is a meaningful description of the impact of NSF investments in undergraduate STEM education R&D projects, based on Principal Investigators' (PIs)' perspectives?

a) What claims do PIs make about the impact of their NSF-funded projects? How do PIs support their claims about the impact of their work? (Qualitative)
b) How do PIs' perspectives of impact align with existing impact frameworks found in the literature to form a preliminary description of the impact of NSF investments in undergraduate STEM education projects? (Interpretation)

Research Question Guiding Phase II: In what ways do NSF Program Officers' (POs) perspectives on the impact of NSF investments in undergraduate engineering education R&D projects align with or differ from PIs' perspectives on impact?

a) To what extent do POs agree with PIs' perspectives on impact? (Quantitative)
b) How do POs talk about the impact of a NSF-funded R&D projects? (Qualitative)
c) Are there consistencies in how PIs on NSF-funded R&D projects and POs overseeing NSF's R&D programs talk about the impact of NSF-funded R&D projects? (Interpretation)

Research Design & Epistemological Assumptions

"Mixed method research is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis and the mixture of qualitative and quantitative approaches in many phases of the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches, in combination, provides a better understanding of research problems than either approach alone" (J. W. Creswell & Plano Clark, 2007, p. 5).

When discussed in the context of conducting research, worldviews are a set of ideas that influence various philosophical elements of a study (i.e., perspectives on the nature of reality, the relationship between the researcher and what is being researched, the role of values, the process of research, and the language used in the writing). A pragmatism paradigm aligns with mixed methods research (J. Creswell & Plano Clark, 2011; Tashakkori & Teddlie, 2003). This worldview draws on ideas like using "what works", using diverse approaches, and valuing both objective and subjective knowledge (J. Creswell & Plano Clark, 2011). Tashakkori and Teddlie (2003) formally argue the following points that link pragmatism and mixed methods research:

- A single study may include both quantitative and qualitative and methods.
- The research questions should be of primary importance, not the methods or philosophical paradigms that underlie them.
- The "forced-choice dichotomy" between constructivism and postpositivism should be abandoned.
- Metaphysical terms such as "truth" and "reality" should not be used.
- Methodological choices should be informed from a practical and applied research philosophy.

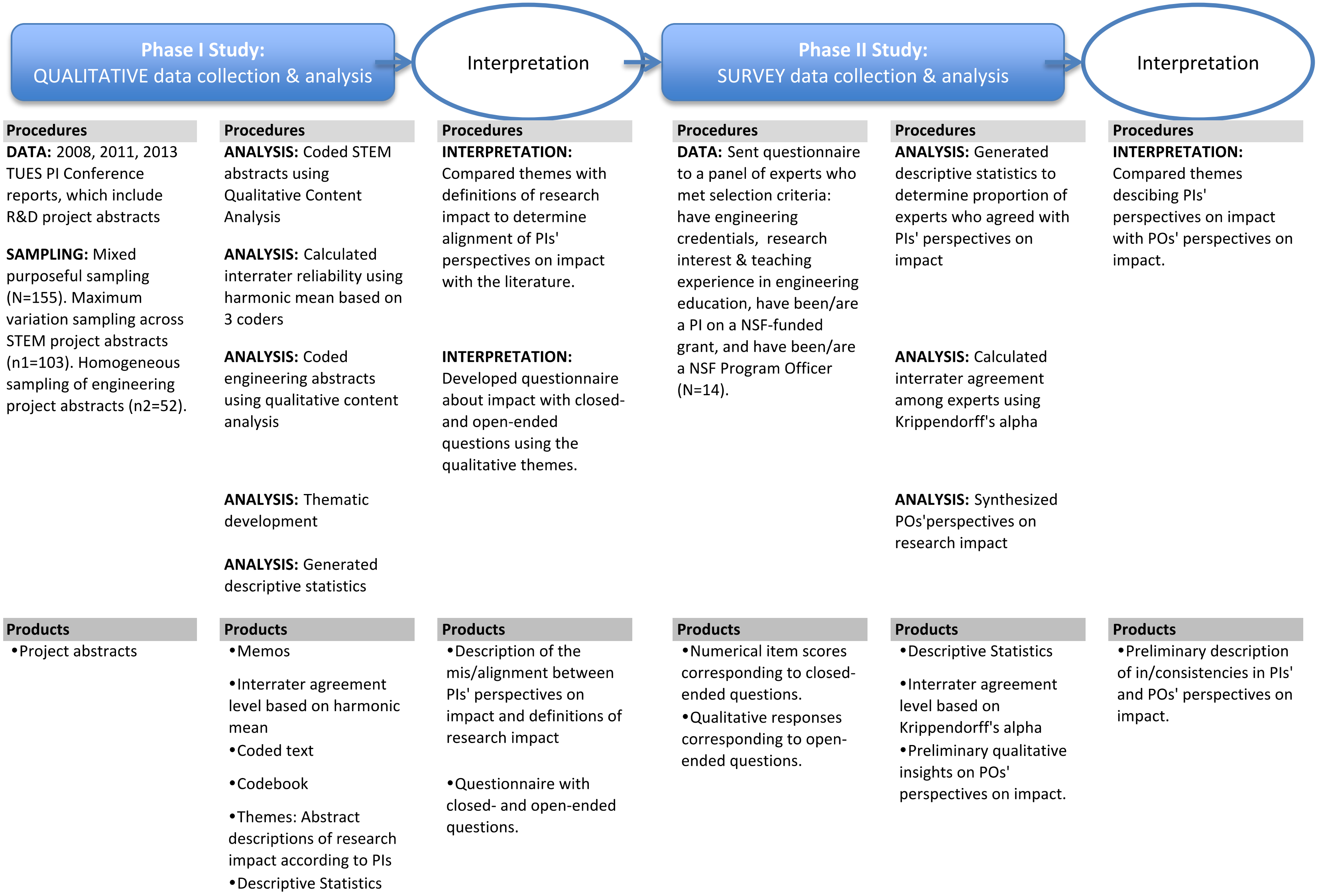
A multiphase mixed methods research design was used in this study. Multiphase mixed methods research designs provide an overarching methodological framework that includes sequentially aligned qualitative and quantitative studies to address a central research objective (J. Creswell & Plano Clark, 2011). In this study, this design consists of two distinct, sequential phases. The first phase is an exploratory phase where qualitative (text-based) data was collected and analyzed. The two phases are connected in the intermediate stage of the study: the results from the first phase informed the development of the questionnaire that will be used to collect data in the second phase. The second phase employed a questionnaire with open- and close-ended questions; thus the analysis and results includes quantitative and qualitative findings. The initial qualitative phase was prioritized in this study. The reason for mixing qualitative and quantitative methods was for the purpose of triangulation—to explore two groups' (PIs' and POs') perspectives on impact using different approaches, and to determine if there is corroboration between PIs and POs' perspectives on the impact of NSF-funded research (A. Bryman, 2006; Alan Bryman, 2006; J. Creswell & Plano Clark, 2011; Greene, Caracelli, & Graham, 1989).

The use of mixed methods in engineering education research has increased in prominence over recent years and engineering education researchers, Borrego, Douglas, and Amelink (2009), maintain that it is a research design approach that will be "essential" to advancing scholarship in this discipline. Since there is little literature on research impact, in general, and this topic has not been explored in the context of undergraduate engineering or STEM education research, an exploratory, qualitative study approach was an appropriate strategy for the first phase (Miles & Huberman, 1994). The survey research associated with the second phase is used to garner the perspectives of a select sample of NSF Program Officers for the purpose of determining the degree of consistency in PIs' and POs' opinions on research impact.

The six main data collection and analysis steps corresponding to this study's research design are (adapted from J. Creswell & Plano Clark, 2011, pp. 218-219):

1. Collect the qualitative data.
2. Analyze the qualitative data qualitatively using analytic approaches best suited for the qualitative research questions.
3. Develop and pilot test the questionnaire for the phase II study using the qualitative results.
4. Collect the questionnaire data.
5. Analyze the questionnaire data using analytic approaches best suited for the research questions.
6. Interpret how the connected results answer the research questions.

Figure 8 depicts these steps in relation to this study. The next chapter, Chapter 5, includes an elaboration on the methods associated with Phase I while Chapter 7 elaborates on the Phase II methods.



Phase I Study: QUALITATIVE data collection & analysis → Interpretation → **Phase II Study:** SURVEY data collection & analysis → Interpretation

Procedures	Procedures	Procedures	Procedures	Procedures	Procedures
DATA: 2008, 2011, 2013 TUES PI Conference reports, which include R&D project abstracts **SAMPLING:** Mixed purposeful sampling (N=155). Maximum variation sampling across STEM project abstracts (n1=103). Homogeneous sampling of engineering project abstracts (n2=52).	**ANALYSIS:** Coded STEM abstracts using Qualitative Content Analysis **ANALYSIS:** Calculated interrater reliability using harmonic mean based on 3 coders **ANALYSIS:** Coded engineering abstracts using qualitative content analysis **ANALYSIS:** Thematic development **ANALYSIS:** Generated descriptive statistics	**INTERPRETATION:** Compared themes with definitions of research impact to determine alignment of PIs' perspectives on impact with the literature. **INTERPRETATION:** Developed questionnaire about impact with closed- and open-ended questions using the qualitative themes.	**DATA:** Sent questionnaire to a panel of experts who met selection criteria: have engineering credentials, research interest & teaching experience in engineering education, have been/are a PI on a NSF-funded grant, and have been/are a NSF Program Officer (N=14).	**ANALYSIS:** Generated descriptive statistics to determine proportion of experts who agreed with PIs' perspectives on impact **ANALYSIS:** Calculated interrater agreement among experts using Krippendorff's alpha **ANALYSIS:** Synthesized POs'perspectives on research impact	**INTERPRETATION:** Compared themes descibing PIs' perspectives on impact with POs' perspectives on impact.
Products	**Products**	**Products**	**Products**	**Products**	**Products**
•Project abstracts	•Memos •Interrater agreement level based on harmonic mean •Coded text •Codebook •Themes: Abstract descriptions of research impact according to PIs •Descriptive Statistics	•Description of the mis/alignment between PIs' perspectives on impact and definitions of research impact •Questionnaire with closed- and open-ended questions.	•Numerical item scores corresponding to closed-ended questions. •Qualitative responses corresponding to open-ended questions.	•Descriptive Statistics •Interrater agreement level based on Krippendorff's alpha •Preliminary qualitative insights on POs' perspectives on impact.	•Preliminary description of in/consistencies in PIs' and POs' perspectives on impact.

Figure 8. Diagram of This Study's Multiphase Mixed Methods Research Design

CHAPTER 5: PHASE ONE METHODS

The overarching research question addressed in the qualitative phase of this study was: *What is a meaningful description of the research impact of NSF investments in undergraduate STEM education R&D projects, based on Principal Investigators' (PIs') perspectives?* Qualitative research is often referred to as *interpretivist research* (Stake, 2010a). According to Stake, "[I]nterpretivist research is investigation that relies heavily on observers defining and redefining the meanings of what they see and hear" (p. 36). Qualitative researchers often use social science theories and triangulate their data to minimize flaws in observations and assertions (J. W. Creswell, 2007; Stake, 2010a). *Toulmin's Model* (1958) and the *Common Guidelines for Education Research and Development* (Earle et al., 2013) are the two frameworks that guided the data collection, sampling, and analysis in this phase of the study.

The important criterion for identifying what constitutes data in a content analysis study are: the text provide useful evidence for answering the research questions, and that they communicate a message from a sender to a receiver (White & Marsh, 2006). Abstracts about NSF-funded undergraduate STEM education projects are the data that were used in this study. *Mixed purposeful sampling* (Patton, 1987 as cited by Johnson & Christensen, 2012, p. 237) was used to select which abstracts were analyzed. *Qualitative content analysis* (Elo & Kyngas, 2007; Krippendorff, 2004b; White & Marsh, 2006) is the analytic technique that was used to conduct the analysis in this study. The remaining sections of this chapter provide additional details about the data collection and analysis associated with this phase of the study.

Data Collection

One of the largest DUE programs was entitled "Transforming Undergraduate Education in STEM" (TUES). Its predecessor program was called "Curriculum, Course, and Laboratory Improvement" (CCLI). The CCLI/TUES program supported research and development projects focused on improving the quality of undergraduate STEM education for all students, with particular interest in projects that have the potential to transform undergraduate STEM education. (TUES was discontinued in 2013 and is now the predecessor of another DUE program called "Improving Undergraduate STEM Education" (IUSE)).

Periodically, DUE hosts a conference that all PIs on the program's active grants are encouraged to attend. The purpose of the PIs' conference is to provide PIs an opportunity to share updates on their projects and to exchange ideas with colleagues. In some cases, one of the outputs of the event is a conference report. The CCLI/TUES PIs conference reports from the most recent three consecutive PI conferences are publicly available online (NSF, 2008, 2011, 2013d) and were downloaded in Fall 2013. The project abstracts in these three reports are the population of data from which this study's sample data was selected to address the overarching research question proposed in this phase of the study.

Overview of TUES PI Conference Report Abstracts

Each report includes the abstracts of R&D projects presented at the conference. With little variations across reports, each abstract includes two major sections. One section includes basic project information: PI(s) name, Institution, Project Title, Project Number, Project Type, Target Discipline, and Focus. In addition to basic PI and project information, the abstract template provides space for PIs to add content corresponding to six topics: goals/goals & intended outcomes, methods/methods & strategies, evaluation/evaluation methods & results, dissemination, impact, and challenges. (The (/) denotes the little variations in labels across conference year reports.) While most of these labels are intuitive, some need additional explanation.

BASIC PROJECT INFORMATION- PROJECT TYPE: The CCLI/TUES program primarily funded four types of R&D projects. Project types vary by funding level and scope. The four main project types are: (1) Type 1 – Exploratory, (2) Type 2 – Expansion, (3) Type 3 – Comprehensive, and (4) Central Resource Project. (See Section C of the 2010 TUES program solicitation online for detailed descriptions of the project types.) Basic information about the CCLI/TUES projects by PI conference year is captured in Table 8. (The values in Table 8 are based on the tally provided by searching the project abstracts database provided on the CCLI/TUES PIs' conference website.)

Table 5. Number of CCLI/TUES Projects Included in PI Conference Reports, by Project Type

Project Type	**Conference Year**		
	2008	**2011**	**2013**
CCLI/TUES Type 1 - Exploratory	179	262	244
CCLI/TUES Type 2 - Expansion	78	102	114
CCLI/TUES Type 3 - Comprehensive	14	13	18
TUES Central Resource Project	0	2	3
Total Project Abstracts	**271**	**379**	**379**

BASIC PROJECT INFORMATION- TARGET DISCIPLINE: PIs on all CCLI/TUES projects must specify a target discipline. The eight STEM disciplines funded by DUE are: Biological Sciences, Chemistry, Computer Science, Engineering, Geological Sciences, Mathematics, Physics/Astronomy, and Social Science. Sub-disciplines within each of these disciplines (i.e., Industrial Engineering, Mechanical Engineering) is not uniquely specified; however, reading other contents of the abstract (e.g., title, goals) usually provide insights on which, if any- sub-discipline is of most interest. In light of this, engineering education projects are usually classified as "Engineering". When PIs perceive that their projects do not fit within either of these disciplines, they have the option to specify the discipline as: Interdisciplinary, Research/Assessment of Research, or Other.

Figure 9 displays the number of projects included in the three reports; the bars are organized by STEM discipline, project type, and conference year. One of the most salient insights that emerges from this visual representation of the data is Engineering awards account for the largest number of awards funded by the TUES program. This is more than twice as many as the second highest number of awards—Interdisciplinary awards. An equal number of projects have a specified discipline of Biological Sciences or Computer Science. The four disciplinary specifications with the fewest number of awards are: Geological Sciences, Social Science, Research/Assessment of Research, and Other.

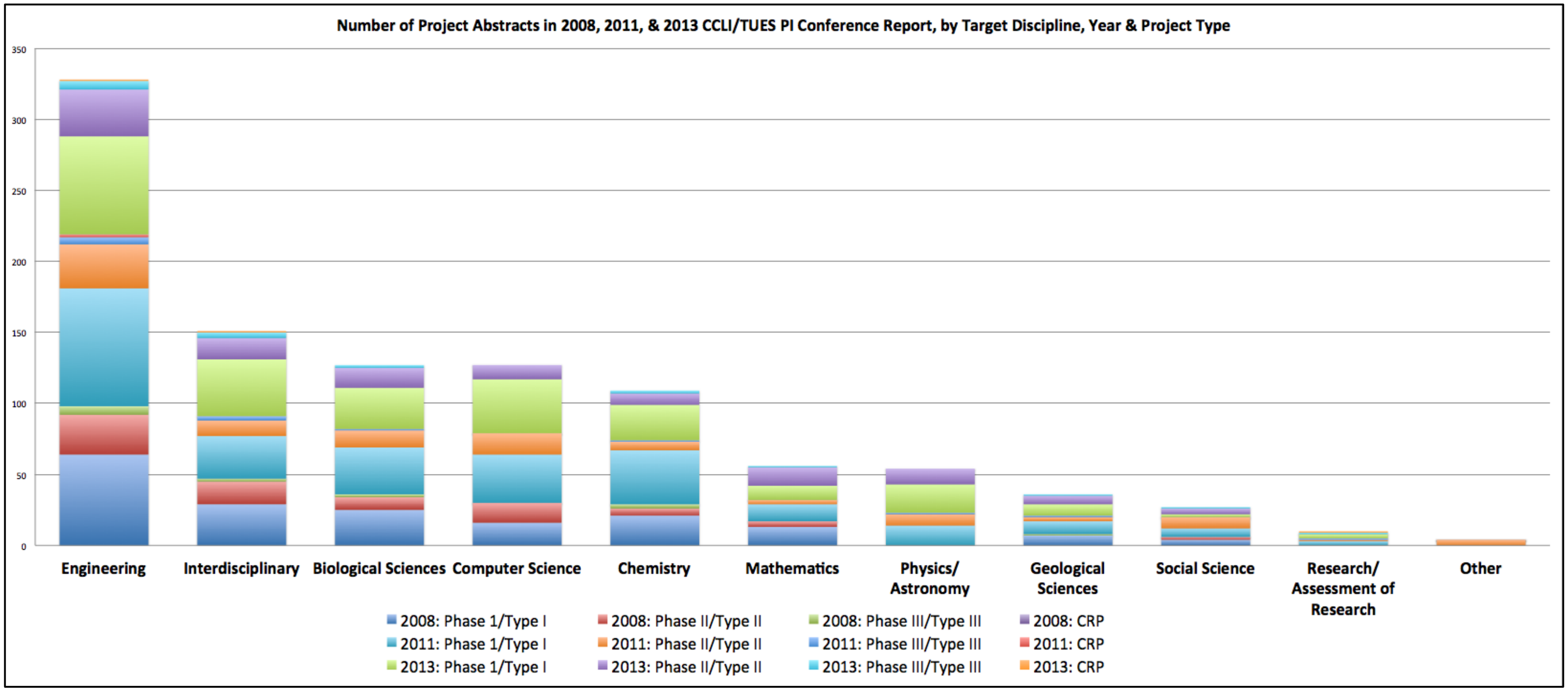
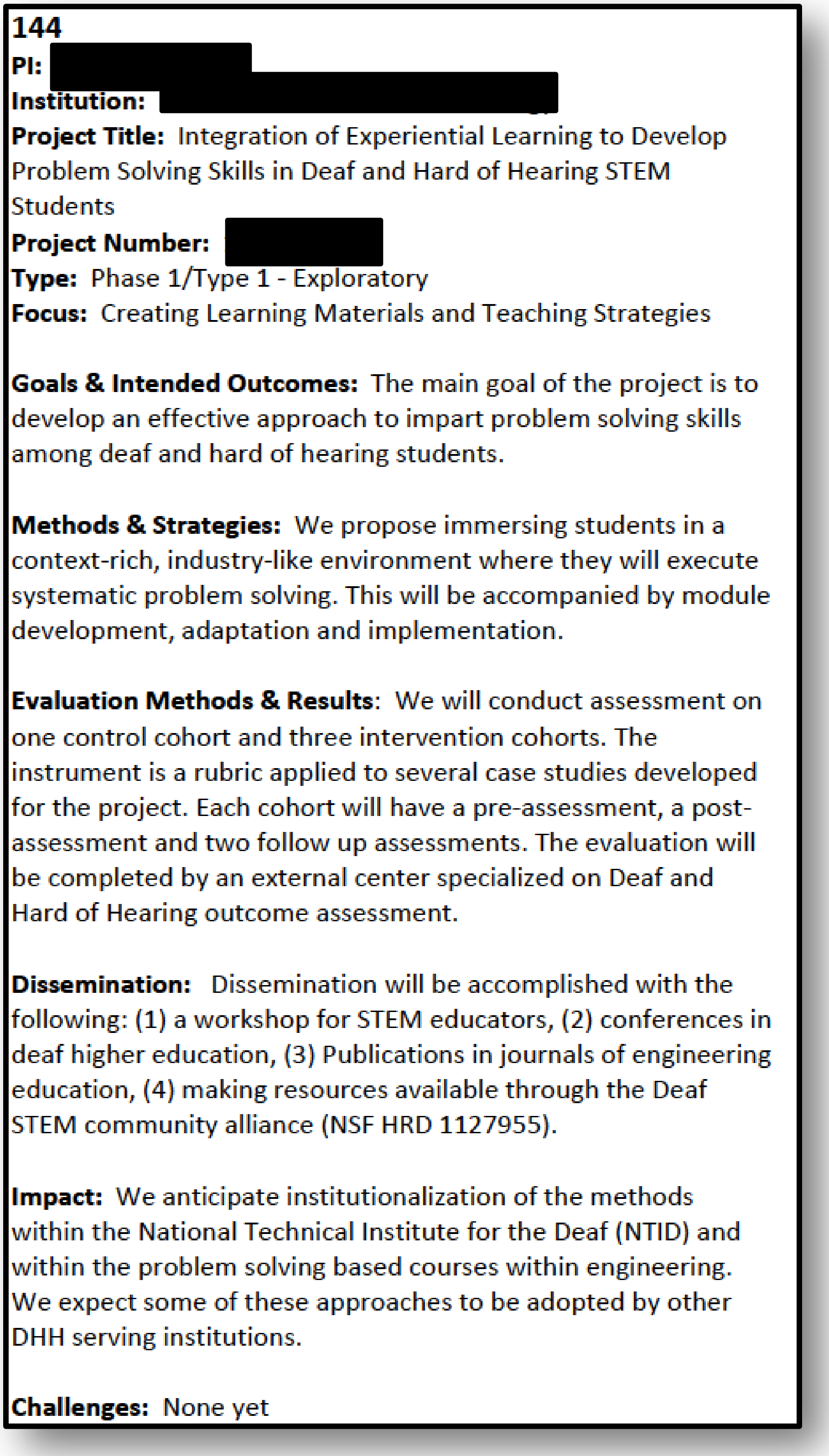


Figure 9. Project Abstracts in Conference Reports by Discipline, Year, Project Type

BASIC PROJECT INFORMATION- PROJECT FOCUS: PIs on all CCLI/TUES projects must specify a focus, an area of emphasis within STEM education. The five project foci specified in the CCLI/TUES solicitation are: Creating Learning Materials and Strategies, Implementing New Instructional Strategies, Developing Faculty Expertise, Assessing and Evaluating Student Achievement, and or Conducting Research on Undergraduate STEM Education. (Visit http://www.nsf.gov/pubs/2010/nsf10544/nsf10544.htm for 2010 TUES program solicitation, which includes detailed project foci descriptions.)

DETAILED PROJECT INFORMATION: In addition to basic PI and project information, the abstract template provides space for PIs to add content corresponding to six headings: goals/goals & intended outcomes, methods/methods & strategies, evaluation/evaluation methods & results, dissemination, impact, and challenges. (The (/) denotes the little variations in labels across conference year reports.) Figure 10 provides an example of an abstract randomly selected from the 2013 TUES PIs' Conference report (NSF, 2013a). (See Appendix A for a sample abstract from each of the three conference reports.)



144
PI: [redacted]
Institution: [redacted]
Project Title: Integration of Experiential Learning to Develop Problem Solving Skills in Deaf and Hard of Hearing STEM Students
Project Number: [redacted]
Type: Phase 1/Type 1 - Exploratory
Focus: Creating Learning Materials and Teaching Strategies

Goals & Intended Outcomes: The main goal of the project is to develop an effective approach to impart problem solving skills among deaf and hard of hearing students.

Methods & Strategies: We propose immersing students in a context-rich, industry-like environment where they will execute systematic problem solving. This will be accompanied by module development, adaptation and implementation.

Evaluation Methods & Results: We will conduct assessment on one control cohort and three intervention cohorts. The instrument is a rubric applied to several case studies developed for the project. Each cohort will have a pre-assessment, a post-assessment and two follow up assessments. The evaluation will be completed by an external center specialized on Deaf and Hard of Hearing outcome assessment.

Dissemination: Dissemination will be accomplished with the following: (1) a workshop for STEM educators, (2) conferences in deaf higher education, (3) Publications in journals of engineering education, (4) making resources available through the Deaf STEM community alliance (NSF HRD 1127955).

Impact: We anticipate institutionalization of the methods within the National Technical Institute for the Deaf (NTID) and within the problem solving based courses within engineering. We expect some of these approaches to be adopted by other DHH serving institutions.

Challenges: None yet

Figure 10. Abstract Example from 2013 TUES PIs' Conference Report

Aligning Abstract Sections and Guiding Frameworks

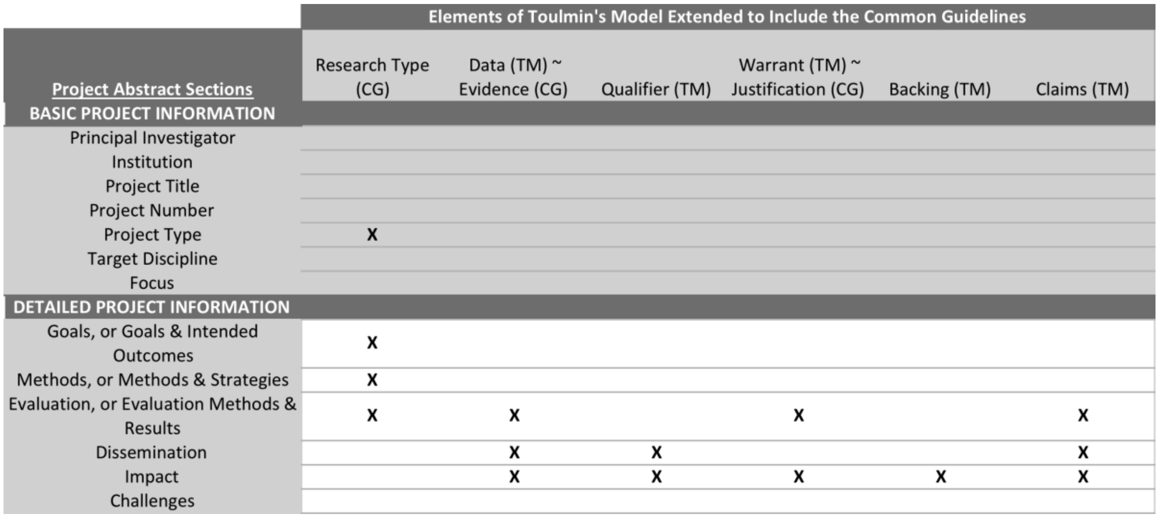
Tables 9 and 10 show how I (the researcher) conceptualized the alignment between the project abstract components and the integrated *Toulmin's Model/Common Guidelines* framework depicted in Figure 7. The scope of the four Project Types loosely aligns with the three research types in the *Common Guidelines* (Earle et al., 2013); this alignment is shown in Table 9. More specifically, the description of the scope of TUES Type 1, 2, and 3 projects align with the description of the three research types in the *Common Guidelines* (Earle et al., 2013), respectively; and the scope of a TUES Central Resource project aligns with the third research type in the *Common Guidelines* (Earle et al., 2013).

Table 6. Alignment of Common Guidelines Framework and CCLI/TUES Project Type

Common Guidelines (Earle et al., 2013) Research Type(s)	**CCLI/TUES Program**
	Project Type
1-2: Foundational Research and Early Stage or Exploratory Research	CCLI/TUES Type 1 – Exploratory
3: Design and Development Research	CCLI/TUES Type 2 – Expansion
4-6: Efficacy, Effectiveness, and Scale-Up Research	CCLI/TUES Type 3– Comprehensive TUES Central Resource Project

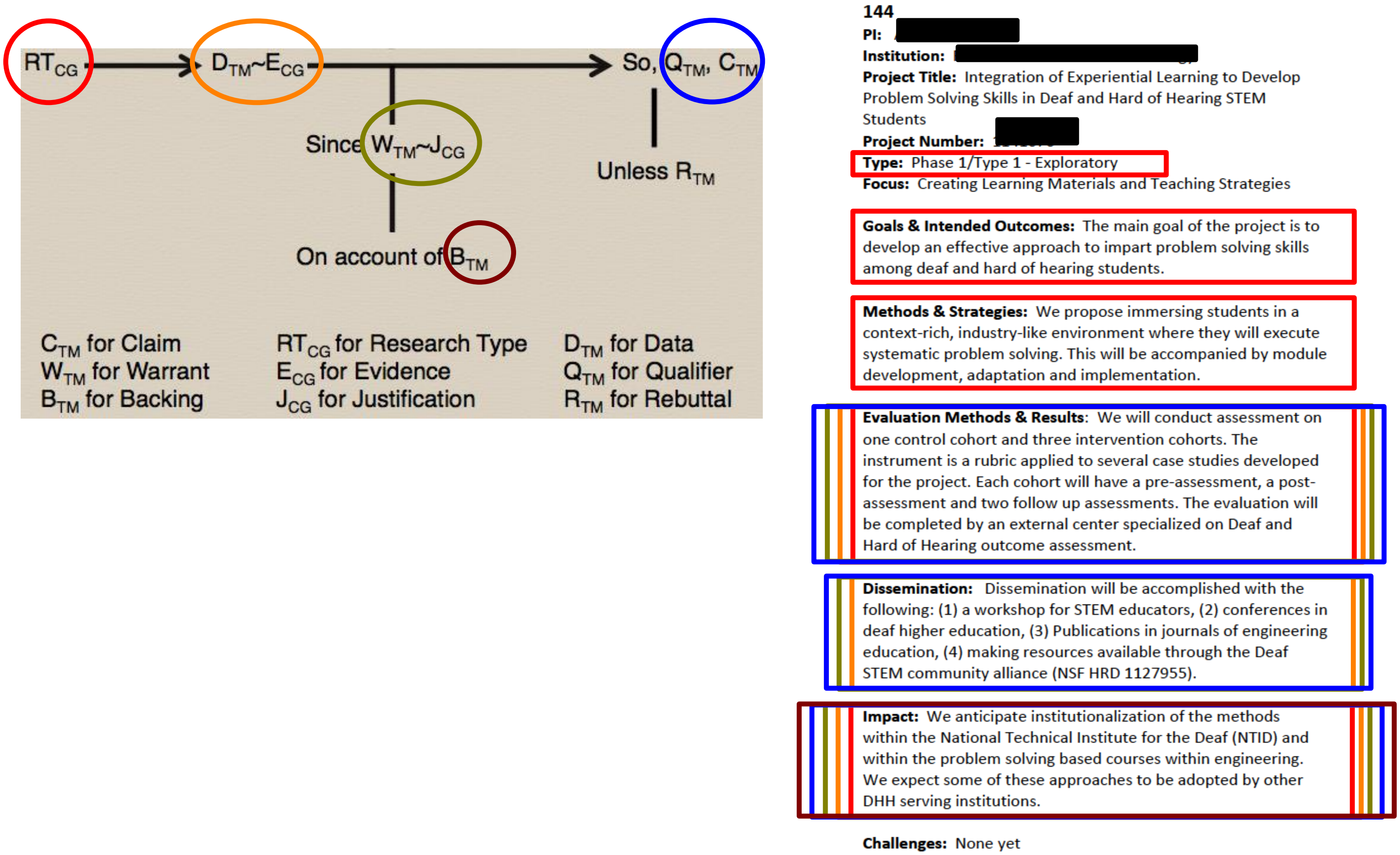
Since *Toulmin's Model* (Toulmin, 1958) is useful for understanding verbal reasoning, in general, and the *Common Guidelines* (Earle et al., 2013) provides additional details on the contents of the verbal reasoning found in project abstracts, an integration of the two is useful for understanding claims about research impact in the context of interest. Table 10 shows how I conceptualized the alignment between the elements of the abstracts and the two guiding frameworks; Figure 11 shows the alignment using an abstract example.

Table 7. Anticipated Alignment of Project Abstract and Guiding Frameworks



	Elements of Toulmin's Model Extended to Include the Common Guidelines					
Project Abstract Sections	Research Type (CG)	Data (TM) ~ Evidence (CG)	Qualifier (TM)	Warrant (TM) ~ Justification (CG)	Backing (TM)	Claims (TM)
BASIC PROJECT INFORMATION						
Principal Investigator						
Institution						
Project Title						
Project Number						
Project Type	X					
Target Discipline						
Focus						
DETAILED PROJECT INFORMATION						
Goals, or Goals & Intended Outcomes	X					
Methods, or Methods & Strategies	X					
Evaluation, or Evaluation Methods & Results	X	X		X		X
Dissemination		X	X			X
Impact		X	X	X	X	X
Challenges						

The “Challenges” section of the abstract is a place for PIs to discuss any practical difficulties they have faced while trying to complete the study (e.g., delays in getting access to a student population). These ideas are not related to reasoning surrounding the research itself, and as a result, do no map to the theoretical and conceptual frameworks guiding this study. (The *Data Analysis* section provides details on how this alignment informed the analysis.)



144
PI: [redacted]
Institution: [redacted]
Project Title: Integration of Experiential Learning to Develop Problem Solving Skills in Deaf and Hard of Hearing STEM Students
Project Number: [redacted]
Type: Phase 1/Type 1 - Exploratory
Focus: Creating Learning Materials and Teaching Strategies

Goals & Intended Outcomes: The main goal of the project is to develop an effective approach to impart problem solving skills among deaf and hard of hearing students.

Methods & Strategies: We propose immersing students in a context-rich, industry-like environment where they will execute systematic problem solving. This will be accompanied by module development, adaptation and implementation.

Evaluation Methods & Results: We will conduct assessment on one control cohort and three intervention cohorts. The instrument is a rubric applied to several case studies developed for the project. Each cohort will have a pre-assessment, a post-assessment and two follow up assessments. The evaluation will be completed by an external center specialized on Deaf and Hard of Hearing outcome assessment.

Dissemination: Dissemination will be accomplished with the following: (1) a workshop for STEM educators, (2) conferences in deaf higher education, (3) Publications in journals of engineering education, (4) making resources available through the Deaf STEM community alliance (NSF HRD 1127955).

Impact: We anticipate institutionalization of the methods within the National Technical Institute for the Deaf (NTID) and within the problem solving based courses within engineering. We expect some of these approaches to be adopted by other DHH serving institutions.

Challenges: None yet

Figure 11. Conceptual Alignment Between the Integrated Guiding Framework & Abstract Sections

Rationale for Data Source

There were many reasons for using the abstracts in the CCLI/TUES PI conference reports as data—many of which are pragmatic, but some are more meaningful. Since the reports include abstracts of over 1,000 STEM education R&D projects funded by DUE, they provided a generous amount of data from which to sample, both from a qualitative and quantitative methods perspective. Additionally, the reports correspond to all of the CCLI/TUES PI conferences that have occurred over the last five years; this presents a longitudinal perspective on the central topic of research impact. The last pragmatic reason is because this data is conveniently available online; this translated to no delay in collecting data. While the annual and final project reports would have been a more comprehensive source of information about the impact of NSF-funded projects, this information is stored in a repository that is only viewable by NSF employees, and thus not available to the public for review, critique, or research use.

Apart from pragmatism, there were more meaningful reasons for using this data. This is appropriate data source given that this study focuses on engineering education research, and the CCLI/TUES program was the largest funder of this research. Moreover, projects funded by the CCLI/TUES program focused on a wide range of needs and challenges that exist across undergraduate STEM education; as a result, insights about impact that result from this study have the potential to be just as broad in their applicability. Furthermore, there is an alignment between the project types, abstract elements and the frameworks guiding this study. This alignment helped with data analysis and interpretation. Also, the abstracts are written by the PIs on CCLI/TUES projects, and as a result, provides the data will answer the overarching research question (on how PIs talk about the impact of their research.) Moreover, the "Impact" section of the abstract includes PIs' perceptions of the *realized* impact of their project, not the proposed impact described in grant proposals. This is one distinction between this study and the previous NSF-focused studies mentioned in the literature review (Kamenetzky, 2013; Roberts, 2009). Another rationale for this data source is related to one of the difficulties with studying impact: illusive data. Impact narratives in PI conference reports are one of the few instances where the impact of a project is explicitly documented, and thus available for study. Finally, and possibly most importantly, project abstracts are the primary source of public information about NSF-awards (especially those funded before www.research.gov, a resource that provides the research community with information on federal grants, was launched in 2013). The audience of NSF abstracts includes Program Officers, the PI community, Congress, the media, and the public at large. As a result, project abstracts are key to advancing transparency about NSF's investments in undergraduate STEM education. Conducting a study on the claims about impact being made in this primary source of information is an appropriate first step toward improving understanding of the impact of federal investments in undergraduate engineering education.

Sampling

Mixed purposeful sampling (Johnson & Christensen, 2012; Patton, 2002) was used to select the abstracts analyzed in this study. Mixed purposeful sampling is the mixing of more than one sampling strategy, in this case, two sampling strategies. The two sampling strategies used are maximum variation sampling and homogeneous sampling (Johnson & Christensen, 2012; Miles & Huberman, 1994; Patton, 2002).

Maximum variation sampling is purposefully selecting a wide range of cases. It is useful for identifying "a central theme or pattern that exists across cases" (Johnson & Christensen, 2012, p. 236). For the purposes of this study, this meant selecting abstracts across conference years, project types, project foci, and STEM disciplines. The reason for sampling across STEM disciplines was because engineering is rarely discussed in isolation. The field, as it is currently conceived, is part of the larger set of disciplines of national interest, commonly referred to as STEM. The focus of this study is engineering education research, with the expectation that the results can be easily translated to other STEM disciplines. In light of this, the majority of the abstracts specified a target discipline across STEM fields. Additionally, there is no evidence to suggest that the claims PIs make about impact and information used to support it would vary significantly across STEM disciplines. On the other hand, since the focus of this study is engineering education, homogeneous sampling was also used to select a smaller set of abstracts where Engineering was the target discipline.

Sample size is just as important as the sampling strategy. In total, the sample included approximately fifteen percent of the total abstracts (155 of the 1029 abstracts). Ten percent of the abstracts were selected using maximum variation sampling. An additional five percent of the total abstracts were selected based on the homogeneous sampling strategy. Although the objective was the review 15% of the total abstracts, the ultimate number of abstracts was determined by the point of saturation, the point at which no new or relevant information emerged as a result of analyzing more data (J. W. Creswell, 2007). (See Appendix B for a table that indicates which abstracts were included in the analysis.) Figures

12-14 present the proportion of abstracts analyzed in this study based on project type, project focus, and STEM discipline, respectively.

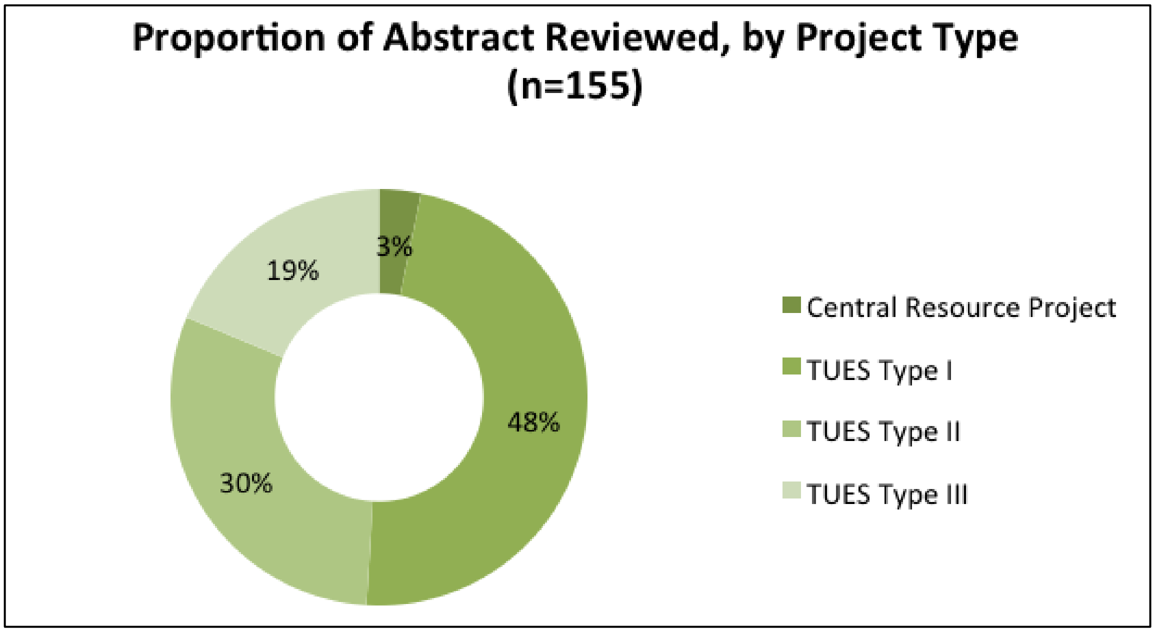


Figure 12. Abstracts Reviewed by Project Type

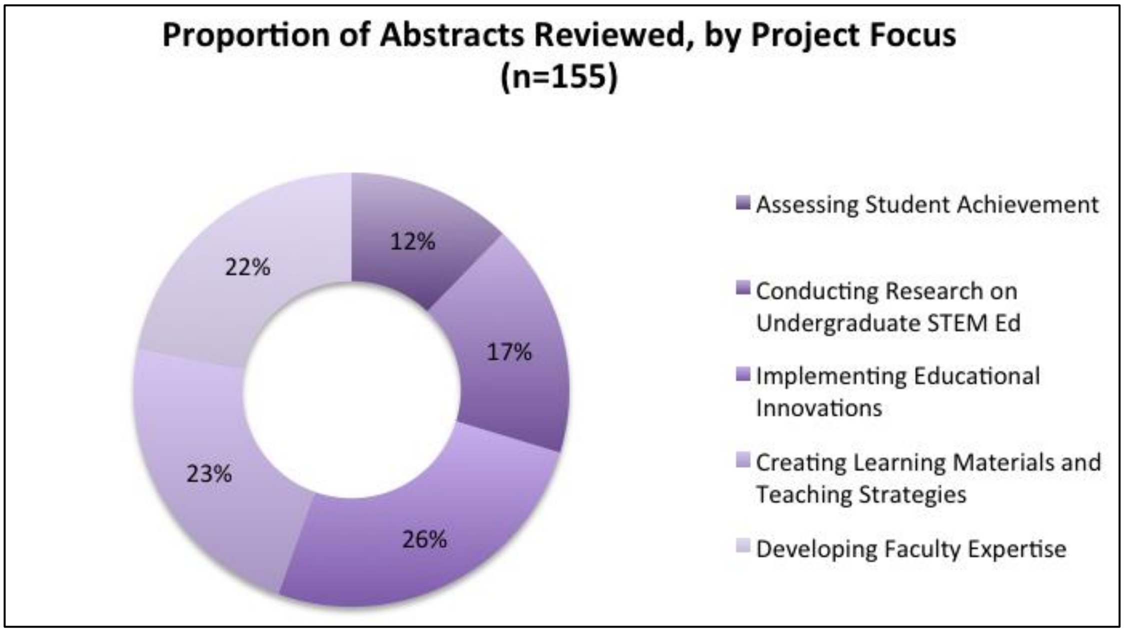


Figure 13. Abstracts Reviewed by Project Focus

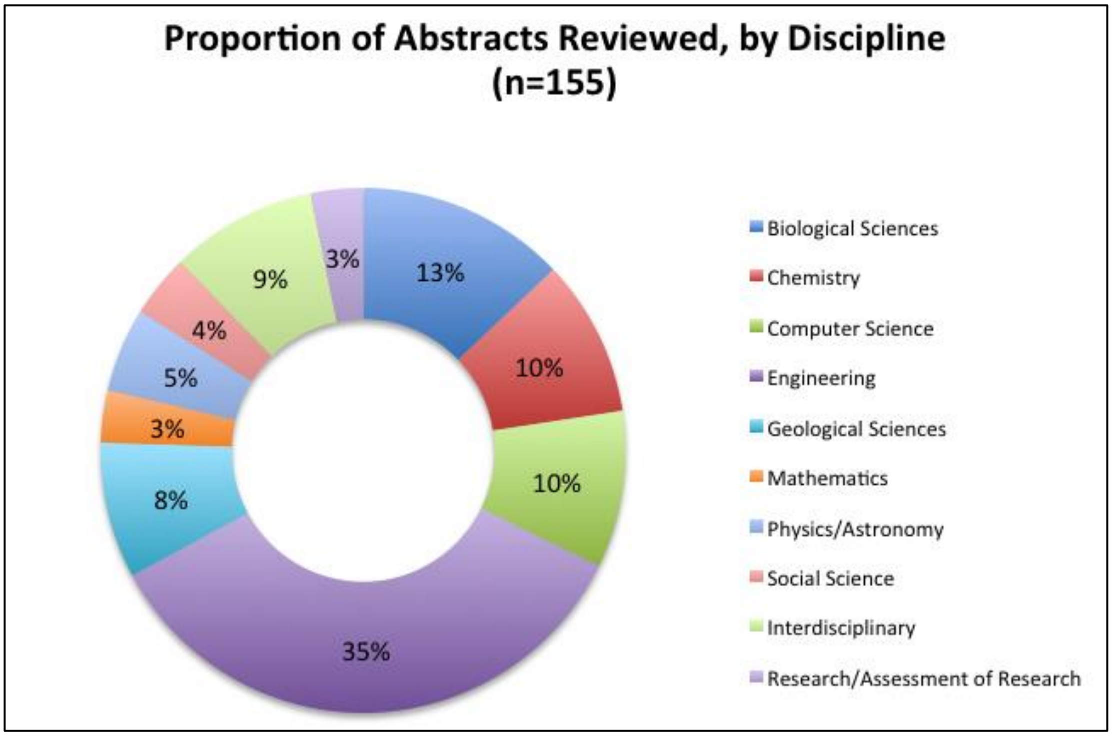


Figure 14. Abstracts Reviewed by Discipline

Data Analysis

The two sub-questions corresponding to an overarching research question proposed in this phase are:

a) What claims do PIs make about the impact of their NSF-funded projects? How do PIs support their claims about the impact of their work? (Qualitative)

b) How do PIs' perspectives of impact align with the three dimensions of research impact found in the literature (i.e., scientific, societal, and domain-specific impact) to form a preliminary description of the impact of NSF investments in undergraduate STEM education projects? (Interpretation)

The first sub-question was answered using content analysis, while the second was addressed by comparing the results of the first question with the definitions of research impact in the literature (See Figure 2).

Content analysis is a "research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (Krippendorff, 2004b, p. 18). According to Elo and Kyngas (2007), content analysis can be used as an inductive or deductive research strategy, depending on the purpose of your research. Inductive content analysis is typically used if there is little former knowledge about the phenomenon of interest or if the knowledge is fragmented (Lauri & Kyngas, 2005 as cited by Elo & Kyngas, 2007). In light of this, inductive content analysis was used in this study to understand the claims (assertions) PIs make about the impact of their work and how they support those claims.

There are three main phases in the data analysis: preparation, organizing and reporting. Atlas.ti, qualitative data analysis software, was used to perform the majority of the analysis. In the preparation phase, the sample of abstracts were copied from the individual PI conference reports and pasted into a single document. Although I read the entire abstract when conducing the analysis, the unit of analysis were the two sections that are most likely to contain claims about research impact: the Dissemination section and the Impact section. (Figure 11 above shows the alignment between abstract components and the guiding frameworks.) Abstract sections that include research "results" were not analyzed because although the contents could potentially be related to the project's impact, the outcomes of this analysis would have been a duplication of efforts currently being conducted by engineering education researchers developing a taxonomy of engineering education research (Finelli & Borrego, 2014).

Ideas mentioned in the literature review –namely the three dimensions of research impact, and the 76 impact categories in other research impact frameworks-- served as *sensitizing concepts* (Bowen, 2006) and *provisional codes* (Saldaña, 2009). Sensitizing concepts are "interpretive devices" that provide "a starting point for a qualitative study" (Glaser, 1978; Padgett, 2004; Patton, 2002, 2004 as cited by Bowen, 2006, p. 2). They are "ideas in the background that inform the overall research problem" and offer ways of organizing and understanding the text (Bowen, 2006). In qualitative research, "codes" are labels used to describe a segment of text. Provisional codes are those that were established before the analysis began, and were possible categories that might have been reflected in the text (Dey, 1993; Miles & Huberman, 1994 as cited by Saldaña, 2009). See Table 11 for the list of initial codes. These codes were put into the Atlas.ti codebook before the analysis began. However, these categories were not imposed on the text as *a priori* codes (as it would be in a deductive content analysis); I used open and axial coding to allow concepts and patterns discussed in the research findings to emerge from the data.

Table 8. Starting List of Codes Used in Atlas.ti

Research Impact Dimensions
Research Impact- Scientific
Research Impact- Domain-specific
Research Impact-Societal
Impact Dimensions from Existing Frameworks
Awareness, knowledge or understanding (of) STEM concepts, processes, or careers
Engagement or interest (in) STEM concepts, processes, or careers
Attitudes (towards) STEM-related topics or capabilities
Behavior (related to) STEM concepts, processes, or careers
Skills (based on) STEM concepts, processes or careers
Science: knowledge, research activities, training
Technology: products and processes, services, know-how
Economy: production, financing, investments, commercialization, budget
Culture: Knowledge, know-how, attitudes, values
Society: welfare, discourses and actions of groups
Policy: policy-makers, citizens, public programs, national security
Organization: planning, work organization, administration, human resources
Health: public health, health systems
Environment: management of natural resources and the environment, climate & meteorology
Training: curricula, pedagogical tools, qualifications, graduates, insertion into the job market, fitness of training/work, career, use of acquired knowledge
Symbolic: legitimacy/credibility/visibility, notoriety
Academic impact
Impact on public knowledge creation
Impact on policy
Impact on the preservation of heritage
Economic impact on the wider society
Impact on leisure and entertainment
Direct economic impact
Impact Dimensions from Existing Frameworks (Continued)
Impact through teaching
Cross-cutting findings
Knowledge
Benefits of future research and research use
Benefits from informing policy and product development
Broader economic benefits
Health and health sector benefits
Research-related impacts: Type of problem/knowledge
Research-related impacts: Research methods
Research-related impacts: Publications and papers
Research-related impacts: Products, patents, and translatability potential
Research-related impacts: Research networks
Research-related impacts: Leadership and Awards
Research-related impacts: Communication
Policy impacts: Level of policy-making
Policy impacts: Type of policy
Policy impacts: Nature of policy impact
Policy impacts: Policy networks
Policy impacts: Political capital
Societal Impacts: Types of service: health/intersectoral
Societal Impacts: Evidence-based practice
Societal Impacts: Quality of care
Societal Impacts: Information systems
Societal Impacts: Services management
Societal Impacts: Cost-containment and cost-effectiveness
Service Impacts: health and intersectoral: Knowledge, attitudes, and behavior
Service Impacts: health and intersectoral: Health literacy
Impact Dimensions from Existing Frameworks (Continued)
Service Impacts: health and intersectoral: Health status
Service Impacts: health and intersectoral: Equity and human rights
Service Impacts: health and intersectoral: Macroeconomic/related to the economy
Service Impacts: health and intersectoral: Social capital and empowerment
Service Impacts: health and intersectoral: Culture and art
Service Impacts: health and intersectoral: Sustainable development outcomes
Increasing the stock of useful knowledge
Creating new scientific instrumentation and methodologies
Creating new firms
Training skilled graduates
Forming networks and stimulating social interactions
Increasing the capacity for scientific and technological problem solving
Changes in knowledge and understanding
Changes in access to research
Changes in the extent to which research is considered, referred to, or read
Citations in documents
Changes in attitudes and beliefs
Changes in behavior
Advances in knowledge
Additional investment
Financial return
Economic impact
Social impact
Environmental impact
More effective teaching
Intangible impacts

The second step in inductive content analysis is the organizing phase (Elo & Kyngas, 2007), or what is commonly referred to as the coding process (J. W. Creswell, 2008; Miles & Huberman, 1994). The coding process is “a qualitative research process in which the researcher makes sense out of text data, divides it into text or image segments, labels the segments, examines codes for overlap and redundancy, and collapses these codes into broad themes” (J. W. Creswell, 2008, p. 251). Coding is an iterative process, and one the first outputs of it is a codebook. A codebook includes an organized list of the codes resulting from the analysis, along with descriptions of the codes, examples corresponding to the codes, and coding instructions on how to apply the codes to text segments to ensure that the coding process is structured and applied consistently across abstracts. In this study, the process of coding the abstracts (in Atlas.ti) and developing the codebook (in Microsoft Word) happened in parallel; this process is depicted in Figure 15.

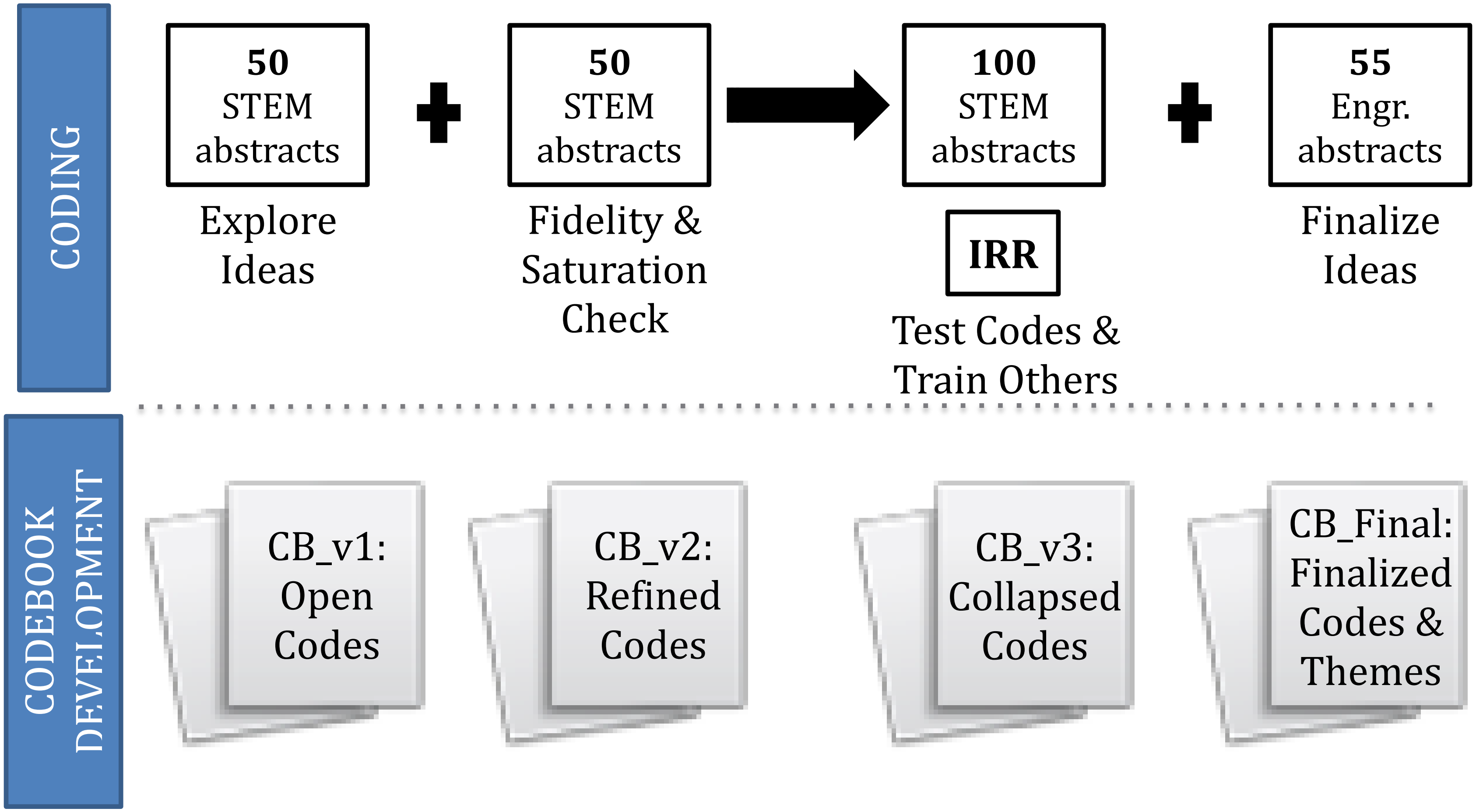


Figure 15. Overview of the Coding & Codebook Development Process

The initial coding was an exploratory analysis to get a general sense of the data (J. W. Creswell, 2008; Miles & Huberman, 1994) about one-third of the abstracts. After reading the abstract, two types of coding occurred: Attribute coding, and Provisional or Descriptive coding (Saldaña, 2009). Attribute coding is the way to denote basic information about the text (Bazeley, 2003; DeWalt & DeWalt, 2002; Gibbs, 2002; Lofland et al., 2006 as cited by Saldaña, 2009). In this study, the following attribute codes were assigned to the abstract: project type, project focus, and STEM discipline. Next, either provisional or descriptive codes were assigned to the ideas in the "Dissemination" and "Impact" sections of the abstract. While coding, I asked myself questions like: *What claims (assertions) are PIs making about impact of their research? What types of things are PIs referencing? How are they supporting or discussing those claims? Are there consistencies in the ideas they use to articulate impact?* Elements of the two guiding frameworks also served as sensitizing concepts when looking for answers to these questions. To the extent that the ideas in these sections aligned with the existing categories of impact, a provisional code was assigned (see Table 8). For instances where new ideas emerged that do not align with those reflected in the literature or could benefit from additional clarification, the segment was open coded using descriptive codes— short statements that summarize the topic of the text (Miles & Huberman, 1994; Saldaña, 2003; Wolcott, 1994 as cited by Saldaña, 2009). The Notes feature in Atlas.ti and handwritten notes in a research journal were used to keep track of memos (i.e. reflective remarks) during the coding process (Miles & Huberman, 1994). Figure 16 shows an example of how an abstract was coded at this stage in the coding process, where "PC" indicates the provisional codes assigned to the text segments, and "OC" indicates the open codes assigned to the segment.

144
PI: Andres Carrano
Institution: Rochester Institute of Technology
Project Title: Integration of Experiential Learning to Develop Problem Solving Skills in Deaf and Hard of Hearing STEM Students
Project Number: 1141076
Type: Phase 1/Type 1 - Exploratory
Focus: Creating Learning Materials and Teaching Strategies

Goals & Intended Outcomes: The main goal of the project is to develop an effective approach to impart problem solving skills among deaf and hard of hearing students.

Methods & Strategies: We propose immersing students in a context-rich, industry-like environment where they will execute systematic problem solving. This will be accompanied by module development, adaptation and implementation.

Evaluation Methods & Results: We will conduct assessment on one control cohort and three intervention cohorts. The instrument is a rubric applied to several case studies developed for the project. Each cohort will have a pre-assessment, a post-assessment and two follow up assessments. The evaluation will be completed by an external center specialized on Deaf and Hard of Hearing outcome assessment.

Dissemination: Dissemination will be accomplished with the following: (1) a workshop for STEM educators, (2) conferences in deaf higher education, (3) Publications in journals of engineering education, (4) making resources available through the Deaf STEM community alliance (NSF HRD 1127955).

Impact: We anticipate institutionalization of the methods within the National Technical Institute for the Deaf (NTID) and within the problem solving based courses within engineering. We expect some of these approaches to be adopted by other DHH serving institutions.

Challenges: None yet

PT: TUES Type 1
PF: Creating Learning Materials & Teaching Strategies
STEM Discipline: Engineering

PC: Research Impact: Scientific Impact
PC: Science: knowledge, research activities, training
PC: Academic impact
PC: Impact on public knowledge creation
PC: Research-related impacts: publications and papers
PC: Research-related impacts: Communication
PC: Increasing the useful stock of knowledge
PC: Changes in knowledge or understanding
PC: Advances in knowledge
OC: Host a Workshop
OC: Participate in a Conference
OC Share Resources via a Partnership
Memo: “Dissemination will be accomplished” denotes future tense

PC: Research Impact: Domain-specific impact
PC: Training: curricula, pedagogical tools…
PC: Academic impact
PC: Impact on policy
PC: Benefits from informing policy and product development
PC: Policy impact: level of policy-making
PC: Policy impact: nature of policy-making
PC: Changes in behaviors
OC: Widespread use of research outputs by a specified group (x2)
OC: Adoption of research outputs at similar institutions

Memo: “We anticipate” and “We expected” does not denote realized impact

Figure 16. Coding Example using Codebook, Version 1

Miles and Huberman (1994) have commented on the importance of having a quality codebook. During the exploratory phase, I used Microsoft Word to create a duplicate codebook as codes are assigned to segments of text in Atlas.ti. This is the time during the coding process when the provisional and open codes were added to the document, sample segments of text were supplied, and coding guidelines were drafted to promote consistency in the application of codes across abstracts. Once the coding of this set of abstracts was complete in Atlas.ti, the MS word version of the codebook (CB_v1:Open Codes) was used to code the second set of STEM abstracts. When necessary, new codes, examples, and guidelines were added to the codebook.

Once the coding for all of the abstracts identified in the maximum variation sampling was complete (i.e., from projects across STEM), the MS Word version of the codebook will be refined and preliminary categories were described. This act of going from codes to categories (to abstractions) is more iterative than linear. Using an analytic technique commonly referred to among grounded theory researchers as a "constant comparative" approach (J. W. Creswell, 2007; Strauss & Corbin, 1998), I collapsed together codes that conveyed redundant ideas, and began to identify a category/label for the group of codes that remained. Memos documented during the coding process were used to facilitate the refinement of the codebook. One of the most critical decisions that was made in revising the codebook was the created a category and cluster of codes for impacts that were discussed in the impact narratives, but had not been realized. In an attempt to present an accurate description of how PIs talk about impact, it seemed imperative to distinguish between the types of impact that actually occurred, but what the PIs hoped would happen at some point in the future. These changes resulted in the second version of the codebook (CB_v2:

Refined Codes). Table 9 is an example of how the list of codes used in the coding example was reduced at this point of refining the codebook. There were a total of 114 codes in the codebook: 19 attribute codes, and 95 descriptive and provisional codes. There were seven themes corresponding to the first sub-questions (about the types of claims PIs make), and three corresponding to second sub-question (about how PIs support claims about impact).

Table 9. Sample of Original vs. Refined Codes Across First Two Versions of the Codebook

CB_v1: Open Codes	CB_v2: Refined Codes
Research Impact- Scientific	Research Impact- Scientific
Science: knowledge, research activities, training	Highlights of research findings (new data, methods, insights)
Academic impact	Plans for future dissemination - written communications
Impact on public knowledge creation	Plans for future dissemination - oral communications
Research-related impacts: Publications and papers	
Research-related impacts: Communication	
Increasing the stock of useful knowledge	
Changes in knowledge and understanding	
Advances in knowledge	
Host a workshop	
Participate in a conference	
Share resources via a Partnership	
Research Impact- Domain-specific	Research Impact- Domain-specific
Training: curricula, pedagogical tools…	Adoption of research outputs (curricula, pedagogy)
Impact on policy	Changes in policy
Benefits from informing policy and product development	Expected/Anticipated Impact
Policy impacts: Level of policy-making	
Policy impacts: Nature of policy impact	
Changes in behavior	
Widespread use of research outputs by a specific group	
Adoption of research outputs at similar institutions	

An interrater reliability (IRR) check was performed before coding the engineering-only set of abstracts; the timing is consistent with Miles and Huberman (1994) recommendation to conduct the IRR testing when approximately two-thirds of the coding is complete.

Interrater Reliability Check

Interrater reliability is “near the heart of content analysis; if the coding is not reliable, the analysis cannot be trusted” (Singletary, 1993, p. 294). Interrater reliability (also referred to as intercoder reliability) is a term that is commonly used to describe “the extent to which the different judges tend to assign exactly the same rating to each object” (Tinsley & Weiss, 2000, p. 98). Despite its importance, there are few standards or “rules of thumb” on how to properly calculate and report IRR, and it is difficult to find information on the few IRR software tools that are available (Lombard, Snyder-Duch, & Bracken, 2002). Some of the most commonly used methods to calculate IRR are: percent agreement, Holsti’s Method (Holsti, 1969), Scott’s Pi (π) (W. Scott, 1955), Cohen’s Kappa (κ) (J. A. Cohen, 1960; J.A. Cohen, 1968), Fliess’ Kappa (κ) (Fleiss, 1971), and Krippendorff’s Alpha (α) (Krippendorff, 1970, 2004a). When summarizing the literature on how IRR approaches compare to one another, Lombard et al. (2002) states “there is a general agreement that indices which do not account for chance agreement are too liberal while those that do are too conservative” (Lombard et al., 2002, p. 593).

While these traditional methods address issues of chance and what to do if there are multiple raters, they are only useful for texts that belong to mutually exclusive categories. They do not account for instances when multiple codes may be assign to the same text–which is the case in this study because one project abstract almost always includes more than one form of research impact. Typically, the F_1 score (Devore, 2012) is used in contexts where categories are not mutually exclusive. A F_1 score is the harmonic mean between two data sets.

Two coders agreed to participate in the IRR process. The harmonic mean was calculated to determine the extent of agreement between the researcher and each of the two independent coders, and between the two independent coders. More specifically, the IRR calculation is as follows: assume there are a total of N abstracts coded as part of the IRR analysis. For the i-th abstract: x_{1i} represents the number of codes assigned to the abstract by the coder A; x_{2i} denotes the number of codes assigned to the abstracts by coder B; and s_i represents the number of codes that are agreed upon between researchers A and B. Let $p_{1i} = s_i/x_{1i}$, and $p_{2i} = s_i/x_{2i}$, then

$$F_1 = \frac{1}{N}\sum_{i=1}^{N}\frac{2p_{1i} \cdot p_{2i}}{p_{1i} + p_{2i}}$$

Thus, the F_1 score is 1 when the code assignments between the two coders' data sets are exactly the same, and a F_1 score of 0 indicates no agreement. In this context, this measure represents the closeness of two sets of codes assigned to the same set of N abstracts by different researchers.

To prepare for the IRR testing, I invited the coders to participate, identified abstracts to be included in the testing, and prepared notes for discussion that would proceed the coding. The two independent coders were fellow-PhD students in the Engineering Education program, who have academic backgrounds in engineering disciplines, and have been involved in the coding of at least three qualitative research projects in the past. I randomly selected seven project abstracts (based on project type, project focus, and STEM discipline) that would be coded during the IRR session.

The IRR session lasted 2.5 hours and was comprised of three main parts: training and practice session; individual coding; and a brief discussion on how to improve the codebook.

During the training and discussion, I described the guidelines for coding abstracts and definitions of codes. Next, I modeled the process for analyzing an abstract by assigning codes to sections of texts and discussing the rationale for the code assignments. Then, the two independent coders were given instructions on how to complete the IRR analysis, and given an abstract to code independently as a practice coding session. Once the practice coding was complete, we discussed the codes that were assigned to texts and the rationale for the assignments, reconciled discrepancies in coding assignments along with any outstanding questions related to the codebook or IRR-related task ahead. Next, each coder analyzed the remaining five abstracts independently, and submitted the documents including their analysis to me. Once all of the documents were turned in, the session concluded with a discussion on how to improve the codebook. The IRR analysis yielded positive results. There were a total of 161 codes assigned across all three coders and five abstracts that were analyzed independently after the training session. Table 10 presents the F_1 scores associated with the IRR check.

Table 10. Harmonic Mean Among Interrater Reliability Raters

	Coder 1	**Coder 2**
Researcher	0.894	0.879
Between Coders	0.804	

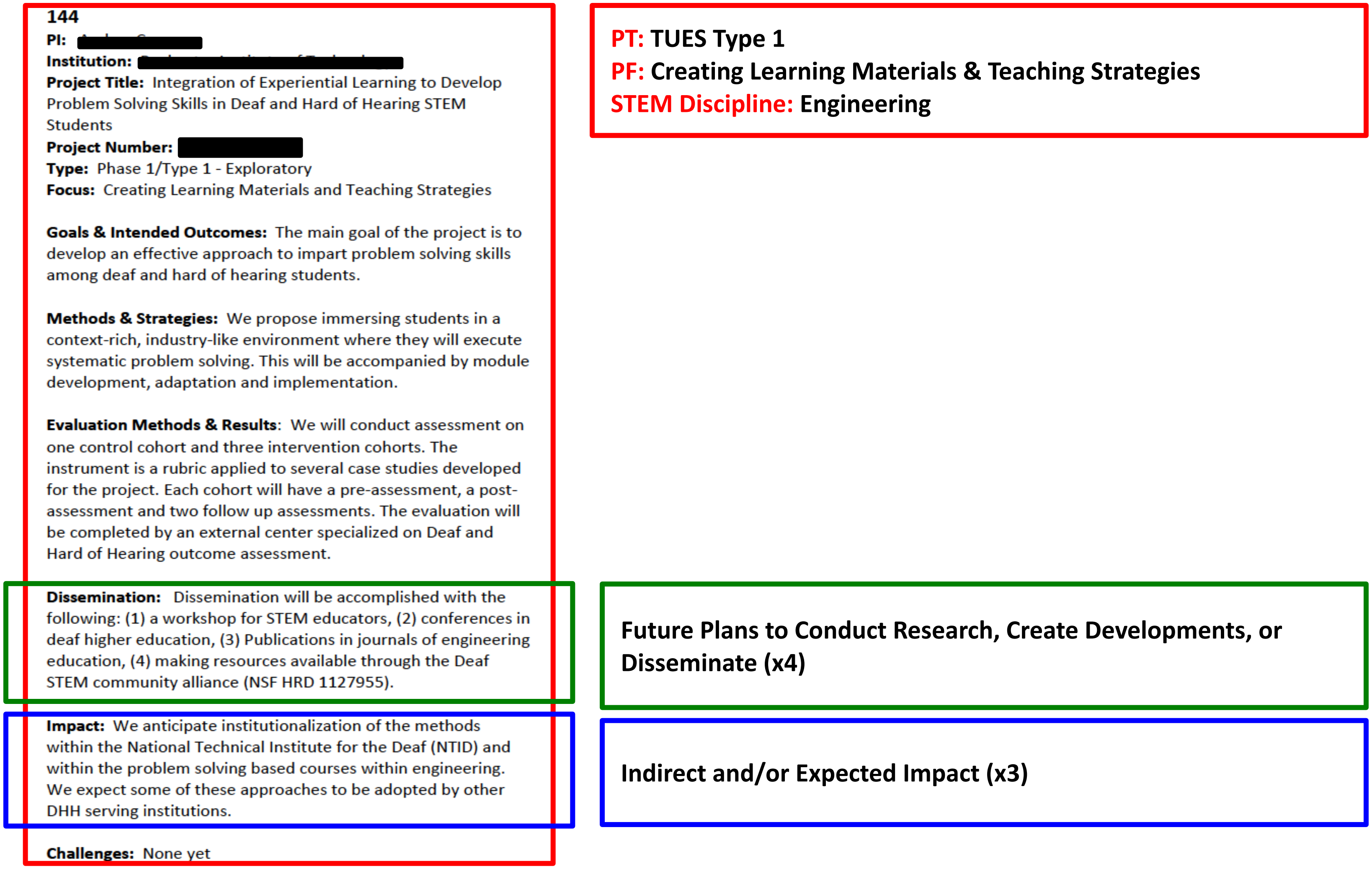
Again, an F_1-score of 1 denotes perfect agreement among the coders, while an F_1-score of 0 indicates no agreement. As the values indicate, there is a 12-22% difference between the F_1-scores calculated and the ideal F_1-score.

As a result of the IRR analysis, modifications were made to the clarity of the descriptions in the codebook based on the feedback from the IRR session. Additionally, codes representing examples of a larger code were simply adding to the description associated with the code, and demoted from being a separate code. Lastly, Atlas.ti was used to identify the provisional codes that had not been used as part of coding the first two-third of the abstracts in the sample. All provisional codes that had not been used were deleted from the codebook. These modifications led to the third version of the codebook (CB_v3: Collapsed Codes), which included 19 Attribute Codes and 50 Provisional/Descriptive Codes. Once the modifications to the Microsoft Word version of the codebook were complete, the Atlas.ti codebook was updated to match.

Final Round of Coding

The final round of coding included two parts. In the first part, the code assignments associated with the 100 abstracts analyzed before the IRR analysis were checked, and re-coded when the code assignment did not align with the codes and/or guidelines in the current version of the codebook. Next, the refined codebook (CB_v3: Collapsed Codes) was used to code the abstracts in the homogeneous sample (i.e., the engineering-only abstracts) using the same approach described above.

Once all of the coding was complete, the constant comparative method was used to refine the codes one last time (J. W. Creswell, 2007; Strauss & Corbin, 1998). After aligning the Microsoft Word and Atlas.ti versions of the codebook, I performed a spot check on 20 randomly selected abstracts to ensure that the segments associated with each revised code were still appropriately assigned. Figures 17 and 18 provide examples of how two abstracts were coded, according to the codes in the final version of the codebook (CB_v4: Finalized Codes and Themes).



144
PI: [redacted]
Institution: [redacted]
Project Title: Integration of Experiential Learning to Develop Problem Solving Skills in Deaf and Hard of Hearing STEM Students
Project Number: [redacted]
Type: Phase 1/Type 1 - Exploratory
Focus: Creating Learning Materials and Teaching Strategies

Goals & Intended Outcomes: The main goal of the project is to develop an effective approach to impart problem solving skills among deaf and hard of hearing students.

Methods & Strategies: We propose immersing students in a context-rich, industry-like environment where they will execute systematic problem solving. This will be accompanied by module development, adaptation and implementation.

Evaluation Methods & Results: We will conduct assessment on one control cohort and three intervention cohorts. The instrument is a rubric applied to several case studies developed for the project. Each cohort will have a pre-assessment, a post-assessment and two follow up assessments. The evaluation will be completed by an external center specialized on Deaf and Hard of Hearing outcome assessment.

Dissemination: Dissemination will be accomplished with the following: (1) a workshop for STEM educators, (2) conferences in deaf higher education, (3) Publications in journals of engineering education, (4) making resources available through the Deaf STEM community alliance (NSF HRD 1127955).

Impact: We anticipate institutionalization of the methods within the National Technical Institute for the Deaf (NTID) and within the problem solving based courses within engineering. We expect some of these approaches to be adopted by other DHH serving institutions.

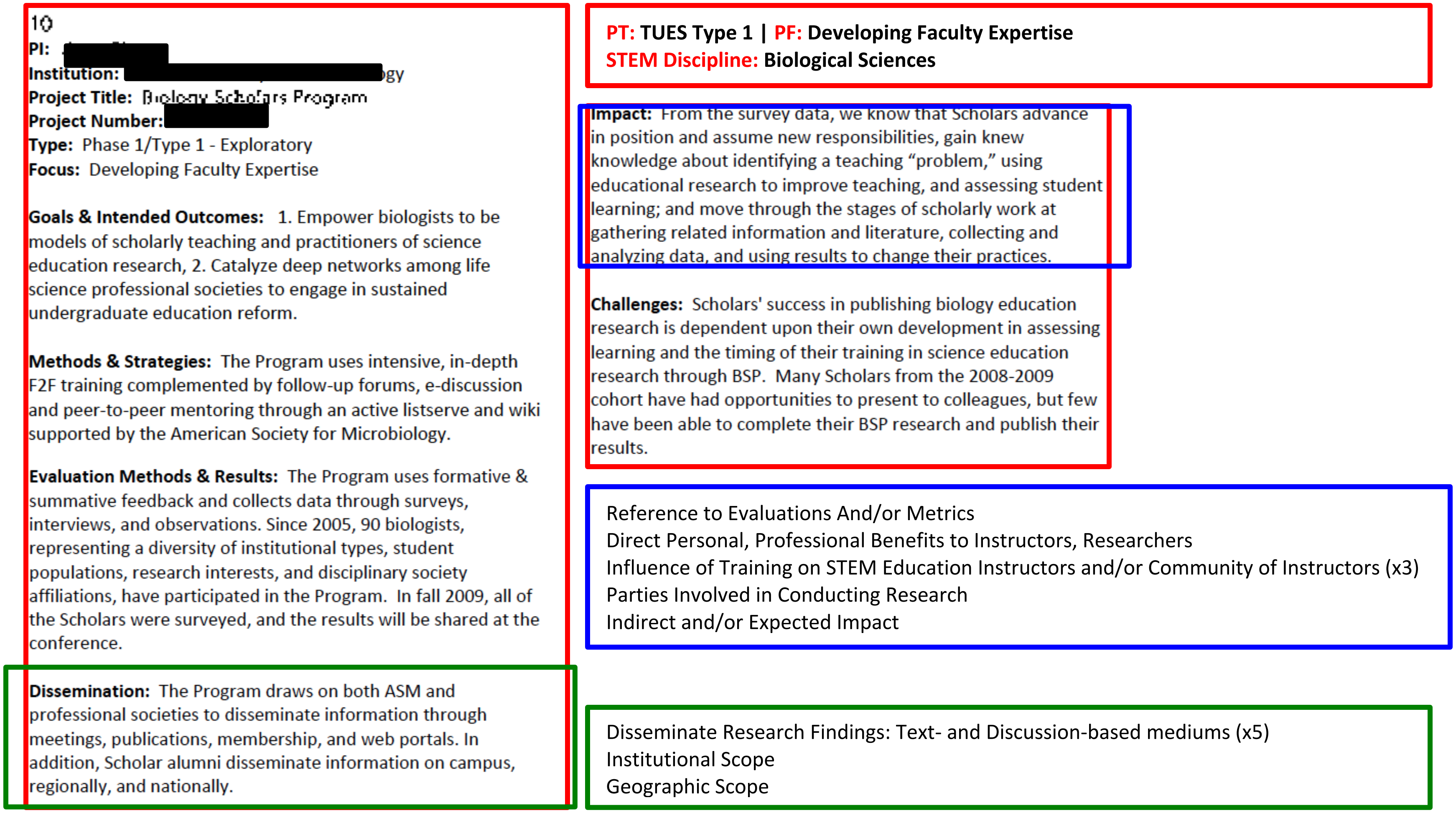
Challenges: None yet

PT: TUES Type 1
PF: Creating Learning Materials & Teaching Strategies
STEM Discipline: Engineering

Future Plans to Conduct Research, Create Developments, or Disseminate (x4)

Indirect and/or Expected Impact (x3)

Figure 17. Coding Example 1 Using Codebook, Final Version



10
PI: [redacted]
Institution: [redacted]
Project Title: Biology Scholars Program
Project Number: [redacted]
Type: Phase 1/Type 1 - Exploratory
Focus: Developing Faculty Expertise

Goals & Intended Outcomes: 1. Empower biologists to be models of scholarly teaching and practitioners of science education research, 2. Catalyze deep networks among life science professional societies to engage in sustained undergraduate education reform.

Methods & Strategies: The Program uses intensive, in-depth F2F training complemented by follow-up forums, e-discussion and peer-to-peer mentoring through an active listserve and wiki supported by the American Society for Microbiology.

Evaluation Methods & Results: The Program uses formative & summative feedback and collects data through surveys, interviews, and observations. Since 2005, 90 biologists, representing a diversity of institutional types, student populations, research interests, and disciplinary society affiliations, have participated in the Program. In fall 2009, all of the Scholars were surveyed, and the results will be shared at the conference.

Dissemination: The Program draws on both ASM and professional societies to disseminate information through meetings, publications, membership, and web portals. In addition, Scholar alumni disseminate information on campus, regionally, and nationally.

PT: TUES Type 1 | PF: Developing Faculty Expertise
STEM Discipline: Biological Sciences

Impact: From the survey data, we know that Scholars advance in position and assume new responsibilities, gain knew knowledge about identifying a teaching "problem," using educational research to improve teaching, and assessing student learning; and move through the stages of scholarly work at gathering related information and literature, collecting and analyzing data, and using results to change their practices.

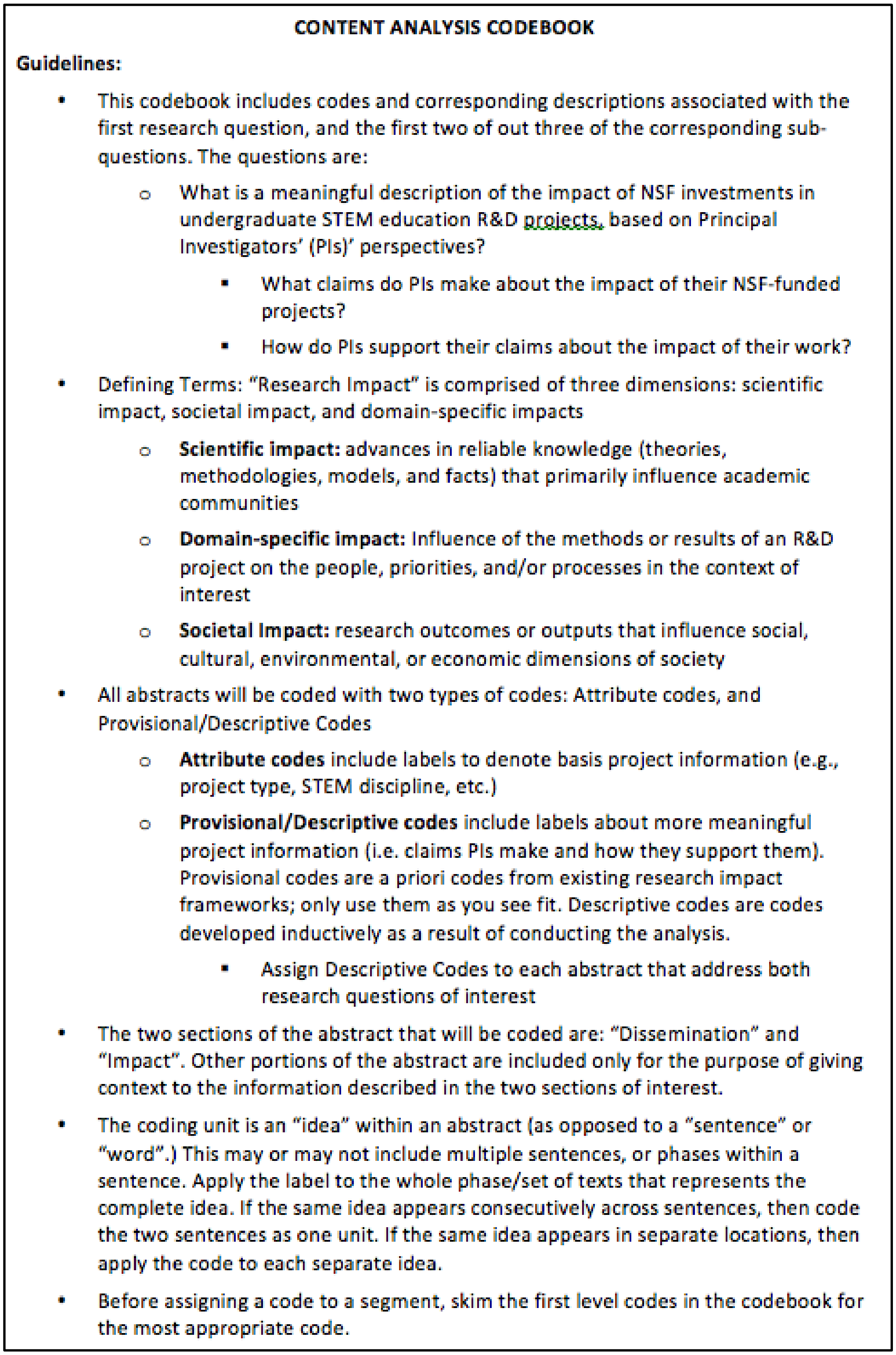
Challenges: Scholars' success in publishing biology education research is dependent upon their own development in assessing learning and the timing of their training in science education research through BSP. Many Scholars from the 2008-2009 cohort have had opportunities to present to colleagues, but few have been able to complete their BSP research and publish their results.

Reference to Evaluations And/or Metrics
Direct Personal, Professional Benefits to Instructors, Researchers
Influence of Training on STEM Education Instructors and/or Community of Instructors (x3)
Parties Involved in Conducting Research
Indirect and/or Expected Impact

Disseminate Research Findings: Text- and Discussion-based mediums (x5)
Institutional Scope
Geographic Scope

Figure 18. Coding Example 2 Using Codebook, Final Version

After aligning both versions of the codebook and ensuring consistency in the coding, the final step in the analysis –articulating abstractions— began. "Abstraction means formulating a general description of the research topic through generating categories" (Marshall & Rossman, 1995 as cited by Elo & Kyngas, 2007); these are commonly referred to as themes by other qualitative research scholars (J. W. Creswell, 2008; Miles & Huberman, 1994). Provided are two excerpts from the final codebook in Figure 19, corresponding to the guidelines used to guide the coding process, and also one of the themes that resulted from this analysis. See Appendix C for the complete version of the final codebook (CB_v4: Finalized Codes & Themes).

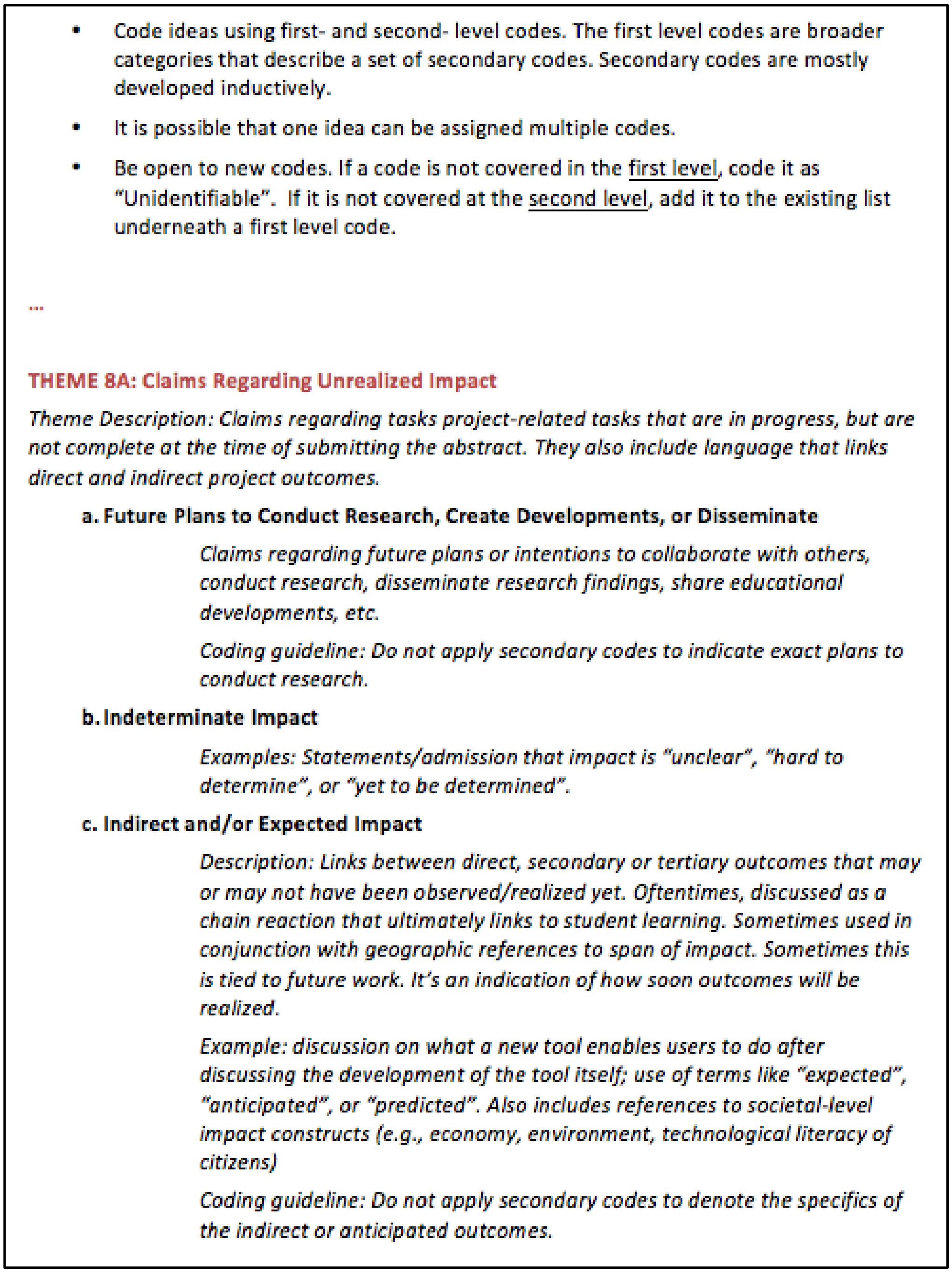


CONTENT ANALYSIS CODEBOOK

Guidelines:

- This codebook includes codes and corresponding descriptions associated with the first research question, and the first two of out three of the corresponding sub-questions. The questions are:
    - What is a meaningful description of the impact of NSF investments in undergraduate STEM education R&D projects, based on Principal Investigators' (PIs)' perspectives?
        - What claims do PIs make about the impact of their NSF-funded projects?
        - How do PIs support their claims about the impact of their work?
- Defining Terms: "Research Impact" is comprised of three dimensions: scientific impact, societal impact, and domain-specific impacts
    - **Scientific impact:** advances in reliable knowledge (theories, methodologies, models, and facts) that primarily influence academic communities
    - **Domain-specific impact:** Influence of the methods or results of an R&D project on the people, priorities, and/or processes in the context of interest
    - **Societal Impact:** research outcomes or outputs that influence social, cultural, environmental, or economic dimensions of society
- All abstracts will be coded with two types of codes: Attribute codes, and Provisional/Descriptive Codes
    - **Attribute codes** include labels to denote basis project information (e.g., project type, STEM discipline, etc.)
    - **Provisional/Descriptive codes** include labels about more meaningful project information (i.e. claims PIs make and how they support them). Provisional codes are a priori codes from existing research impact frameworks; only use them as you see fit. Descriptive codes are codes developed inductively as a result of conducting the analysis.
        - Assign Descriptive Codes to each abstract that address both research questions of interest
- The two sections of the abstract that will be coded are: "Dissemination" and "Impact". Other portions of the abstract are included only for the purpose of giving context to the information described in the two sections of interest.
- The coding unit is an "idea" within an abstract (as opposed to a "sentence" or "word".) This may or may not include multiple sentences, or phases within a sentence. Apply the label to the whole phase/set of texts that represents the complete idea. If the same idea appears consecutively across sentences, then code the two sentences as one unit. If the same idea appears in separate locations, then apply the code to each separate idea.
- Before assigning a code to a segment, skim the first level codes in the codebook for the most appropriate code.

Figure 19. Codebook Section



- Code ideas using first- and second- level codes. The first level codes are broader categories that describe a set of secondary codes. Secondary codes are mostly developed inductively.
- It is possible that one idea can be assigned multiple codes.
- Be open to new codes. If a code is not covered in the first level, code it as "Unidentifiable". If it is not covered at the second level, add it to the existing list underneath a first level code.

…

THEME 8A: Claims Regarding Unrealized Impact

Theme Description: Claims regarding tasks project-related tasks that are in progress, but are not complete at the time of submitting the abstract. They also include language that links direct and indirect project outcomes.

a. Future Plans to Conduct Research, Create Developments, or Disseminate

Claims regarding future plans or intentions to collaborate with others, conduct research, disseminate research findings, share educational developments, etc.

Coding guideline: Do not apply secondary codes to indicate exact plans to conduct research.

b. Indeterminate Impact

Examples: Statements/admission that impact is "unclear", "hard to determine", or "yet to be determined".

c. Indirect and/or Expected Impact

Description: Links between direct, secondary or tertiary outcomes that may or may not have been observed/realized yet. Oftentimes, discussed as a chain reaction that ultimately links to student learning. Sometimes used in conjunction with geographic references to span of impact. Sometimes this is tied to future work. It's an indication of how soon outcomes will be realized.

Example: discussion on what a new tool enables users to do after discussing the development of the tool itself; use of terms like "expected", "anticipated", or "predicted". Also includes references to societal-level impact constructs (e.g., economy, environment, technological literacy of citizens)

Coding guideline: Do not apply secondary codes to denote the specifics of the indirect or anticipated outcomes.

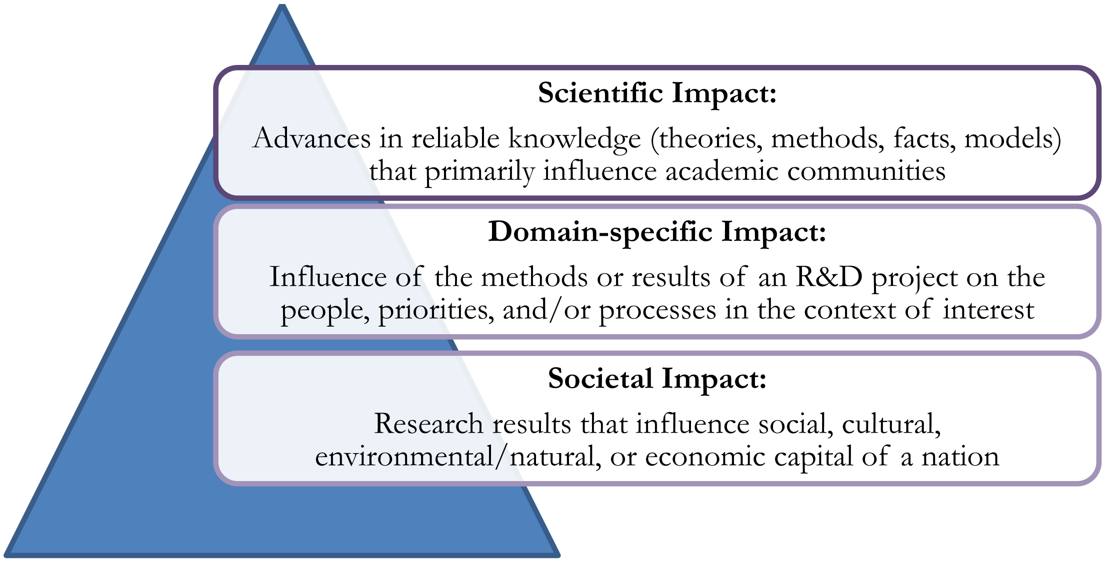
Figure 19 (continued.) Codebook Section

Generating Descriptive Statistics

Qualitative research oftentimes uses some form of counting to make judgments about patterns/consistencies (Miles & Huberman, 1994). Numbers, when used in qualitative data analysis, allow a researcher to quickly make sense of the data (when there is a large dataset), verify or disprove an intuition or hypothesis, and helps protect against bias (Miles & Huberman, 1994). Atlas.ti's analysis tools were used to generate a co-occurrence table; this data was used to describe how often the themes occur in relation to one another. Additionally, Microsoft Excel was used in this study to generate descriptive statistics surrounding the themes that resulted from the analysis. This provides insights on the patterns of themes based on the project type, focus, and STEM discipline.

Interpreting the Qualitative Results

The second sub-question proposed in this phase of the study was: *How do PIs' perspectives of impact align with existing impact frameworks found in the literature to form a preliminary description of the impact of NSF investments in undergraduate STEM education projects?* Essentially, this research question is asking for an interpretation of the research findings in light of the past literature on research impact. This type of interpretation is commonly done in both qualitative and quantitative studies (J. W. Creswell, 2008). In this study, the codes corresponding to the themes that emerged from the coding analysis were compared with the three definitions of research impact and the dimensions of impact in existing research impact frameworks (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010; Rymer, 2011; Salter & Martin, 2001; Walter et al., 2003). The constructs in the existing frameworks are in Tables 4-6. For the sake of convenience, Figure 2 is repeated below.



Scientific Impact:

Advances in reliable knowledge (theories, methods, facts, models) that primarily influence academic communities

Domain-specific Impact:

Influence of the methods or results of an R&D project on the people, priorities, and/or processes in the context of interest

Societal Impact:

Research results that influence social, cultural, environmental/natural, or economic capital of a nation

Figure 2. Defining the Three Dimensions of Research Impact

This analysis describes the ways in which PIs perspective on impact align, or do not align, what research impact means according to the literature.

Acknowledging Bias

Qualitative research is sometimes referred to as *interpretive research* (Miles & Huberman, 1994; Stake, 2010b). When conducting qualitative research, the qualitative researcher is the primary instrument for collecting data and making meaning of it (Stake, 2010b). As a result of this, it was important to recognize that "the perceptions we have of objects and events and relationships are simultaneously interpretive. They can get continuing reinterpretation" (Stake, 2010b, p. 37). While it is hopeful that my experience and understanding will benefit the goals of making sense of what research impact means in this context, it was important to recognize that interpretations can be faulty and that my bias could potentially influence the process and outcomes of the study.

I (the researcher conducting this study) hold B.S. and M.S. degrees in Industrial Engineering, and am currently pursuing a Ph.D. in Engineering Education—all from the same institution. My academic background in Industrial Engineering undergird my research interests in studying systems-level issues and also facilitates a systems engineering approach to solving problems (i.e. understanding the individual components in socio-technical systems, how they fit & work together, and pursuing actions to enhance the overall system's efficiency.) Lastly, I have conducted two mixed methods studies during two consecutive summers as an intern in NSF's Division of Undergraduate Education (London, 2012, 2013), and am currently leading to a third research project (London & Young, 2014).

These experiences provide me with rich, experiential understanding of the overlapping contexts in which this study is situated—the contexts of undergraduate engineering education in the U.S., federal investments in undergraduate engineering education research, and stakeholders desiring better information to inform better funding decisions. On the other hand, I am also aware of how these experiences can contribute to bias, which Scriven summed up well as "the lack of objectivity,... a predisposition to error" (Stake, 2010b, pp. 164-165). An example of this might include expecting systems-level impact from an individual project where a more modest expectation is more reasonable. I monitored these biases and was intentional about engaging in research activities that should increase confidence in the results; examples of this includes keeping reflective notes and triangulating results using methods that involve other STEM education researchers.

CHAPTER 6: PHASE ONE RESULTS

This chapter includes results corresponding to the qualitative questions posed in the first phase of this study: *What is a meaningful description of the impact of NSF investments in undergraduate STEM education R&D projects, based on PIs' perspectives?* It includes three sections. The first section describes the themes surrounding the claims PIs make when discussing the impact of their work in PI conference reports, and how they support them. The second section includes statistics about the frequency of the claims, by project type, project focus, and STEM discipline. The last section includes an interpretation of how PIs' perspectives on impact align with existing literature on research impact, and concludes with a preliminary description of what impact means in this context.

Making & Supporting Claims About Research Impact

Making Claims About Impact

This section describes results of the first sub-question: *What claims do PIs make about the impact of their NSF-funded projects?* When discussing impact of a research project, PIs tend to make claims about eight thematic ideas. Table 12 provides a summary of the themes. The remainder of this section provides additional details on each theme, and quotations corresponding to each theme.

Table 11. Summary of the Types of Claims PIs Make about Impact

Theme	Description
Conducting Research	Claims about people involved in conducting the research and the major steps in the research process.
Research- and Education-focused Developments	Claims about the development of artifacts that imply permanence and sustainability of the research topic beyond the current study, and tangible, educational materials informed from the current study.
Disseminating Research Findings and Propagating Developments	Claims about how research findings and/or educational developments are being shared with other researchers and/or practitioners.
Influence on Individuals and/or Communities	Claims about ways in which individual or communities of learners, instructors, or researchers are being affected by the outcomes of the study.
Influence on Environmental/Structural Decisions, Metrics	Claims about how insights from the current study inform administrative decisions that ultimately influence the actions of others, and how the current study contributes to assessments and/or metrics of interest to administrators.
Scope of Influence	Claims about the span associated with their project outcomes.
Symbols of Impact	Claims about the receipt of public affirmation as a result of connections to the current study.
Unrealized Impact	Claims about activities, events, and outcomes that have not yet happened, but are either future plans or anticipated outcomes that will be realized at a later time.

THEME: CONDUCTING RESEARCH

When given an opportunity to discuss the impact of their work, PIs make claims about people involved in conducting the study, and claims related to the major steps of the research process in their impact narratives. References are usually made to collaborators on the project, undergraduate student researchers, graduate research assistants, and post-doctoral research staff. Making claims regarding *conducting research* also includes connections between the current study and existing literature or work that serve as a motivation for the study. This may also include connections to prior research and developments that serve as the foundation for the current work. Highlights of current research activities are other discussion topics associated with this theme. Impact narratives may also include a succinct statement on the key research findings, with emphasis on the new contribution to the body of literature. Lastly, impact narratives sometimes mention the submission of applications for additional funding and/or references to securing funding to continue the study. Provided are quotes from three abstracts that provide evidence in support of this theme; the code among those assigned that corresponds to this theme is **bold.**

Project Attributes: Central Resource Project | Engineering | Conducting Research on Undergraduates in STEM Education

> ***"The project has already helped and will continue to build a community of engineering education scholars by training and mentoring twelve graduate and post-doctoral researchers in both qualitative and quantitative data collection and analysis. …"***
>
> Assigned Codes: **Parties Involved in Conducting Research;** Influence of Research on STEM Education Researchers and/or Research Community

Project Attributes: TUES Type I | Computer Science | Assessing Student Achievement

> *"A goal of many CS education projects is determine the extent to which an instructional intervention has impacted student attitudes. A challenge is that valid and reliable instruments that measure the necessary constructs are not currently available. Instead, each project is left to develop its own, resulting in problems. First, most computer scientists are not training in measurement, and therefore, are not familiar with psychometric principles. This could result in questionable instruments and interpretations. Second, without a common set of instruments, valid comparisons cannot be made across project[s]. This project seeks to address this need for a valid survey in CS."*
>
> Assigned Codes: **Connections between the current study & existing literature;** References to what motivated the study; Curricular materials, training resources, and pedagogy

Project Attributes: TUES Type I | Geological Sciences | Creating Learning Materials and Teaching Strategies

> *"...An internal grant was funded to upgrade and modernize the surface monitoring station."*
>
> Assigned Codes: **Applying for and securing additional funding to continue research;** Affirmation from within the academic community

THEME: RESEARCH- AND EDUCATION-FOCUSED DEVELOPMENTS

Impact narratives may include claims regarding the formation of artifacts that imply permanence and sustainability of the research topic beyond the current study, but were motivated by the current study. There are three types of research-focused developments: text-based entities, discussion-based entities, and facilities or technology developed primarily for the purpose of conducting research. Examples of each of these may include: the establishment of a new scholarly journal or an annual research symposium focused on a niche research area, and the installation of a new research center dedicated to specific research areas. For example, for one of the abstracts (Project Attributes: TUES Type I |

Interdisciplinary | Creating Learning Materials and Teaching Strategies), they mentioned four types of dissemination activities—one of which was *"Launching 'Journal of IT Education Discussion Cases' through Informal Science Institute to publish cases"*

Additionally, impact narratives also include claims about the development of tangible, educational materials that were informed from the current study and designed to benefit individuals (e.g., learners, instructors) or groups in an educational setting. This includes curricular and pedagogical materials, as well as resources for training instructors. Technology developments and instruments purchased (e.g. lab equipment) for educational purposes are also included among the examples of education-focused developments. Consider the following quote from an abstract as an example.

Project Attributes: TUES Type II | Physics/Astronomy | Conducting Research on Undergraduates in STEM Education

> *"The curriculum is in use at our institution and has been tested at a handful of pilot sites. As noted above, our materials are available to potential adopters as well."*
>
> Assigned Codes: **Curricular materials, training resources, and pedagogy;** Curricular changes

When discussing both the research- and education-focused developments, PIs sometimes elaborate on the affordances of the new development. The use of the term "affordances" is borrowed from Gibson (1977), who first used it in an ecological context to describe properties of the environment), and Conole (2013), who is among the researchers who now uses it to discuss the contributions information and communications technology can make to the learning process. In this context, this may include discussion on the utility of the development, economic value associated with it, and ways in which it provides the capacity to engage in a new set of learning and/or research experiences. Provided is one final quote from the Dissemination section of an abstract that provides evidence to support this theme.

Project Attributes: TUES Type III | Interdisciplinary | Assessing Student Achievement

> *"The project has yielded very effective interdisciplinary tools to assess students' critical thinking and real-world problem solving skills. This tool is useful for assessing both program improvement efforts, and for evaluating sponsored research project outcomes related to critical thinking. …"*
>
> Assigned Codes: **Curricular materials, training resources, and pedagogy; Affordances of the developments**

THEME: DISSEMINATING RESEARCH FINDINGS AND PROPAGATING DEVELOPMENTS

When PIs discuss the ways in which they are sharing research findings and propagating their education- and research-focused developments, they make claims about the mediums used to circulate the findings, along with activities and outcomes surrounding propagation. The mediums used to disseminate research findings are text-based and/or discussion-based mediums. Examples of text-based mediums include conference proceedings, journal publications, and research briefs. Discussion-based mediums include sharing research findings during a conference presentation or by participating in an expert panel discussion. Another form of disseminating research findings includes sharing insights informed from research in various venues for teaching and training (e.g., workshop, seminar, consulting, demonstrations at a research conference). Oftentimes, the audience for the text- and discussion-based mediums is the same: STEM education researchers and/or practitioners.

In addition to discussing mediums and venues for disseminating research findings, PIs also discuss ideas related to the spread of developments resulting from their study. These topics include highlights of current activities they are engaged in to propagate developments—such as activities that contribute to the establishment of partnerships, marketing and commercialization of materials, and instituting mailing lists to keep track of educators and/or vendors who have expressed interest in their developments when they become available. Provided are quotes from three abstracts that provide evidence in support of this theme; the code among those assigned that corresponded to this theme is **bold.**

Project Attributes: TUES Type II | Research/Assessment of Research | Assessing Student Achievement

> *"To date, more than 15 peer-reviewed conference presentations; several posters; 1 publication; 2 under review; 2 in preparation by team members. Several campus visits and workshops."*
>
> Assigned Codes: **Text- and/or Discussion-based Mediums; Dissemination via Venues for Teaching, Training;** Highlights of Current Activities; Quantifying Outcomes

Project Attributes: TUES Type I | Physics/Astronomy | Assessing Student Achievement

> *"We have published five peer-reviewed conference proceedings, one peer-reviewed journal paper, and about 20 contributed talks or posters. We are presenting a workshop at the National Meeting of the American Association of Physics Teachers in August, and plan an additional 2-4 papers to be submitted next year."*
>
> Assigned Codes: **Text- and/or Discussion-based Mediums; Dissemination via Venues for Teaching, Training;** Future Plans to Conduct Research, Create Developments, or Disseminate; Quantifying Outcomes

Project Attributes: TUES Type III | Social Sciences | Creating Learning Materials and Teaching Strategies

> *"Through the efforts of our project, we have authored a workbook, performed several webinars and workshops, and created a website in collaboration with Carleton College's Science Education Resource Center to disseminate materials developed through the project. We have also made these available through NSDL."*
>
> Assigned Codes: Curricular materials, training resources, and pedagogy; **Text- and Discussion-based mediums; Dissemination via Venues for Teaching, Training; Outcome of activities supporting propagation of developments**

THEME: INFLUENCE ON INDIVIDUALS, COMMUNITIES

When given an opportunity to discuss the impact of their work, PIs make claims about ways in which individuals or communities of people are affected by outcomes of the current study. The individuals most commonly mentioned are learners, instructors, and researchers. As it relates to learners, PIs may make claims about how participation in an experience associated with the current study leads to the development and application of knowledge, skills, and ways of thinking relevant to STEM concepts and careers. It also includes undergraduate students' changes in interest in pursuing graduate studies, as well as improvements in STEM literacy among those who participate in outreach activities associated with the current study. Consider the following examples of this idea.

Project Attributes: TUES Type II | Computer Science | Implementing Educational Innovations

> *"Students who use write their own tests for their own software and are graded by Web-CAT produce 28% fewer bugs per thousand lines of code. …"*
>
> Assigned Codes: **Influence of Teaching on STEM Learners;** Quantitative Evaluation and/or Metrics

The use of the term "instructors" is intended to encapsulate a cross-section of individuals, for example, K-12 teachers, graduate teaching assistants, post-doctoral staff, and/or STEM education faculty in higher education. When making claims about instructors, PIs assert that participation in a set of activities contributes to the development and application of knowledge, skills, and ways of thinking surrounding improved pedagogical practices. The activities include attending a workshop hosted by the PI's research team, or joining a virtual network designed by the PIs' research team to facilitate interactions among instructors and to exchange resources. The latter of these two activities (i.e., providing a virtual venue) is an example of how PIs make claims about their influence on a community of instructors. Provided is an example of an impact narrative that includes this idea.

Project Attributes: TUES Type II | Geological Sciences | Conducting Research on Undergraduates in STEM Education

> *"Participating faculty modified their teaching approach to include active strategies involving ConcepTests. Several of these participants have leveraged this experience to expand to more sophisticated active learning approaches. ..."*
>
> Assigned Codes: **Influence of Training on STEM Education Instructors and/or Community;** Curricular Materials, Training Resources, and Pedagogy

The last category of individuals that are commonly referenced in impact narratives are researchers. Researchers may include undergraduate researchers, graduate student research assistants, post-doctoral researchers, and faculty in higher education. When mentioning these groups, PIs discuss ways in which participation in the current study is contributing to the development of new data collection and analysis skills, or influencing the quality of the research-related documents (e.g., grant proposals, research publications). PIs also make claims about the development or expansion of the research community who share interests in their area of expertise. This theme also captures instances where teaching-or learning-related insights resulting from the current study serves as the basis for a new set of research-focused activities/projects. Provided are quotes from two abstracts that include influences to at least two groups of interest to this theme (i.e., learners, instructors, or researchers); the code among those assigned that corresponded to this theme is **bold.**

Project Attributes: TUES Type II | Geological Sciences | Implementing Educational Innovations

> *"This project has had a significant impact on our undergraduate female students: several students have pursued senior thesis projects stemming from grant activities; stating that the field activities were the highlight of their semester. Some students love the experience and wan more. Others decide that they may want to pursue a different career. All learn how science is conducted and have a better foundation to understand concepts like sampling, uncertainty and variability, which are important in many fields. …"*
>
> Assigned Codes: **Influence of Teaching on STEM Learners;** Qualifying Claims; Scope via Target Populations; **Influence of Research on STEM Education Researchers and/or Research Community**

Project Attributes: TUES Type I | Chemistry | Implementing Education Innovations

> *"Even though the course has been offered only once, there has been a measured positive impact on student's laboratory abilities and confidence. This has been the first funded CCLI grant in the department in about 5 years which has had a positive impact on submission by colleagues. While scheduling of the course has been supported, it has been tricky."*
>
> Assigned Codes: **Influence of Teaching on STEM Learners; Influence of Research on STEM Educations Researchers and/or Research Community;** Affirmation from within the Academic Community; Curricular Changes

The last collection of ideas that correspond to *claims regarding influence on individuals and/or communities* are claims about direct personal, professional benefits to instructors and/or researchers. This includes, but is not limited to, expansions in the number of contacts in the PIs' professional network as a result of conducting research or hosting a training session for faculty. It also entails claims about how the inclusion of research activities in faculty's promotion and tenure package contributed to positive professional outcomes for the individual. Here's one example of this idea:

Project Attributes: TUES Type III | Engineering | Implementing Education Innovations

> *"…Through invited presentations at conferences, workshops, and a variety of academic institutions, the PI has established well over 200 contacts from dozen of engineering programs across the country."*
>
> Assigned Codes: Dissemination via Venues for Teaching, Training; Affirmation from within the Academic Community; **Direct Personal, Professional Benefits to Instructors, Researchers;** Institutional Scope; Geographic Scope;

THEME: INFLUENCE ON ENVIRONMENTAL/STRUCTURAL DECISIONS, METRICS

The use of the term "environment/structure" is adopted from a framework for facilitating instructional change in undergraduate STEM education (Beach et al., 2012). The framework is comprised of four types of strategies: Disseminating Curriculum and Pedagogy; Developing Reflective Teachers; Enacting Policies; and Developing Shared Vision. While the first two strategies focus on individuals (i.e. STEM educators), the last two focus on the environment and structural elements of the STEM education system. "Environment/structure" is a label for rules and policies that govern an environment, the reward system, reporting requirements, and supporting structures.

Using these definitions, PIs make claims about ways in which their research informs administrative decisions, which ultimately influence the actions of others—either at the PIs' home institution or elsewhere. Curricular changes such as modifications to an existing course, new course offerings, and changes in the set of courses students are advised to take are examples of influences on the educational environment. Informing or enacting new policies by participating in policy-related discussions at local- or national-level gatherings is another way that PIs influence the decision making process. This theme also captures claims about how adding the PIs' research project to their promotion and tenure package affects administrators' decisions.

Furthermore, this theme also includes ways in which the current study influences assessments and/or metrics of interest at the level of the undergraduate departments. This may include how the current study highlights latent environmental/structural issues that need to be addressed. It also includes claims about how the research insights affect aggregate student outcomes—such as enrollment, retention, and "Drop, Fail, Withdraw"-rates within an undergraduate STEM education department. Lastly, the impact narrative may reference instances where the inclusion of research data or findings contributed to an accreditation evaluation.

An example of how research may influence insights about the environment/structure is a TUES Type II project (Project Type) focused on Assessing Student Achievement (Project Focus). The study concluded that after one institution used the assessments that were developed to evaluate their program, they "found infrastructural weaknesses that led to motivation problems". As a result of using the assessments, the department identified environmental/structural issues that were influencing their students' motivation; such insights were not evident before conducting the assessment.

Provided are quotes from two abstracts that provide evidence in support of this theme; the codes among those assigned that corresponded to this theme are **bold.**

Project Attributes: TUES Type II | Computer Science | Creating Learning Materials and Teaching Strategies

> *"CS1 enrollment doubled after integration of course in freshman program. ..."*
>
> Assigned Codes: **Department-level Assessments and/or Outcomes;** Curricular Materials, Training Resources, and Pedagogy; Quantifying Outcomes

Project Attributes: TUES Type I | Social Sciences | Implementing Educational Innovations

> *" ...A substantial CNS learning community is growing on campus and we have developed a new CNS minor and are in the process of developing a BS is Cognitive Science."*
>
> Assigned Codes: Qualitative Evaluation and/or Metrics; Influence of Teaching on STEM Learners; **Curricular Changes; Informing or Enacting Policy**

THEME: SCOPE OF INFLUENCE

In many of the impact narratives reviewed in study, PIs make claims about the span associated with their projects outcomes. Span may be geographic, disciplinary, or institutional in scope. References to states in the United States, and use of the term "international" are words that suggest geographic scope. Highlighting other disciplines involved in the study besides the specific target discipline serve as an indicator of disciplinary scope. Citing the names of other institutions using the R&D resulting from the study is what is meant by institutional scope. *Scope of influence* may also include span via target populations. This might include the use of terms like "at risk" students, underrepresented minorities, and women as qualifiers of the types of learners affected by the study. Finally, this theme also captures scope in the form of non-academic partnerships—with vendors, industry, or professional societies—usually for the purpose of advancing the dissemination of research findings or propagation of research developments. What follows are quotes from abstracts that references various dimensions of the scope associated with this theme.

Project Attributes: TUES Type III | Interdisciplinary | Developing Faculty Expertise

> *"To date, more than 350 institutions in 45 states 13 foreign nations have been intensively involved. In the last three years, participants estimate their efforts have impacted 145,000 students. …"*
>
> Assigned Codes: Influence of Training on STEM Education Instructors and/or Community of Instructors; **Geographic Scope; Institutional Scope;** Less Specific; Indirect or Expected Outcomes

Project Attributes: TUES Type I | Geological Sciences | Conducting Research on Undergraduates in STEM Education

> *"The project has already made a big impact on the instrumental capability of the department's hydrology lab. Many new pieces of equipment have been purchased, which triggered research interests among undergraduate as well as graduate students. …"*
>
> Assigned Codes: Instruments, Technology used for Educational Purposes; Influence of Research on STEM Education Researchers and/or Research Community; **Scope via Target Populations**

Project Attributes: Type III | Biological Sciences | Developing Faculty Expertise

> *"Analysis of the program's impact from 2002-2007 shows that we have trained 224 participants who developed 54 teachable units to engage students in learning biology and the nature of science. Participants have taught the units to over 52,000 undergrads at 50 universities. While the units are developed for biology courses, the interdisciplinary topics include microbiology, ecology, cell biology, neurobiology, statistics, molecular biology, and imaging. In 200, we trained 36 faculty and instructors to lead ST-type workshops at 13 universities"*
>
> Assigned Codes: Quantitative Evaluation and/or Metrics; Quantifying Outcomes; Influence of Training on STEM Education Instructors and/or Community of Instructors; Curricular Materials, Training Resources, and Pedagogy; Indirect or Expected Outcomes; **Disciplinary Scope;** Dissemination via Venues of Teaching, Training

THEME: SYMBOLS OF IMPACT

At times, PIs make claims about the forms of public affirmation they have received as a result of the research they have conducted. Such affirmations vary depending on whether it comes from within the academic community or from outside. In most cases, the symbol is bestowed on the PI by other experts in the research community because of their unique contribution to the body of knowledge or perceptions of the extent to which their work advances the discipline. By doing such, those affirming the research bolster its credibility and visibility.

The forms of affirmation that come from within the academic community are often in the form of receiving special recognition/awards, being labeled an exemplar, leader, or fellow, or being asked to give a keynote address related to the PIs' area of expertise. These are all symbols that carry significance among the members of their research community. Examples of the forms of affirmation that come from outside the research community may be in the form of press coverage, news media reports, featured stores on television or in written publications, or having a project featured in the *NSF Highlights.* (The *NSF Highlights* are a source of information about NSF investments in R&D projects stored on the NSF intranet. NSF Program Officers submit "highlights" to the internal database at will. Each highlight is a brief summary of the funded project, along with its transformative results or outcomes.) Provided are quotes from two abstracts that provide evidence in support of this theme; the code among those assigned that corresponded to this theme is **bold.**

Project Attributes: TUES Type I | Biological Sciences | Implementing Educational Innovations

> *"The impact that this project has had on students has been positive and we have been successful in achieving our goals for them based on gains in the classroom, from survey results and CAT gains. Each of the classes has generated research-quality data for the collaborating research faculty. Some of the data have resulted in a research publication with the students as authors. The success of our project has been recognized nationally and within our department with several faculty members approaching us with ideas for future-based classes."*
>
> Assigned Codes: Influence of Teaching on STEM Learners; Quantitative Evaluation and/or Metrics; Highlights of Research Findings; Parties Involved in Conducting Research; Text- and Discussion-based Mediums; Geographic Scope; Institutional Scope; **Affirmation from within the Academic Community**

Project Attributes: Type I | Geological Sciences | Creating Learning Materials and Teaching Strategies

> *"Over 750 undergraduate students enrolled in the introductory Geology Laboratory for non-majors have completed two exercises at the GetWET measuring surface and groundwater quantity and quality. Over 65 undergraduate majors have participated in various field exercises and experiments through participation in five courses within the Geology curriculum. Forty teachers have been trained and 390 K-12 students have visited the GetWET on field trips. Press coverage resulted in a major gift from a local groundwater equipment manufacturer, allowing student use of cutting edge, professional technology. An internal grant was funded to upgrade and modernize the surface water monitoring station."*
>
> Assigned Codes: Department-level Assessments and/or Outcomes; Quantifying Outcomes; Influence of Teaching on STEM Learners; Scope via Target Populations; Instruments, Technology, used for educational purposes; Influence of Training on STEM Education Instructors and/or Community of Instructors; **Affirmation from outside the academic community; Affirmation from within the academic community**

THEME: UNREALIZED IMPACT

The final type of claim PIs make when discussing the impact of their work is *claims about unrealized impact*. These are claims about activities, events, and outcomes that have not yet happened, but are either future plans or anticipated outcomes the PI expects will happen at a later time. More specifically, this includes three types of ideas. One set of ideas is future plans to conduct research, create developments, disseminate findings, or propagate developments.

A second set of ideas includes indirect and/or expected outcomes. More specifically, these are claims that create links between direct project outcomes and secondary, or tertiary outcomes that may or may not have been realized yet. Oftentimes, this is discussed in form of a chain of events that ultimately leads to effects on student learning. For example, PIs make connections between faculty participation in their workshop to changes in their pedagogical practices, which will ultimately influence how their students learn.

Sometimes, claims are made to describe anticipated outcomes in conjunction with references to an anticipated geographic span of influence. In some cases, PIs posit that the findings of their work have implications for teachers and learners on a national or international level. However, this is usually limited to a non-specific use of the term national or international, and often does not include a detailed explanation of how and where the implications will take effect. In a few instances, PIs also connect their project outcomes to societal-level impact constructs of interest—such as workforce development, technological literacy of citizens, or the economy.

The third type of idea associated with this theme are indeterminate outcomes: statements of admission that the impact is “unclear”, “hard to determine”, or “yet to be determined.” Provided is an impact narrative from a Central Resource Project (Project Type) focused on Developing Faculty Expertise (Project Focus):

> *Unfortunately, the actual impact of the project with respect to its overall goals is unclear. That is, we have not determined if, as a result of the webinar series, the quality of proposals submitted by the workshop participants is higher and the number* [of proposal submissions] *is greater. We do know that we have introduced the TUES Program to a significant number of faculty who had never submitted a proposal to the Program.*

This shows an example of how PIs may perceive some aspects of their project's impact are indeterminate while also highlighting impacts about which they are more certain. This section of the chapter concludes with one last example of abstracts that include evidence to support this theme.

Project Attributes: TUES Type II | Engineering | Creating Learning Materials

> ***Over the past years, several state-of-the-art laboratories and new courses have been developed by the NSF grant. Creation of the Internet-based laboratories significantly contributes to the development of technologically literate students and workforce that could be in great demand not only in the tri-state area but also nationwide. Information-based technology has become the new realm of manufacturing and mechanical engineering technology graduates. The NSF project helps the AET program to prepare faculty and students to: apply discipline-specific theory, conduct experiments, and use real world experience to interpret, analyze, and solve current and emerging technical problems. Annual workshop has been held for faculty development. The successful implementation of the NSF project is crowned by the Applied Engineering Technology program's successful accreditation by the TAC of ABET. The AET program was granted accreditation and the ABET evaluation team found no deficiencies, concerns or weaknesses. The NSF project has been well-performed by the PIs as planned in the early stage.***
>
> Assigned Codes: Instruments, Technology used for Educational Purposes; Influence on STEM Learners; MISC–Reference to Societal Level Impact; Geographic Scope; **Indirect or Expected Outcomes;** Influence of Training on STEM Instructors and/or Communities of Instructors; Affirmation from within the academic community; Department-level Assessments and/or Outcomes

Supporting Claims about Research Impact

This section describes results of the second sub-question: *How do PIs support claims about the impact of their work?* It elaborates on how PIs qualify and back their claims, irrespective of the types of impact being discussed. Two themes describe the mechanisms PIs used to support their claims (see Table 12); they are: clarifying claims using degrees of specificity; and supporting claims by establishing credibility and/or relevance. The remainder of this section is an elaboration on these two themes.

Theme	Description
Clarifying Claims Using Degrees of Specificity	The use of more or less specific language when discussing claims about impact. Examples of more specific language include quantifying project outcomes, and using qualifying term to describe the extent of impact. Examples of less specific language are succinct, broad, vague ideas that could easily describe a variety of projects.
Supporting Claims by Establishing Credibility and/or Relevance	References to ideas that suggest reasons why the reader should perceive the study as trustworthy and/ore closely related to the priorities of the STEM education discipline associated with the project.

Table 12. Summary of Ways PIs Support Claims About Impact

THEME: CLARIFYING CLAIMS USING DEGREES OF SPECIFICITY

When discussing claims, PIs often do so by using more or less specific language. When the language is more specific, the impact narrative includes a quantification of project outcomes, such as the number of: publications that resulted from the study, workshops hosted, participants who attended the workshop, institutions represented among the technology- or curriculum-adopters. It can also include references to the duration of time associated with how long research- and education-focused developments have been in use. Another way in which PIs clarify claims using degrees of specificity is by using qualifying terms, such as “possibly”, “likely” or “potentially”. The use of these terms limits the extent to which the claim can be considered credible and/or applicable. In some instances, PIs begin the impact narrative with sentences that can be perceived as "cautionary disclaimers". The use of language sets parameters around the claim so that readers know that it is not relevant in all cases, but can be applied in particular circumstances.

Provided are two examples of impact narratives that begin with the "cautionary disclaimers":

> *"Our project impact is by necessity limited but our results are suggestive. In our pilot study…"*

> *"While it is too early to determine the impact of the workshop series, it is expected to increase the number and quality of engineering proposals submitted to the TUES program."*

The use of the term "suggestive" in the first quote implies information about the extent to which the results can be used. On the other hand, "while it is too early to determine the impact" sets limits on the timeframe when more conclusive statements about impact can be made.

In some instances, the impact narratives are very short and much less specific. They are usually broad, vague claims that arguably, would not be considered unique to the project being reported on. The following quotation is the complete impact narrative from a TUES Type I project focused on Creating Learning Materials in computer science:

> *" Anticipated impacts are to transform the way intro programming courses are taught and provide material for the instructors to use. A new model was created. Significant student learning is being demonstrated."*

In this example, the impact narrative includes little insight about what is meant by the use of words like "significant", "student learning", and "demonstrated". As a result, readers are not given enough information to know how the results of the study might be applicable beyond the study itself.

Provided are quotes from three abstracts presented above; consider them in light of how PIs use degrees of specificity to support claims about impact. The code among those assigned that corresponded to this theme is **bold.**

Project Attributes: TUES Type II | Research/Assessment of Research | Assessing Student Achievement

> *"To date, more than 15 peer-reviewed conference presentations; several posters; 1 publication; 2 under review; 2 in preparation by team members. Several campus visits and workshops."*
>
> Assigned Codes: Text- and/or Discussion-based Mediums; Dissemination via Venues for Teaching, Training; Highlights of Current Activities; **Quantifying Outcomes**

Project Attributes: TUES Type III | Interdisciplinary | Developing Faculty Expertise

> *"To date, more than 350 institutions in 45 states 13 foreign nations have been intensively involved. In the last three years, participants estimate their efforts have impacted 145,000 students. ..."*
>
> Assigned Codes: Influence of Training on STEM Education Instructors and/or Community of Instructors; Geographic Scope; Institutional Scope; **Less Specific;** Indirect or Expected Outcomes

Project Attributes: TUES Type II | Engineering | Creating Learning Materials

> *"All of project outcomes have been presented by talks and posters at international and national education conferences in the ASEE, IERC, IEEM, IJIE, ASME, IMECE, MSEC, MES, EEET, and annual NSF workshop every year. More than 15 articles have been published by journals such as CED, JSysCI, JAMS, RCIM, JEE, JCE, and IJAMT. Other activities include lab tour for high school students and program advisory board."*
>
> Assigned Codes: Text- and/or Discussion-based Mediums; Geographic Scope; Scope via Target Populations; **Qualifying Claims; Quantifying Outcomes;**

THEME: SUPPORTING CLAIMS BY ESTABLISHING CREDIBILITY AND/OR RELEVANCE

The last theme that will be discussed in this study relates to how PIs support claims about impact by establishing credibility and/or relevance. These are ideas that convey reasons the reader should perceive the study as trustworthy and/or closely connected to the priorities of the STEM education discipline. References to a survey of the literature and any gaps in the literature, which the current study seeks to fill, may be an example of establishing credibility and/or relevance. Connections between the current study and national-level policy reports or discussions surrounding STEM education are yet another example. By extension, this also includes references to societal impacts (i.e., social, cultural, environmental and economic dimensions of society). Provided is a quote that was mentioned in references to the types of claims PIs make; consider it now in light of how the claims is being supported. The code among those assigned that corresponds to this theme is **bold.**

Project Attributes: TUES Type I | Computer Science | Assessing Student Achievement

> *"A goal of many CS education projects is determine the extent to which an instructional intervention has impacted student attitudes. A challenge is that valid and reliable instruments that measure the necessary constructs are not currently available. Instead, each project is left to develop its own, resulting in problems. First, most computer scientists are not training in measurement, and therefore, are not familiar with psychometric principles. This could result in questionable instruments and interpretations. Second, without a common set of instruments, valid comparisons cannot be made across project[s]. This project seeks to address this need for a valid survey in CS."*
>
> Assigned Codes: Connections between the current study & existing literature; **References to what motivated the study;** Curricular materials, training resources, and pedagogy

Brief references to the type of evaluation conducted as part of the study are another way to establish credibility. The evaluations may result in quantitative data (e.g., course performance indicators, pre-post tests results, adoption rates); qualitative data (e.g., anecdotal remarks from workshop participants, feedback from instructors using the curriculum); or a mix of both quantitative and qualitative data (e.g., survey results; course evaluation data). Anyone familiar with the type of data resulting from the evaluation would be able to make inferences about the types of claims a researcher can make based on it. In some cases, PIs make claims that suggest some type of evaluation has occurred, but there is no explicit reference to what type of evaluation it was. (Since there are other sections of the abstract that focuses on Methods and Evaluation, it is likely that such information can be found there.) Provided are quotes from three abstracts that were presented before when discussing the types of claims PIs make; now consider it with this theme in mind**.**

Project Attributes: TUES Type II | Computer Science | Implementing Educational Innovations

> *"Students who use write their own tests for their own software and are graded by Web-CAT produce 28% fewer bugs per thousand lines of code. ..."*
>
> Assigned Codes: Influence of Teaching on STEM Learners; **Quantitative Evaluation and/or Metrics**

Project Attributes: TUES Type I | Biological Sciences | Implementing Educational Innovations

> *"The impact that this project has had on students has been positive and we have been successful in achieving our goals for them based on gains in the classroom, from survey results and CAT gains. Each of the classes has generated research-quality data for the collaborating research faculty. Some of the data have resulted in a research publication with the students as authors. The success of our project has been recognized nationally and within our department with several faculty members approaching us with ideas for future-based classes."*

Assigned Codes: Influence of Teaching on STEM Learners; **Quantitative Evaluation and/or Metrics;** Highlights of Research Findings; Parties Involved in Conducting Research; Text- and Discussion-based Mediums; Geographic Scope; Institutional Scope; Affirmation from within the Academic Community

Descriptive Statistics Surrounding Results on Making & Supporting Claims

The last section of this chapter presents descriptive statistics on how often the themes appear in the set of abstracts reviewed. A total of 1,454 codes about making and supporting claims were assigned to texts in the 155 abstracts reviewed in this study. Each code maps to one of the 10 themes about making and supporting claims about impact. (See Appendix C for codebook.) Figure 20 shows the proportion of claims discussed in abstracts.

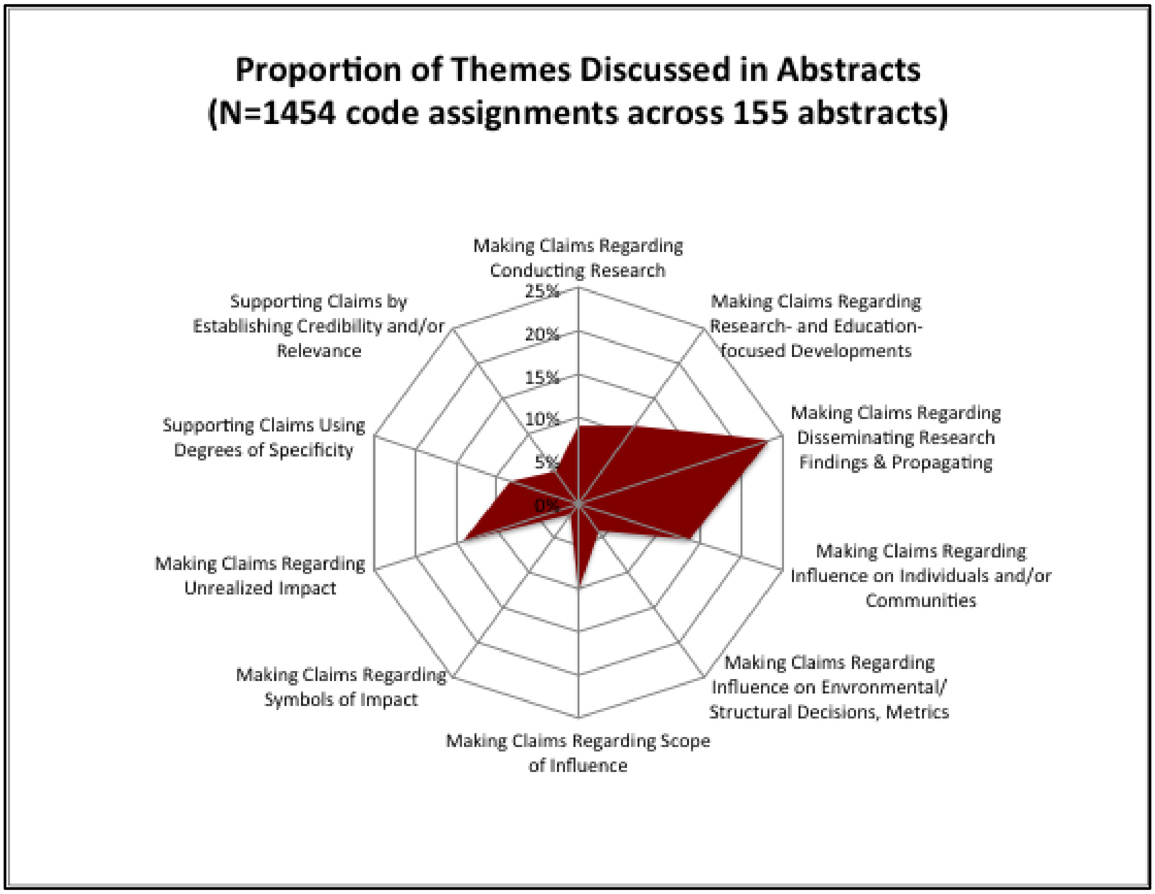


Figure 20. Proportion of Themes Discussed in Abstracts

The three themes that account for the largest proportion of claims that are made about impact are: 1) *claims regarding disseminating research findings and propagating developments;* 2) *claims about unrealized impact;* and 3) *claims about influence on individuals and/or communities.* On the other hand, there are fewer instances of *claims about influence on environmental/structural decisions, metrics;* and *claims about symbols of impact* are practically negligible when looking across abstracts. With respect to ways PIs support claims about impact, there are fewer instances of supporting them by *establishing credibility and/or relevance* than using *degrees of specificity*.

Table 13 shows the co-occurrence of claims PIs make and how they are supported.

Table 13. Co-occurrence of Themes in Abstracts

	Support Using Degrees of Specificity	Support by Establishing Credibility and/or Relevance
Conducting Research	♦♦♦♦🔷	♦♦♦♦
Research- and Education-focused Developments	♦♦♦♦♦ ♦♦🔷🔷🔷	♦♦♦♦🔷
Disseminating Research Findings and Propagating Developments	♦♦♦♦♦ ♦♦♦♦♦ ♦♦♦🔷🔷 🔷	♦🔷🔷
Influence on Individuals and/or Communities	♦♦♦♦♦ ♦♦♦♦♦ ♦♦🔷🔷	♦♦♦♦♦ ♦♦♦🔷
Influence on Environmental/Structural Decisions, Metrics	♦♦♦♦🔷 🔷	♦
Scope of Influence	♦♦♦♦♦ ♦♦♦♦♦ ♦♦🔷🔷🔷 🔷	♦♦🔷🔷🔷
Symbols of Impact	♦🔷🔷🔷🔷	

Frequency Legend	♦ 5 instances	🔷 1 instance

One of the first observations to note at first glance is that PIs tend to support claims *using varying degrees of specificity* significantly more often than by *establishing credibility and/or relevance.* More specifically, when making claims about dissemination & propagation, scope, and influence in individuals, PIs tend to be use (or less) specific language. On the other hand, they are least likely to use this form of support when discussing *symbols of impact* or ideas related to *conducting research.* On the contrary, PIs do not tend to support claims about *symbols of impact* by *establishing credibility and/or relevance*, and rarely is this form of support used when discussing influence on *environmental/structural decision, metric*. It is used, however, when making claims about the influence of a project on *individuals and/or communities*. Now that the types of claims have been discussed in relation to how they are commonly supported, the next set of descriptive statistics will present information on the types of claims made in light of the project parameters.

Figure 21-23 are split-stacked graphs that depict the average number of impact claims made based on project type, project focus, and discipline, respectively. In Figure 21, the themes are arranged in descending order based on the sum total across project types.

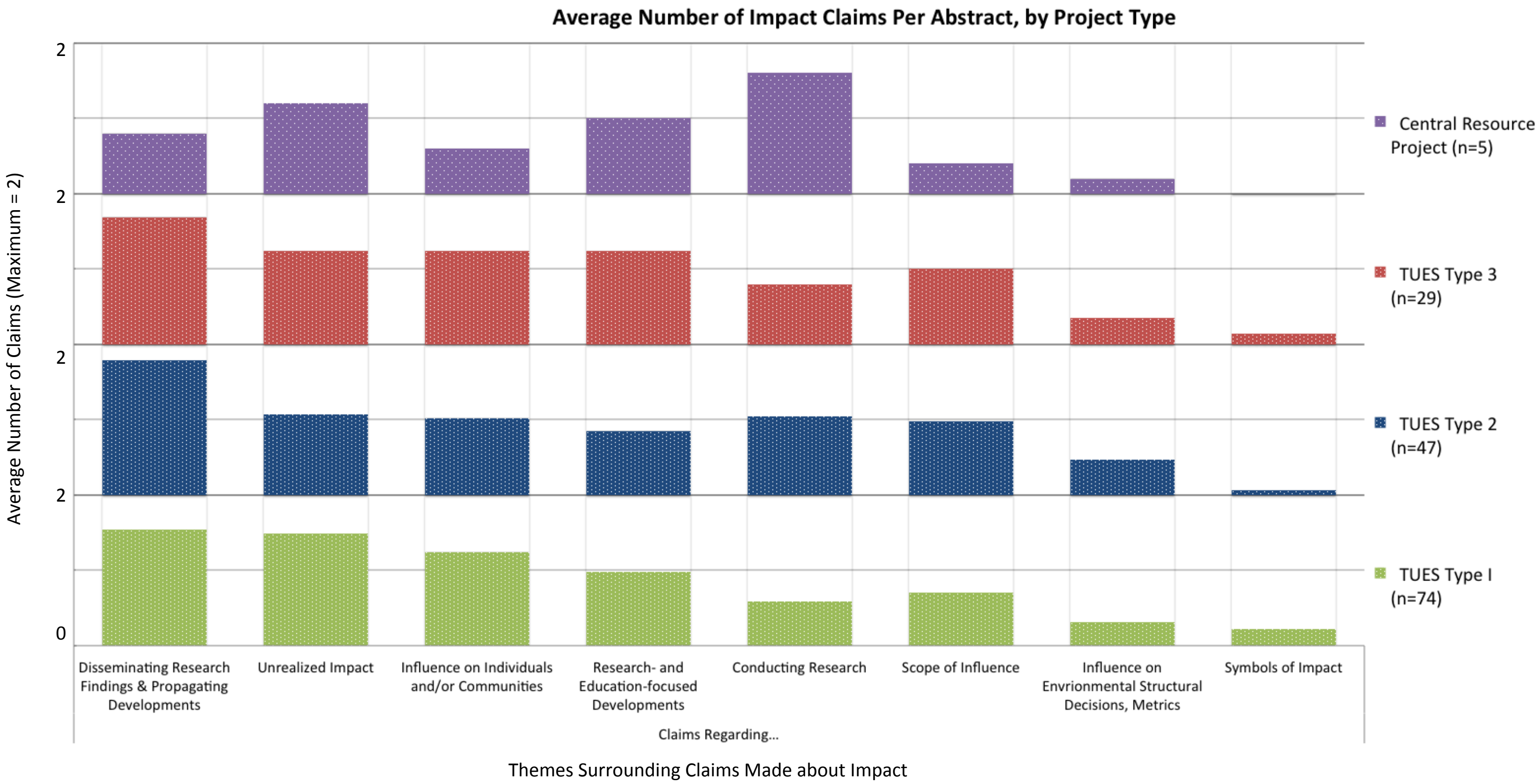


Figure 21. Split Stacked Graph Depicting Proportion of Impact Claims by Project Type

The data shows that there are instances when the average number of claims does not vary much by project type, but there are instances when there are noticeable differences. The two themes with the highest total average (i.e., *dissemination of research findings and propagation of developments*; and *unrealized impact*) are also the two most often mentioned in Type I and Type II projects. The same is true for the two types of claims that are mentioned the least often (i.e., *symbols of impact* and *influence on environment/structural decisions, metrics*). The deviation from the highest and lowest overall sum occurs with the Type III and Central Resource Project (CRP) projects. For Type III, *influence on individuals and communities* along *with education- and research-focused developments* are mentioned just as often as *unrealized impacts*. For CRP projects, *dissemination of research findings and propagation of developments* are not among the claims mentioned the most often; *conducting research* is. Similarly, for the two types of claims mentioned the least often across abstracts (i.e., *symbols of impact* and *influence on environment/structural decisions, metrics*), Type III abstracts have even fewer claims regarding *conducting research* than *influences on environmental/structural decisions, metrics*. Furthermore, the CRP projects did not include any claims regarding *symbols of impact*.

When comparing claims made by projects receiving the largest funding (CRP) to projects with the smallest funds (Type I), there are a few notable differences. Type I project abstracts tend to include five types of claims more often than CRPs: claims regarding *dissemination of research findings and propagation; unrealized impact; influence on individuals and/or communities; scope of influence;* and *symbols of impact.* On the contrary, CRP projects mention claims about *conducting research* more often than Type I projects. The two project types are comparable in the average number of claims about *research- and education-focused developments*, and *influence on environmental/structural decisions, metrics*. The differences in claims made are less apparent when comparing Type II and Type III projects.

A look at the average number of claims based on project focus reveals another set of insights; Figure 22 presents this information. The themes are arranged in descending order based on the sum total across project foci. The project foci are stacked in descending order as well.

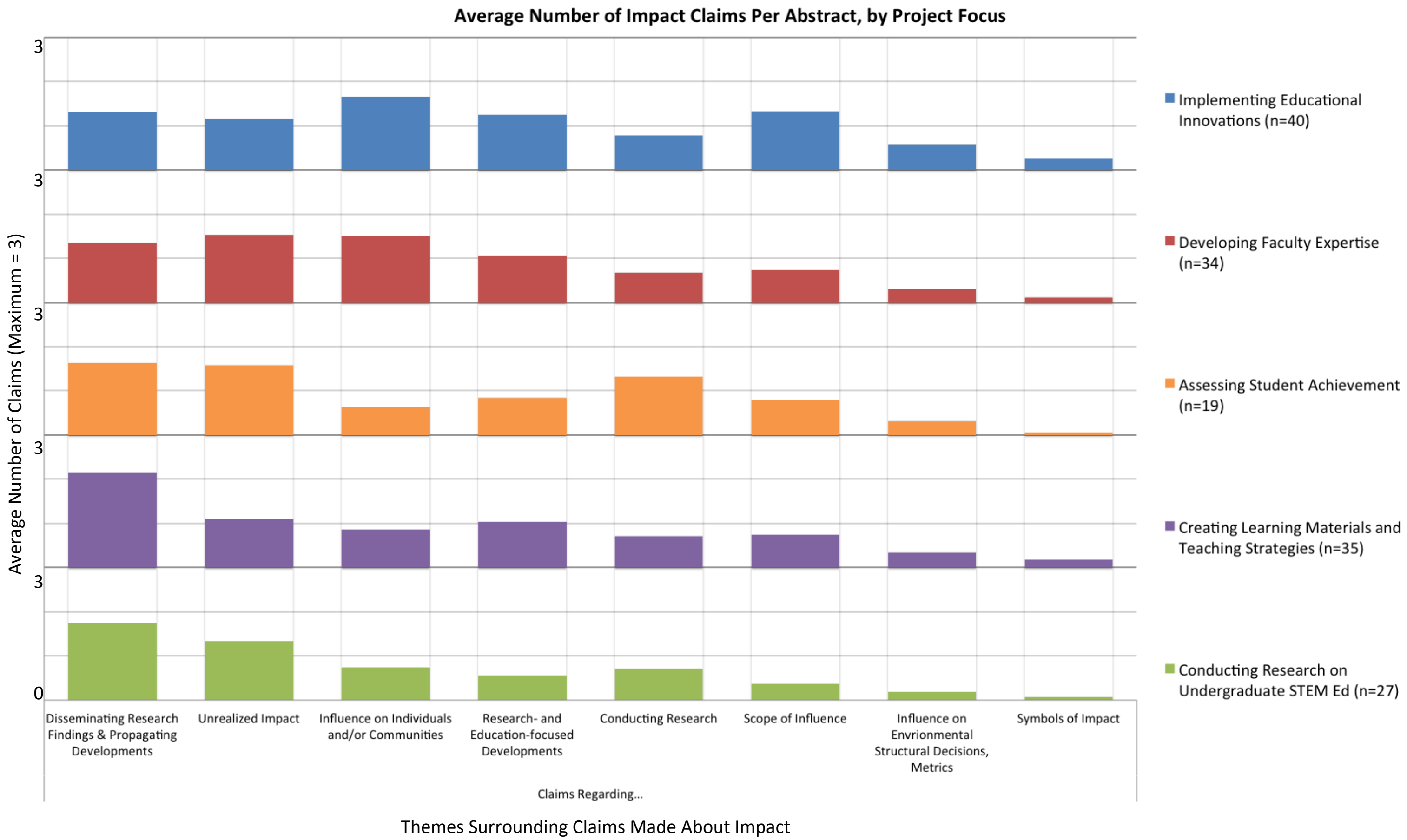


Figure 22. Split Stacked Graph Depicting Proportion of Impact Claims by Project Type

Tables 14 and 15 are an interpretation of the stacked graph in Figure 22. They summarize the most and least commonly mentioned themes, respectively, by project focus.

Table 14. Summary of Themes Most Commonly Mentioned in Abstracts, by Project Focus

Most Commonly Mentioned Themes					
	Project Focus				
Rank	Implementing Educational Innovations	Developing Faculty Expertise	Assessing Student Achievement	Creating Learning Materials	Conducting Research on Undergraduate STEM Education
1	Claims Regarding Influence on Individuals and/or Communities	Claims Regarding Unrealized Impact	Claims Regarding Disseminating Research Findings and Propagating Developments	Claims Regarding Disseminating Research Findings and Propagating Developments	Claims Regarding Disseminating Research Findings and Propagating Developments
2	Claims Regarding Scope of Influence	Claims Regarding Influence on Individuals and/or Communities	Claims Regarding Unrealized Impact	Claims Regarding Unrealized Impact	Claims Regarding Unrealized Impact
3	Claims Regarding Disseminating Research Findings and Propagating Developments	Claims Regarding Disseminating Research Findings and Propagating Developments	Claims Regarding Conducting Research	Claims Regarding Research- and Education-focused Developments	Claims Regarding Influence on Individuals and/or Communities

Table 15. Summary of Themes Least Commonly Mentioned in Abstracts, by Project Focus

Least Commonly Mentioned Themes					
	Project Focus				
Rank	Implementing Educational Innovations	Developing Faculty Expertise	Assessing Student Achievement	Creating Learning Materials	Conducting Research on Undergraduate STEM Education
7	Claims Regarding Influence on Environmental/Structural Decisions, Metrics	Claims Regarding Influence on Environmental/Structural Decisions, Metrics	Claims Regarding Influence on Environmental/Structural Decisions, Metrics	Claims Regarding Influence on Environmental/Structural Decisions, Metrics	Claims Regarding Influence on Environmental/Structural Decisions, Metrics
8	Claims Regarding Symbols of Impact	Claims Regarding Symbols of Impact	Claims Regarding Symbols of Impact	Claims Regarding Symbols of Impact	Claims Regarding Symbols of Impact

With three exceptions, the three themes most commonly mentioned in the abstracts (i.e., *disseminating researching findings and propagating developments; unrealized impact; and influence on individuals and/or communities*) are consistently among the most commonly mentioned across project foci. The first exception, *claims regarding scope of influence,* is second among the most commonly mentioned in projects focused on Implementing Educational Innovations. The second exception, *claims regarding conducting research,* is third in the list of the most commonly mentioned themes for projects that focus on Assessing Student Achievement. The third exception, *claims regarding education-and research-focused developments*, is among the most commonly mentioned in projects focused on Creating Learning Materials.

Unlike the claims most commonly mentioned from the perspective of project focus, there is no variation in the least commonly mentioned impact claim. For all five project foci, the two least commonly mentioned claims are: *claims regarding influence on environmental/structural decisions, metrics* and *claims regarding symbols of impact.* However, projects focused on Implementing Educational Innovations tend to include more claims about *influences on environmental/structural decisions, metrics* than any other project focus.

Finally, analyzing the average number of impact claims based on disciplinary differences leads to other patterns in the data as well. See Figure 23 for the split stacked graph depicting the average number of claims per abstract, organized by discipline. The themes are arranged in descending order based on the sum total across disciplines. The disciplines are stacked in descending order as well.

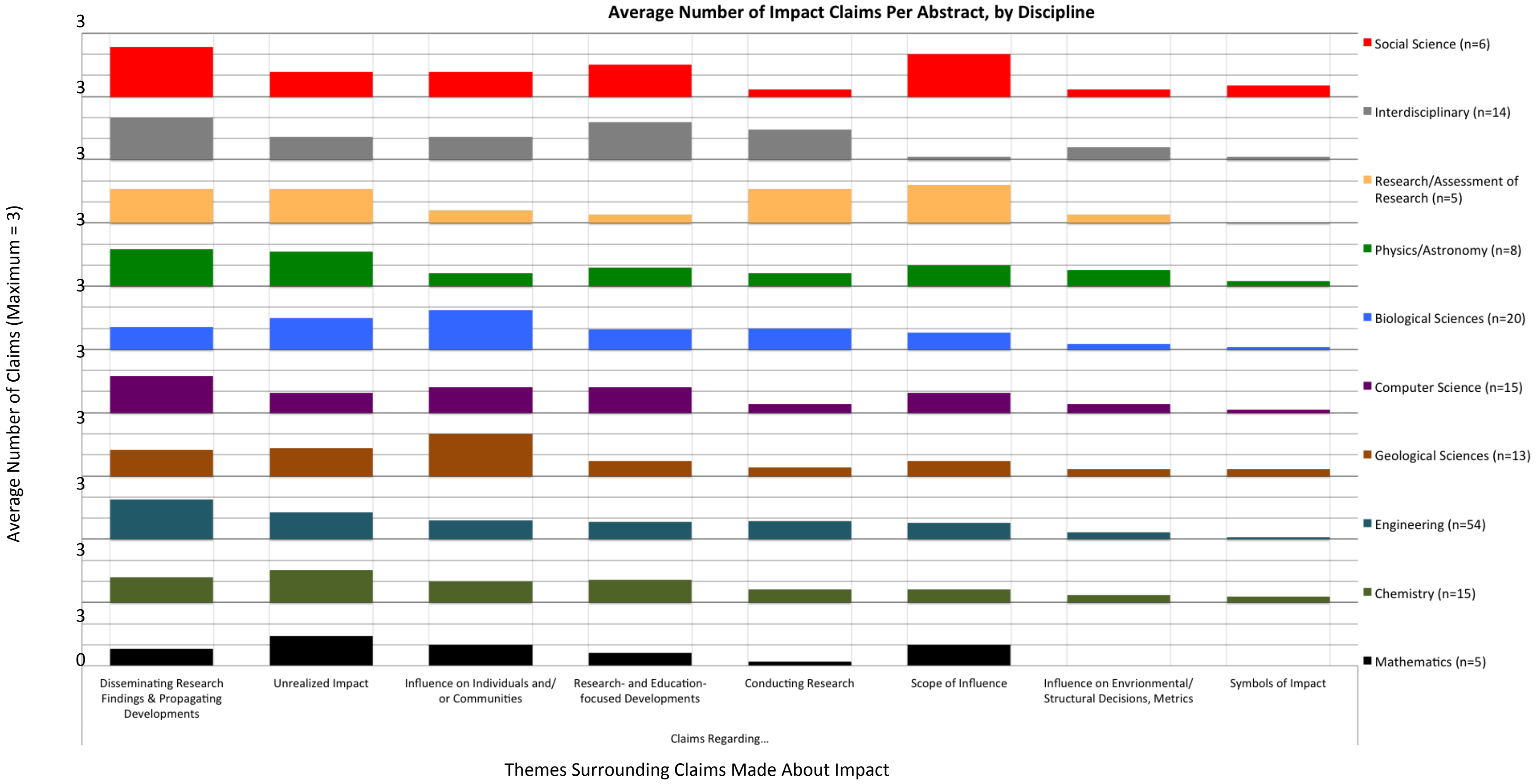


Figure 23. Split Stacked Graph Depicting Proportion of Impact Claims by Discipline

The order of the stacks provides information on the average number of claims mentioned in each abstract based on discipline. Thus, on average, PIs reporting on Social Science projects tend to include more claims about impact than PIs reporting on Mathematics projects. Additionally, Engineering, Geological Sciences, and Computer Science projects tend to include the same number of claims in each abstract, on average.

Table 16 summarizes one interpretation of the stacked graphs. It summarizes the association between the eight themes on making claims about impact and the two disciplines that mention it the most, and the discipline that mentions the theme the least. The color associated with each discipline is the same in both Table 16 and Figure 23 to facilitate ease of referencing.

Table 16. Summary of Themes Most & Least Commonly Mentioned in Abstracts, by STEM Discipline

Claims Regarding…	Rank	Highest Two Disciplines	Rank	Lowest Two Disciplines
Disseminating Research Findings and Propagating Developments	1	Social Science		
	2	Interdisciplinary		
			9	Biological Sciences
			10	Mathematics
Unrealized Impact	1	Physics/Astronomy		
	2	Research/ Assessments of Research		
			9	Interdisciplinary
			10	Computer Science
Influence on Individuals and/or Communities	1	Geological Sciences		
	2	Biological Sciences		
			9	Research/ Assessments of Research
			10	Physics/Astronomy
Research- and Education-focused Developments	1	Interdisciplinary		
	2	Social Science		
			9	Mathematics
			10	Research/ Assessments of Research
Conducting Research	1	Research/ Assessments of Research		
	2	Interdisciplinary		
			9	Social Science
			10	Mathematics
Scope of Influence	1	Social Science		
	2	Research/ Assessments of Research		
			9	Chemistry
			10	Interdisciplinary
Influence on Environmental/ Structural Decisions, Metrics	1	Physics/Astronomy		
	2	Interdisciplinary		
			9	Biological Sciences
			10	Mathematics
Symbols of Impact	1	Social Science		
	2	Geological Sciences		
			9	Research/ Assessments of Research
			10	Mathematics

Provided are some of the most salient insights reflected in the summary table. Social science and Interdisciplinary projects include the most number of claims about *dissemination of research findings and propagation of developments* and *research- and education-focused developments.* Although *claims regarding dissemination of research findings and propagation of developments* represent the largest number of impact claims in abstracts on average across disciplines, they are the least commonly mentioned in Mathematics and Biological Sciences project abstracts. Computer science projects are the least likely to include *claims regarding unrealized impact.* The two disciplines associated with projects that make the most claims about *influence on individuals and/or community* are Geological Sciences and Biological Sciences. Claims about *influence on environmental/structural decisions, metrics* are mentioned most often in abstracts that specify Physics and Interdisciplinary as the discipline. The only discipline that does not appear in Table 16 is Engineering. This may imply that the impact narratives focused on engineering education have a balance of each of the types of claims, without an overemphasis of one over another. While this concludes the presentation of descriptive statistics associated with the themes, the last section of this chapter presents the interpretation of the themes in light of existing literature on research impact.

Alignment of Themes & Existing Literature on Research Impact

The first research question proposed in this study was: *What is a meaningful description of the impact of NSF investments in undergraduate STEM education R&D projects, based on Principal Investigators' (PIs)' perspectives?* The two questions supporting this overarching question were: *What claims do PIs make about the impact of their NSF-funded projects?* and *How do PIs' perspectives of impact align with existing impact frameworks found in the literature to form a preliminary description of the impact of NSF investments in undergraduate STEM education projects?* Using Toulmin's Model (Toulmin, 1958) and the Common Guidelines (Earle et al., 2013), this question was addressed by performing a content analysis of what PIs describe in the Dissemination and Impact section of project abstracts included in PIs conference reports. The results of that analysis reveal eight types of claims that are usually made and two ways of supporting them.

The focus of each theme supports inferences about how some STEM education researchers perceive research project outcomes as impact. Some of these findings are consistent with the current literature on research impact, many that are not, and others that are unique contributions to the scholarly discussion on this topic. Each theme will be briefly discussed with reference to the points of continuity and discontinuity with existing literature.

One idea related to *claims regarding conducting research* aligns with the most familiar dimension of research impact. When PIs highlight unique contributions to the body of literature, this is an expression of the study's scientific impact. With the exception of the framework developed for informal science education projects (Allen et al., 2008) and focused primarily on societal impacts (Molas-Gallart et al., 2002), every framework included in the literature review includes a construct of related to advances in knowledge (Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010; Rymer, 2011; Salter & Martin, 2001; Walter et al., 2003). References to collaborations that developed as part of or as a result of the study is similar to what Salter and Martin (2001) refer to as "forming networks and stimulating social interaction" in their framework. Although securing additional funding to continue research is a specific example of the symbolic impact mentioned in the framework characterizing the impact of science research (Godin & Doré, 2005), merely applying for additional funding is not mentioned as form of impact in the literature on research impact, and does not align with the descriptions associated with any of the three dimensions of research impact. Similarly, there are no references in existing literature to highlights of current research activities as forms of impact –unless these activities involve the creation of new scientific instruments or methodologies (Salter & Martin, 2001).

There are many points of continuity between the forms of impact mentioned in existing literature and ideas supporting the theme *claims regarding education- and research-focused developments.* More specifically, most of the existing frameworks reference the creation of tangible artifacts, instruments or products meant to be most useful to researchers and practitioners in the domain associated with the research that led to the development (Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006). On the other hand, the idea of PIs highlighting the affordances of developments is comparable to the use of qualifying phrases like “more effective” training in one framework (Rymer, 2011) and “increasing the capacity for” scientific and technological problem solving in another (Salter & Martin, 2001).

As it relates to ideas supporting *claims regarding disseminating research findings and propagating developments,* there are two areas of overlap with studies that have already been done. Disseminating research finding through text- and/or discussion-based mediums is part of advancing reliable knowledge— scientific impact. Dissemination through venues oftentimes used for training researchers, practitioners, or students is usually not mentioned in tandem with scientific impact, however. Thus, simply hosting a workshop for instructors is not an example of impact. On the other hand, shifts in the attendees’ epistemology or pedagogical practice as a result of the workshop seems more consistent with what impact means. Furthermore, existing literature on this topic does not reference activities that may lead to propagating developments as forms of impact, but outcomes of such activities (e.g., established partnerships, commercialized materials) could possibly be perceived as being consistent with the formation of networks mentioned in the framework on the impact of research, in general, developed by Salter and Martin (2001).

Ideas related to *claims regarding influence on individuals and/or communities* are significant in this context because they are directly connected to the mission of the university—the setting where learners, instructors, and most STEM education researchers converge. Many of the ideas that support this theme also found in the existing literature on impact. For example, ideas related to the influence of teaching on STEM learners is consistent with nearly all of the frameworks mentioned in the review of the literature (Allen et al., 2008; Godin & Doré, 2005; Levitt et al., 2010; Walter et al., 2003). While influence on learners is commonly mentioned in the literature (e.g., Allen et al., 2008), there are very fewer references to influences on instructors. There may be a variety of reasons for this. For example, this may simply be an oversight on the part of the PI documenting the impact of the study, or could be a reflection of the priorities of program funding the grant. Ironically, although the focus of this study is on research impact, frameworks in the literature rarely mention influence on impact on communities of researchers. The two frameworks that do mention research networks were related to research, in general (Walter et al., 2003) and the impact of health sciences research (Kuruvilla et al., 2006).

The next two types of claims that will be discussed were briefly referenced in the literature. *Claims regarding influence on environmental/structural decisions, metrics* relates to how research may influence administrative decisions or metrics of interest. Although this does not capture the focus on metrics, some aspects of this theme connect with the existing studies on how research informs policy at various levels (Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010). Additionally, ideas supporting *claims regarding symbols of impact* is consistent with the symbolic impact mention in the framework on the impact of science (Godin & Doré, 2005). Apart from mentioning the symbol of affirmation itself (i.e., name of the award, special recognition), information on what lead to the affirmation or the criteria used to determine why the study (or PI) received the affirmation would add more insights on what types of impact is associated with the symbol.

The last two of the eight types of claims PIs tend to make in impact narratives that are not mentioned in the literature are *claims regarding scope of impact* and *claims regarding unrealized impact*. To the extent that other stakeholders would perceive the span of reach as a valid form of impact, this finding would be considered a unique contribution to the body of literature on the impact of research. Five dimensions of scope were identified on this study (i.e., geographic, disciplinary, institutional, target populations, and non-academic partnerships), but additional ways in which a project may have scope could be a topic of a future study. On the other hand, a plausible reason by the literature does not include ideas consistent with *claims regarding unrealized impact* is because from a temporal perspective, impact is usually focused on what has occurred, not what will occur. Moreover, projections about anticipated impact is very different from realized impact.

In light of these points of alignment with existing literature, Figure 24 presents a preliminary description of the impact of undergraduate STEM education R&D.

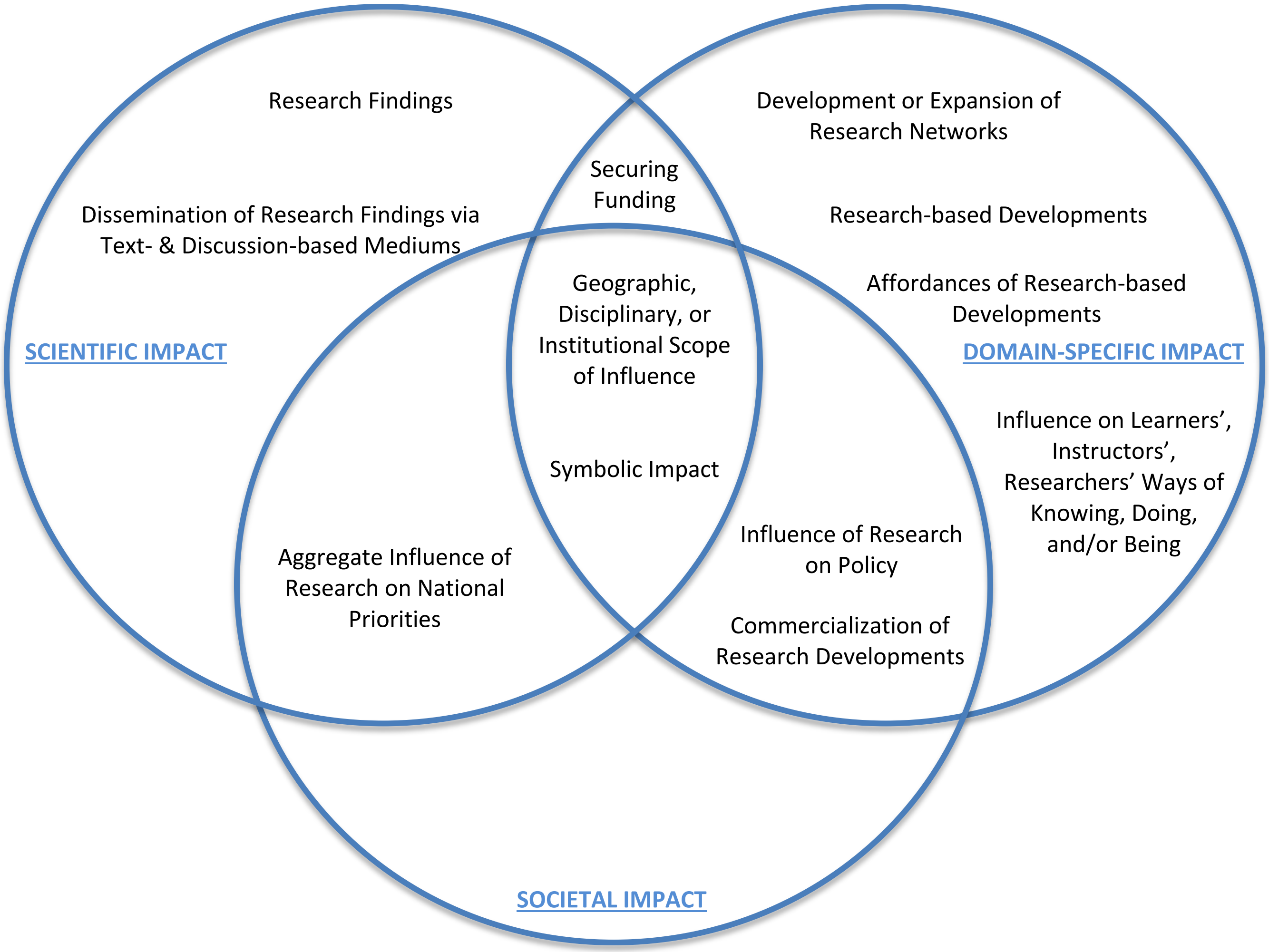


Figure 24. Preliminary Description of Impact of Undergraduate STEM Education R&D

In short, this chapter includes results to the overarching question about a description of the impact of NSF-funded STEM education R&D, according to PIs' perspectives. As a result of this analysis, eight themes emerged to describe the types of claims PIs make in impact narratives, and two themes to denote how these claims are supported. Descriptive statistics on the frequency of these claims among the sample reviewed indicate that PIs tend to support claims about impact using various degrees of specificity more often than by including ideas that help establish credibility and/or relevance. The statistics also reveal that the types of claims PIs make differ more based on project focus and discipline than project type (which corresponds to levels of funding and intended scope). The final section of this chapter described the ways in which PIs claims about impact align and do not align with the three dimensions of research impact. It included a preliminary description of what impact looks like in this context, according to PIs' perspectives, and shows that the ways in which impact is revealed does not always map cleanly to one dimension of research impact. This concludes the qualitative results and interpretation that correspond to the research question posed in the first phase of this study. The next chapter includes the methods associated with the question proposed in the second phase of this study; this will be followed by the corresponding results.

CHAPTER 7: PHASE TWO METHODS

The first phase of this study resulted in themes surrounding the types of claims PIs make about the impact of their work. In this phase, a survey was used to conduct an exploratory study on how the PIs' perspectives on impact compare to NSF Program Officers'. Again, the guiding research question addressed in this phase was: *In what ways do POs' perspectives on the impact of NSF investments in undergraduate engineering education R&D projects align with or differ from PIs' perspectives on impact?* The results of the survey, including closed- and open-ended questions, provided responses to the following sub-questions: *To what extent do POs agree with PIs' perspectives on impact? How do POs talk about the impact of a NSF-funded R&D projects?* The survey results were compared to the phase I result to address the final sub-question proposed in this phase: *Are there consistencies in how PIs on NSF-funded R&D projects and POs overseeing NSF's R&D programs talk about the impact of NSF-funded R&D projects?*

There are a few primary reasons why a survey of POs was used in this study. The first reason is because it served as a form of triangulation. This is especially important because all the data in the abstracts were self-reported, and one researcher worked on data analysis in this study. Research methods surrounding questionnaires and other *self-report inventories* come with numerous ways in which this type of data has the potential to be inaccurate (Johnson & Christensen, 2012). In this context, for example, a PI may report his/her impact in a particular way because of a desire to be perceived favorably by the agency funding their research. Additionally, a lack of insight about the impact of their work or how to communicate it can also lead to inaccurate data. Moreover, it is also plausible for the researcher conducting the analysis to misinterpret a PI's responses. One final reason for surveying POs is because has not been a study to date that includes insights on impact from both the perspective of those conducting the research and those making decisions to fund it. In light of this, a survey was used to garner the perspectives of another set of people who have the potential to understand and respond to PIs' insight on the impact of STEM education research.

Participants

People with "general expertise" on impact in STEM education research were invited to participate in this study by taking a survey. Since the 1980s, cognitive science and cognitive ergonomics researchers continue to study the topic of expertise in light of the individual, interpersonal, and social aspects (Bailey, 1996; Ericsson & Smith, 1991; Garrett, Caldwell, Harris, & Gonzalez, 2009; Sternberg, 1997). As part of proposing a six-dimensional expertise framework that accounts for group level performance on complex tasks, Garrett et al. (2009) summarized some of the most salient research perspectives on the attributes of experts. In short, experts reason differently, process tasks quicker than others (because of extensive practice and skill), have a more complete and organized knowledge of a domain, are more 'intelligent' than others (where intelligence is measured by mental ability and/or creativity), and have more experience to help organize knowledge (Hoffman et al., 1997; Sternberg, 1997 as cited by Garrett et al., 2009). This understanding of expertise is what guided the selection of participants in this study.

Before developing the selection criteria, it was important to remember that the training and experience of those who would agree to take the survey would influence how PIs' statements on impact would be evaluated (J. Grant & Kinney, 1992; J. S. Grant & Davis, 1997). Ideally, experts on *both* research impact and undergraduate engineering education research would have been most suitable. However, it was difficult to identify experts on a niche research area in which there is so little existing literature. Researchers who have published reviews of literature related to research impact (Bornmann, 2013; Jonathan Grant et al., 2010; Salter & Martin, 2001; Walter et al., 2003) and/or developed research impact frameworks for other domains (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010) may possess expertise on research impact, but have much less understanding of undergraduate STEM education research. Since impact looks different in different research disciplines and institutions (Bornmann, 2013; Molas-Gallart et al., 2002), the specialized, disciplinary expertise in relations to federal investments in STEM education R&D was prioritized over domain knowledge of research impact in this study. As this field of research impact becomes more established, future studies that call for "experts" should find it easier to identify people with expertise in both research impact and undergraduate STEM education research.

For the purposes of this study, researchers with academic training and professional experience in undergraduate engineering education and NSF investments in undergraduate STEM education R&D were invited to take the survey. More specifically, the five inclusion criteria for those who were invited were:

(1) Must possess at least one degree in an engineering discipline;

(2) Must have a minimum of two years of professional experience as a NSF Program Officer, preferably in the Division of Undergraduate Education;

(3) Must have professional experience as a PI on at least one NSF-funded research grant, preferably awarded by the Division of Undergraduate Education;

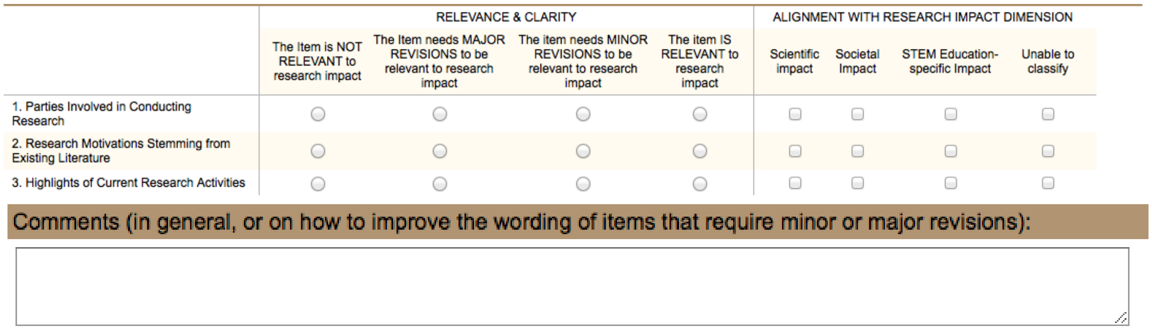
(4) Must have professional experience as a faculty member at a U.S. institution, preferably teaching engineering or engineering education courses;

(5) Must demonstrate research expertise in engineering education, as evidenced by at least one peer-reviewed publications related to engineering education topics published within the past two years.

Starting with a list of 7 people who met the selection criteria, a snowball sampling approach (Johnson & Christensen, 2012) was used to identify and invite others to participate in the study. In total, 14 people were invited; all started the survey; some dropped out before completing it.

Data Collection

Qualtrics™ survey software was used to develop and disseminate the survey. The 33 codes corresponding to the themes in phase I were used as statements in the survey. Participants were asked to rate 33 topical statements for relevance, clarity, comprehensiveness and alignment with the three dimensions of research impact. More specifically, the participants used a Likert-scale to address questions of relevance, and to provide qualitative comments in response to the questions on clarity and comprehensiveness. The labels on the Likert-scale were: (1) This item is *not relevant* to research impact; (2) This items needs *major revisions* to be relevant to research impact; (3) This items needs *minor revisions* to be relevant to research impact; and (4) This items *is relevant* to research impact. Space was provided to comments on the style of individual statements and comprehensiveness of the collection of statements. To evaluate the alignment of a statement with the three dimensions of research impact or “unable to classify”, the participants had the option to select at least one response. The format of the survey was modeled after the content expert validation survey used to validate an instrument that would be used by healthcare professionals to measure the burden on a caregiver (J. S. Grant & Davis, 1997). One person who fit the selection criteria piloted the survey, and provided feedback on how to improve it. The reason for piloting with one person is because the population of people who fit the selection criteria is small and the researcher did not want to reduce the number of possible people who would complete the survey when invited. Figure 25 presents a snapshot of a Qualtrics™ survey question.



	RELEVANCE & CLARITY				ALIGNMENT WITH RESEARCH IMPACT DIMENSION			
	The Item is NOT RELEVANT to research impact	The Item needs MAJOR REVISIONS to be relevant to research impact	The item needs MINOR REVISIONS to be relevant to research impact	The item IS RELEVANT to research impact	Scientific impact	Societal Impact	STEM Education-specific Impact	Unable to classify
1. Parties Involved in Conducting Research	○	○	○	○	☐	☐	☐	☐
2. Research Motivations Stemming from Existing Literature	○	○	○	○	☐	☐	☐	☐
3. Highlights of Current Research Activities	○	○	○	○	☐	☐	☐	☐

Comments (in general, or on how to improve the wording of items that require minor or major revisions):

Figure 25. Sample Survey Items

Apart from the closed-ended questions, participants were asked three general questions to solicit their perspectives on impact. The open-ended questions were:

1) Apart from the items listed above, brainstorm a list of at least three other ways a NSF-funded undergraduate engineering education project may have impact.

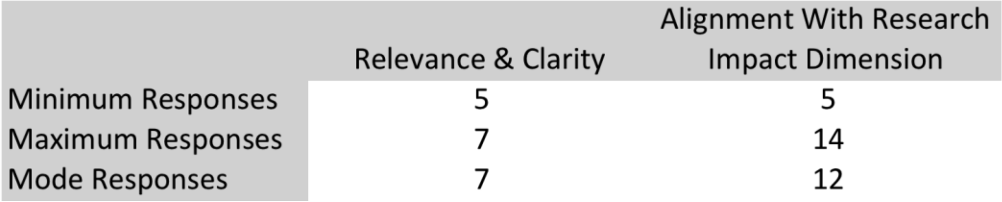
2) Please list any statements that seem unique to undergraduate engineering education R&D.
3) Please list any statements that are not applicable to undergraduate STEM education R&D.

After agreeing to participate, each PO was emailed the following documents: a link to a document summarizing the overarching study, survey instructions, and brief explanations of the 33 statements; and a link to the Qualtrics™ questionnaire..

Survey Response Rate

According to Qualtrics™, all fourteen people started the survey and nine completed it. The minimum, maximum, and modal number of responses per statement varied based on the two types of ratings (i.e. "Relevance and Clarity" vs. "Alignment with Research Impact Dimension"). Table 17 summarizes the response data.

Table 17. Survey Response Data



	Relevance & Clarity	Alignment With Research Impact Dimension
Minimum Responses	5	5
Maximum Responses	7	14
Mode Responses	7	12

A minimum of five experts rated all of the statements for relevance and clarity and for alignment with research impact dimensions. At most, half of the experts rated the statements for relevance and clarity. There are instances were all of the experts rated a statements' alignment with the three dimensions of research impact. The inconsistency in the number of responses despite having both types of ratings side-by-side (see Figure 25) suggests a potential lack of clarity in the instructions participants were given or misinterpretation of the instructions.

Data Analysis

Three forms of analysis occurred in this phase of the study. Percent agreement was used to determine the extent to which POs agreed with PIs' perspectives on impact. A PO was considered to be in agreement with the statement if the item was rated: "This Item IS RELEVANT to research impact" or "This item needs MINOR REVISIONS to be relevant to research impact". After calculating percent agreement, the Likert-scale data was used to determine the extent to which POs agreed with one another on the relevance of the items. The interrater agreement for relevance was evaluated using Krippendorff's α (Krippendorff, 1970, 2004a). SPSS statistical software and a SPSS macro for generating this value (described elsewhere (Hayes & Krippendorff, 2007)) was used to perform the calculation of the α–value. The rational for using Krippendorff's α is because it is a measure of agreement for nominal data, it accounts for chance agreements among raters, and can be used regardless of the number of raters. Apart from calculating percent agreement and the alpha-value, the POs' qualitative responses were synthesized into preliminary findings on their perspectives of impact. These responses were then compared to those of the PIs to determine if there were any consistencies. The next chapter presents the results of the survey study conducted in this phase.

CHAPTER 8: PHASE TWO RESULTS

This chapter begins with the results corresponding to the research question posed in the second phase of this study: *In what ways do NSF Program Officers' (POs) perspectives on the impact of NSF investments in undergraduate STEM education R&D projects align with or differ from PIs' perspectives on impact?* It includes the percent agreement between PIs and POs on the statements related to research impact, along with the statistic on the extent to which the POs agreed with one another when answering the survey questions. It closes with the POs' qualitative insights on impact that were included in the survey.

Survey Results

Table 18 presents the participants' opinions on each statements' relevance, and clarity, and alignment with the dimensions of research impact. Since the survey responses were anonymous, there is no way to know exactly which participant provided which rating. However, the symbols in the table correspond to the data Qualtrics provides on the total number of responses for each item. The items are arranged in descending order, starting with the items all seven experts who rated as: "This Item IS RELEVANT to research impact" or "This item needs MINOR REVISIONS to be relevant to research impact".

Table 18. Summary of Survey Responses

Statement	Expert							% Agree	Research Impact Dimension				N
	1	2	3	4	5	6	7		Sc-I	So-I	DM-I	UTC	
Influence of Training on STEM Education Instructors and/or Community of Instructors	X	X	X	X	X	X	X	100.0%	8%	42%	50%	0%	12
Department-level Assessments and/or Outcomes	X	X	X	X	X	X	X	100.0%	10%	20%	60%	10%	10
Curricular Changes	X	X	X	X	X	X	X	100.0%	17%	33%	50%	0%	12
Geographic scope	X	X	X	X	X	X	X	100.0%	15%	38%	31%	15%	13
Disciplinary scope	X	X	X	X	X	X	X	100.0%	29%	21%	50%	0%	14
Informing or Enacting Education Policy	X	X	X	X	X	X	-	85.7%	20%	20%	50%	10%	10
Informing Promotion & Tenure Decisions	X	X	X	X	X	X	-	85.7%	20%	10%	50%	20%	10
Activities Supporting Propagation of Developments	X	X	X	X	X	X	-	85.7%	33%	25%	33%	8%	12
Outcomes of Activities Supporting Propagation of Developments	X	X	X	X	X	X	-	85.7%	29%	36%	29%	7%	14
Influence of Teaching on STEM Learners and/or Community of Learners	X	X	X	X	X	X	-	85.7%	8%	46%	46%	0%	13
Influence of Research on STEM Education Researcher and/or Research Community	X	X	X	X	X	X	-	85.7%	27%	27%	36%	9%	11
Institutional scope	X	X	X	X	X	X	-	85.7%	17%	33%	42%	8%	12
Scope via target populations	X	X	X	X	X	X	-	85.7%	9%	45%	36%	9%	11
Non-academic partnerships	X	X	X	X	X	X	-	85.7%	20%	60%	20%	0%	10

Legend
X: Item Rated 3 or 4 on 4-point Relevance Scale
(-): Item Rated 1 or 2 on a 4-point Relevance Scale
NR: Item Not Rated by the Expert

Legend
Sc-I: Scientific Impact
So-I: Societal Impact
DM-I: Domain-specific Impact
UTC: Unable to Classify

Table 18 (continued). Summary of Survey Responses

Statement	Expert							% Agree	Research Impact Dimension				N
	1	2	3	4	5	6	7		Sc-I	So-I	DM-I	UTC	
Instruments, Technology Used for Educational Purposes	X	X	X	X	X	-	-	71.4%	23%	23%	46%	8%	13
Dissemination via Venues for Teaching, Training	X	X	X	X	X	-	-	71.4%	18%	27%	55%	0%	11
Research Findings Lead to Insights about Structural Issues	X	X	X	X	X	-	-	71.4%	22%	22%	33%	22%	9
Curricular materials, training resources, and pedagogy	X	X	X	X	X	-	-	71.4%	17%	33%	42%	8%	12
Research Motivations Stemming from Existing Literature	X	X	X	X	-	-	-	57.1%	23%	31%	31%	15%	13
Highlights of Current Research Activities	X	X	X	X	-	-	-	57.1%	33%	25%	25%	17%	12
Text-based mediums	X	X	X	X	-	-	-	57.1%	46%	23%	31%	0%	13
Text-based entities	X	X	X	X	-	-	-	57.1%	44%	0%	33%	22%	9
Discussion-based entities	X	X	X	X	-	-	-	57.1%	27%	18%	27%	27%	11
Facility Instituted, Technology Developed for conducting research	X	X	X	X	-	-	-	57.1%	50%	20%	20%	10%	10
Discussion-based mediums	X	X	X	X	-	-	-	57.1%	42%	25%	25%	8%	12
Indirect and/or Expected Outcomes	X	X	X	X	-	-	NR	66.7%	36%	27%	18%	18%	11
Parties Involved in Conducting Research	X	X	X	-	-	-	-	42.9%	25%	25%	42%	8%	12
Indeterminate Outcomes	X	X	X	-	-	-	NR	50.0%	33%	22%	22%	22%	9

Legend
X: Item Rated 3 or 4 on 4-point Relevance Scale
(-): Item Rated 1 or 2 on a 4-point Relevance Scale
NR: Item Not Rated by the Expert

Legend
Sc-I: Scientific Impact
So-I: Societal Impact
DM-I: Domain-specific Impact
UTC: Unable to Classify

Table 18 (continued). Summary of Survey Responses

Statement	**Expert**							**% Agree**	**Research Impact Dimension**				**N**
	1	**2**	**3**	**4**	**5**	**6**	**7**		**Sc-I**	**So-I**	**DM-I**	**UTC**	
Direct, Personal Benefits of Instructors, Researchers	X	X	-	-	-	-	-	28.6%	14%	14%	29%	43%	7
Applying for and Securing Additional Funding to Continue Research	X	X	-	-	-	-	-	28.6%	11%	11%	44%	33%	9
Affordances of the Developments	X	X	-	-	-	-	NR	33.3%	20%	20%	20%	40%	10
Future plans to conduct research, create developments, and/or disseminate	X	X	-	-	-	-	NR	33.3%	33%	17%	17%	33%	6
Highlights of Research Activities	X	-	-	-	-	NR	NR	20.0%	0%	20%	0%	80%	5

Legend

X: Item Rated 3 or 4 on 4-point Relevance Scale

(-): Item Rated 1 or 2 on a 4-point Relevance Scale

NR: Item Not Rated by the Expert

Legend

Sc-I: Scientific Impact

So-I: Societal Impact

DM-I: Domain-specific Impact

UTC: Unable to Classify

The percent agreement results provide some interesting preliminary findings on PIs' versus POs' perspectives on the impact of research. Among the five statements that have 100% agreement that they are relevant to research impact, two are related to the theme about the *scope* of a project (i.e., geographic scope, disciplinary scope). In fact, all of the statements corresponding to the theme about scope had at least an 85% agreement. As it was previously mentioned in the interpretation of the Phase I results, scope of influence is the dimension of research impact that emerged from this study's qualitative analysis, and was not in existing literature on research impact. The survey results presents a preliminary finding that suggest there is an agreement between PIs and POs that the scope of influence of an R&D project is an indicator of impact – where scope may be geographic, disciplinary, include target populations, or reach across institutional types (e.g., academic & non-academic partnerships).

Meanwhile, two other statements out of five with 100% agreements are related to the theme about the influence of a project *on environmental/structural decisions, metrics*. Apart from these two themes, the two other things that are the most commonly represented among the statements with 85% agreement or better are: *influence on individuals and/or communities* and *disseminating research findings and propagating developments.* More POs agreed with PIs that impact comes in the form of *influence on instructors and/or instructor communities* (100% agreement) rather than *influence on STEM leaners* (85% agreement). This finding is somewhat surprising because the impact narratives in abstracts include many more claims about the influence of research on STEM learners and rather than STEM instructors.

Among the statements with the lowest percent agreement between POs and PIs, those corresponding to the theme about *conducting research* are the most common (e.g., future plans to conduct research – 33%; applying for and securing additional funding – 28%). The POs disagreement with PIs that such ideas count as impact is consistent with the existing literature; references to the mundane details associated with conducting research does not align with the dimensions of impact. Based on the number of abstracts that included claims related to conducting research, this is potentially another point of disconnect between PIs and POs perspectives on impact. Lastly, more than half of those who rated the statement on related to *unrealized impact* (e.g., indirect and/or expected outcomes) agreed with PIs that this is relevant to impact; although the two stakeholders agree with one another, this is not consistent with existing literature on the topic. This is another area worthy of future study.

As it relates to POs perspectives on the alignment of the statements with the three dimensions of research impact (i.e., scientific, societal, and domain-specific), other findings emerge. What is apparent at first glance of the values is that the responses are spread across the dimensions. In some instances, two out of three dimensions are significantly higher than the third, but there is no instance where there is a one-to-one-mapping between a statement about impact and a dimension of impact. This idea of little to no one-to-one mapping is consistent with the interpretation of the Phase I results, which show, for the most part, that facets of impact do not map cleanly to one dimension of impact or another.

There are a few notable mappings that need to be discussed. Among the few most notable ratings on the alignment of a statement with an impact dimension, *influence on instructors and/or instructor communities* was almost evenly split between domain-specific impact and societal impact. Not surprisingly, two statements that map the most clearly to domain-specific impacts were *department level outcomes* and *curriculum changes*. On the other hand, *non-academic partnerships* maps the most cleanly to societal impact—which also is not surprising, since this dimension of research impact includes the influence of research on the economic capital of a nation.

After using percent agreement to determine the extent to which POs agreed with PIs' perspectives on impact, it was interesting to determine the extent to which the POs agreed with one another. Using Krippendorff's alpha (Hayes & Krippendorff, 2007; Krippendorff, 1970) to determine the interrater agreement among experts, the analysis resulted in analysis an **α- value of 0.0166.** This dismal value indicates that there was little agreement among participants. The lack of consistency in responses is part of the reason for such a low alpha-value; other possible reasons are discussed in the next chapter. Before going there, however, the last section of this chapter will include a synthesis of this participants' qualitative responses to the open-ended items in the survey.

Synthesis of the Experts' Qualitative Feedback

The comments in the survey provide meaningful insights about the survey design and a small glimpse of the experts' perspectives on research impact. Various comments in the survey expressed confusion about how to complete the survey. Apart from comments related to the survey design, there were a variety of comments that express discontentment with some of the types of impact claims PIs tend to make in abstracts—at least, as they are depicted in the survey items. Consider the following comment:

> *"Conducting the activity by itself does not equate to impact. The quality and efficacy of the activities is important to determine impact."*

This comment was mentioned in response to the set of statements corresponding to the theme, claims regarding *disseminating research findings and propagating developments.* It seems to express what additional information PIs could provide when discussing dissemination and propagation-related activities. This additional information is not necessarily impact either, but it would help "to determine impact."

In reference to statements corresponding to the theme, claims regarding *influences on individuals and/or communities,* someone said the following:

> *"The types of influences need to be described to make these statements more relevant."*

This statement is consistent with responses from several experts who commented on the vagueness of the claims they were evaluating, and the need for more specificity in the PIs descriptions of impact. Provided is one other example of an expert commenting on the vagueness of the statements corresponding to PIs' claims, and also sharing guidance on how to improve them.

> *"I found items 1 and 2 very vague and my ranking is heavily influenced by how I interpret them. …These statements need to be more fully developed to include evidence generated to support the quality and efficacy of the developments."*

One comment that was, arguably, the most aggressive among the comments was:

> *"Do you really think given the power relationship inherent between most PIs and NSF that they will report 'unrealized' impacts?"*

In this rhetorical question, there is an underlying sentiment of quarrel with the idea of including in an impact narrative the ideas associated with the theme *unrealized impact.* The response does not provide any additional information on what PIs should provide instead, but it does seem to imply that including ideas like future plans to conduct research and anticipated impacts should not be included in the narrative on the impact of a project.

In one instance, one participant largely agreed with the statements corresponding to a theme. In response to the statements corresponding to *claims regarding scope of impact,* one expert said:

> *"These for the most part seem pretty clear. It seems to be useful to define ranges for each of these. The target population one is not clear as written."*

Thus, in addition to affirming that the statements corresponding to this theme seem relevant to research impact, they shared feedback on how to make them even clearer.

In addition to providing comments on the survey design and disagreements with the statements reflecting the types of claims PIs make in impact narrative, there were also comments reflecting the participants' conceptual understanding of impact. One person stated, *"impact is very context-dependent",* while another shared suggestions on what a research impact framework might consist of:

> *"I might suggest you think about first and second order impacts. I could better use this reductionist framework if you clarified the closeness of impacts or interactions."*

Another began to tease out differences between outcomes, outputs and impact. The following comment was written in response to the set of statements corresponding to claims regarding *research- and education-focused developments:*

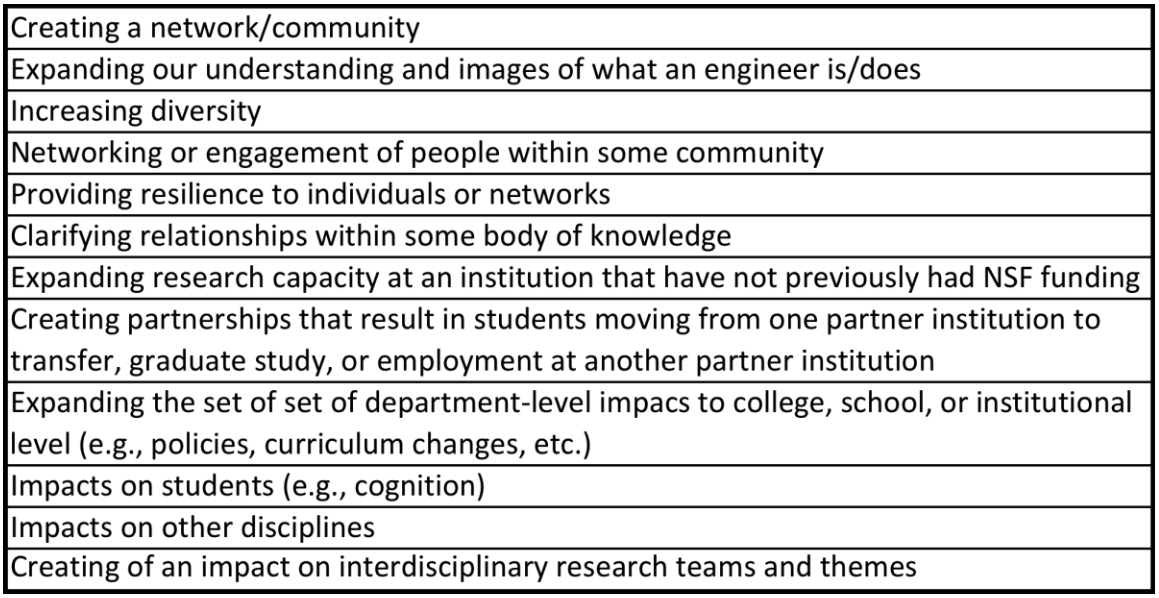
> *"For items 1-8, a key question is on 'Outputs' versus 'Outcomes' in judging 'impact'. If the research is to determine if I can produce a given type of curricular material, then its creation is a noteworthy research outcome. If the question is the effect of that curriculum on student learning, then the materials are an intermediate product in a longer study. Similarly, creating a journal does not appear to demonstrate substantial impact,* UNLESS *getting to that stage catalyzes a new community the publishes in the journal."*

This is specific example of the types of *first and second order impacts* mentioned in the previous comment and is concrete example of how some of the participants conceptualize impact.

One of the last survey questions asked participants whether they thought the statements about impact seemed unique to undergraduate engineering education research and could be not extended to undergraduate STEM education research, in general. The majority of those who responded said that they did not find the statements to be unique to undergraduate engineering education research; but one stated that *"the nature of impact on educational practices seems unique to me."* This preliminary information is helpful for considering the extent to which the findings might be useful to other STEM disciplines besides engineering education.

The last set of comments was in response to a final survey question inviting the experts to brainstorm ways a NSF-funded undergraduate engineering education project may have impact. Although there were only a few participants who responded, Table 19 is the list of ideas they submitted in response.

Table 19. List of Examples of Impact Generated by POs



Creating a network/community
Expanding our understanding and images of what an engineer is/does
Increasing diversity
Networking or engagement of people within some community
Providing resilience to individuals or networks
Clarifying relationships within some body of knowledge
Expanding research capacity at an institution that have not previously had NSF funding
Creating partnerships that result in students moving from one partner institution to transfer, graduate study, or employment at another partner institution
Expanding the set of set of department-level impacs to college, school, or institutional level (e.g., policies, curriculum changes, etc.)
Impacts on students (e.g., cognition)
Impacts on other disciplines
Creating of an impact on interdisciplinary research teams and themes

In short, the closed-ended survey responses help to provide insight on how PIs and POs perspectives on impact compare, while the survey participants' comments on what impact means in this context serve as a fortuitous contribution to the overarching goal of this study – to investigate what it means for a federal-funded STEM education R&D project to have impact.

CHAPTER 9: DISCUSSION & IMPLICATIONS

Although there is a lot of interest among stakeholders at various levels of governance in the impact of investments in research, there have been few scholarly studies on this topic. The current literature on research impact highlights three dimensions: scientific impact, societal impact, and domain-specific impact. There is a substantial body of work on scientific impact, while societal impact and domain-specific impact are less understood. Over the last decade, a number of frameworks have been developed to characterize research in general (Molas-Gallart et al., 2002; Rymer, 2011; Walter et al., 2003) and in specific research domains (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010). This scholarly work is not only the first to weave the existing bodies of literature together, it add to the body of knowledge by investigating how researchers on NSF-funded undergraduate STEM education R&D projects discuss the impact of their work, and exploring POs' perspectives on the PIs' perspectives of impact – with hope that the field of engineering education will develop a valid, comprehensive framework for characterizing impact in this context. While some of the sub-questions associated with the two overarching research questions proposed in this study included interpretations of some of the research findings, this chapter weaves together additional interpretation of the results findings in light of existing literature, study limitations and presents corresponding suggestions for future research directions. It concludes with implications for policy and practice.

Discussion

Discussion on the Research Design

It is only fitting to start by discussing how the study was framed. The dearth of research on how to study impact in a scholarly way translated to little guidance on to frame this study. What is commonly among the few studies seeking to understand and characterize what the impact of research looks like for a specific domain (Allen et al., 2008; Donovan & Hanney, 2011; Godin & Doré, 2005; Kuruvilla et al., 2006; Levitt et al., 2010) is the use of method designs that start with an exploratory approach, and are oftentimes supplemented with quantitative analysis. This study was framed similarly. The ideas in the two frameworks that were used to guide the qualitative analysis (Earle et al., 2013; Toulmin, 1958) served as sensitizing concepts—more specifically, they provided ideas on what to expect in terms of how researchers may construct arguments about impact and what type of evidence may be associated with certain types of research. Although existing literature on the use of Toulmin's model discussed its limitations in terms of using it to evaluate arguments (Klumpp, 2006; Tans, 2006; Voss, 2006), it was an useful analytic tool for trying to understand PIs' verbal reasoning surrounding the impact of their work. (Connections between the research findings and the frameworks will be discussed next—after completing the discussion on how the study was framed.) The second phase of the study included a survey of the opinions of a select group of people, people very familiar with both engineering education research and federal investments in it. The use of different selection criteria would have lent itself to a larger sample size and an ultimately, an increase in the number of responses. On the other hand, the survey results suggest that an exploratory approach could have also been an appropriate choice for garnering POs' perspectives on impact. (A discussion on the Phase I research findings and future research directions will be presented in the next sections.)

There is one final point of discussion related to how the study was framed, and it relates to the data that was used in the first phase of this study. The inclusion of "Broader Impacts" in NSF's review criteria sends a clear message that the agency cares about impact. Early on in the stages of the grant life cycle process (see Figure 1), PIs are expected to write about the potential impact of their project in their grants, and by extension, the panel review process is one mechanism for evaluating (potential) impact. While conducting the research, PIs submit progress reports (i.e., annual and final reports), which include rich information about project-related decisions, rationale, outcomes, and accomplishments. If at all, the contents of these reports are rarely assessed in a way that holds PIs accountable for what they proposed to do when they requested funding; and, until the recently, no version of this information was publicly-available for other stakeholders to review and/or study. This was one of the first scholarly studies on the realized impact of a collection of NSF-funded projects, and it relied on small sections of project abstracts in the conference report of a particular NSF program because it is what was available. To the extent that is it is true that one of the best sources of information on the impact of projects in which millions of taxpayers' dollars are invested is buried in small sections of project abstracts in the PI conference reports of some NSF programs, this may imply that impact is largely invisible from the current reporting processes, and speaks to the the urgent need for improved infrastructure.

Discussion on the Research Findings

The goal of the research question proposed in the first phase of the study was to understand and characterize how PIs talk about the impact of their NSF-funded R&D projects. The themes that resulted form the abstract analysis showed what is talked about—and what is not. What PIs discuss are a breadth of topics ranging from research activities, research outputs, and influence on populations. One example of an idea that is not discussed, however, is the negative impact of a project. Moreover, in some ways, these findings support the idea that impact is commonly used interchangeably with terms like outputs and outcomes (Brewer, 2011; Martin, 2007). The lack of focus and coherence in the ideas that are discussed may also provide evidence in support of the idea that researchers may really struggle with knowing how to communicate the impact of their work or evaluate the contents of Broader Impact statements in NSF proposals (Holbrook & Frodeman, 2011).

One of the most surprising themes that emerged from the qualitative analysis conducted in the first phase was *unrealized impact.* Although the ideas supporting this theme are not consistent with any of the dimensions of impact mentioned in existing research impact frameworks, it does help shed light on how researchers may think about research impact. Given how frequently this theme appeared in the sample of abstracts reviewed, it seems as if PIs are so certain about what might potentially happen in the future, they see no problem with reporting it as impact now; said differently, PIs may perceive that there is no need to distinguish between realized impact and expected/anticipated impact. Furthermore, PIs only have the option of writing about the potential and expected impact of a project in their NSF grant proposals; however, the prevalence of this practice might suggest many perceive an implicit expectation to continue hypothesizing about impact when reporting project outcomes and impact. In some ways, this phenomenon is consistent with the biggest problem with studying impact—*the attribution problem* (Bornmann, 2013; Godin & Doré, 2005; Jonathan Grant et al., 2010; Martin, 2007; Rymer, 2011; J. E. Scott et al., 2011; Spaapen & van Drooge, 2011). As this scholarship on impact continues to grow, it will continue to be difficult to make proper attribution if claims about realized and unrealized impact are made together and are oftentimes indistinguishable.

While the types of claims PIs make about the impact of their work were closely related to the dimensions of impact in existing literature, the way they supported their claims is connected to elements of Toulmin's model (Toulmin, 1958). To support their claims about impact, PIs often use various *degrees of specificity* and add information that helps to *establish credibility and/or relevance.* Using Toulmin's model (Toulmin, 1958) as a tool for interpretation, the quantification of claims is a form of *data* that supports the claims, whereas the use of qualifying statements that limit the scope of the claim is similar to the *qualifiers* mentioned in Toulmin's model. Hints at evaluations used in the study tend to imply certain types of evidence is, in some ways, related to the use of *warrants* in Toulmin's model since warrants are used to make connections between claims and data. There are limits to this interpretation of the results, but it is possible to draw similar inferences from what the PIs writes when a reader is presented with a warrant and/or evaluation information in an impact narrative.

The descriptive statistics show how often different types of claims are mentioned based on the project type, focus and STEM discipline. It reasonable to assume that impact may vary based on the amount of funding awarded to a project. Surprisingly, this data shows that there were few notable differences in the types of claims PIs make when comparing across projects with different levels of funding, however. Part of the reason for this may be because the differences in impact may not occur until many years or decade after the project is completed (Bornmann, 2012). Additional studies are needed to explore the extent to which this finding is consistent across programs and to trace the impact of projects after the grant lifecycle—since a longer horizon will provide more time for the differences impact to be evident.

The research findings also indicate that the project focus and discipline seem to have more influence on the types of impacts that are observed than the amount of funding award (at least in the early stages of the project). While differences in impact based on the focus of the project might seem intuitive, further research explorations into the interesting differences in the claims based on disciplines are also needed. Furthermore, this study took the perspective that there was no reason to assume that PIs reasoning about the impact of engineering education R&D projects would significantly vary from the ways PIs on other STEM education R&D projects since engineering is an integrative discipline among the STEM disciplines. By doing so, the findings and implications of this study have the potential to be applicable to all of the STEM disciplines – although it does not give the finest resolution to impact of research in any particular STEM discipline. However, the absence of engineering in the table on the most and least commonly mentioned themes (Table 16) suggests that this assumption may need to be revisited. It might be valuable to conduct a future study that does not make this assumption, and in turn, focuses solely on the impact of federally funded engineering education R&D projects—especially in light of the engineering education community's research agenda ("The Research Agenda for the New Discipline of Engineering Education," 2006).

Another finding from this study was the connection between the types of claims PIs make about the impact of their work and how they support them; the table on the co-occurrence of themes (Table 13) was one way of representing this connection. One of the most salient findings was that PIs tend to support claims by providing additional details about the claims they are making more often than providing information that helps bolster credibility and/or provide evidence of relevance. Although the purpose of this analysis is not to suggest that the ways of support claims about impact needs to be equally represented in each abstract, it is reasonable to suggest that PIs could do a better job at the latter of the two. Two specific ways to help establish credibility and/or relevance in the minds of the readers include: hinting at the evidence that supports their claims (e.g., changes in students' grades, evaluation data) and adding statements which explicitly connect the results of the current study to issues of importance in the PIs' research community or priorities of the program funding the study.

Phase I of this study concluded with a preliminary description of what it means for a federally-supported STEM education project to have impact (see Figure 24). This interpretation of the themes in light of existing literature revealed that there areas of consistency (and inconsistency) in how PIs talk about impact, and the existing literature on the topic. Additionally, some of the forms of impact that are also mentioned align with the literature on how change happens in education (e.g., via publishing, teaching, policies) (e.g., Beach et al., 2012; Burkhardt & Schoenfeld, 2003). One the unique contributions to the literature is the idea that one form of impact is the *scope of influence* associated with project, where scope may have multiple dimensions spanning geography, disciplines, or populations of people. The three dimensions of research impact provided a nice framework for not only organizing the forms of impact associated with the themes, but also show how impact can be "messy" and does not always cleanly map to one dimension of research impact.

The research questions posed in the second phase of this study focused on how POs' perspectives on impact compared to PIs'. In some ways, this preliminary study suggests that there may be points of overlap in how the two groups talk about the impact of research. An example of consistency in the perspectives of the two groups is their agreement on the scope of influence of a research project as a form of impact. This was an exploratory study on how these two groups' perspectives compare, but a larger scale study is needed to provide more evidence of the alignment, differences, and potential mismatches in PIs' and POs' perspectives on the impact of a federally-funded project. As part of such a follow-up study, it would be interesting to investigate the extent to which the differences in immediate contexts of PIs (i.e., academic & disciplinary) and POs (i.e., government) might influences what might be observed.

One of the puzzling results of the survey study was the very low interrater agreement value. Krippendorff and Hayes (2007; 2011) provide multiple ways to interpret an extremely low alpha value. In short, extremely high or extremely low alpha may indicate measurement error or floor/ceiling effects. An extremely low alpha value may indicate that the participants did not agree on the ratings provided (e.g., "This Item is Relevant to Research Impact", "This Item is Not Relevant to Research Impact"). As a result, it is possible that the rating scale provided may have been unclear and that each participant interpreted the items differently. Thus, the survey would induce a level of measurement error, because the scale used to collect the data may have flaws that will affect the quality of the data collected. Because reliability statistics are affected by sample size and variation, another cause of low reliability could be because of too few participants or too few items. Smaller samples will tend to have less variation and are more likely to have lower reliability, and vice versa. Another way to interpret a low interrater reliability is that it is an indication of floor effects. Floor (and/or ceiling) effects occur when all participants select the same answer/rating. Floor and ceiling effects reduce the variation in the scale and, as a result, can negatively affect its reliability.

The expertise of the participants and format of the survey items is another point of consideration when discussing the low agreement among them. Oftentimes, participants with dissimilar academic training and professional experiences that are too dissimilar can lead to significant differences in how items are interpreted, and thus rated. In this study, academic training in an engineering discipline linked the experts, and oftentimes NSF POs usually have at least one advanced degree, but the specific disciplines in which they earned their degrees varies greatly. Moreover, all of them shared professional experience as an engineering education researcher and NSF program officer at some point in their career paths. However, a myriad of other experiences make up the professional background of the participants in this study. Moreover, it is possible that the topical statements about impact were too decontextualized, and as a result, also led to dissimilar interpretations, and by extension, low consistency among them.

Although some of the results of the survey data were puzzling, the results did provide insights on how to better construct the survey in the future to get more meaningful data, and has also led to more ideas about new approaches to collect data on the impact of projects. (Some of these ideas are discussed in the section of this chapter labeled, *Study Limitation and Future Research Directions.*) For example, instead of revising the survey and redistributing it more program officers, a follow-up study that involves an exploratory approach to understanding various stakeholders' perspectives on what impact means in the context of undergraduate engineering education may be more appropriate. Stakeholders might include not only PIs and POs, but also personnel at agencies like the National Academy of Engineering (NAE), Office of Science and Technology Policy (OSTP), and the Accrediting Board for Engineering and Technology (ABET). Methods that involve interviews and/or focus groups would allow participants to provide open-ended responses, ask clarifying questions, and elaboration on their initial thoughts. Appendix E includes an example of an interview protocol that could be used in a follow-up study.

This section of the Discussion chapter concludes with some general reflections on impact. Writing reflective notes is an important part of the qualitative research process (Miles & Huberman, 1994), and two recurring ideas in the memos written while conducting this study relate to personal reflections about what impact means, and what the findings of this work means for PIs, POs, and others interested in the impact of federal investments in STEM education R&D. The last section of this chapter present implications, but a proposed definition of impact is presented here.

Currently, there is no consensus in the literature on what impact means (Brewer, 2011). As a result of completing this study and reflecting on the meaning of the findings, the following definition of impact is proposed:

> Impact is a time-sensitive interpretation of the extent to which change has occurred in (and/or beyond) the context in which the change originated.

The rationale for saying that impact is "time sensitive" for two reasons: 1) the interpretation that is articulated is significantly influenced by when the change is observed (Bornmann, 2013; Martin, 2007); and 2) the length of time it takes to achieve some extent of change largely influences how it is interpreted as impact (e.g., large changes that happen in short time may be perceived as more "impactful" than small changes over the same amount of time.) Descriptions of impact tend to include explicit and implicit links between a set of activities, changes resulting from these activities, and an interpretation of what these outcomes mean. Thus, impact is being defined as an "interpretation because it is a way someone (with a particular perspective) makes sense of something that has happened (or for some, what will happen), and is framed in light of abstract constructs of interest in both the immediate context and in broader contexts. Examples of such constructs that are relevant to this study include: teaching and learning, STEM education communities; institutional and national priorities related to STEM education. Defining impact as interpretation may help provide an explanation for why impact looks so different in different contexts, and why a set of activities can be perceived as having both positive and negative impact—depending on who's reporting it. While the focus of this study is research impact, this definition of impact can be extended to other contexts as well, and continue the construction of a meeting place on the proverbial "impact terrain" (Brewer, 2011).

Implications for Policy and Practice

The findings of this study have implications for policy and practice. From both perspectives, university stakeholders and funding agencies alike will benefit from an improved understanding of the impact of federal investments in STEM education research since both parties are interested in positive transformations and outcomes in undergraduate engineering education. Researchers on NSF-funded projects may use the research findings to more effectively communicate the impact of their work in the document that is viewed by program officers, the PI community, congress, the media, and the public at large. This is an important skill for researchers to practice because the stiff economic climate necessitates that research that demonstrates impact using concrete evidence is the research that will continue to be publicly supported. Additionally, PIs' ability to provide definitive insights on the impact of their projects have the potential to aid STEM education colleagues in better selecting and utilizing existing resources, and provide a stronger basis upon which future studies can build.

Provided are guidelines for PIs to consider when writing about the impact of their NSF-funded projects:

1. Discuss the scientific impact of the study by highlighting advances in knowledge, or ways in which the current study clarifies existing ways of thinking about a topic. Limit generic references to mundane steps in the research process and lists of publications resulting from the study.
2. Discuss the societal impact of the study by making connections between the outcomes of the current study and national priorities and/or salient discipline-

specific issues (e.g., increasing the quantity of engineering graduates and improving the quality of undergraduate engineering education).

3. Apart from mentioning the outcomes of the study, discuss the domain-specific impact of the study by mentioning the unique ways people (e.g., learners, instructors, administrators, networks of researchers, parents, industry partners), priorities (e.g., effective teaching, meaningful learning) and processes (e.g., in classrooms, departments, institutions) are affected by the outcomes of the study.
4. Make concrete statements about the impact of your work and briefly mention evidence that supports the claims.
5. Make clear distinctions between realized and anticipated impacts.

From a policy perspective, this study presents a description of what PIs are saying about the impact of their work. There are many instances when the impact narratives are very short (i.e., less than 25 words), vague, and sometimes cursory, at best. In light of this presentation on what PIs are saying, these results speak to the need for NSF to seek answers to questions about whether this is the type of information they wants or not. If it is not, there may be a need to provide PIs with guidance on how to construct impact narratives in order to get the desired information. Such guidance might be communicated via what is presented to PIs when they are submitting electronic progress reports, during training workshops, or some other venue. On the other hand, the preliminary findings of the survey study might indicate the need for more conversations among people making funding decisions about what impact means in this context since a better understanding of impact should promote better alignment between projects, program outcomes and national priorities. Two examples of venues in which such a conversation might take place are: during the training associated with hiring a NSF Program Officer or periodically during meetings among program officers. The findings on the types of claims that are made and how they are supported also speak to the need for a better process for vetting the claims PIs make in research reports. Improvements in existing processes in this regard will not only lead to more definitive insights on the impact of project throughout the grant cycle, it will also reduce the likelihood of PIs self-plagiarizing the contents of their impact narratives by reporting the exact same impacts across years of reporting. Mechanisms that enable this should be added to the suite of the current accountability mechanisms (e.g., annual reports, final reports). Lastly, this study begs the question: to what extent can policymakers be asked to make causal claims about the impact of federal investments in research if the current documenting procedures does not permit it? Thus, one of the notable implications of this work is it provides a starting point for NSF to develop a reporting structure that would allow program officers and the wider STEM education community to get better data –and ultimately, a better understanding—of what it means for a research project to have impact.

Study Limitations and Future Research Directions

Although the goal is to propose a research methodology that is sound, rigorous and valid, there are still limitations to this study. Four of them are as follows. One, the data used in this study originally comes from a context where PIs were reporting project outcomes to the agency funding their research. It is likely that this context has implications on how they report the impact of their work; in survey research, this phenomenon is referred to as response bias (i.e. when participant's responses are overly negative or positive) (John W.

Creswell, 2008). Another limitation of this study is associated with the primary data source in this study: project abstracts. By nature, abstracts are intended to be a brief overview of the main parts of a study, and by extension, are not as comprehensive as conference posters, project reports, etc. Because the data, by nature, is brief, this has the potential to limit the extent to which project outcomes are described. Third, the data included in this study is only from one NSF program, and the goals of this particular program could have implications on the research impact that is observed. Thus, the comprehensiveness and generalizability of the resulting description of impact will need to be explored beyond this program upon completion of this study. Another NSF program to consider for an exploration still focused on engineering education might be the Engineering Education Centers program, which is funded by the Engineering Directorate within NSF. Lastly, the resulting description of research impact may or may not capture the research impact of a project that occurs beyond the life of the grant. (The *evaluation timescale* and *temporality problem* are among challenges commonly cited among scholars interested in studying research impact (Bornmann, 2012, 2013; Martin, 2007; Rymer, 2011; J. E. Scott et al., 2011; Spaapen & van Drooge, 2011).)

Additional studies that involve different forms of data, variations in frameworks, and different methodologies are critical to advancing the scholarship of impact. For example, future studies using a theoretical lens focused on organizational change (Grusky & Miller, 1970; Katz & Kahn, 1978) or change and transformation within the context of undergraduate STEM education (e.g., Henderson, Beach, & Finkelstein, 2011) will lead to different insights. Apart from using project abstracts as data for analyzing research impact, other documents that might be useful include: program objectives mentioned in grant solicitations; grant proposals, annual reports, and final reports of expired research projects voluntarily shared by PIs; research briefs, executive summaries, and other short synopsis of research written for lay audiences; and media reports that highlight research projects. Other methods to consider include other qualitative approaches (e.g., grounded theory), quantitative methods (e.g., correlational studies), other mixed methods research designs, and computational approaches (e.g., data mining) (J. Creswell & Plano Clark, 2011; J. W. Creswell, 2008; Hey, Tansley, & Tolle, 2009; Johnson & Christensen, 2012). Another way to gauge how impact is defined may be to ask the STEM education research community to identify what they think has been the most impactful projects and why. Lastly, future studies that trace the impact of research beyond the grant cycle are vital to understand the immediate and long-term impact of a project.

CHAPTER 10: CONCLUSION

Each year, NSF invests millions of dollars in undergraduate engineering education research and development projects. Given the importance of the role of engineers in society and the size of these investments, an understanding of the extent to which these investments are leading to desired outcomes is necessary for better programming and more informed decision making among practitioners and policymakers. A scholarly understanding of what impact means, in general, and a comprehensive understanding of what it looks in this context would facilitate such insights. This study is the first systematic attempt to explore what it means for a federally funded project to have impact, adds clarity to the topic impact, and provides recommendations for where to go next.

This study provides a synthesis of the fragmented scholarship of impact to form a unified starting point for the conversation on this topic, and builds on this body of knowledge. The findings from this study provide a strong foundation for future studies on the research impact of studies focused on other STEM disciplines and education contexts as well. As a result of mixing qualitative and quantitative methods in this study, the findings reveal how researchers on NSF-funded projects talk about the impact of their work, and how often PIs discuss various types of claims in light of the amount of funding awards, the project focus, and STEM discipline. The findings reveal that researchers talk about a vast number of topics in this section of project abstracts. While some of the topics discussed align with current scholarly perspectives on research impact, much of them do not. Thus, this study reveals many insight on what research impact *is not* from the perspective of what PIs write in abstracts as it provides on what it is. The qualitative finding speaks to the need to continue to pursue the goal of figuring out what the multifaceted, complex facets of impact *are,* and to develop a valid conceptual framework that will facilitate a shared understanding of what impact means in this context. In the short term, it might be useful for funding agencies to provide PIs with more guidance on the types of information to include in an impact narrative. Improving the quality of information PIs include in projects abstracts is critical to advancing transparency about NSF's investments in undergraduate STEM education and possibly reducing criticisms of publicly-support R&D.

The descriptive statistics on the type of claims PIs indicate that the project focus and discipline have more to do with the type of impact that are realized than the amount of funding allocated to a project—at least when exploring the impact of projects within the life of the grant cycle. This finding might inform how funding agencies develop programs, design logic models, and ultimately write RFPs. More specifically, since RFP describe the project foci to which PIs respond with project proposals, federal funding agencies like NSF can be more strategic in using grant solicitations as levers for fostering impact.

Finally, the results of the survey study allude to a clash of perspectives on what impact means among stakeholders making funding decisions and researchers conducting the studies. The preliminary description of what impact looks like in this context may help alleviate some of this miscommunication, but this is an area of study that deserves additional exploration and insight. To the extent that this is true that PIs and POs do not know what impact means or now to communicate it, the current research findings might uncover a societal issue. Basic disconnects regarding what impact means to people in positions making decisions about what research to fund and those receiving the funds to conduct research will not be improved by simply modifying the format of the survey used to compare the two perspectives. To the extent that this is true, this could be an indication of the huge wall between researchers and stakeholders of publicly funded research that needs to be scaled or broken down as we attempt to be more efficient in the use of taxpayers' funds allocated to engineering education research. Such a pursuit is necessary as we seek to support studies that lead to notable impacts on the number, quality, and diversity of engineers equipped to address society's more pressing social and technological needs.

REFERENCES

REFERENCES

Allen, Sue, Campbell, Patricia B., Dierking, Lynn D., Flagg, Barbara N., Friedman, Alan J., Garibay, Cecilia, . . . Ucko, David A. (Eds.). (2008). *Framework for Evaluating Impacts of Informal Science Education Projects*. (Available at: http://informalscience.org/evaluations/eval_framework.pdf).

Bailey, R W. (1996). *Human Performance Engineering: Using Human Factors/Ergonomics to Achieve Computer System Usability* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Beach, Andrea L., Henderson, Charles, & Finkelstein, Noah. (2012). Facilitating Change in Undergraduate STEM Education. *Change,* 52-59.

Besterfield-Sacre, Mary, Cox, Monica F., Borrego, Maura, Beddoes, Kacey, & Zhu, Jiabin. (2014). Changing Engineering Education: Views of U.S. Faculty, Chairs, and Deans. *Journal of Engineering Education, 103*(2), 193-219.

Blau, John. (2011). Germany Faces a Shortage of Engineers. *IEEE Spectrum.* http://spectrum.ieee.org/at-work/tech-careers/germany-faces-a-shortage-of-engineers

Booth, Wayne C., Colomb, Gregory G., & Williams, Joseph M. (2008). *The Craft of Research* (3rd ed.). Chicago: University of Chicago.

Bornmann, Lutz. (2012). Measuring the Societal Impact of Research. *EMBO, 13*(8), 673-676.

Bornmann, Lutz. (2013). What is Societal Impact of Research and How Can it Be Assessed? A Literature Survey. *Journal of the American Society for Information Science and Technology, 64*(2), 217-233.

Bornmann, Lutz, & Marx, Werner. (2014). How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons. *Scientometrics, 98*(1), 211-219.

Borrego, Maura, Douglas, Elliot P., & Amelink, Catherine T. (2009). Quantitative, Qualitative, and Mixed Research Methods in Engineering Education. *Journal of Engineering Education, 98*(1), 53-66.

Bowen, Glenn A. (2006). Grounded theory and sensitizing concepts. *International Journal of Qualitative Methods, 5*(3).

Boyer Commission on Educating Undergraduates in the Research University. (1998). Reinventing Undergraduate Education: A Blueprint for America's Research Universities. http://naples.cc/sunysb/edu/pres/boyer.nsf/

Boyer, E. L. (1990). *Scholarship Reconsidered: Priorities of the Professorate*. San Francisco: Jossey-Bass.

Bozeman, B, & Sarewitz, D. (2011). Public value mapping and science policy evaluation. *Minerva, 49*(1), 1-23.

Brewer, John D. (2011). The impact of impact. *Research Evaluation, 20*(3), 255-256. doi: 10.3152/095820211x12941371876869

Bryman, A. (2006). *Qualitative Research, 6*(1), 105-107.

Bryman, Alan. (2006). Integrating Quantitative and Qualitative Research: How Is It Done? *Qualitative Research, 6*(1), 97-113.

Burkhardt, Hugh, & Schoenfeld, Alan H. (2003). Improving Educational Research: Toward a More Useful, More Influential, and Better-Funded Enterprise. *Educational Researcher, 32*(9), 3-14.

Bush, Vannevar. (1945). Science, the Endless Frontier: A Report to the President on a Program for Postwar Scientific Research. Washington, D.C.: United States Government Printing Office.

Butz, William P., Bloom, Gabrielle A., Gross, Mihal E., Kelly, Terrence K., Kofner, Aaron, & Rippen, Helga E. (2003). Is There a Shortage of Scientists and Engineers? How Would We Know? In M. A. Fox (Ed.), *Pan-Organizational Summit on the U.S. Science and Engineering Workforce: Meeting Summary* (pp. 167-179). Washington, D.C.: National Academies Press.

Charette, Robert N. (2013). The STEM Crisis is a Myth. *IEEE Spectrum.* http://m.spectrum.ieee.org/at-work/education/the-stem-crisis-is-a-myth

Cohen, J. A. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(37-46).

Cohen, J.A. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement of partial credit. *Psychological Bulletin, 70*, 213-220.

Conole, Grainne. (2013). Affordances. In J. M. Spector & S. LaJoie (Eds.), *Designing for learning in an open world.* New York, NY: Springer.

Creswell, J, & Plano Clark, V L. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J W. (2007). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches* (2nd ed.). Thousand Oaks: Sage Publications.

Creswell, J W. (2008). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (3rd ed.). Upper Saddler River, NJ: Pearson Education, Inc.

Creswell, J W, & Plano Clark, V L. (2007). *Designing and Conducting Mixed Methods Research.* Thousand Oaks, CA: Sage Publications.

Dall'Alba, Gloria. (2009). Learning Professional Ways of Being: Ambiguities of Becoming. *Educational Philosophy and Theory, 41*(1), 34-45.

Dancy, Melissa H., & Henderson, Charles. (2008). Barriers and Promises in STEM Reform.

Devore, J.L. (2012). *Probability and Statistics for Engineering and the Sciences*: Duxbury Press.

Donovan, Claire. (2011). State of the art in assessing research impact: Introduction to a special issue. *Research Evaluation, 20*(3), 175-179.

Donovan, Claire, & Hanney, Stephen. (2011). The 'Payback Framework' Explained. *Research Evaluation, 20*(3), 181-183.

Doz, Yves, Santos, Jose, & Williamson, Peter. (2001). *From Global to Metanational: How Companies Win in the Knowledge Economy*. Boston: Harvard Business School Press.

Earle, Janice, Maynard, Rebecca, Neild, Ruth Curran, Easton, John Q., Ferrini-Mundy, Joan, Albro, Elizabeth, . . . Winter, Susan. (2013). Common Guidelines for Education Research and Development. Washington, DC: IES, DOE, and NSF.

Education and Human Resources Directorate. (2013). FY 2014 NSF Budget Request to Congress. Arlington, VA: National Science Foundation.

Eiseman, Jeffrey W., Fairweather, James S., Rosenblum, Sheila, & Britton, Edward. (1996). Evaluation of the National Science Foundation Undergraduate Course and Curriculum Development Program: Final Report. In R. f. t. N. S. Foundation (Ed.). Washington, D.C.: SRI International.

Elo, Satu, & Kyngas, Helvi. (2007). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115.

Ericsson, K Anders, & Smith, J (Eds.). (1991). *Toward a General Theory of Expertise*. New York: Cambridge University Press.

Fackler, Martin. (2008). High-Tech Japan Running Out of Engineers, *The New York Times*. Retrieved from http://www.nytimes.com/2008/05/17/business/worldbusiness/17engineers.html?pagewanted=all&_r=1&

Fairweather, James. (2010). Linking Evidence and Promising Practices in Science, Technology, Engineeering and Mathematics (STEM) Undergraduate Education: The National Academies National Research Council Board of Science Education.

Fealing, Kaye Husbands, Lane, Julia I, Marbuger III, John H, & Shipp, Stephanie S. (2011). Editor's Introduction. In K. H. Fealing, J. I. Lane, J. H. M. III & S. S. Shipp (Eds.), *The Science of Science Policy: A Handbook*. Stanford, CA: Stanford Business Books.

Fealing, Kaye Husbands, Lane, Julia I, Marburger III, John H, & Shipp, Stephanie S (Eds.). (2011). *The Science of Science Policy: A Handbook*. Redwood City, CA: Stanford University Press.

Finelli, Cynthia, & Borrego, Maura J. (2014). *An inclusive process for developing a taxonomy of keywords for engineering education research*. Paper presented at the American Society for Engineering Education, Indianapolis, IN.

Finelli, Cynthia, Daly, Shanna R, & Richardson, Kenyon M. (2014). Bridging the Research-to-Practice Gap: Designing an Institutional Change Plan Using Local Evidence. *Journal of Engineering Education, 103*(2), 331-361.

. *FIRST Act of 2014* (2014). Retrieved from http://science.house.gov/sites/republicans.science.house.gov/files/documents/HR4186.pdf.

Fleiss, Joseph L. (1971). Measuring Nominal Scale Agreement Among Many Raters. *Psychological Bulletin, 76*(5), 378-382.

Froyd, J. E., Wankat, P. C., & Smith, K. A. (2012). Five Major Shifts in 100 Years of Engineering Education. *Proceedings of the IEEE, 100*(Special Centennial Issue), 1344-1360. doi: 10.1109/JPROC.2012.2190167

Garrett, S. K., Caldwell, B. S., Harris, E. C. , & Gonzalez, M. C. (2009). 'Six dimensions of expertise: a more comprehensive definition of cognitive expertise for team coordination',. *Theoretical Issues in Ergonomics Science, 10*(2), 93-105. doi: 10.1080/14639220802059190

Gibson, J.J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Percieving, acting, and knowing*. Hillsdale: Lawrence Erlbaum.

Gillespie, Stephanie M, McKenna, Ann F., & Pimmel, Russell. (2011). *Analyzing the Transformative Nature of Engineering Education Proposals*. Paper presented at the American Society for Engineering Education, Vancouver, B.C., Canada.

Godin, Benoît, & Doré, Christian. (2005). Measuring the Impacts of Science: Beyond the Economic Dimension. http://www.csiic.ca//pdf/godin_dore_impacts.pdf

Government Accountability Office. (2010). *Science, Technology, Engineering, and Mathematics Education: Strategic Planning Needed to Better Manage Overlapping Programs Across Multiple Agencies, GOA-12-108*. Washington, D.C.: Retrieved from http://www.gao.gov/assets/590/587839.pdf.

Government Accountability Office. (2012). Science, Technology, Engineering, and Mathematics: Strategic Planning Needed to Better Manage Overlapping Programs across Multiple Agencies (pp. 65). Washinton, D.C.: United States Government Accountability Office.

Grant, J., & Kinney, M. . (1992). Using the Delphi technique to examine the content validity of nursing diagnoses. *Nursing Diagnosis, 3*, 12-22.

Grant, Joan S, & Davis, Linda L. (1997). Selection and Use of Content Experts in Instrument Development. *Research in Nursing & Health, 20*, 269-274.

Grant, Jonathan, Brutscher, Philipp-Bastian, Kirk, Susan Ella, Butler, Linda, & Wooding, Steven. (2010). Capturing Research Impacts: A Review of International Practice: RAND.

Greene, Jennifer C., Caracelli, Valerie J., & Graham, Wendy F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis, 11*, 255-274.

Grusky, Oscar, & Miller, George A (Eds.). (1970). *The Sociology of Organizations: Basic Studies*. New York: The Free Press.

Guess, Andy. (2008). Challenging Conventional Wisdom on STEM Supply. *Inside Higher Ed.* http://www.insidehighered.com/news/2008/09/17/pcast - ixzz2UtYiQMFC

Haghighi, Kamyar. (2005). Quiet No Longer: Birth of a New Discipline. *Journal of Engineering Education, 94*(4), 351-353.

Halse, Christine, & Mowbray, Susan. (2011). The Impact of the Doctorate. *Studies in Higher Education, 36*(5), 513-525.

Hayes, Andrew F, & Krippendorff, Klaus. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures, 1*(1), 77-89.

Henderson, Charles, Beach, Andrea, & Finkelstein, Noah. (2011). Facilitating Change in Undergraduate STEM Instructional Practices: An Analytic Review of the Literature. *Journal of Research in Science Teaching, 48*(8), 952-984.

Henderson, Charles, & Dancy, Melissa H. (2010). Increasing the Impact and Diffusion of STEM Education Innovations. http://www.nae.edu/File.aspx?id=36304

Hey, T, Tansley, S, & Tolle, K. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Redmond, Washington: Microsoft Research.

Hitchcock, David. (2006). Good Reasoning on the Toulmin Model. In D. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation* (Vol. 10, pp. 203-218). Dordrecht, The Netherlands: Springer.

Hitchcock, David, & Verheij, Bart. (2006). Introduction. In D. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation* (Vol. 10, pp. 1-23). Dordrecht, The Netherlands: Springer.

Holbrook, J Britt. (2005). Assessing the science–society relation: The case of the US National Science Foundation's second merit review criterion. *Technology in Science, 27*, 437-451.

Holbrook, J Britt. (2012). Re-assessing the science – society relation: The case of the US National Science Foundation's broader impacts merit review criterion (1997 – 2011). http://www.peerev.org/data/7cce53cf90577442771720a370c3c723/PE_doc_28403.pdf

Holbrook, J Britt, & Frodeman, Robert. (2011). Peer review and the ex ante assessment of societal impacts. *Research Evaluation, 20*(3), 239-246. doi: 10.3152/095820211x12941371876788

Holsti, O R. (1969). *Content Analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

IBM. (2003). Gray Matter Matters: Preserving Critical Knowledge in the 21st Century. http://www-935.ibm.com/services/us/gbs/bus/pdf/g510-3314-gray-matter-matters.pdf

Jackson, Shirley Ann. (2002). The Quiet Crisis: Falling Short in Producing American Scientific and Technical Talent. San Diego, CA: Building Engineering and Science Talent (BEST).

Johnson, Burke, & Christensen, Larry. (2012). *Educational Research* (4th ed.). Los Angelos: Sage Publications.

Kamenetzky, Julia R. (2013). Opportunities for impact: Statistical analysis of the National Science Foundation's broader impacts criterion. *Science and Public Policy, 40*(1), 72-84. doi: 10.1093/scipol/scs059

Katz, Daniel, & Kahn, Robert L. (1978). *The Sociology of Organizations* (2nd ed.). New York: John Wiley & Sons.

Kennedy, Donald, Taylor, Crispin, Urquhart, Kirstie, & Austin, Jim. (2004). Supply Without Demand. *Science Magazine*.

Klumpp, James F. (2006). Warranting Arguments, the Virtue of Verb. In D. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation* (Vol. 10, pp. 103-114). Dordrecht, The Netherlands: Springer.

Krippendorff, Klaus. (1970). Estimating the reliability, systematic error and random error of interval data. *Educational and Psychological Measurement, 30*, 61-70.

Krippendorff, Klaus. (2004a). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications.

Krippendorff, Klaus. (2004b). *Content Analysis: An Introduction to its Methodology* (2 ed.). Thousand Oaks, CA: Sage.

Krippendorff, Klaus. (2011). Agreement and Information in the Reliability of Coding. *Communication Methods and Measures, 5*(2), 93-112.

Kuruvilla, Shyama, Mays, Nicholas, Pleasant, Andrew, & Walt, Gill. (2006). Describing the Impact of Health Research: a Research Impact Framework. *BMC Health Services Research, 6*(134), 1-17.

Lande, Micah, Adams, Robin, Chen, Helen, Currano, Becky, & Leifer, Larry. (2007). *"Scholarship of Impact" Framework in Engineering Education Research: Learnings from the Institute of Scholarhip on Engineering Education*. Paper presented at the ASEE/IEEE Frontiers in Education Conference, Milwaukee, WI.

Lane, Julia. (2009). Assessing the Impact of Science Funding. *Science, 324,* 1273-1275.

Lane, Julia. (2010). Let's Make Science Metrics More Scientific. *Nature, 464*(25), 488-489.

Lane, Julia, & Bertuzzi, Stefano. (2011). Measuring the Results of Science Investments. *Science Magazine, 331,* 678-680.

Levitt, Ruth, Celia, Claire, Diepeveen, Stephanie, Chonaill, Siobhán Ní, Rabinovich, Lila, & Tiessen, Jan. (2010). Assessing the impact of arts and humanities research at the University of Cambridge: Rand Corporation.

Lombard, Matthew, Snyder-Duch, Jennifer, & Bracken, Cheryl Campanella. (2002). Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. *Human Communication Research, 28*(4), 587-604.

London, Jeremi. (2012). *Exploring Cyberlearning through a NSF Lens.* Paper presented at the American Society for Engineering Education, San Antonio, TX.

London, Jeremi. (2013). *Key Aspects of Cyberlearning Resources with Compelling Results.* Paper presented at the American Society for Engineering Education, Atlanta, GA.

London, Jeremi, & Young, Cynthia. (2014). *Toward a Research Agenda Around MOOCs in Engineering Education: Findings from Two NSF-funded Workshops (Work in Progress Paper).* Paper presented at the MOOCs in STEM: Exploring New Educational Techologies Conference, San Jose, CA.

Lowell, Lindsay, & Salzman, Hal. (2008). Into the Eye of the Storm: Assessing the Evidence on Science and Engineering Education, Quality, and Workforce Demands. Washington, D.C.: Urban Institute.

Lynn, Leonard, & Salzman, Hal. (2006). Collaborative Advantage. *Issues in Science and Technology*(Winter), 74-82.

Lynn, Leonard, & Salzman, Hal. (2009). The 'New' Globalization of Engineering: How the Offshoring of Advanced Engineering Affects Competitiveness and Development. *Economics, Management, and Financial Markets, 4*(1), 11-46.

Madhavan, Krishna, Vorvoreanu, Mihaela, Elmqvist, Niklas, Jorhri, Aditya, Ramakrishan, Naren, Wang, G Alan, & McKenna, Ann. (2012). Portfolio Mining. *Computer, 45*(10), 95-99.

Martin, Ben R. (2007). *Assessing the impact of basic research on society and the economy.* Paper presented at the Rethinking the impact of basic research on society and the economy (WF-EST International Conference), Vienna, Austria. http://www.science-impact.ac.at/documentation/pdf/Session_C_Martin.pdf

Martin, Ben R. (2011). The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? *Research Evaluation, 20*(3), 247-254. doi: 10.3152/095820211x13118583635693

McKenna, Ann F, Froyd, Jeffrey, King, Judson, Litzinger, Thomas, & Seymour, Elaine. (2011). The Complexities of Transforming Engineering Higher Education: Report on Forum on Characterizing the Impact and Diffusion of Transformative Engineering Education Innovation.

McKenna, Ann F, Froyd, Jeffrey, & Litzinger, Thomas. (2014). The Complexities of Transforming Engineering Higher Education: Preparing for Next Steps. *Journal of Engineering Education, 103*(2), 188-192.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis.* Newbury Park: Sage.

Miller, Barbara, & Pasley, Joan. (2012). What do we know and how well do we know it? Identifying practice-based insights in education. *Evidence & Policy, 8*(2), 193-212.

Molas-Gallart, Jodi, Salter, Ammon, Patel, Pari, Scott, Alister, & Duran, Xavier. (2002). Measuring Third Stream Activities *Final Report to the Russell Group of univerisities.* Bighton, United Kingdom: Science and Technology Policy Research Unit: University of Sussex.

National Academies Press. (2011). Measuring the Impacts of Federal Investments in Research: A Workshop Summary. Washington, D.C.: National Academies Press.

National Academy of Engineering. (2002). *Raising Public Awareness of Engineering*. Washington, D.C.: The National Academies Press.

National Academy of Engineering. (2004). *The Engineer of 2020: Visions of Engineering in the New Century*. Washington, D.C.: National Academies Press.

National Academy of Engineering. (2005a). *Educating the Engineer of 2020: Adapting Engineering Education to the New Century*. Washington, D.C.: National Academies Press.

National Academy of Engineering. (2005b). Engineering Research and America's Future: Meeting the Challenges of a Global Economy. Washington, D.C.: The National Academies Press.

National Academy of Engineering. (2008a). *Changing the Conversation: messages for Improving Public Understanding of Engineering*. Washington, D.C.: National Academies Press.

National Academy of Engineering. (2008b). Grand Challenges for Engineering. Washington D.C.: The National Academies Press.

National Academy of Engineering. (2010). *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5*: The National Academies Press.

National Academy of Sciences. (2014). STAR METRICS Working Group. *Federal Demonstration Partnership: Redefining the Government & University Research Partnership.* Retrieved September 18, 2014, from http://sites.nationalacademies.org/PGA/fdp/PGA_057189

National Research Council. (1999a). Transcforming Undergraduate Education in Science, Mathematics, Education, and Technology. Washington, D.C.: National Academies Press.

National Research Council. (1999b). *Transforming Undergraduate Education in Science, Mathematics, Engineering, and Technology*. Washington, D.C.: The National Academies Press.

National Research Council. (2002). Scientific Research in Education. Washington, D.C.: National Academies Press.

National Research Council. (2003). *Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics*: The National Academies Press.

National Research Council. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington, DC: The National Academies Press.

National Research Council. (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. Washington, DC: The National Academies Press.

National Research Council. (2014a). Science of Science and Innovation Policy: Principal Investigators' Conference Summary. Washington, D.C.: National Academies Press.

National Research Council. (2014b). Science of Science and Innovation Policy: Principal Investigators' Conference Summary. In K. H. F. teering Committee on the Science of Science and Innovation Policy Principal Investigators' Conference, A.S. Beatty, and C.F. Citro, Rapporteurs. (Ed.), *Committee on National Statistics, Division of Behavioral and Social Sciences and Education*. Washington, D.C.: The National Academies Press.

National Science and Technology Council. (2008a). Finding out What Works: Agency Efforts to Strengthen the Evaluation of Federal Science, Technology, Engineering, and Mathematics (STEM) Education Programs. Washington, D.C.: Executive Office of the President of the United States.

National Science and Technology Council. (2008b). The Science of Science Policy: A Federal Research Roadmap. http://www.whitehouse.gov/files/documents/ostp/NSTC Reports/39924_PDF_Proof.pdf

National Science and Technology Council. (2013). Federal Science, Technology, Engineering and Mathematics (STEM) Education 5-Year Stategic Plan. Washington, D.C.: Executive Office of the President.

National Science Board. (2011). National Science Foundation's Merit Review Criteria: Review and Revisions. Washington, D.C.: National Science Board.

National Science Foundation/National Center for Science and Engineering Statistics. (2008). *Diversity in Science and Engineering Employment in Industry*. Retrieved from: http://www.nsf.gov/statistics/infbrief/nsf12311/ - tab1

NSF.). About DUE. Retrieved Sept 8, 2014, from http://www.nsf.gov/ehr/due/about.jsp

NSF.). About the National Science Foundation. Retrieved Sept. 8, 2014, from http://www.nsf.gov/about/

NSF.). NSF Division of Undergradaute Education: About Undergraduate Education (DUE). Retrieved May 8, 2014, from http://www.nsf.gov/ehr/due/about.jsp

NSF. (2008). *Invention and Impact 2: Building Excellence in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education.* Paper presented at the A Conference of Course, Curriculum, and Laboratory Improvement (CCLI) Program, National Science Foundation, Division of Undergraduate Education, Washington, D.C.

NSF. (2011). *Transforming Undergraduate Educatio in STEM: Making and Measuring Impacts.* Paper presented at the 2011 CCLI/TUES Principal Investigators (PIs) Conference, Washington, D.C.

NSF. (2013a). 2013 TUES Principal Investigators (PIs) Conference. http://ccliconference.org/files/2013/01/2013-Tues-Conference-Program.pdf

NSF. (2013b). NSF FY13 Budget Request to Congress. Retrieved June 30, 2014, from http://www.nsf.gov/about/budget/fy2013/pdf/16-EHR_fy2013.pdf

NSF. (2013c). Science and Engineering Degrees: 1966-2010. http://www.nsf.gov/statistics/nsf13327/

NSF. (2013d). *Transforming Undergraduate Education in STEM: Building a Community to Transform Undergraduate STEM Education.* Paper presented at the 2013 TUES Principal Investigators (PIs) Conference, Washington, D.C.

Office of Science and Technology Policy. (2010). Science of Science Policy: SciSIP Central. Retrieved January 5, 2014

Patton, Michael Quinn. (2002). *Qualitative Research and Evaluation Methods*. Thousand Oaks, CA: Sage Publications.

President's Council of Advisors on Science and Technology (Producer). (2010). Prepare and Inspire: K-12 Education in Science, Technology, Engineering and Math (STEM) for America's Future. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

President's Council of Advisors on Science and Technology (Producer). (2012). Engage to Excel: Producing One Million Additional College Graduates With Degrees in Science, Technology, Engineering, and Mathematics. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

Raghavan, Vijay, Xie, Ying, Johnsten, Tom, Benton, Ryan, Lemoine, Blake, & Difallah, Djellel. (2011). CISE and SBE AC Subcomittee on Discovery in a Research Portfolio: Tools for Structuring, Analyzing, Visualizing, and Interacting with Proposal and Award Portfolios. Arlington, VA: National Science Foundation.

The Research Agenda for the New Discipline of Engineering Education. (2006). *Journal of Engineering Education, 95*(4), 259-261.

Richardson, Hannah. (2012, September 30). Warning Over Shortage of Engineering Graduates, *BBC News*. Retrieved from http://www.bbc.co.uk/news/education-19760351

Roberts, Melanie R. (2009). Realizing societal benefit from academic research: Analysis of the National Science Foundation's broader impacts criterion. *Social Epistemology, 23*(3-4), 199-219.

Rossi, Peter H., Lipsey, Mark W., & Freeman, Howard E. (2004). *Evaluation: A Systematic Approach* (7th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Rymer, Les. (2011). Measuring the impact of research - the context for metric development. Turner, Australia: The Group of Eight.

Saldaña, J. (2009). *The Coding Manual for Qualitative Researchers*. Washington, DC: Sage Publications Ltd.

Salter, Ammon J, & Martin, Ben R. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy, 30*, 509-532.

Salzman, Hal, & Lowell, Lindsay. (2008). Making the Grade. *Nature, 453*(May), 28-30.

Salzman, Hal, & Lynn, Leonard. (2010). *Engineering and Engineering Skills: What's really needed for global competitiveness*. Paper presented at the Association for Public Policy Analysis and Management, Boston, MA.

Sargent Jr, John F. (2013). *Federal Research and Development Funding: FY2013*. Washington, D.C.: Retrieved from http://fas.org/sgp/crs/misc/R42410.pdf.

Schwarz, Andrew, Mehta, Manjari, Johnson, Norman, & Chin, Wynne W. (2007). Understanding Frameworks and Reviews: A Commentary to Assist us in Moving Our Field Forward by Analyzing Our Past. *Advances in Information Systems, 38*(3), 29-50.

Scott, Jack E, Blasinsky, Margaret, Dufour, Mary, Mandai, Rachel J, & Philogene, G Stephanae. (2011). An evaluation of the Mind-Body Interactions and Health Program: assessing the impact of an NIH program using the Payback Framework. *Research Evaluation, 20*(3), 185-192. doi: 10.3152/095820211x12941371876661

Scott, W. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly, 17*, 312-325.

Seymore, Elaine. (2011). *Determining Progress in Improving Undergraduate STEM Education: The Reformers' Tale*. Paper presented at the Characterizing the Impact and Diffusion of Engineering Education Innovations.

Seymour, Elaine. (2001). Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education, 86*(1), 79-105. doi: 10.1002/sce.1044

Seymour, Elaine, & DeWelde, Kris. Resistance in Sustaining Pedagogical Innovations: Lessons for Sociology from STEM Innovators.

Siddiqui, Junaid A., & Adams, Robin. (2013). *The Challenge of Change in Engineering Education: Is it the Diffusion of Innovations or Transformative Learning?* Paper presented at the American Society for Engineering Education, Atlanta.

Silverman, Hannah. (2010). Education to Tackle Shortage of 20,000 Engineers in Australia, *The Advertiser*. Retrieved from http://www.adelaidenow.com.au/news/south-australia/education-to-tackle-shortage-of-20000-engineers-in-australia/story-e6frea83-1225907964866

Singletary, M W. (1993). *Mass Communication Research: Contemporary Methods and Applications*. Boston: Addison-Wesley.

Spaapen, Jack, Dijstelbloem, Huub, & Wamelink, Frank. (2007). Evaluating research in context: A method for comprehensive assessment. The Hague, The Netherlands: Consultative Committee of Sector Councils for Research and Development.

Spaapen, Jack, & van Drooge, Leonie. (2011). Introducing 'productive interactions' in social impact assessment. *Research Evaluation, 20*(3), 211-218. doi: 10.3152/095820211x12941371876742

Srinivas, Rajulapudi. (2013). India Faces Shortage of Engineers in S&T: DRDO DG, *The Hindu*. Retrieved from http://www.thehindu.com/news/cities/Vijayawada/india-facing-shortage-of-engineers-in-st-drdo-dg/article4356570.ece

Stake, Robert E. (2010a). Interpretation: The Person as Instrument *Qualitative Research: Studying How Things Work* (pp. 36-55). New York: The Guilford Press.

Stake, Robert E. (2010b). *Qualitative Research: Studying How Things Work*. New York: The Guilford Press.

Sternberg, R J. (1997). Cognitive conceptions of expertise. In P. J. Feltovich, K. M. Ford & R. R. Hoffman (Eds.), *Expertise in Context: Human and Machine* (pp. 149-162). Menlo Park, CA: The MIT Press.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (2nd ed.). Thousand Oaks, CA: Sage.

Tans, Olaf. (2006). The Fluidity of Warrants: Using the Toulmin Model to Analyze Practical Discourse. In D. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation* (Vol. 10, pp. 219-230). Dordrecht, The Netherlands: Springer.

Tashakkori, Abbas, & Teddlie, Charles (Eds.). (2003). *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks: Sage.

Teitelbaum, Michael S. (2004). *Do We Need More Scientists?* Paper presented at the The U.S. Scientific and Technical Workforce: Improving Data for Decisionmaking, Arlington, VA.

Teitelbaum, Michael S. (2007). Testimony of Michael S. Teitelbaum before the Subcommittee on Technology and Innovation. Washington, D.C.: U.S. House of Representatives.

The National Commission on Mathematics and Science Teaching for the 21st Century. (2000). Before It's Too Late: A Report to the Nation from the National Commision on Mathematics and Science Teaching for the 21st Century. Washington, D.C.: Department of Education.

Tinsley, H E A, & Weiss, D J. (2000). Interrater reliability and agreement. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 95-124). San Diego, CA: Academic Press.

Toulmin, Stephen E. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

U.S. Department of Education. (2007). Report of the Academic Competitive Council. Washington, DC.

UK Parliament. (2012). UK Must Produce More High Calibre Science Graduates to Secure Economic Growth. Retrieved Sept 8, 2014, from http://www.parliament.uk/business/committees/committees-a-z/lords-select/science-and-technology-committee/news/stem-report-published/

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2012). UNESCO Engineering Initiative. Retrieved Oct. 20, 2013, from http://www.unesco.org/new/en/natural-sciences/science-technology/engineering/unesco-engineering-initiative/

Voss, James F. (2006). Toulmin's Model and the Solving of Ill-Structured Problems. In D. Hitchcock & B. Verheij (Eds.), *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation* (Vol. 10, pp. 303-312). Dordrecht, The Netherlands: Springer.

Wadhwa, Vivek. (2007, October 26, 2007). The Science Education Myth. *BusinessWeek*.

Wadhwa, Vivek. (2011, September 1, 2011). Mr. President, there is no engineer shortage, *The Washington Post*. Retrieved from http://www.washingtonpost.com/national/on-innovations/president-obama-there-is-no-engineer-shortage/2011/09/01/gIQADpmpuJ_story.html

Wadhwa, Vivek, Gereffi, Gary, Rissing, Ben, & Ong, Ryan. (2007). Where the engineers are. *Issues in Science and Technology*(Spring), 73-84.

Walter, Isabel, Nutley, Sandra, & Davies, Huw. (2003). Research Impact: A Cross Sector Review Literature Review. Research Unit for Research Utilisation: University of St. Andrews.

White, Marilyn Domas, & Marsh, Emily E. (2006). Content Analysis: A Flexible Methodology. *Library Trends, 55*(1), 22-45.

Wooding, Steve, Hanney, Steve, Buxton, Martin, & Grant, Jonathan. (2004). The Returns from Arthritis Research Volume 1: Approach, Analysis, and Recommendations. Cambridge: RAND Europe.

APPENDICES

Appendix A: Data Sample: Three Abstract from 2008, 2011, 2013 PI Conference Reports

Example Abstract 1 – Conference Year 2008 (NSF, 2008, p. A95)

Poster 149
PI: Monica Cox
Institution: Purdue University
Title: Development of a Pedagogically Focused Course for Engineering Graduate Teaching Assistants
Project #: 0632879
Type: Phase I – Exploratory
Target Discipline: Engineering
Focus: Implementing Educational Innovations

Goals: This project explores whether engineering graduate teaching assistants (GTAs) pedagogical perceptions changed after their enrollment in an engineering education course based on elements of the "How People Learn" framework. Course effectiveness will be examined via course materials, GTA interviews, and online undergraduate surveys.

Methods: The research group will conduct one semi-structured interview per GTA to identify GTAs' perceptions about their instruction. Interviews will occur face-to-face and will be audio-recorded. Researchers will distribute online surveys to undergraduates in GTAs' courses. Researchers will map responses to elements of the "How People Learn" framework.

Evaluation: After the interview questions have been asked, researchers will transcribe the interviews. A general sense of the results will be obtained by continuous reading and rereading of the data and an examination of the reflexive notes. The research group will note significant comments, organize the statements into segments, pool the segments together, and assign codes. Then researchers will test the codes and categorize the codes based on repeated patterns. Differences and similarities in patterns for control and treatment groups will be examined. Informal and formal observations of GTA laboratories and results from student surveys will be used to triangulate results obtained from GTA interviews.

Dissemination: A poster and paper describing the GTA course has been accepted for the 2007 American Society for Engineering Education conference. Results from the qualitative and quantitative studies will be disseminated in peer-reviewed engineering education journals.

Impact: Through the creation of a professional development course for GTAs that focuses on understanding the science and principles of learning and teaching, we have exposed future engineering faculty to the importance of creating learning environments that promote all students' higher-level learning and retention. Results from this study are being used in the redesign of a first-year course at Purdue University and will be used to understand the roles that engineering GTAs play in the development of undergraduate engineering students.

Challenges: One of the biggest challenges within this project was the recruiting engineering GTAs to participate in the effective teaching seminar. In the future, researchers anticipate combining course content with current GTA training content so that all students can be exposed to elements of the "How People Learn" framework and to innovative problems called Model-Eliciting-Activities.

Example Abstract 2 – Conference Year 2011 (NSF, 2011, p. A97)

Poster 156
PI: Sean Brophy
Institution: Purdue University
Project Title: Graphical Representations to Assess System Performance (GRASP) Assessments for Engineering Education
Project #: 0817486
Type: Phase 2/Type 2 -- Expansion
Focus: Assessing Student Achievement

Goals & Intended Outcomes: The aim of this project is to construct and research an automated dynamic formative assessment environment that develops and measures students' ability to represent, identify and explain solutions to complex problems. Specifically, the project focuses on representational tools used to facilitate the early stages of design including problem comprehension (e.g. system block diagrams, concept maps), evaluation of design alternatives (e.g. House of quality, morphological charts) and analysis of systems to inform design decisions (e.g. free body diagrams).

Methods & Strategies: The automated assessment system will contain a general set of graphical construction tools used to represent systems. The targeted tools are those associated with engineering activities of design and analysis of a system's performance. We are developing an epistemological framework to characterize problem types to inform the design of quality problem sets. These problem sets will lead to students' comprehension of the disciplinary knowledge and how to use the tool strategically. In addition this framework and research on students approach to the problem will inform the design of effective diagnosis and feedback for our system and instructors.

Evaluation Methods & Results: We are using qualitative methods to capture and describe students' approach to constructing a representation of complex systems (e.g. cruise control, flying craft) they can use to explain how the system works in multiple situations. The results from these experiments will inform the diagnostic and feedback mechanisms used to support students' learning. Eventually quantitative methods will be used as pre/post measures of students' learning gains related to solving problems in multiple situations. To date we have conducted interviews with first and second year engineering students as they construct diagrams of complex systems and use the diagrams to answer "what if" questions. The context for these challenges relate to either evaluating a design, troubleshooting a problem or analyzing a subcomponent or factors associate with the systems' performance.

Dissemination: The project is still in the development phase. The goal will be to make these assessment tools publicly available within digital libraries associated with engineering content. In addition, we are working talking with a publisher interested in using our program as complement to their engineering related text books. Further a dedicated web site will be constructed to describe the potential of formative assessment and feedback and provide access to a large assortment of problem sets associated with engineering activities from multiple disciplines.

Impact: The system currently has two representational formats under development, vector analysis, and node link graphs. The vector construction tool has been used with great success with engineering students in second year biomechanics and first year engineering students learning basic mechanics. Simple paper and pencil mockups are being used to simulate the graphical modeling of systems with highly inter-related components. We are investigating the conjecture that we need problem sets that co-develop tools knowledge and domain knowledge simultaneously. Further we anticipate learners will become more proficient at solving novel problem after given sufficient training with our tools.

Challenges: A challenge is identifying problem sets with sufficient complexity to engage learners in the content without overwhelming them. This was not an unexpected challenge, but identifying a larger range of problems requires more interactions with domain experts. This summer we will be working on outlining a larger set of problems based on what we have learned about the interaction of content and tools.

Example Abstract 3 – Conference Year 2013 (NSF, 2013d, pp. A120-A121)

Poster 225
PI: Ruth Streveler
Institution: Purdue University
Project Title: Collaborative Research: Expanding and Sustaining Research Capacity in Engineering and Technology Education: Building a Successful Program for Faculty and Graduate Students
Project #: 0817461
Type: Phase 3/Type 3 - Comprehensive
Focus: Developing Faculty Expertise

Goals and Intended Outcomes:
Three project goals:

1. Design and deliver a new generation of programs to educate engineering and engineering technology faculty and graduate students to conduct and use educational research which are effective, flexible, inclusive, and sustainable after funding ends.
2. Foster a virtual community of engineering and engineering technology education researchers through the use of Purdue HUBzero technology.

3. Evaluate the impact of these programs on individuals who participate and on the participants' students and institutions.

Methods & Strategies:
Goal 1: Creating face-to-face workshops and short courses that help engineering faculty learn about educational research methods. We take the approach that faculty already know about quality research in their technical area – we ask faculty to compare and contrast research in technical areas with research on teaching and learning.

Goal 2: Create online educational materials (mainly video of face-to-face workshops) that is useful for a remote audience.

Goal 3: Impact is evaluated through short surveys (for short courses) and through observation, focus groups and follow-up interviews for multiple-day workshops.

Evaluation Methods & Results: Evaluation is accomplished through the following methods.

1. Usage statistics – face-to-face attendance [for short courses and workshops] and Google Analytics for virtual community
2. Pre- and post- knowledge surveys for short courses
3. For multiple-day workshops – observation during the workshops, focus groups, interviews, analysis of products.

Dissemination: Dissemination through workshops at ASEE and FIE, online through CLEERhub. Publications and conference papers as continued and continuing dissemination. Continued dissemination through cyberinfrastructure (CLEERhub.org) which continues after funding for this project ends.

Impact: Usage Statistics are our most immediate level of impact. About 200 people have participated in face-to-face activities. CLEERhub.org has about 5000 visitors to date. Pre- and post- knowledge surveys do show increase in knowledge by participants. Analysis is ongoing for longer-term impacts.

Challenges: Creating a cyberinfrastructure (Collaboratory for Engineering Education Research or CLEERhub.org) was a new venue for us. Method for dealing with this was including collaborators who were experts in online delivery and keeping an open mind.

Appendix B: Descriptive Information about Abstracts Reviewed

This table provides the identifying characteristics of the 155 abstracts that were analyzed in this study.

Conference Year	Project Type
2008	Type I - TUES Type I
2011	Type II - TUES Type II
2013	Type III - TUES Type III
	CRP - Central Resource Project

Discipline
Biological Sciences (BS)
Chemistry (Ch)
Computer Science (CS)
Engineering (E)
Geological Sciences (GS)
Interdisciplinary (I)
Mathematics (M)
Physics/Astronomy (/A
Research/Assessment of Research (R/AR)
Social Science (SS)

Project Focus
Assessing Student Achievement (ASA)
Creating Learning Materials and Teaching Strategies (CLM)
Conducting Research on Undergraduate STEM Education (CR)
Developing Faculty Expertise (DFE)
Implementing Educational Innovations (IEI)

Analysis #	Year	Project Type	Discipline	Project Focus	Abstract Number
1	2011	CRP	E	CR	213
2	2011	CRP	E	DFE	243
3	2013	CRP	E	DFE	213
4	2013	CRP	I	IEI	294
5	2013	CRP	R/AR	IEI	385
6	2008	Type I	BS	ASA	32
7	2011	Type I	BS	ASA	36
8	2011	Type I	BS	CR	15
9	2011	Type I	BS	CR	43
10	2011	Type I	BS	DFE	10
11	2013	Type I	BS	DFE	39
12	2008	Type I	BS	IEI	17
13	2011	Type I	BS	IEI	38
14	2013	Type I	BS	IEI	28
15	2011	Type I	Ch	ASA	90
16	2011	Type I	Ch	CLM	53
17	2011	Type I	Ch	CR	57
18	2011	Type I	Ch	DFE	58
19	2011	Type I	Ch	DFE	61
20	2011	Type I	Ch	DFE	92
21	2013	Type I	Ch	DFE	67
22	2011	Type I	Ch	DFE	58
23	2008	Type I	Ch	IEI	105
24	2011	Type I	Ch	IEI	66
25	2008	Type I	CS	ASA	118

Analysis #	Year	Project Type	Discipline	Project Focus	Abstract Number
26	2008	Type I	CS	CLM	130
27	2011	Type I	CS	CLM	96
28	2011	Type I	CS	CLM	97
29	2013	Type I	CS	CLM	104
30	2011	Type I	CS	CR	115
31	2008	Type I	CS	IEI	132
32	2011	Type I	CS	IEI	104
33	2011	Type I	CS	IEI	121
34	2013	Type I	CS	IEI	100
35	2008	Type I	E	ASA	165
36	2013	Type I	E	ASA	184
37	2011	Type I	E	ASA	214
38	2013	Type I	E	CLM	134
39	2011	Type I	E	CLM	272
40	2013	Type I	E	CLM	170
41	2008	Type I	E	CR	156
42	2011	Type I	E	CR	244
43	2013	Type I	E	CR	212
44	2011	Type I	E	CR	158
45	2011	Type I	E	CR	187
46	2013	Type I	E	CR	229
47	2011	Type I	E	DFE	218
48	2013	Type I	E	DFE	163
49	2013	Type I	E	DFE	233
50	2011	Type I	E	IEI	149
51	2011	Type I	E	IEI	154
52	2008	Type I	GS	ASA	256
53	2008	Type I	GS	CLM	254
54	2011	Type I	GS	CLM	214
55	2011	Type I	GS	CLM	286
56	2013	Type I	GS	CLM	260
57	2008	Type I	GS	CR	249
58	2011	Type I	GS	CR	277
59	2013	Type I	GS	CR	252
60	2011	Type I	GS	DFE	280
61	2008	Type I	GS	DFE	251
62	2011	Type I	I	ASA	329
63	2013	Type I	I	CLM	279
64	2008	Type I	I	CR	257
65	2013	Type I	I	IEI	278
66	2013	Type I	M	CLM	345
67	2008	Type I	M	DFE	321
68	2008	Type I	M	IEI	331
69	2013	Type I	M	IEI	325
70	2013	Type I	P/A	ASA	376
71	2008	Type I	P/A	IEI	365
72	2013	Type I	P/A	IEI	359
73	2013	Type I	P/A	IEI	365
74	2013	Type I	R/AR	ASA	383
75	2013	Type I	R/AR	CR	386
76	2008	Type I	SS	CLM	378
77	2008	Type I	SS	DFE	230
78	2008	Type I	SS	IEI	386
79	2013	Type I	SS	IEI	392
80	2011	Type II	BS	CLM	5
81	2011	Type II	BS	CLM	11
82	2008	Type II	BS	DFE	37
83	2013	Type II	BS	DFE	44
84	2013	Type II	BS	DFE	41
85	2011	Type II	BS	DFE	18

Analysis #	Year	Project Type	Discipline	Project Focus	Abstract Number
86	2011	Type II	Ch	ASA	70
87	2011	Type II	Ch	CLM	59
88	2011	Type II	Ch	CR	86
89	2011	Type II	CS	CLM	107
90	2011	Type II	CS	CLM	139
91	2008	Type II	CS	IEI	107
92	2011	Type II	CS	IEI	136
93	2011	Type II	E	ASA	156
94	2011	Type II	E	ASA	194
95	2013	Type II	E	CLM	194
96	2011	Type II	E	CLM	163
97	2008	Type II	E	CLM	223
98	2008	Type II	E	CLM	210
99	2011	Type II	E	CLM	178
100	2013	Type II	E	CLM	144
101	2011	Type II	E	CLM	166
102	2013	Type II	E	CLM	160
103	2013	Type II	E	CR	140
104	2013	Type II	E	CR	209
105	2013	Type II	E	CR	195
106	2013	Type II	E	CR	218
107	2011	Type II	E	CR	166
108	2013	Type II	E	DFE	153
109	2011	Type II	E	DFE	230
110	2008	Type II	E	IEI	114
111	2008	Type II	E	IEI	141
112	2008	Type II	E	IEI	232
113	2013	Type II	E	IEI	149
114	2011	Type II	GS	CR	273
115	2008	Type II	GS	IEI	248
116	2011	Type II	GS	IEI	298
117	2013	Type II	GS	IEI	261
118	2013	Type II	I	ASA	269
119	2013	Type II	I	CR	35
120	2013	Type II	I	DFE	284
121	2008	Type II	I	IEI	263
122	2013	Type II	P/A	CR	370
123	2011	Type II	P/A	DFE	373
124	2013	Type II	P/A	IEI	368
125	2011	Type II	R/AR	ASA	394
126	2011	Type II	SS	IEI	396
127	2008	Type III	BS	CR	16
128	2008	Type III	BS	DFE	23
129	2013	Type III	BS	DFE	13
130	2013	Type III	BS	IEI	18
131	2008	Type III	Ch	CLM	63
132	2008	Type III	Ch	DFE	88
133	2013	Type III	Ch	IEI	296
134	2011	Type III	E	ASA	236
135	2013	Type III	E	ASA	206
136	2008	Type III	E	ASA	219
137	2008	Type III	E	CLM	220
138	2008	Type III	E	CR	203
139	2008	Type III	E	CR	190
140	2008	Type III	E	CR	171
141	2011	Type III	E	CR	177
142	2011	Type III	E	CR	245
143	2013	Type III	E	DFE	225
144	2013	Type III	E	DFE	227
145	2011	Type III	E	IEI	226
146	2013	Type III	E	IEI	224
147	2008	Type III	I	ASA	302
148	2011	Type III	I	ASA	392
149	2011	Type III	I	DFE	292
150	2013	Type III	I	DFE	295
151	2013	Type III	I	DFE	297
152	2011	Type III	I	IEI	322
153	2013	Type III	M	DFE	326
154	2008	Type III	P/A	IEI	366
155	2013	Type III	SS	CLM	389

Appendix C: Codebook for Content Analysis

THE IMPACT OF NSF INVESTMENTS IN UNDERGRADUATE ENGINEERING EDUCATION RESEARCH ~CODEBOOK~

Guidelines:

- This codebook includes codes and corresponding descriptions associated with the first research question, and the first two of out three of the corresponding sub-questions. The questions are:
    - What is a meaningful description of the impact of NSF investments in undergraduate STEM education R&D projects, based on Principal Investigators' (PIs)' perspectives?
        - What claims do PIs make about the impact of their NSF-funded projects?
        - How do PIs support their claims about the impact of their work?
- Defining Terms: "Research Impact" is comprised of three dimensions: scientific impact, societal impact, and domain-specific impacts
    - **Scientific impact:** advances in reliable knowledge (theories, methodologies, models, and facts) that primarily influence academic communities
    - **Domain-specific impact:** Influence of the methods or results of an R&D project on the people, priorities, and/or processes in the context of interest
    - **Societal Impact:** research outcomes or outputs that influence social, cultural, environmental, or economic dimensions of society
- All abstracts will be coded with two types of codes: Attribute codes, and Provisional/Descriptive Codes
    - **Attribute codes** include labels to denote basis project information (e.g., project type, STEM discipline, etc.)
    - **Provisional/Descriptive codes** include labels about more meaningful project information (i.e. claims PIs make and how the support them). Provisional codes are a priori codes from existing research impact frameworks. Descriptive codes are codes developed inductively as a result of conducting this analysis.
        - Assign Descriptive Codes to each abstract that address both research questions of interest
- The two sections of the abstract that will be coded are: "Dissemination" and "Impact". Other portions of the abstract are included only for the purpose of giving context to the information described in the two sections of interest.
- Before assigning a code to a segment, skim the first level codes in the codebook for the most appropriate code.
- The coding unit is an "idea" within an abstract (as opposed to a "sentence" or "word".) Apply the label to the whole phase/set of texts that represents the complete idea. If the same idea appears consecutively across sentences, then code the two sentences as one unit. If the same idea appears in separate locations, then apply the code twice.
- Codes ideas using first- and second- level codes. The first level codes are themes that describe a set of secondary codes. Secondary codes are mostly developed inductively.
- It is possible that one idea can be assigned multiple codes.
- Be open to new codes. If a code is not covered in the first level, code it as "Unidentifiable". If it is not covered at the second level, add it to the existing list underneath a first level code.

Overview of Attribute and Descriptive Codes

Attribute Codes – Basic Project Information

- Project Type
- Target STEM Discipline
- Project Focus

Descriptive Codes – Meaningful Project Information

What claims do PIs make about the impact of their research? **- Themes**

- Claims Regarding Conducting Research
- Claims Regarding Research- and Education-focused Developments
- Claims Regarding Disseminating Research Findings and Developments
- Claims Regarding Influence on Individuals and/or Communities
- Claims Regarding Influence on Environmental/ Structural Decisions, Metrics
- Claims Regarding Scope of Influence
- Claims Regarding Symbols of Impact
- Claims Regarding Unrealized Impact

How do PIs support their claims about impact? **- Themes**

- Clarifying Claims Using Degrees of Specificity
- Supporting Claims by Establishing Credibility and/or Relevance

ATTRIBUTE CODES

All Attribute codes *are labels used to indicate basic information about each abstract.*

PROJECT TYPE

- **A.** TUES Type I – Exploratory
- **B.** TUES Type II – Expansion
- **C.** TUES Type III – Comprehensive
- **D.** TUES Central Resource Project

TARGET DISCIPLINE

- **A.** Biological Sciences
- **B.** Chemistry
- **C.** Computer Science
- **D.** Engineering
- **E.** Geological Sciences
- **F.** Interdisciplinary
- **G.** Mathematics
- **H.** Physics/Astronomy
- **I.** Research/Assessment of Research
- **J.** Social Sciences

PROJECT FOCUS

- **A.** Assessing Student Achievement
- **B.** Conducting Research on Undergraduate STEM Education
- **C.** Creating Learning Materials and Teaching Strategies
- **D.** Developing Faculty Expertise
- **E.** Implementing Educational Innovations

DESCRIPTIVE CODES

What claims do PIs make about the impact of their research?

THEME 1A: Claims Regarding Conducting Research

Theme Description: Claims about people involved in conducting the research and/or the major steps in the research projecess.

a. **Parties Involved in Conducting Research**

   Examples: working with collaborators on the project; engaging new people in conducting research (e.g., undergraduate student researchers)

b. **Connections between the Current Study & Existing Literature**

   Examples: Links to existing literature/bodies of work, other research, prior work; references to prior tools that may have inspired or been the foundation for the current work. This also includes connections to national level policy reports suggesting research priorities for readers to consider.

c. **Highlights of current research activities**

d. **Highlights of research findings**

   Examples: new insights, unique contributions to the body of literature; data that resulted from the study that can be analyzed as part of a future study

e. **Applying for and securing additional funds to continue research**

   Examples: submitting a grant proposal that builds on the current work, and/or being awarded funding in response to submitting a grant proposal

THEME 2A: Claims Regarding Research- and Education-focused Developments

Theme Description: Claims about the development of artifacts that imply permanence and sustainability of the research topic beyond the current study; and tangible, educational materials informed from the current study.

f. **Research-focused Developments**

   i. Text-based entities

      Example: new scholarly journal focused on a niche research area

   ii. Discussion-based entities

      Example: new research symposium; new research conference or new strand of activities within an existing conference; interactive website for engaging in discourse with others who share similar interest; new journal club focused on a niche research area; venue for people who share similar research interests

   iii. Facility Instituted, Technology Developed for Conducting Research

      Example: new research center/facility, or learning center/facility; data mining tool developed to benefit a group within the research and/or education community

g. **Education-focused Developments**

   iv. Curricular materials, training resources, and pedagogy

      Description: The development of text-based curricular materials, digital resources useful for teaching and/or training educators; evidence-based pedagogical practices

      Examples: modules; educational lessons; textbook; assessment tools; static tutorials; workbook; how-to guides; training materials on DVD

   v. Instruments, Technology used for educational purposes

Examples: media-rich technology; purchasing lab equipment; dynamic tutorials; interactive websites; repository for developments

h. Affordances of the Developments

Examples: discussion on the utility of the development as part of conducting research, teaching, or in promotion & tenure activities/documents; discussion on what the development enables; references to economic value/cost saving associated with the development. It also includes the capacity to engage in a new set of learning and/or research experiences.

THEME 3A: Claims Regarding Disseminating Research Findings and Propagating Developments

Theme Description: Claims about how research findings and/or educational developments are being shared with other researchers and/or practitioners.

i. Disseminating Research Findings

vi. Text- and Discussion-based mediums

Description: A passive way of informing others about research findings. Examples: conference poster or proceeding; journal manuscript or publication; research brief; guidebook/manual; book about the research, that is not a textbook; website – if it is static/only used to post content.

Description: An interactive way of informing others about research findings. Examples: presentation about research findings-- at a particular research symposium, conference, or similar gathering; participation in a panel discussion.

Coding guideline: When a statement lists multiple types of texts, code each text-based medium individually.

j. Dissemination Via Venues for Teaching, Training

Examples: classroom; workshop; webinar; seminar; demonstration at a conference or museum; consulting services; one-on-one training/coaching.

k. Propagation of Developments

vii. Activities supporting propagation of developments

Examples: engaging in activities to promote the establishment of partnerships, marketing, commercialization; establishing mechanisms (e.g., mailing list) to keep track of interests from educators and/or vendors; following up on interest from vendors, industry.

viii. Outcome of activities supporting propagation of developments

Examples: established partnerships; marketed materials; commercialized materials; paid subscribers

THEME 4A: Claims Regarding Influence on Individuals and/or Communities

Theme Description: Claims about ways in which individual or communities of learners, instructors, and researchers are being affected by the outcomes of the study.

l. Influence of Teaching on STEM Learners

Description: participation in experiences that lead to the acquisition and application of knowledge and skills, changes in epistemology (ways of knowing), changes in ontology (ways of being) (Dall'Alba, 2009).

This also includes undergraduate students' interest in pursuing graduate studies.

This also includes changes in STEM literacy among individuals outside of undergraduate education.

m. **Influence of Training on STEM Education Instructors and/or Community of Instructors**

"Instructors" include Teaching Assistants, Post-doctoral Staff, Faculty, K-12 Teachers

Description: participation in experiences that lead to the acquisition and application of knowledge and skills, changes in epistemology (ways of knowing), changes in ontology (ways of being) (Dall'Alba, 2009). *This also incudes participation a virtual venue designed for instructors to interact with and learn from each other.*

Coding guideline: This may overlap with education-focused developments.

n. **Influence of Research on STEM Education Researchers and/or Research Community**

"Researchers" include, but is not limited to: undergraduate & graduate research assistants; post-doctoral researchers.

Examples: The development of new skills associated with conducting STEM/Education research; effects on grant writing, such as improved quality of proposals; increased interest among colleagues in applying for funding and/or conducting research; significant expansions in the research community associated with the PIs' expertise; development of new community of researchers. This also includes instances where teaching- learning-related results that stem from the current project lead to new research-focused activities/projects.

o. **Direct Personal, Professional Benefits to Instructors, Researchers**

Examples: expansion of the PIs' professional network as a result of conducting research, teaching, training session; inclusion of research activities in promotion and tenure package that lead to positive professional outcomes

THEME 5A: Claims Regarding Influence on Environmental/Structural Decisions, Metrics

Theme Description: Claims about how insights from the current study inform administrative decisions that ultimately influence the actions of others, and how the current study contributes to assessments and/or metrics of interest to administrators.

p. **Department-Level Decisions**

ix. Curricular Changes

Example: Administrative support for … within a department; modifications to a course; new course offering; changes in courses students are advised to take.

x. Informing or Enacting Education Policy

Description: participation in policy-related discussion taking place at local and national –level gatherings that have implications for the local department

xi. Informing Promotion & Tenure Decisions

q. **Department-Level Insights, Assessments and Metrics**

Example: research findings that lead to insights about structural issues; changes in student enrollment, student retention, and other department level/aggregate student outcomes of interest (e.g., "Drop, Fail, Withdraw-rates); curricular changes that contributed to department accreditation; inclusion of research data or findings in ABET evaluation.

THEME 6A: Claims Regarding Scope of Influence

Theme Description: Claims about the span associated with their project outcomes.

r. **Geographic scope**

Examples: use of terms like "international", "nation" (usually in a broad, vague sense); mention of specific states in the U.S.; references to particular types of states (e.g, "EPSCor" state)

s. Disciplinary scope

Example: reference to a discipline other than the target discipline associated with the study

t. Institutional scope

Description: References to type of institutions making use of or have expressed an interest in making use of research developments resulting from the study of interest.

Example: List of institutional types includes, but is not limited to, Historically Black Colleges & University (HBCU); Hispanic-serving institution (HSI); Tribal College or University (TCU); Community College. This also includes institutions comprised of formal & informal learning environments (e.g., K-12 school, museum), and funding agencies (e.g. NSF).

u. Scope via Target Populations

Example: reference to undergraduate students in particular STEM disciplines (e.g., engineering), "at risk" students, underrepresented minorities; particular grade levels (e.g., freshman, first-year, sophomore, junior, senior, capstone course). This also includes activities that target administrators; program officers at funding agencies.

v. Non-academic Partnerships

Example: reference to partnerships with vendors, industry, or professional societies to advance dissemination.

THEME 7A: Claims Regarding Symbols of Impact

Theme Description: Claims about the receipt of public affirmation as a result of connections to the current study.

w. Affirmation from within the academic community

Example: receiving and award or special recognition; being labeled/recognized as a model or exemplar; presenting a keynote address; securing funding to continue research or participated in a specialized professional development opportunity. This also includes bestowing an award or special recognition; identifying a model or exemplar; appointing a "fellow" (leaders to carry on the work)

x. Affirmation from outside the academic community

Example: press coverage; media reports/news articles; featured stores on TV or in written publications; featured in the NSF Highlights

THEME 8A: Claims Regarding Unrealized Impact

Theme Description: Claims about activities, events, and outcomes that have not yet happened, but are either future plans or anticipated outcomes that will be realized at a later time.

y. Future Plans to Conduct Research, Create Developments, or Disseminate

Claims regarding future plans or intentions to collaborate with others, conduct research, disseminate research findings, share educational developments, etc.

Coding guideline: Do not apply secondary codes to indicate exact plans to conduct research.

z. Indeterminate Impact

Examples: Statements/admission that impact is "unclear", "hard to determine", or "yet to be determined".

aa. Indirect and/or Expected Impact

Description: Links between direct, secondary or tertiary outcomes that may or may not have been observed/realized yet. Oftentimes, discussed as a chain reaction that ultimately links to student learning. Sometimes used in conjunction with geographic references to span of impact. Sometimes this is tied to future work. It's an indication of how soon outcomes will be realized.

Example: discussion on what a new tool enables users to do after discussing the development of the tool itself; use of terms like "expected", "anticipated", or "predicted". Also includes references to societal-level impact constructs (e.g., economy, environment, technological literacy of citizens)

Coding guideline: Do not apply secondary codes to denote the specifics of the indirect or anticipated outcomes.

How do PIs support their claims about impact?

THEME 2A: Clarifying Claims Using Degrees of Specificity

Theme Description: Clarifying claims using language that is either more or less specific.

a. More specific

xii. Quantifying outcomes

Example: number of documents (publications, etc.), workshops hosted, workshop attendees/participants; adopters/users (if its an online system), institutions represented among the participants/users/adopters; references to durations of time

xiii. Qualifying claims

Description: A way to clarify and/or limit the extent to which a claim can be taken.

Examples: use of terms like "possibly", "likely", "potentially".

b. Less specific

Description: broad, vague descriptions of "impact"

THEME 2B: Supporting Claims by Establishing Credibility and/or Relevance

c. References to what motivated the current R&D project

Examples: References to prior work or gaps in the literature upon which the current study is built; connections to national-level priorities or discussions.

d. Reference to societal-level impact construct

Description: influence on social, cultural, environmental, or economic dimensions of society

e. Reference to evaluations and/or metrics

i. Quantitative Evaluation and/or Metrics

Example: Counting number of institutions served, involved; Course performance indicators; National data across institutions; Adoption rate and/or Usage statistics; Students' course performance indicators

ii. Qualitative Evaluation and/or Metrics

Example: Counting: Anecdotal Remarks; Indicators of Participation at events to gauge interest; Misc. indicators of interest; Teaching assistants' feedback; Online reflection tool; Topics of discussion at national meetings and conferences

iii. Mixed -Quantitative and Qualitative- Evaluation and/or Metrics

Example: Counting: Survey; Formative and/or summative evaluations from an external evaluator; Pre-post test; Course evaluation data; Students' work

iv. Unable to determine

Description: Judgment statements that suggest that some sort of evaluation has occurred, but no explicit reference to the type of evaluation. (Information may be found in other sections of the abstract.) This code is oftentimes linked with "Less specific" support for claims about impact

Appendix D: Recruitment Email to Survey Participants

Cover Letter

Dear [Participant's Name],

My name is Jeremi London. I am Ph.D. Candidate in the School of Engineering Education at Purdue University and an intern in the Division of Undergraduate Education at the National Science Foundation. With the increasing pressure on federal funding agencies to allocate funds more wisely to proposed research that shows great promise and to advance studies with compelling results, now is the time to articulate what the impact of research looks like. In light of this need, my dissertation will result in the development of a conceptual framework that characterizes the impact of NSF-funded undergraduate engineering education research and development (R&D) projects. This study is approved by IRB #1406014915.

The purpose of this email is to invite you to serve as a content expert on the statements that will be included in the framework. You are asked to serve because of your past experience as a NSF program officer, principal investigator on at least one NSF-funded project focused on engineering education, and academic training in an engineering discipline. Your participation in the review process is a valuable step toward future studies on ways to evaluate research impact.

The survey consists of items that were developed as a result of qualitatively analyzing the *claims* Principal Investigators (PI) make about the impact of their NSF-funded work in the "Dissemination" and "Impact" sections of project abstracts included in PI conference reports. If you choose to participate, you will have the opportunity to provide feedback on the relevance and clarity of each item in relation to definitions of research impact. You will also have an opportunity to evaluate the comprehensiveness of the set of items, and make suggestions for the addition or deletion of items.

This invitation is being sent to people who meet the criteria provided above. Thus, you are welcome to participate in this study via an online survey that can be found at: https://purdue.qualtrics.com/SE/?SID=SV_6Ve1ZCl49kclr81. The survey will take approximately 30-60 minutes to complete, and will close **Friday, June 27**. All survey responses are anonymous. Your participation is voluntary, and you can withdraw from this study any time before completing the survey. Since it is an anonymous survey, we cannot identify how to remove a participant who has completed the survey.

If you are interested in obtaining a copy of the final report that results from this study, the last survey question includes the link to another survey that allows you to add submit your email address to the researchers on this study. As a result of using two distinct surveys, your responses in one will not be linked to responses in the other.

If you have questions or concerns, please contact me, Jeremi London (jslondon@purdue.edu), or my Ph.D. research advisor, Dr. Monica Cox (mfc@purdue.edu). Thank you in advance for your cooperation and input.

Regards,

Jeremi London
Ph.D. Candidate, School of Engineering Education, Purdue University
Summer Scholar, Division of Undergraduate Education, National Science Foundation

Monica Cox, Ph.D.
Associate Professor, School of Engineering Education, Purdue University

Appendix E: Sample Protocol Questions for a Follow-up Study

One example of a follow-up study might include an exploration of various stakeholders' perspectives on what impact means in the context of undergraduate engineering education. Stakeholders might include not only PIs and NSF Program Officers, but also personnel at agencies like the National Academy of Engineering (NAE), Office of Science and Technology Policy (OSTP), and the Accrediting Board for Engineering and Technology (ABET). Methods that involve interviews and/or focus groups would allow participants to provide open-ended responses, ask clarifying questions, and elaboration on their initial thoughts. Sample interview protocol questions include:

1. Questions About the Participant's Work:
    a. Tell me the story of how you started doing engineering education-related work.
    b. In what ways is your current work at [NSF, NAE, OSTP, ABET] connected to engineering education?
2. Questions about the Agency's Work:
    a. What are the overarching mission and primary objectives of [your current agency of employment: NSF, NAE, OSTP, ABET]?
    b. How does the work of [NSF, NAE, OSTP, ABET] fit within the larger context of engineering education in the U.S.?
3. Questions Linking the Agency's Work and Impact:
    a. In ways does [NSF, NAE, OSTP, ABET] impact engineering education in the U.S.?

   b. What factors promote the impact of the work that [NSF, NAE, OSTP, ABET] does? What impedes it?

4. General Questions About Impact:
   a. More broadly, how would you define impact?
   b. What is an example of a project, program, or initiative related to engineering education that has been impactful? Why?

VITA

VITA

Jeremi S. London

jslondon@purdue.edu

ENGINEER AND ENGINEERING EDUCATION SCHOLAR WITH A DYNAMIC BACKGROUND IN POLICY, RESEARCH, INDUSTRY, TEACHING AND SERVICE

HIGHLIGHTS

Keen insight on relationships between government and academic sectors as part of educating engineers.
Makes scholarly contributions to research on STEM education policy, industrial engineering, and engineering education.
Knowledge of various analytic software (e.g., SAS, SPSS, Atlas.ti, Arena Simulation).
Mixed methods researcher with interests in: STEM education policy, cyberlearning, and agent-based simulation modeling.

EDUCATION

Purdue University

Ph.D., Engineering Education	GPA: 3.9 (Winter 2014)
M.S., Industrial Engineering	Spring 2013
B.S., Industrial Engineering	Spring 2008

HONORS

Purdue School of Engineering Education Graduate Student Researcher Award, Honorable Mention (2014)
Purdue College of Engineering Outstanding Service Award (2013)
Gerald I. Gilbert Memorial Scholarship (2011, 2012, 2013)
Indiana Space Grant Consortium Graduate Fellowship (2012)
National Science Foundation Summer Scholar Internship Program Intern (2011, 2012, 2013)
National Science Foundation S-STEM Fellowship Scholar (2010)
National Science Foundation REU (Research Experiences for Undergraduates) Student (2007 – 2008)
INROADS Frank C. Carr Community Service Award (2005 & 2006); Leadership Award (2005 & 2006)

POLICY

National Science Foundation, Washington, D.C.
Summer Scholar, Division of Undergraduate Education (DUE) 2011, 2012, 2013

Explored ways cyberlearning tools can provide more equitable learning experiences for all learners by:

- Presented descriptive statistics to summarize 800 awards in the DUE portfolio funded between 2001-2011
- Created a taxonomy of cyberlearning after synthesizing the perspectives of 18 NSF Program Officers
- Proposed 17 recommendations to NSF Program Directors on research directions for cyberlearning

Provided a snapshot of the current landscape of cyberlearning research; and studied 16 exemplars by:

- Reviewed the outcomes of 100 noteworthy cyberlearning projects; organized them using a taxonomy
- Interviewed the leader of 16 exemplars to garner insights on development, dissemination & sustainability
- Summarized highlights of exemplars, recommendations to future developers, and untapped opportunities

Contributing to DUE, colleagues within DUE & the engineering education research community by:

- Developing a framework DUE Program Officers can use to evaluate the impact of projects
- Mentored a fellow-intern through the execution of a research idea that led to a poster presented at the 2012 Society for the Advancement of Chicanos and Native Americans in Science (SACNAS) conference

Developing a NSF research agenda around the role of MOOCs in engineering education:

- Wrote a travel grant to obtain support four workshops taking place at engineering conferences; managing a $40,000 budget
- Bringing together a panel of experts on Massive Open Online Courses (MOOCS), learning science, and engineering faculty to discuss the potential for MOOCs, and stimulating conversations and research collaborations among engineering faculty
- Writing a research agenda that will influence NSF's investments in MOOCs in engineering education

RESEARCH

Purdue University, West Lafayette, IN
Ph.D. Dissertation: **"The Research Impact of National Science Foundation Investments in Undergraduate Engineering Education: An Exploratory Mixed Methods Study".** Major Professor: Dr. Monica Cox

Master's Thesis: **"Analysis and Modeling of Learning Outcome Mappings in Engineering Education".** Major Professor: Dr. Barrett Caldwell

Graduate Research Assistant, School of Engineering Education
Fall 2010-current
Exploring the knowledge, skills & attributes engineering Ph.D.s need for successful careers in industry or academia by:

- Interviewing 40 engineering Ph.D.s, synthesizing responses, and disseminating research findings
- Translating qualitative findings into assessment instruments, and professional development opportunities for doctoral engineering students

Assessing and providing feedback on the effectiveness of graduate teaching assistants' pedagogy by:

- Developing a multi-dimensional, real-time assessment tool based on the *How People Learn* framework
- Providing graduate teaching assistants with actionable feedback based on profiles generated by the tool

Undergraduate Research Assistant, School of Engineering Education Fall 2007- Spring 2008
Carried out research and administrative tasks as part of learning about the research process by:

- Creating a project plan for a study that included three data collection methods and human subjects
- Preparing an interview protocol and conducting interviews

Graduate Student Researcher on NSF-funded Projects

- Project Title: "Engineering Education Pioneers and Trajectories of Impact". Co-PIs: Cynthia Atman, KenYashura, Jennifer Turns (Spring 2014)
- Project Title: "Workshops to Create a Taxonomy for Engineering Education Research and Prioritize Areas of Research". Co-PIs: Cynthia Finelli, Maura Borrego (Summer 2013)

INDUSTRY

GE Healthcare, Wauwatosa, WI
Quality Assurance Specialist, Invasive Cardiology (ICAR) Division 2008-2009

Reviewed 100% of complaints about ICAR medical devices weekly; 600-900 complaints by:

- Responded immediately with corrective action to complaints about patient safety, quality or compliance.
- Prepared weekly trend reports & presentation for FDA audit that resulted in zero federal observations.
- Led an in-depth investigation with nurses and design engineers to examine unusual data trends, leading to fact-based discoveries and methodical direction for the management team.
- Initiated training and guidelines for field engineers to improve communication with QA Specialists. Resulted in a 50% reduction in the number of complaints requiring follow-up with field engineers.

Medallion Entry Systems, Indianapolis, IN
Industrial Engineering Senior Design Student
Spring 2008

- Work on a team to design a software application that enabled the company to accurately estimate project lead times, inventory levels, and provide precise quotes for their customers

Anheuser-Busch, St. Louis, MO
Logistics Coordinator
Summer 2006

- Planned logistics to ensure that wholesalers had at least three days of inventory
- Created weekly trucking schedules & monitored supply chain to ensure low cost, on-time deliveries

Corporate Quality Assurance Analyst
Summer 2005

- Consolidated vendor relationship by identifying those that could serve as national and regional suppliers; thereby saving the company money and strengthening supplier relations
- Streamlined process for vendor order by developing a database to store equipment purchasing data and a Standard Operating Procedure

Research Pilot Brewery Lab Technician
Summer 2004

- Evaluated brew samples using various lab equipment, and generated trend reports weekly

- Presented lab observations and peculiar data to team members, and Director of Research Pilot Brewery

TEACHING

Course: Preparing Future Professionals, Purdue Graduate School Fall 2012

Restructured critical elements of course instruction and prepared content for online mediums by:

- Aligning course content, assessment, and pedagogy for 5 lectures on topics relevant to graduating students
- Translating content for five 90 min lectures into a collection of 10-15 minute online modules

Facilitator, Purdue Teaching Assistant Orientation
Fall 2011

Facilitated workshops on classroom management and gave personalized feedback on pedagogy by:

- Led interactive sessions on "Managing Learning Environments", "Academic Integrity" and "Managing TA Responsibilities"
- Provided 9 Teaching Assistants with personalized feedback on the effectiveness of their teaching practices after their 7-min Micro Teaching Session

Summer Camp Leader, Purdue School of Biomedical Engineering
Summer 2007

Taught 4th & 5th grade students Biomedical Engineering concepts using everyday occurrences by:

- Engaging the students in hands-on learning activities
- Using materials to simulate the whiplash that can occur during a car accident and the impact of riding a bike without a helmet

Tutor, WyzAnt Home Tutoring Services, Inc. Fall 2009-Spring 2010

Tutored high school and undergraduate students in math and science by:

- Established and maintained rapport with students while sharing study skills necessary for academic success
- Evaluated students academic performance and provided ways to bridge gaps in understanding

SERVICE

Graduate Student Representative, Program Committee
Cyberlearning Summit 2014 Fall 2013 –current

Reviewer, Journal of Engineering Education Spring 2014

Inaugural Chair, Professional Development Committee
Graduate Engineering Education Consortium for Students (GEECS)
2011-2012

Reviewer, Education and Research Methods Division
American Society for Engineering Education Fall 2013

Junior Member
Graduate Committee, Purdue School of Engineering Education 2011-2013
Student Representative for School of Engineering Education
Purdue Graduate Student Advisory Council 2011 - 2012

Treasurer
Adventist Collegiate Fellowship at Purdue 2012 - 2014

Member
American Society for Engineering Education 2010-current

Human Factors and Ergonomics Society 2010-current

Prior to 2010 Mentor & Mentee, Women in Engineering Program at Purdue
Mentor, Big Brothers Big Sisters of Lafayette, IN
Tutor, Chapter Senator, National Society for Black Engineers (Purdue Chapter)
FUNFest Coordinator, INROADS/St. Louis, MO

PUBLICATIONS

Thesis

London, J.S. (2013). *Analysis and Modeling of Learning Outcome Mappings in Engineering Education.* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1544419).

Journal Publications

London, J., Cox, M.F., Ahn, B., Branch S., Torres-Ayala, A., Zephirin, T., Zhu, J. (In Review). Motivations for Pursuing an Engineering Ph.D. and Perceptions of its Added Value. *International Journal of Doctoral Studies.*

Ahn, B., Cox, M.F., **London, J.,** Cekic, O. and Zhu, J. (2014). Creating an Instrument to Measure Leadership, Change, and Synthesis in Engineering Undergraduates. *Journal of Engineering Education, 103*(1), 115-136.

Zhu, J., Li, Y., Cox, M., **London, J.,** Hahn, J., Ahn, B. (2013). Validation of a Survey for Graduate Teaching Assistants: Translating Theory to Practice. *Journal of Engineering Education 102*(3), 426-443.

Cox, M.F., **London, J.,** Zhu, J., Ahn, B., Zephirin, T., Taylor, K. (Paper Accepted, 2012). Curriculum Vitae Analyses of Engineering Ph.D.s Working in Academia and Industry. *International Journal of Engineering Education*

Cox, M. F., Hahn, J., McNeill, N., Cekic, O., Zhu, J, & **London, J.** (2011). Enhancing the Quality of Engineering Graduate Teaching Assistants through Multidimensional Feedback. *Advances in Engineering Education*, *2*(3), 1-20.

Cox, M.F., Zhu, J., Cekic, O., Chavela, R., & **London , J.** (2010). Knowledge or Feelings: First-year Students' Perceptions of Graduate Teaching Assistants in Engineering. *Journal of Faculty Development, 24*(1), 27-34.

Book Chapter

Cox, M.F., Zhu, Jiabin, Z., **London, J.,** Hahn, J., Ahn., B. (2012). Feedback about Graduate Teaching Assistants' Pedagogical Practices: Content Validation of a Survey Informed from Principles of the "How People Learn" Framework. In G. Gorsuch (Ed.), *Working Theories for Teaching Assistants and International Teaching Assistant Development.* (pp. 63-82). Stillwater, OK: New Forums Press, Inc.

Conference Proceedings

London, J., Young, C. (Abstract Accepted, 2014). *Developing a Research Agenda Around MOOCs in Engineering Education (Work in Progress).* Abstract Accepted to the MOOCs in STEM: Exploring New Educational Technologies Conference, San Jose, CA.

Berdanier, C., Cox, M., Ahn, B., **London, J.** (Paper Accepted, 2013). *Survey Analysis of Engineering Graduate Students' Perspectives on the Skills Necessary for Career Success in Industry and Academia.* Paper accepted to the 2014 American Society for Engineering Education Conference, Indianapolis, IN.

Ahn, B., Cox, M., **London, J.,** Zhu, J. (Paper Accepted, 2013). *Investigating the Attributes and Expectations of Engineering Ph.D.s Working in Industry.* Paper Accepted to the 2013 Frontiers in Education Conference, Oklahoma City, OK.

Zhu, J., Cox, M., Branch, S., Ahn, B., **London, J.** (Paper Accepted, 2013). *Recommendations for Engineering Doctoral Education: Design of an Instrument to Evaluate Change.* Paper Accepted to the 2013 Frontiers in Education Conference, Oklahoma City, OK.

London, J. (2013). *Highlighting Cyberlearning Tools with Compelling Results.* Paper presented at the 2013 American Society for Engineering Education Conference, Atlanta, GA.

London, J., Caldwell, B.S., Patsavas, K. (2013). *Aligning Learning Outcomes and the Engineering Educational Accreditation Expectations.* Paper presented at the 2013 International Systems Engineering and Research Conference, San Juan, Peurto Rico.

London, J., Caldwell, B.S., Patsavas, K. (2013). *Modeling Expertise Development in Undergraduate Engineering*

Education. Paper presented at the 2013 International Systems Engineering and Research Conference, San Juan, Peurto Rico.

Sambamurthy, N., Hahn, J., **London, J.,** Cox, M. (2013). *Reliability of the Global Real-time Assessment Tool for Teaching Enhancement (G-RATE).* Paper presented at the 2013 American Society for Engineering Education Conference, Atlanta, GA.

London, J. (2012). *Exploring Cyberlearning through a NSF Lens.* Paper presented at the annual American Society for Engineering Education Conference, San Antonio, Texas.

London, J., Ahn, B., Cox, M.F. (2012). *LSAMP X: Lessons Learned from a Diversity Program Serving Underrepresented Minority Students.* Paper presented at the annual American Society for Engineering Education Conference, San Antonio, Texas.

Zhu, J., **London, J.,** Ahn, B., Cox, M.F. (2012). *Recommendations for Promoting Desirable Characteristics in Engineering Ph.D.s: Perspectives from Industry and Academia.* Paper presented at the annual American Society for Engineering Education Conference, San Antonio, Texas.

Coso, A., Louis, R., **London, J.,** Ngambeki, I., Sattler, B. (2012). *Exploring the Reasons for Collaboration and Cooperation among Graduate Student Researchers.* Paper presented at the annual American Society for Engineering Education Conference, San Antonio, Texas.

Cox, M.F., **London, J.,** Ahn, B., Frazier, S., Zhu, J., Torres-Ayala, A. (2011). *Attributes of Success for Engineering Ph.D.s: Perspectives from Academia and Industry.* Paper presented at the annual American Society for Engineering Education Conference, Vancouver, Canada.

Cox, M.F., Ahn, B., **London, J.,** Frazier, S., Zhu, J., Torres-Ayala, A. (2011). *Choices for Ph.D.s in Engineering: Analysis of Career Paths in Academia and Industry.* Paper presented at the annual American Society for Engineering Education Conference, Vancouver, Canada.

McNeill, N., Cox, M.F., Diefes-Dux, H., **London (Hayes), J.,** & Medley, T. (2008). *Development of an Instrument to Collect Pedagogical Data from Graduate Teaching Assistants within Engineering Laboratories.* Paper presented at the annual American Society for Engineering Education Conference, Pittsburg, PA.

Reports

London, J., Millard, D., Piotrowski, V., Zia, L. (2012). *A Snapshot of Cyberlearning: Highlights and Implications.* NSF Internal Report

London, J., Millard, D., Piotrowski, V., Zia, L. (2011). *Exploring Cyberlearning through a NSF Lens.* NSF Internal Report

Workshops & Invited Presentations

London, J., Young, C. (Workshop Request Accepted, 2013). *Developing a NSF Research Agenda Around MOOCs in Engineering Education.* Workshop request accepted to the 2014 ASME International Mechanical Engineering Education Leadership Summit, San Juan, PR.

London, J., Young, C. (Workshop Request Accepted, 2013). *Developing a NSF Research Agenda Around MOOCs in Engineering Education.* Workshop request accepted to the 2014 Electrical and Computer Engineering Department Heads Association Annual Meeting, Napa, CA.

London, J., Young, C. (Workshop Request Accepted, 2013). *What is the Role of MOOCs in Engineering Education?* Workshop request accpeted to the 2014 American Society for Engineering Education Conference, Indianapolis, IN.

London, J., Young, C. (Workshop Request Accepted, 2013). *What is the Role of MOOCs in Engineering Education?: NSF Research Agenda Development Workshop.* Workshop request submitted to the 2014 Biomedical Engineering Society Annual Conference, San Antonio, TX.

London, J., Cox, M., Ahn, B. (Workshop Request Accepted, 2013). *Convincing the Non-believers: Selling Engineering Education Experiences on the Job Market.* Workshop request accepted to the 2014 American Society for Engineering Education Conference, Indianapolis, IN.

London, J., Herman, G.L., Thomas, L.D. (Special Session Request Submitted, 2013). *Benefits of Being in GEECS.* Workshop request submitted to the 2014 American Society for Engineering Education Conference, Indianapolis, IN.

London, J. (2012). Exploring Cyberlearning Through a NSF Lens. Invited presentation on Jan. 24, 2012 in course: "Research, Design, and Evaluation of Learning Experiences and Environments for Discipline-Based Computational Thinking" (Instructor: Alejandra Magana, Ph.D.) [Invited Presentation]

London, J. (2013). Is Graduate School Right for Me? – Part 1. Louis Stokes Alliance for Minority Participation in STEM – Purdue University Campus. Live webinar. [Invited Presentation]

London, J., Ahn, B. (2013) Is Graduate School Right for Me? – Part II. Louis Stokes Alliance for Minority Participation in STEM – Purdue University Campus. Live webinar. [Invited Presentation]

GRANT CONTRIBUTIONS

Learning Outcomes Assessment Grant 2012. Purdue University. Principal Investigator: Dr. Barrett Caldwell

"What is the Role of MOOCs in Engineering Education?". National Science Foundation Travel Grant. Principal Investigator: Dr. Cynthia Young