ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE
Paper ID #36830

Catalyzing U.S. Innovation and Entrepreneurship: Approaching the Evaluation of the National Science Foundation's I-Corps Program

Nathalie Duval-couetil (Associate Professor and Director)

Nathalie Duval-Couetil is the Director of the Certificate in Entrepreneurship and Innovation Program, Associate Director of the Burton D. Morgan Center, and a Professor in the Department of Technology Leadership and Innovation at Purdue University. Nathalie's education and research activities focus on entrepreneurship pedagogy and assessment; entrepreneurship and STEM; student intellectual property policy; graduate student professional development; and women and entrepreneurship. Prior to academia, Nathalie worked in the field of market research and strategy consulting in Europe and the U.S. with Booz Allen and Hamilton and Data and Strategies Group. She received a BA from UMass Amherst, an MBA from Babson College, and MS and PhD degrees from Purdue University.

Alanna Epstein

Assistant Research Scientist

Aileen Huang-saad (Associate Professor)

Dr. Huang-Saad is an Associate Professor of Bioengineering at Northeastern University and the Director of Life Sciences and Engineering Programs at Northeastern's Roux Institute in Portland, Maine. Dr. Huang-Saad is Deputy Editor-in-Chief of Springer's Biomedical Engineering Education and Division Chair for the American Society of Engineering Education's Biomedical Engineering Division. Dr. Huang-Saad's current research areas are entrepreneurship, innovation, and transforming higher education. She is funded by the NSF to explore the influence of the microenvironment of entrepreneurship education on minoritized populations, entrepreneurial ecosystems, and fostering graduate student professional development.

Catalyzing U.S. Innovation and Entrepreneurship: Approaching Assessment and Evaluation of the National Science Foundation's I-Corps Program

The National Science Foundation's (NSF) Innovation Corps (I-Corps) Program was created to stimulate innovation and benefit society by training academic entrepreneurs to participate in technology commercialization and entrepreneurship. Defining outcomes and measuring the impact of a national initiative such as I-Corps is complex for many methodological and practical reasons. This study describes the I-Corps program structure and highlights the role that regional I-Corps Nodes play in training and evaluation activities. Interviews with Node administrators highlight the diversity of approaches to evaluation activities and the existence of key evaluation challenges. They also offer insights into how these challenges could be overcome with more support and coordination nationally and across regional ecosystems.

Keywords: technology commercialization; technology transfer; entrepreneurship education; program evaluation; assessment

INTRODUCTION

The National Science Foundation's (NSF) Innovation Corps (I-Corps) Program was created to "stimulate innovation and benefit society through the translation of fundamental research" (National Science Foundation, 2019, p. 4) by delivering entrepreneurship and commercialization training to university researchers in fields such as science and engineering. The training is based on Lean Startup principles (Blank & Dorf, 2012), and is intended to enable scientific researchers to assess the commercialization potential of their innovations by identifying potential customers and markets (Huang-Saad et al., 2017; Nnakwe et al., 2018). Since many consider education to be key in fostering entrepreneurial activity (Wright et al., 2009), training is viewed as a key component of university entrepreneurial ecosystems (Bischoff et al., 2018; Boh et al., 2016; Goethner & Wyrwich, 2020).

Until very recently, the program was administered through regional networks of universities referred to as I-Corps Nodes, which played a critical role in generating awareness of the I-Corps program; identifying academic scientists with marketable inventions; offering workshops and training to increase awareness of an involvement in the national I-Corps Teams program; and activating regional entrepreneurial ecosystems to identify industry partners able to guide and mentor commercialization teams. Not surprisingly, given the objectives and structure of the program, program evaluation and assessment activities are complex for several methodological and practical reasons that are similar to those identified in prior literature. First, there is heterogeneity in training format, content and instructors. Second, educational and technology commercialization outcomes are often confounded within academic technology entrepreneurship programs (Nelson & Monsen, 2014). Third, outcomes can be measured in the short- and long-term at the individual, program, and institutional levels (Wright et al., 2009). Fourth, there is also variability in the resources and expertise available to carry out program evaluation activities (Rossi et al., 2018). These and other issues are important to address

given the expansion of I-Corps, NSF's desire for more precise evaluation and assessment data to improve regional programs, and demand for research on entrepreneurship education practices that generalize nationwide.

The purpose of this paper is to examine the program evaluation and assessment activities that were undertaken at I-Corps Nodes. To do so, we conducted interviews with program administrators at each of the nine Nodes, who were familiar with regional evaluation activities, to explore the diversity of approaches being used and key challenges. We were also interested in how challenges might be overcome with more resources and coordination. It is clear that many Nodes and Sites lack funding and expertise necessary to do evaluation research, and the sharing of instruments and evaluation protocols could enhance evaluation and assessment activities. As I-Corps enters a new phase of expansion, findings from this study will inform our own evaluation work going forward, and allow us to share best practices with I-Corps educators and administrators across the U.S.

BACKGROUND

Academic Entrepreneurship

Many institutions are actively investing in entrepreneurial ecosystems and initiatives to boost the involvement of faculty and graduate students in commercialization activity given evidence that commercializing university research is beneficial for economic growth both regionally and nationally (Calcagnini et al., 2016; Cesaroni & Piccaluga, 2016; Fitzgerald & Cunningham, 2016; Guerrero et al., 2015; Huyghe & Knockaert, 2015; Welter et al., 2015). Although involvement in academic entrepreneurship has grown considerably in recent years (Abreu et al., 2016; Chang et al., 2016), scholars view further expansion of these activities as desirable (Hayter, Link, et al., 2018; Siegel & Wright, 2015). Therefore, scholars are interested in opportunities for growth in academic entrepreneurship, through a better understanding of the predictors of more involvement in technology commercialization by scientific researchers (Guerrero et al., 2016; Guo et al., 2019; Hayter, Nelson, et al., 2018; Klofsten et al., 2019; Liu, 2018).

In a review of academic entrepreneurship ecosystems worldwide, Hayter and colleagues (2018) summarized the research on the predictors of activity in academic entrepreneurship into eight independent variable categories; (1) characteristics of academic entrepreneurs, (2) human capital, (3) social networks, (4) entrepreneurial environment, (5) financial resources, (6) scientific, technical, and product characteristics, (7) academic entrepreneurship programs, and (8) university management and policies. The authors concluded that researchers should not rely on "the linear, patent-focused technology transfer context" (p. 1073) and instead conceptualize it as a more complex system occurring within and among individual, organizational, and regional levels. Klofsten and colleagues (2019) summarized key "strategic challenges" for academic entrepreneurship having to do with: (1) internal institutional factors, (2) external and environmental factors, (3) entrepreneurial curricula and pedagogy, (4) support for diverse entrepreneurial pathways, and (5) impact measures. This is demonstrated in the varying levels of support of entrepreneurial activities (Klofsten et al., 2019; Rasmussen & Wright, 2015), and attitudes towards the expansion of traditional faculty and administrator roles (Etzkowitz, 2016; Wright & Phan, 2018), both of which may influence tenure and promotion decisions (Sanberg et al., 2014).

Other challenges include time constraints, conflict of interest concerns, and changes to relationships between graduate students and their faculty advisers when they become "employees" of academic startups, in addition to their roles as students (Duval-Couetil et al., 2021).

Expanding education and training for university scientists has been proposed as means of increasing interest and involvement in technology commercialization and entrepreneurship activity (Duval-Couetil et al., 2021; Gianiodis & Meek, 2019; Huang-Saad et al., 2017; Klofsten et al., 2019; Maritz et al., 2016). However, the research linking entrepreneurship education to technology commercialization outcomes is somewhat disconnected, as is the literature of the two fields, i.e., "entrepreneurship education typically focuses on students, whereas technology commercialization describes the professor's experience," and both tend to be studied "in relative isolation" (Gianiodis and Meek, 2020, p. 1169). Similarly, as Fayolle et al. (2020) observed, "there has traditionally been only a weak link between the fields of research, entrepreneurship education and technology transfer and commercialization" (p. 278).

The National Science Foundation I-Corps Program

The NSF I-Corps program is a nationwide initiative designed to "accelerate the development of new technologies, products and processes that arise from fundamental research" by addressing challenges "inherent in the early stages of the innovation process" (National Science Foundation, 2019, p. 6). Specifically, I-Corps addresses the lack of knowledge academic researchers have about technology entrepreneurship given that few receive such training as part of their doctoral programs (Duval-Couetil, et al., 2020). As a result, more are likely to conform to conventional cultural norms in academia which incentivize obtaining funding to support research and laboratories, rather than pursuing commercialization opportunities (Huyghe & Knockaert, 2015). Hayter et al. (2018), stated that "limited attention has been placed on the role of national programs and policies and their critical role supporting academic entrepreneurship" (p. 1059). Considering this, I-Corps offers a valuable context for examining the ability of a program to promote and grow involvement in academic entrepreneurship.

At a national level, the NSF describes four overarching objectives for I-Corps, including: 1) leveraging federal research investments by advancing commercialization of research outcomes; 2) transforming the culture at the nation's institutions of higher education by preparing scientists and engineers to extend their focus beyond the research laboratory; 3) assisting STEM (science, technology, engineering, math) researchers to translate laboratory discoveries into product technologies with benefits for the economy and society; and, 4) increasing the economic impact of federally-funded research (National Science Foundation, 2019, p. 6). To achieve this, I-Corps was initially administered through a network of regional Nodes and Sites designed to build, draw upon, and sustain a national innovation ecosystem (National Science Foundation, 2016). Originally, these three network elements were funded by three separate NSF grants. Going forward, Site and Node grants will be consolidated into a new regional "Hub" model.

The Midwest I-Corps Node

The Midwest I-Corps Node (now rolled into the Great Lakes I-Corps Hub) was one of the nine NSF-funded, regional Nodes in the United States. It was comprised of The University of Michigan, University of Illinois, Purdue University, and University of Akron. These research-intensive institutions have significant infrastructure, programming and talent to support entrepreneurship education and technology commercialization activities (Huang-Saad, Duval-Couetil & Park, 2018). The I-Corps Node supports national I-Corps programming, education, and network-building initiatives through the regional activities described below:

Introduction to Customer Discovery (ICD): ICD is a 5-week introduction to the customer discovery methodology designed to get participants to think about their innovations from a market perspective, and learn how to conduct interviews with prospective customers. The course covers the core elements of NSF I-Corps training, including value propositions, customer segments, and understanding whether a technology addresses a specific and significant market need.

Industry Connect (IC): IC events offer the opportunity for university innovators to network with industry leaders. They leverage industry tradeshows and conferences in specific industries, giving research teams the opportunity to introduce themselves, pitch ideas to technology decision makers, arrange follow-on meetings to get feedback on the best paths forward, and identify key milestones.

Workshops: These focus on topics such as writing effective Small Business Innovation Research (SBIR) proposals to participating national funding agencies. They familiarize participants with funding opportunities and offer best practices for applications.

Evaluating Large Scale Education and Training Programs

Large scale education and training programs require an evaluation component to understand impact, which can include changes in student learning over time or other measures of effectiveness. A primary purpose of evaluation is to improve programs through the use of its results. Other reasons are to demonstrate accountability to stakeholders, justify funding, focus attention on critical issues, enhance the image of a program, provide direction for staff and administrators, and to identify training or technical assistance needs.

This manuscript will not cover all aspects of program evaluation since there are many excellent resources available on the topic, however, we will describe some basic standards that have long been considered good practice regardless of the evaluation model chosen. For example, program evaluation research designs should be customized or adapted to each context, and in many cases will require a variety of evaluation criteria to arrive at accurate measures of impact (Mertens & Wilson, 2018). Pertinent stakeholders should be involved in the design of an evaluation protocol to ensure the effective use of results (Rossi et al., 2018). Reasons why programs do not follow through with evaluation activity typically have to do with insufficient funding, resources and/or a lack of staff with expertise to support data collection and evaluation activities.

From a methodological perspective, there are several challenges to evaluation, particularly for large-scale programs. The "gold standard" is the use of randomized controlled trials using quantitative measures and with low levels of sample attrition so that "competing hypotheses can be ruled out and external validity increased" (Frechtling, 2010, p. 32). However, it is recognized that these are often impossible to undertake in real-world settings because they are resource intensive, and unfeasible given challenges associated with identifying equivalent treatment and comparison groups (Mertens & Wilson, 2018; The W.K. Kellogg Foundation, 2017). Even studies that meet lesser standards, including interventions with control and comparison groups, such as quasi-experimental designs with group equivalence can be difficult to undertake in certain settings (What Works Clearinghouse, 2020). Some scholars have even critiqued experimental designs as reductionist, due to the assumption of linear causal relationships between program elements and outcomes (Frye & Hemmer, 2012).

Given these viewpoints, the NSF recommends that evaluations based on experimental design be limited to specific contexts "where the intervention is stable, well-specified, and likely to stay bounded during the evaluation period; where the control or comparison groups are passive, unlikely to seek conditions similar to the treatment; and where the primary interest is estimation of the net value-added of the intervention" (Frechtling, 2010, p. 32). They also suggest using a combination of quantitative and qualitative approaches given, qualitative data may be perceived as more informative by some stakeholders in certain cases (Frechtling, 2010; Shadish, 1993). The NSF's guidelines for evaluators of multi-site programs are outlined in Table 1.

Table 1
NSF evaluation guidelines (source Frechtling, 2010)

	NSF Recommendation
Procedure	1) Development of a conceptual model of the program and identification of key evaluation points, 2) Development of evaluation questions and definition of measurable outcomes, 3) Development of an evaluation design, 4) Collection of data, 5) Analysis of data, 6) Provision of information to interested audiences
Design	Use both quantitative and qualitative data, Use experiments or other controls
Sources of data	Use multiple: Surveys, Interviews, Focus groups, Social Network Analyses, Observations, Tests, Documents, Key Informants, Case Studies
High quality data	Random sampling, Gather both pre and post data, Cause minimal disruption to the program, (Survey) response rate of at least 70%, (Qualitative) Determine inter-rater reliability
Multi-site recommendations	1) Standardize goals, 2) Pretest data collection, 3) Monitor implementation fidelity, 4) Centralize training for data collection, 5) Assess data collection reliability

Entrepreneurship Education Evaluation and Assessment

Researchers have identified many challenges associated with measuring the success or impact of entrepreneurship education. Practical barriers include: a lack of consensus on learning objectives, inconsistent terminology; variations in program models and

approaches to teaching, and the resources necessary to evaluate long term outcomes (Thrane et al., 2016). Methodological barriers include self-selection bias, a lack of comparison groups, a lack of longitudinal data, and a focus on specific psychological constructs such as entrepreneurial self-efficacy, entrepreneurial orientation, and entrepreneurial intention (Rideout, 2012; Yi & Duval-Couetil, 2021). Therefore, scholars have called for greater consensus on the use of conceptual frameworks (Pittaway & Cope, 2007; Wahid et al., 2016), as well as definitions and methods (Klofsten et al., 2019; Miranda et al., 2020). Huang-Saad and colleagues (2018) reviewed 359 research papers on the topic of entrepreneurship education, concluding that validated measures and references to theories and were usually lacking.

Meeting varying stakeholder needs has also been highlighted as a challenge in assessing the impact of entrepreneurship education (Bischoff et al., 2018; Gianiodis & Meek, 2019; Siegel & Wright, 2015). Duval-Couetil (2013) noted that key stakeholders include university administrators, program administrators, students, faculty, university donors, and regional communities, with each group potentially having different questions about entrepreneurship education, implying a need for different types of data collection. For example, data that is necessary for basic program operations, such as tracking enrolment and course evaluations, differ from the type of data necessary for scholarship, which might require experimental designs and validated measures based on theory.

Scholars also note that methods used to assess entrepreneurship education programs within business schools may not translate successfully to those targeting engineers and scientists (Gilmartin et al., 2016; Klofsten et al., 2019; Linton & Xu, 2020; Turner & Gianiodis, 2018). Menold et al. (2015) and Purzer et al. (2016) concluded that the quality of research in this area has been negatively impacted by a lack of consistency in research goals, methods, and data collection instruments. The researchers reviewed 52 assessment instruments from 29 journal articles about engineering entrepreneurship. They concluded that there was an over-reliance on self-report surveys, validity evidence was not provided, and the differences between engineering contexts and other entrepreneurship contexts was often not examined.

I-Corps Program Evaluation Metrics

The issues above highlight the challenges associated with developing more sustainable assessment methods and protocols, connecting theory to assessment practice, and examining programs targeting academic scientists and engineers, specifically. First and foremost, however, is arriving at a consensus on outcomes, which is a significant barrier to conducting program evaluation. In the field of technology commercialization, commercialization metrics are often confounded with education metrics, as can be the case with I-Corps as described below.

Commercialization Metrics

As stated, a primary objective of I-Corps is the translation of laboratory discoveries into technologies that benefit society, which requires the disclosure and protection of intellectual property (IP). Therefore, commercialization metrics such as invention disclosures, patents, licenses and startups are common assessment measures for academic entrepreneurship programs. These metrics are compiled and reported on

annually by the Association of University Technology Managers (AUTM), and universities pay for access to these data, allowing them to compare their own commercialization activity to that of their peers. The official AUTM data mitigates certain issues that can occur when measuring technology commercialization activity. For example, in university settings it can be challenging to define what a "startup" truly is when companies may be legally established but not have salaried employees or revenues. Institutions with vibrant technology transfer activity often use AUTM metrics to promote their role in economic development.

Increasingly, it is recognized that commercialization metrics alone may not accurately reflect all of the entrepreneurial activity occurring at universities since not all campus innovation involves IP. It should also be noted that these metrics are not always a function of campus innovation, but instead a function of an institution's ability to hire technology transfer staff, pay for the cost of securing and maintaining patents, as well as the administration necessary to support commercialization activity. For these reasons, scholars have called for the expansion of entrepreneurship metrics beyond outcomes such as number of new ventures and amount of licensing revenue (Gianiodis & Meek, 2019; Huang-Saad et al., 2017; Klofsten et al., 2019).

Education and Learning Metrics

As I-Corps is fundamentally a training program, it is expected that faculty who are exposed to the curriculum will find it effective and subsequently integrate it into their teaching or training activities at the undergraduate and graduate levels, however, this type of impact is difficult to track and measure. While it is straightforward to count new courses, it can be more difficult to track curriculum enhancements that are confined to specific activities within a course, or that involve the addition of entrepreneurship-related content and materials.

Measuring growth in entrepreneurial experience or skills is also a potential outcome but is multifaceted (Duval-Couetil, 2013; Fayolle, 2005; Rideout and Gray, 2013; Yi & Duval-Couetil, 2021). On one end of the spectrum, this can consist of showing heightened interest in entrepreneurship, which can be implied from course participation numbers or other activities. At the other end of the spectrum, it can consist of counting the number of startups created by participants. Increasingly, however, there is consensus that startup metrics provide an incomplete view of the long-term impact entrepreneurship education can have on students and their careers. For example, generating economic and social value in other ways (e.g., through intrapreneurship or better trained researchers) can be outcomes of entrepreneurship education as well.

Methods to evaluate the impact of entrepreneurship education vary and include measuring changes in knowledge (recall of information), skills (observable competency to perform and act), and competencies (the knowledge, skills, attitudes, values, and behaviors that people need to successfully perform a particular activity or task) (Morris, Webb, Fu, and Singhal, 2013). Entrepreneurial competency comprises abilities and attributes including communication, teamwork, cross-cultural skills, productivity, adaptability, initiative, resilience, self-direction, accountability, and leadership (Boyles, 2012; Morris et al., 2013). More recently, developing an entrepreneurial "mindset" in students is receiving attention. This refers to enhancing cognitive ability by involving creativity and innovation skills, curiosity and risk-taking, and higher order thinking and

reasoning ability (Boyles, 2012; Davis, Hall, and Mayer, 2016; Haynie, Shepherd, Mosakowski, and Earley, 2010; Wheadon and Duval-Couetil, 2017).

I-Corps Program Evaluation

Evaluation and assessment activities within the I-Corps program are subject to the methodological and practical challenges described above, within the context of a multi-site program accountable to a variety of stakeholders. At a national and legislative level, the goal of I-Corps is to catalyze entrepreneurship and economic development through technology ventures that emerge from science being developed at research universities. At an institutional level, I-Corps aligns with university objectives of achieving technology transfer targets to show contributions to economic development, enhancing public relations opportunities, and recruiting the best faculty and students. At an individual level, academic researchers and graduate students are concerned with technology commercialization outcomes such as impacts on career development, research, teaching, tenure and promotion, or quality of life. These individual outcomes are often key to their support for, or motivation to, become involved in technology commercialization activities (Balven et al., 2018; Duval-Couetil, Huang-Saad, et al., 2020).

In a recent proposal to expand I-Corps, NSF announced a desire for more precise evaluation and assessment goals. This involves collecting data to improve local and regional programs, as well as conducting research on entrepreneurship education practices that will generalize nationwide. The solicitation requires that 15% of new program budgets be allocated to the following activities: 1) Gather, analyze, evaluate and use the data and insights resulting from the experiences of those participating in local, regional, and national programs; 2) share and leverage effective innovation practices on a national scale; 2) track long-term, promising deep technology ventures that participate only in regional cohorts; 3) advance scholarship on topics related to national support of entrepreneurial ecosystems and dissemination of best practices; 4) link metrics to the program's strategic objectives, and take action to improve their activities based on data and assessment; 5) revise and improve operations continuously in a culture of thoughtful experimentation, rapid feedback, and rigorous analysis (National Science Foundation, 2020).

National I-Corps Teams Evaluation

At a national level, the original and primary goal of I-Corps was the launch of new startup businesses based on university discovery. Over time, education and training both faculty and students also became a goal. Today, three short-term outcomes of I-Corps training have been defined and include: 1) a clear go/no go decision based on an assessment of the viability of the overall business model, 2) substantial first-hand evidence for or against product-market fit, with a concise definition of the customer segments and corresponding value propositions, and 3) a narrative of a compelling technology demonstration for potential partners (National Science Foundation, 2021). Longer-term objectives of participation include: 1) patent applications; 2) patents granted and derived; 3) licensing agreements; 4) company formation, 5) royalties realized, 6) SBIR/STTR proposal submission; 7) third party financing, 8) enhanced entrepreneurial mindset of NSF-funded researchers, and enhanced career trajectories for team members (National Science Foundation, 2021).

For the seven-week, national I-Corps Teams program, evaluation is conducted by the nonprofit organization VentureWell, which is contracted by the NSF to administer surveys to participants before beginning the course (pre-survey), after completion of the course (post-survey), as well as approximately a year after completing the course (longitudinal survey). Participants are informed that completion of these three surveys is an expectation of enrollment. The survey questions were created using expert interviews and elements drawn from the Kauffman Firm Survey (Robb, Alicia, et al., 2009; VentureWell, 2019) and were pilot tested with subject area experts. For cohorts participating in I-Corps between 2012 and 2019, survey completion rates were 95% for the pre-course survey, 73% for the post-course survey, and 60% for the longitudinal survey, with EL team members being the most difficult to contact for the longitudinal follow-up.

The post-course survey focuses on questions relating to: (1) the extent and types of knowledge gained, (2) satisfaction with the quality of the course and instruction, (3) post-course interests and intentions (e.g., I will apply for SBIR funding for my technology in the next 12 months) and (4) project updates and accomplishments (e.g., how many customer interviews were conducted or whether a patent application was filed). The longitudinal survey, which participants completed an average of 13 months after program completion, focuses on questions related to: (1) company/business creation, such as legal status, (2) ownership and employees, (3) financing, (4) licensing, and (5) revenue. Other questions were developed to examine the individual-level outcomes of interest such as (6) influences of the program on participants' careers, (7) the extent to which concepts learned in the course were still used, (8) the development of new types of collaborations (e.g., with industry or community partners), and (9) the development of entrepreneurial curricula as a potential proxy for institutional influence of the program.

These data are reviewed by I-Corps Teams instructors and NSF program officers, and are discussed quarterly by the I-Corps Teams curriculum committee. I-Corps Nodes and Sites have very limited access to the real-time data as it is shared only with the lead instructor of each cohort (personal communication). Periodic release of these survey data through a public use dataset is managed by VentureWell in coordination with the NSF (VentureWell, 2019).

Node and Site Level Evaluation

Evaluation was loosely organized within and across Nodes and Sites, with little coordination by the NSF or other stakeholders. This is not surprising given that each Node and Site was charged with developing programming that met regional ecosystem and institutional needs. Therefore, each tended to focus on areas of interest to program administrators and/or regional stakeholders. Evaluation was an integral aspect of Node activity, which is not the case for Sites.

At our Node, the research and evaluation team was comprised of two faculty members, a post-doctoral researcher, and two program administrators. The team was responsible for evaluating regional Node activity, as well as engaging in generalizable I-Corps research and scholarship. Data collection activities for program evaluation consisted of distributing pre- and post-course surveys to participants in the regional ICD, SBIR, and

IC programs described above. These instruments were developed internally using elements drawn from surveys used in prior research on entrepreneurship education (Duval-Couetil et al., 2010, 2012, 2021), and validated scales of entrepreneurial self-efficacy (Mcgee et al., 2009). Surveys were distributed using an online survey tool, and participants were given time in the classroom on the first and last days of the course to complete them. These data were regularly discussed by the Node curriculum committee to inform improvements in instruction. In addition, this team has conducted research on the Node ecosystem (Duval-Couetil, Huang-Saad & Park, 2018), faculty experiences in I-Corps (Duval-Couetil, Huang-Saad & Wheadon, 2021), and analyses of the I-Corps Teams public use dataset (Epstein, Duval-Couetil & Huang-Saad, 2021).

METHOD

Given the national importance of I-Corps to the commercialization of investments in academic research, challenges associated with multi-site program evaluation, and the decentralized approach to Node program evaluation activities, this study explored the following research questions:

- What program evaluation and assessment activities were being undertaken by I-Corps Nodes?
- What challenges did evaluation practitioners experience?
- What did evaluation practitioners view as best practices?

Exploring the answers to these questions is a first step in uncovering potential efficiencies and opportunities for collaboration across Nodes and sharing of practices in light of the potential I-Corps expansion and more comprehensive evaluation research requirements.

To achieve our goals, we conducted interviews with individuals familiar with program evaluation activities at each of the nine I-Corps Nodes. This required reaching out to program directors who could identify individuals at each location who were most familiar with I-Corps evaluation activities. This resulted in semi-structured interviews conducted with 15 people; 8 interviewees were male, 7 were female. The job titles reflected the variety of individuals involved in evaluation activity across the Nodes, none of which include the word “evaluator” or “evaluation”: Executive Director (3), Associate Director (1), Research Scientist (1), Assistant Director (1), Project Director (1), Program Director (1), Coordinator (1), Program Coordinator (1), Regional Director (1), Project Associate (1), Assistant Professor (1), Doctoral Student (1).

These exploratory interviews included the following questions: 1) What data does your institution keep track of about the participants in your Node programs? 2) How is this data used internally for assessment? 3) How is this data used for scholarly research? 4) What are the largest challenges your institution faces in collecting data about your I-Corps program participants? 5) Do you feel your institution has best practices to share about data collection and management, either for research, program evaluation, or tracking outcomes?

Overall, I-Corps regional programs were administered through a variety of university entities offering diverse types of programming, illustrating one of the challenges often described by researchers in assessing entrepreneurship education. The programming

varied in format, length, and requirements. Target audiences for regional programs also varied (e.g., deep technology versus broad entrepreneurial community involvement) and whether teams received grant funding for participating in the program. These differences were reflected in the desired outcomes for participants, and difficulty in obtaining standardized evaluation data.

RESULTS

Types of data collected

We asked interviewees: What data about participants in your Node programs does your institution track? Interviewees were asked to specify if they use pre-course surveys, post-course surveys, and if they obtained follow-up or outcome data from any source (Table 2). Out of the nine Nodes, two did not use any program surveys. We also asked whether they would be willing to share these surveys with us. Out of the five that used pre-course surveys, we were able to obtain three. Out of the seven that use post-course surveys, we were able to obtain four. Out of the 4 that used longitudinal surveys, we were able to obtain three. Therefore, in total, 11 survey instruments were analyzed. Our interpretation of why others did not share their surveys was that they felt they were propriety and/or they intended to use them for research. Post-course surveys were used most often, followed by pre-course and longitudinal surveys. Two Nodes followed up on the status of ventures by searching documents and databases on the internet.

Table 2
Summary of Data Sources

Node	**Surveys**			
	Pre-course	**Post-course**	**Longitudinal follow-up**	**Systematic Document Study**
Node 1	No	No	No	Yes
Node 2	Yes	Yes	Yes	No
Node 3	No	Yes	Yes	No
Node 4	Yes	Yes	No	No
Node 5	No	No	No	Yes
Node 6	No	Yes	Yes	No
Node 7	Yes	Yes	No	No
Node 8	Yes	Yes	Yes	No
Node 9	Yes	Yes	No	No
Total	5	7	4	2

Survey topics

Most Nodes developed their own surveys, which focused on collecting pre-program data related to commercialization, education and learning, and demographics. A summary of the types of information included in these instruments is included in Table 3. These varied greatly and included items related to satisfaction with Node programming, usefulness, time commitment, prior experience, interest in starting a business, intent to pursue commercialization activity, as described below.

Commercialization metrics: Out of the 3 longitudinal follow-up surveys obtained, all included commercialization metrics such as founding a company, number of employees, and revenue.

Education and learning metrics: Out of the four post-course surveys obtained, all included questions about participants' satisfaction with the course, as well as their intent to become an entrepreneur. Three included measures of self-rated improvement in knowledge or learning, and three included measures of confidence or self-efficacy. Of the seven Nodes that used post-course surveys, five also collected pre-course data. Out of the three programs for which both pre- and post- surveys were obtained, all included a subset of questions that were consistent on the pre- and post-surveys to allow for the assessment of change over time. Out of the 3 longitudinal surveys obtained, 2 included education and learning metrics such as the perceived value of participating in the program.

Demographic data: Out of the 3 pre-course surveys obtained, all included questions about participants' self-identified gender and race. While all programs collected some form of contact and project information from participants through an application or intake form, most required one per team and therefore self-identified demographic affiliation could not be collected for individual team members. Further, two programs reported that they did not attempt to include demographic data on applications or intake forms due to concerns about participant privacy.

Table 3

Summary of Survey Topics

Node	**Pre-course**	**Post-course**	**Longitudinal follow-up**
Node 1	Not used	Not used	Not used
Node 2	Contact Job title and organization Gender Race Traits of project (4) Interest in starting a business (5) Confidence in starting a business (4) Do you know someone who has started a business? Are there opportunities in your area?	Contact Gender Rate the value of sessions (10) Category of research area Stage of venture Number of interviews completed Hours spent on course Overall satisfaction (2) Open-ended feedback (4) Virtual format (7) Instructor ratings (5) TA rating (4)	Refused

		Interest in starting a business (5) Confidence in starting a business (4) Do you know someone who has started a business? Are there opportunities in your area?	
Node 3	Not used	Job title Satisfaction with course (10) BMC Knowledge before and after (4) Open-ended feedback (5) Usefulness (5) Readiness (10) Intent to continue customer discovery Importance of “soft skills” (5) Would you like further training in these “soft skills” (5)	Team role Is project still active Did you complete national program Readiness (10) Funding received Revenue generated Company founded Traits of company Why is project inactive Have you changed your view on value of the program? Usefulness (13)
Node 4	Refused	Refused	Not used
Node 5	Not used	Not used	Not used
Node 6	Not used	Not received	Contact Gender Race Education Occupation Years of experience Industry Technology maturity Business advisors Readiness (21) Details about business model (8) Details about revenue/finance (8) Details about employees (10) Start-up activities (13)
Node 7	Gender Race What is your career objective? Relevance of I-Corps to career (3) Social contacts who are involved in entrepreneurship Current knowledge rating (9) Intent to commercialize List of contacts who have assisted you	Contact What is your career objective? Relevance of I-Corps to career (3) General satisfaction (5) Specific element satisfaction (6) Perceived learning (11) Intent to commercialize (2) Intent to take follow-up program	Not used

		Did participation influence your research? Career plans? List contacts who have assisted you	
Node 8	Not Received	Not Received	Contact Institution What have you done with this technology since I-Corps? [open] Company founded Company information Are company women/people of color led? SBIR/STTR funding Number of employees Capital raised Revenue
Node 9	Contact Gender Race Commercialization timeline Expected hours worked Entrepreneurial motivation Entrepreneurial self-efficacy	Satisfaction with elements of course Commercialization timeline Actual hours worked Entrepreneurial self-efficacy	Not used

When asked about the source of the survey questions, several used items from the national I-Corps Teams surveys, or ones they developed locally.

> *When we started, I looked at the VentureWell instruments from the national [I-Corps] longitudinal survey. We developed the survey mostly based on regional grant requirements.*

> *We haven't found a lot of existing questionnaires. We revise our tool every year. Evaluation is all about continuous improvement and capturing the impact rather than being generalizable.*

Survey Response Rates

Interviewees were asked to approximate survey response rates. Most described significant challenges related to data collection, low survey response rates, and concerns about survey fatigue. Response rates ranged from 10-90% with higher rates for post-course rather than longitudinal surveys, and for those administered in-person (Table 4). Response rates for surveys taken outside of class time were extremely low, about 30% to 40% for surveys soon after the course and about 20% for longer term longitudinal follow-up surveys.

Table 4
Summary of Survey Response Rates

	Post-course	**Longitudinal follow-up**
Node 1	-	-
Node 2	90+ % (in person)	10% to 15%
Node 3	40% to 50% (at home)	20% to 30%
Node 4	30% to 40% (at home)	-
Node 5	-	-
Node 6	90+ % (in person)	Uncertain - "very difficult"
Node 7	50% to 60% (at home)	-
Node 8	90+ % (in person)	Under 20%
Node 9	90+ % (in person)	-

Several interviewees offered details about factors they felt influenced response rates.

> *All participants are asked to complete but it often ends up being one per team. We ask that surveys are completed within a week of finishing the course. Doing it the day of was too much time to keep people there. Mentors don't respond as much. (off-site, post survey)*
>
> *The response rate is better if we get the surveys out quickly.*
>
> *People don't have enough time to complete these data collection activities. The surveys should be shortened considerably.*
>
> *Responses on the post-course survey totally depend on whether the instructors leave [sufficient] class time. For the longitudinal survey, out of all our entrepreneurship center programs, response rates are probably the worst for the I-Corps regional programs. Possibly because it's so early stage, and they are shorter programs (2-3 weeks) so their feeling of being part of it is less, people don't identify with it. People think it's not relevant if they're early stage - Why am I filling out a survey that doesn't apply to me? Among people who go on to the national program the response rate is better because they're farther along. We want more info about how to get better responses over email.*
>
> *Much higher based on [how recently they were in the] program.*
>
> *For a longitudinal survey - Very difficult - we've tried gift cards, raffles. Response rates are higher if I send the request from my personal email than a blanket email.*

Longitudinal Data

Interviewees were asked, "What follow-up or outcome data do you collect?" Most said they learned about the status of I-Corps teams or ventures through word of mouth or interactions with participants. Tracking progress appeared easier for I-Corps programs housed in well-resourced university entrepreneurship centers or departments. These had more events and follow-up resources for I-Corps teams, thereby allowing administrators to stay in better contact with participants and monitor the outcomes of their projects. Three offered that they systematically kept track of this data in a database. Two Nodes had access to longitudinal survey data that was related to larger institutional entrepreneurship survey projects; at one institution, surveys were periodically sent to all university alumni, and at another, surveys were sent to all participants in entrepreneurship center programs. Some follow-up experiences are below:

> *Some of the instructors connect with certain teams, they have a coworking space for teams to drop in.*
>
> *We have follow-up information on over 50% of participants, but it may come from alternate sources. They work with various groups on campus, a large part of follow-on data comes from venture mentoring service.*
>
> *We host an annual meeting with ecosystem members, funders, and top teams have opportunity to pitch, which helps maintain a regular relationship with them.*
>
> *They don't come to seek future resources from us as much as we would like, they come back if they get lineage and want SBIR assistance.*
>
> *We maintain good relationships with them, so they often tell us the outcome of the project, and we put this information in our database.*
>
> *We help teams find additional sources of funding after the national program and end up with relationships with people to be mentors, we track all of this in Salesforce.*
>
> *We have a CRM, so we track our interactions with all our startups, we have a team of a dozen people whose job is to meet with them. We're working to build out our Salesforce CRM to keep all this information in a centralized place and capture secondary information.*

Data Scraping

Given low response rates, the only programs that appeared to have significant success in tracking longitudinal team outcomes were two that used web searches or data scraping on a regular basis. One had created an extensive automated data scraping system which gathered detailed data on a variety of individual and company outcomes. However, this interviewee reported that the system required an investment of hundreds of thousands of dollars and several years to set up. The other institution made systematic use of data

scraping, the interviewee reported spending about 8 hours twice a year to do online searches for I-Corps participant companies.

> *We use a database of all participants to track startup outcomes like which companies are still active, which get NSF, SBIR, and venture capital funding. Two times a year, we look for publicly available data from NSF and SBIR, Google, LinkedIn, and the paid tier of Pitchbook. Once we find the company is inactive, we stop tracking. We're still tracking 10 to 20% of participants. It took my intern about 8 hours.*

> *We are actively tracking 500 plus startups. We have automated systems with "Webspider" and "Velocity" to scrape public sources of information. It took several years and 300k to set this up - licenses, time - I have a staff of 10 and everyone has [access to] Salesforce. We also know for every faculty member who their students are - if they go through I-Corps later we infer that the faculty had an influence. Outcomes that are automatically tracked include scholarly publications, company incorporation, and patents.*

Data Management

Data being "siloed" within or across institutions was a common complaint. For example, data related to programs in which a person took part was not linked to their demographic data or survey data. Even in programs with frequent follow-up interactions, data on project outcomes that were a priority for many interviewees were tracked only occasionally and informally.

Best Practices

Interviewees were asked about the challenges they faced when collecting data about I-Corps participants. Responses fell into the following categories: a lack of guidance and access to data sharing; data management; difficulty interpreting survey data; and a lack of standardization across locations. Examples are provided below.

Resource constraints	- *I don't have a budget to do longitudinal follow up.* - *Nobody has the time to compile all the data into a big spreadsheet, we have interns to clean the data when we can.* - *We have a pretty small team, no assessment team.* - *There's a disconnect mostly because of manpower to follow up. People would respond to email if we had time to send them out.* - *We used to prepare survey data reports at the end of every session, 25 to 30 cohorts per year. This was too much work, we stopped doing that and now do an annual report.*
Lack of guidance and access to data sharing	- *How seriously NSF has emphasized assessment to us hasn't been that strong and not much guidance. We're not held accountable. I don't know what NSF wants to track.* - *VentureWell has the national program follow-up, but we've asked for this information and don't get it, we don't know why.*

	- *I would want to have certain questions in the surveys, so I had ideas about the actual feedback from the participants. I think that VentureWell should be sending out questions that they recommend.* - *It would be nice to have a clearinghouse for the Nodes where we could put in our questionnaires. But, people don't want to be compared with other institutions, it's competitive, a common evaluation problem. Also, data is like gold, people don't want to share.*
Data management	- *All participants fill out the survey, then I've got to go in and delete duplicates on certain questions that refer to the whole project* - *IRB did not approve race data.* - *I would highly recommend that the research and evaluation team manage the overall data so you have the names and demographics to link over time.* - *The data is siloed so I can't click on one person and see everything they've done.* - *We have federal, state, and internal funding and they all want different reporting.* - *A disconnect is that we have 5 different institutions, so they have to send us the lists and I send the surveys. We rely on individuals to provide the names and emails of the team members, but I don't know what instructors are involved or what trainers.*
Difficulty in interpreting survey data	- *I don't put too much stock on self-assessment about knowledge of the [Business Model Canvas] because when I talk to them in office hours they don't actually understand. How can you verify that they learned something?* - *There's a challenge of self-assessment, decreasing self-efficacy the more that people know. We have a "what are your suggestions" question, but they are too general to be helpful to instructors.* - *We have 15 institutions included and they wanted specific information about their own instructors, but we found they didn't differ much from each other.* - *It's not true that you can directly tie participation in a program to a specific outcome.* - *(As opposed to using survey data) Instructors have so many years of experience they are very effective in judging what's working and what's not.*
Lack of data standardization across locations	- *A lot of the problem with assessment across the Nodes is that there's too many independent variables, the courses are so different from each other. There would have to be more structure.* - *Let's get together to figure out what the right data to collect is, and frequency for longitudinal stuff, there's many research themes that could be supported by that kind of data, that type of data collection could form a valuable dataset that other researchers could use. If someone ever wanted to compare how Nodes are doing, it would be nice if you had the same dataset to look at collected in pretty much the same way. It would attract other researchers about entrepreneurship, they could get the same data from multiple locations.*

Some interviewees noted other issues that did not fit neatly into the categories above.

> *What does it mean to be more entrepreneurial? It would be interesting to compare different program structures. The culture change aspect is critical in at the academic/lab level [varies in] in different parts of the university. Can you look at I-Corps PIs [by unit] and whether they have more IP?*
>
> *I care about outcomes and entrepreneurial mindset a lot; we have many conversations about this. I really want to focus on the culture shift and the mindset shift from I-Corps. We use questions around risk and confidence, how prepared are you to launch, etc., have they incorporated concepts into teaching, mentoring, etc.*
>
> *Starting a company is one of the routes. But other outcomes are [as well]. [For example], being licensed to another company, does it positively impact how you do your research, does it positively impact how you approach your career somehow?*
>
> *I wish we knew how it was valuable to the team, SBIR acceptance rate, funding that teams get, how many startups have been founded and how many succeed past a year or two years.*
>
> *The surveys are really narrow in scope, that's not the only measure of success.*
>
> *It would be great to have the instructors evaluate the individual teams with a readiness rubric, and evaluate if they are consistent with how the teams evaluate themselves.*
>
> *We'd like to do surveys with the instructors, asking how confident do they feel with the curriculum and how authentic are the examples to them. I'd like to know what kind of resources or support they might need.*
> *A very prevalent group is international students - what are the barriers they experience? Such as discrimination towards accents, and language barriers?*

General Best Practices

Interviewees were asked about their own best practices that could benefit other Nodes.

> *We've created more standardization between programs which has helped.*
>
> *We do a survey at the beginning, because always at the end people want to know how much growth do we have.*
>
> *My biggest recommendation is to do the survey first. Focus on what you want to learn without burning people out on the survey, [approximately] 10 to 20 minutes long.*

We have the survey come from the lead instructor of [the particular] cohort – [since] this is about the class you took with this person.

Collecting data in a centralized way.

Do the survey at the end of the session, at that point we basically get a 100% response rate, so that's something we definitely recommend.

Having more touch points with participants; a workshop where people meet investors was the carrot to get them to participate.

A substantial investment in a CRM, and managing the relationships with the people you encounter is paramount.

Some pointed to external resources, not supported through I-Corps, that were helpful.

Because we [run]so many programs, we have a marketing team, programming team, and events team.

I try to ensure we have standard data across all of these programs, we include the fact we have this standard in every new application and proposal.

I-Corps is inside [name of entrepreneurship center] and benefits from their other activities.

Summary of Assessment Trade-offs

There are many approaches to assessment across the Nodes. In reviewing best practices, resources and practicality, below we summarize tradeoffs that Nodes make when approaching assessment (Table 5).

Table 5
Summary of Assessment Trade-offs

Issue/Topic	**Potential Benefits**	**Costs/Concerns**
Validated measurement scales and surveys	Generalizability	Not sensitive to context
Qualitative data	Provides more in-depth information	Time consuming to collect in a rigorous manner
Administering both pre- and post-surveys	Necessary to track changes resulting from an intervention	Differing program lengths Some programs are too short Pre-post attrition
Giving pre- and post-survey in person instead of at home or off-site	Improves response rate (from 60% to near 100%)	Complaints about length of time it takes to complete Not all participants attend all sessions Participants unlikely to complete "unless instructor reinforces the survey"

Longitudinal data	Infers causation Offers a larger sample size	Low survey response rates (<30%) Requires resources and systems to collect data Funding timelines do not align with medium and long-term outcomes
Online data scraping	Timely information about venture status	Requires regular attention Comprehensive, tailored systems can take years to create
Linking multiple data sources at institutions (e.g., to get demographics, research activity)	Valuable to research	Requires more stringent IRB to publish
Standardization of assessment and curriculum across sites	Enhances generalizability	Not sensitive to context
Include control or comparison group	Infers causation	Difficult to recruit and match a comparable group Self-selection bias

DISCUSSION

The purpose of this study was to examine I-Corps in light of the rapid growth of programs designed to catalyze campus involvement in technology commercialization and entrepreneurship, and calls for improvements and guidance in assessing the effectiveness of such programs (Gianiodis & Meek, 2020; Huang-Saad et al., 2018; Klofsten et al., 2019; Linton & Xu, 2020). We show that significant variations in structure, content, desired outcomes, and audiences for these programs continue to make the evaluation of these programs difficult. Nevertheless, as a national entrepreneurship training initiative that uses a standardized curriculum, and enrolls a large and fairly homogeneous population of graduate students and faculty members, I-Corps offers a foundation for examining evaluation challenges, while exploring potential best practices to be considered in its upcoming expansion.

While some program assessment is conducted at the national level through quantitative survey data collected from individual I-Corps Team participants, prior to this study we knew little about the outcomes of the training and evaluation activities that take place at the Node level, which served as a pipeline to the national Teams program, and where significant resources are being dedicated to offering faculty, students and mentors a variety of educational activities and experiences. Node level evaluation data could contribute to the success of I-Corps as it can help answer questions such as: Are we attracting the right faculty and students to the program? Is the programming we are delivering generating interest and activity in technology commercialization and entrepreneurship? How well is it preparing them for the I-Corps national program? Are the participants satisfied with the curriculum and instructors? To what extent are Nodes synergistic with, or competing with other university resources? And, what changes

could be made to Node programming to increase the number of participants and enhance their success in entrepreneurship and commercialization activities?

Through interviews with I-Corps administrators familiar with data collection and assessment activities at each Node, we sought to identify the most significant assessment challenges for which programs seek guidance on best practices, and those likely to require additional resources and support to achieve evaluation objectives. Interestingly, none of the individuals we interviewed had the term “evaluation” or “evaluator” in their titles. However, all were implicated in Node evaluation activities to some extent since this is an important aspect of annual reporting to NSF as well as university stakeholders.

As we point out, the NSF has compiled guidelines to be used in program evaluations, including recommended overall evaluation procedures, study designs, and types of data (Frechtling, 2010). However, our research shows that these guidelines are pragmatically difficult to adhere to given the differing priorities and the evaluation tradeoffs that we describe, and the resources allocated to evaluation activity at each Node. The guidelines suggest using randomized controlled trials whenever possible, and using comparison groups or quantitative controls when this is not possible (Frechtling, 2010). Evaluation specialists acknowledged that these were not appropriate and practical in all evaluation settings, and we found that these research designs are not systematically included in any Node-level assessment activity. This is not surprising given the challenges associated with identifying, recruiting, and collecting data from a matched comparison group of faculty and graduate students, who are involved in developing innovative technologies, but who are not taking part in I-Corps.

Instead, most Nodes relied on survey data to measure satisfaction, knowledge acquisition, and commercialization activity related to I-Corps programming. However, this wasn’t without challenges as interviewees reported significant concerns about their reliance on surveys including the limitations of self-assessment and the fact that there were few validated instruments pertinent to I-Corps from which to draw. Nodes requiring participants to complete pre-course and post-course surveys in-person and during class sessions reported very high response rates, although this number was not 100% because not all team members attend all course sessions. Interestingly, some Nodes were unwilling to have surveying disrupt valuable instruction time. Instead, they asked that participants complete surveys outside of class time, resulting in significantly lower response rates. Nodes that used longitudinal follow-up surveys to track participants reported very low response rates of around 20%. One interviewee reported using incentives to encourage longitudinal survey completion, which had minimal success. A barrier to using this method is that faculty, graduate students, and entrepreneurs are busy people who tend not prioritize voluntary surveys. Guidelines state that obtaining at least a 70% response rates is important when using surveys, otherwise, results based on these samples may skewed and not representative of the population being examined.

Interestingly, most Nodes felt the data from post-course and longitudinal surveys was valuable for program marketing. These data showed satisfaction with I-Corps programming which helped attract prospective faculty, graduate students and industry mentors. This indicates a potential area of overlap in the staffing needed for assessment as well as marketing and recruiting purposes.

Evaluation guidelines also recommend using mixed method designs, including both quantitative and qualitative data. Best practices suggest triangulation of data from multiple sources such as surveys, interviews, focus groups, social network analyses, observations, tests, documents, key informants, and case studies. While no Nodes reported regular use of interviews or focus groups, qualitative data collected from open-ended responses on post-course surveys was often cited as a helpful resource for program improvement. Use of interviews was reported occasionally, but not on an ongoing basis. No Nodes reported using inter-rater reliability when analyzing open-ended questions. Tracking information about companies and venture status through Internet searches and online databases, was considered very helpful and a potential best practice to be disseminated to other programs. Notably, assessments of quality of instruction or implementation fidelity using observations were not reported by interviewees. Similarly, no Nodes directly assessed skills using tests.

From the interviews, it is clear that it is difficult to implement and execute evaluation activities at one site, let alone across multiple sites. Guidelines for multi-site programs include completing the following tasks across locations: 1) standardize goals, 2) pretest data collection, 3) monitor implementation fidelity, 4) centralize training for data collection, and 5) assess data collection reliability (Frechtling, 2010). In theory, given the structure of I-Corps programming, these activities could be contemplated given how each Node relates to the national program. However, interviewees reported limited coordination of program goals and data collection guidelines across Nodes. Data collection activities at each institution aligned more closely with the local expertise and organizational resources available at each institution. To some degree this is desired to capture the impact of activities within or import to the local ecosystem, however, it limits the ability to generalize findings.

Putting a sustainable evaluation program into place to track and improve program outcomes requires considerable effort that often is not recognized. The literature shows that that there can be big gaps in how program evaluation should be undertaken and how it is actually implemented, often due to a lack of both knowledge and resources. In this study, we found the primary barriers to conducting evaluation and assessment activities were a lack of time or expertise among I-Corps program staff. Given these staffing and funding constraints, several programs leveraged entities such as the entrepreneurship center or university of which they were a part for help with evaluation activities. Ultimately, the goal of I-Corps and entrepreneurship education programs, more generally, is to generate more interest and involvement in venture creation and economic development. For programs with such outcomes, funding timelines often do not align with the resources necessary to track medium- or long-term impacts.

As we showed, defining outcomes and measuring the impact of a national initiative is complex for many methodological and practical reasons. Nevertheless, I-Corps offers a rare opportunity to evaluate educational interventions and outcomes given its curriculum that is largely standardized across Nodes and institutions across the nation. Drawing on the experiences of I-Corps program administrators and coordinators across the nine regional Nodes, we propose that some of these challenges can be overcome with more resources and coordination. In particular, sharing instruments and evaluation protocols could enhance evaluation and assessment activities, leading to more generalizable data. There are also opportunities to connect theory and research in

education and psychology to the assessment of entrepreneurship education, generally, and to programs targetting engineers and scientists, specifically. Our research is particularly timely data given the proposed expansion of I-Corps. These findings and best practices should be considered by program administrators as they contemplate existing evaluation activities, or the implementation of new ones.

CONCLUSION

Evaluating varying entrepreneurship programs across multiple sites is complex for several practical and methodological reasons that we outline in this manuscript. Nevertheless, as a national entrepreneurship training initiative, I-Corps also offers a foundation for examining the impacts of entrepreneurship education at a very large scale. Highlighting the importance of program evaluation and sharing the experiences of current I-Corps administrators is a first step to more collaboration on this important topic in the future.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation under Grant NSF-IIP-1643280.

REFERENCES

Abreu, M., Demirel, P., Grinevich, V., & Karataş-Özkan, M. (2016). Entrepreneurial practices in research-intensive and teaching-led universities. *Small Business Economics*, *47*(3), 695–717.

Balven, R., Fenters, V., Siegel, D. S., & Waldman, D. (2018). Academic entrepreneurship: The roles of identity, motivation, championing, education, work-life balance, and organizational justice. *Academy of Management Perspectives*, *32*(1), 21–42. https://doi.org/10.5465/amp.2016.0127

Bischoff, K., Volkmann, C. K., & Audretsch, D. B. (2018). Stakeholder collaboration in entrepreneurship education: an analysis of the entrepreneurial ecosystems of European higher educational institutions. *Journal of Technology Transfer*, *43*(1), 20–46. https://doi.org/10.1007/s10961-017-9581-0

Blank, S., & Dorf, B. (2012). *The Startup Owner's Manual Vol. 1: The Step-by-Step Guide for Building a Great Company*. K&S Ranch, Inc.

Boh, W. F., De-Haan, U., & Strom, R. (2016). University technology transfer through entrepreneurship: faculty and students in spinoffs. *Journal of Technology Transfer*, *41*(4), 661–669. https://doi.org/10.1007/s10961-015-9399-6

Calcagnini, G., Favaretto, I., Giombini, G., Perugini, F., & Rombaldoni, R. (2016). The role of universities in the location of innovative start-ups. *The Journal of Technology Transfer*, *41*(4), 670–693.

Cesaroni, F., & Piccaluga, A. (2016). The activities of university knowledge transfer

offices: towards the third mission in Italy. *The Journal of Technology Transfer*, *41*(4), 753–777.

Chang, Y.-C., Yang, P. Y., Martin, B. R., Chi, H.-R., & Tsai-Lin, T.-F. (2016). Entrepreneurial universities and research ambidexterity: A multilevel analysis. *Technovation*, *54*, 7–21.

Duval-Couetil, N. (2013). Assessing the impact of entrepreneurship education programs: Challenges and approaches. *Journal of Small Business Management*, *51*(3), 394–409. https://doi.org/10.1111/jsbm.12024

Duval-Couetil, N., Huang-Saad, A., & Wheadon, M. (2020). Training Faculty in Entrepreneurship and Innovation: An Evaluation of the National Science Foundation Innovation-Corps™ Program. *Entrepreneurship Education and Pedagogy*, 251512742092938. https://doi.org/10.1177/2515127420929383

Duval-Couetil, N., Ladisch, M., & Yi, S. (2021). Addressing academic researcher priorities through science and technology entrepreneurship education. *Journal of Technology Transfer*, 1–31. https://doi.org/10.1007/s10961-020-09787-5

Duval-Couetil, N., Reed-Rhoads, T., & Haghighi, S. (2012). Engineering Students and Entrepreneurship Education: Involvement, Attitudes and Outcomes. *International Journal of Engineering Education*, *28*(2), 425–435. https://www.researchgate.net/publication/286953529

Duval-Couetil, N, Reed-Rhoads, T. and Haghighi, S. (2010), "Development of an assessment instrument to examine outcomes of entrepreneurship education on engineering students," 2010 *IEEE Frontiers in Education Conference* (FIE), 2010, pp. T4D-1-T4D-6, doi: 10.1109/FIE.2010.5673411.

Epstein, A., Huang-Saad, A. and Duval-Couetil, N. (2021). Faculty and Trainee Perceptions of NSF I-Corps Technology Commercialization Training, in *IEEE Engineering Management Review*, vol. 49, no. 1, pp. 97-105, doi: 10.1109/EMR.2020.3036280.

Epstein, A., Duval-Couetil, N. and Huang-Saad, A. (2022), Gender differences in academic entrepreneurship: experience, attitudes and outcomes among NSF I-CORPS participants, *International Journal of Gender and Entrepreneurship*, Vol. 14 No. 1, pp. 117-141. https://doi.org/10.1108/IJGE-10-2020-0166

Etzkowitz, H. (2016). The entrepreneurial university: vision and metrics. *Industry and Higher Education*, *30*(2), 83–97.

Fitzgerald, C., & Cunningham, J. A. (2016). Inside the university technology transfer office: mission statement analysis. *The Journal of Technology Transfer*, *41*(5), 1235–1246.

Frechtling, J. (2010). *The 2010 User-Friendly Handbook for Project Evaluation*. The National Science Foundation Directorate for Education and Human Resources.

Frye, A. W., & Hemmer, P. A. (2012). Program evaluation models and related theories: AMEE Guide No. 67. In *Medical Teacher* (Vol. 34, Issue 5). Med Teach. https://doi.org/10.3109/0142159X.2012.668637

Gianiodis, P. T., & Meek, W. R. (2019). Entrepreneurial education for the entrepreneurial university: a stakeholder perspective. *Journal of Technology Transfer*, 1–29. https://doi.org/10.1007/s10961-019-09742-z

Gianiodis, P. T., & Meek, W. R. (2020). Entrepreneurial education for the entrepreneurial university: a stakeholder perspective. *Journal of Technology Transfer*, *45*(4), 1167–1195. https://doi.org/10.1007/s10961-019-09742-z

Gilmartin, S. K., Chen, H. L., & Estrada, C. (2016). Investigating Entrepreneurship Program Models in Undergraduate Engineering Education. *International Journal of Engineering Education*, *32*(5), 2048–2065.

Goethner, M., & Wyrwich, M. (2020). Cross-faculty proximity and academic entrepreneurship: the role of business schools. *Journal of Technology Transfer*, *45*(4), 1016–1062. https://doi.org/10.1007/s10961-019-09725-0

Guerrero, M., Cunningham, J. A., & Urbano, D. (2015). Economic impact of entrepreneurial universities' activities: An exploratory study of the United Kingdom. *Research Policy*, *44*(3), 748–764. https://doi.org/10.1016/j.respol.2014.10.008

Guerrero, M., Urbano, D., & Fayolle, A. (2016). Entrepreneurial activity and regional competitiveness: Evidence from European entrepreneurial universities. *Journal of Technology Transfer*, *41*(1), 105–131. https://doi.org/10.1007/s10961-014-9377-4

Guo, F., Restubog, S. L. D., Cui, L., Zou, B., & Choi, Y. (2019). What determines the entrepreneurial success of academics? Navigating multiple social identities in the hybrid career of academic entrepreneurs. *Journal of Vocational Behavior*, *112*, 241–254. https://doi.org/10.1016/j.jvb.2019.03.003

Hayter, C. S., Link, A. N., & Scott, J. T. (2018). Public-sector entrepreneurship. *Oxford Review of Economic Policy*, *34*(4), 676–694.

Hayter, C. S., Nelson, A. J., Zayed, S., & O'Connor, A. C. (2018). Conceptualizing academic entrepreneurship ecosystems: a review, analysis and extension of the literature. *Journal of Technology Transfer*, *43*(4), 1039–1082. https://doi.org/10.1007/s10961-018-9657-5

Huang-Saad, A., Duval-Couetil, N. and Park, J. (2018), "Technology and talent: capturing the role of universities in regional entrepreneurial ecosystems", *Journal of Enterprising Communities: People and Places in the Global Economy*, Vol. 12 No. 2, pp. 92-116. https://doi.org/10.1108/JEC-08-2017-0070

Huang-Saad, A., Fay, J., & Sheridan, L. (2017). Closing the divide: accelerating technology commercialization by catalyzing the university entrepreneurial ecosystem with I-Corps™. *Journal of Technology Transfer*, *42*(6).

https://doi.org/10.1007/s10961-016-9531-2

Huang-Saad, A., Morton, C., & Libarkin, J. (2018). Entrepreneurship in Higher Education: A Research Review for Engineering Education Researchers. *Journal of Engineering Education*, *107*(2), 263–290.

Huyghe, A., & Knockaert, M. (2015). The influence of organizational culture and climate on entrepreneurial intentions among research scientists. *The Journal of Technology Transfer*, *40*(1), 138–160.

Klofsten, M., Fayolle, A., Guerrero, M., Mian, S., Urbano, D., & Wright, M. (2019). The entrepreneurial university as driver for economic growth and social change - Key strategic challenges. *Technological Forecasting and Social Change*, *141*, 149–158. https://doi.org/10.1016/j.techfore.2018.12.004

Linton, J. D., & Xu, W. (2020). Research on science and technological entrepreneurship education: What needs to happen next? *Journal of Technology Transfer*, *46*(2), 393–406. https://doi.org/10.1007/s10961-020-09786-6

Liu, W. (2018). Lene Foss and David V. Gibson (eds.): the entrepreneurial university: context and institutional change, Routledge, London, 2015, 285pp. *Higher Education*, *75*(4). https://doi.org/10.1007/s10734-017-0163-8

Maritz, A., Koch, A., & Schmidt, M. (2016). The role of entrepreneurship education programs in national systems of entrepreneurship and entrepreneurship ecosystems. *The International Journal of Organizational Innovation*, *8*, 1–7.

Mcgee, J. E., Peterson, M., Mueller, S. L., & Sequeira, J. M. (2009). Entrepreneurial Self-Efficacy: Refining the Measure. *Journals.Sagepub.Com*, *33*(4), 965–988. https://doi.org/10.1111/j.1540-6520.2009.00304.x

Mertens, D. M., & Wilson, A. T. (2018). *Program evaluation theory and practice*. Guilford Publications.

Miranda, C., Goñi, J., Berhane, B., & Carberry, A. (2020). Seven Challenges in Conceptualizing and Assessing Entrepreneurial Skills or Mindsets in Engineering Entrepreneurship Education. *Education Sciences*, *10*(11), 309. https://doi.org/10.3390/educsci10110309

National Science Foundation. (2016, September). *Factsheet: I-Corps Nodes*. https://www.nsf.gov/news/special_reports/i-corps/pdf/factsheet_nodes.pdf

National Science Foundation. (2019). *2019 NSF Innovation Corps (I-Corps™) biennial report*. https://www.nsf.gov/news/special_reports/i-corps/pdf/I-CorpsReport--6_4_19FINAL_508.pdf

Nelson, A. J., & Monsen, E. (2014). Teaching technology commercialization: introduction to the special section. *The Journal of Technology Transfer*, *39*(5), 774–779. https://doi.org/10.1007/s10961-014-9341-3

Nnakwe, C. C., Cooch, N., & Huang-Saad, A. (2018). Investing in academic technology innovation and entrepreneurship: Moving beyond research funding through the NSF I-CORPS ™ program. *Technology & Innovation*, *19*(4), 773–786. https://doi.org/10.21300/19.4.2018.773

Pittaway, L., & Cope, J. (2007). Entrepreneurship education: a systematic review of the evidence. *International Small Business Journal*, *25*(5), 479–510.

Purzer, Ş., Fila, N., & Nataraja, K. (2016). Evaluation of current assessment methods in engineering entrepreneurship education. *Advances in Engineering Education*, *5*(1). https://eric.ed.gov/?id=EJ1090526

Rasmussen, E., & Wright, M. (2015). How can universities facilitate academic spin-offs? An entrepreneurial competency perspective. *Journal of Technology Transfer*, *40*(5), 782–799. https://doi.org/10.1007/s10961-014-9386-3

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach*. Sage publications.

Sanberg, P. R., Gharib, M., Harker, P. T., Kaler, E. W., Marchase, R. B., Sands, T. D., Arshadi, N., & Sarkar, S. (2014). Changing the academic culture: Valuing patents and commercialization toward tenure and career advancement. *Proceedings of the National Academy of Sciences*, *111*(18), 6542–6547.

Siegel, D. S., & Wright, M. (2015). Academic Entrepreneurship: Time for a Rethink? *British Journal of Management*, *26*(4), 582–595. https://doi.org/10.1111/1467-8551.12116

The W.K. Kellogg Foundation. (2017). *The Step-by-Step Guide to Evaluation*. The W.K. Kellogg Foundation. https://www.wkkf.org/resource-directory/resources/2017/11/the-step-by-step-guide-to-evaluation--how-to-become-savvy-evaluation-consumers

Thrane, C., Blenker, P., Korsgaard, S., & Neergaard, H. (2016). The promise of entrepreneurship education: Reconceptualizing the individual–opportunity nexus as a conceptual framework for entrepreneurship education. *International Small Business Journal*, *34*(7), 905–924.

Turner, T., & Gianiodis, P. (2018). Entrepreneurship Unleashed: Understanding Entrepreneurial Education outside of the Business School. *Journal of Small Business Management*, *56*(1), 131–149. https://doi.org/10.1111/jsbm.12365

VentureWell. (2019, August 28). *Impact of NSF's I-Corps™ national program on individual participants, data file and codebook: Release 2.1*. VentureWell. https://go.venturewell.org/PublicUseData2

Wahid, A., Ibrahim, A., & Hashim, N. B. (2016). The review of teaching and learning on entrepreneurship education in institution of higher learning. *Journal on Technical and Vocational Education*, *1*(2), 82–88.

Welter, F., Smallbone, D., & Pobol, A. (2015). Entrepreneurial activity in the informal economy: a missing piece of the entrepreneurship jigsaw puzzle. *Entrepreneurship & Regional Development*, *27*(5–6), 292–306.

What Works Clearinghouse. (2020). *Standards Handbook, Version 4.1*. https://ies.ed.gov/ncee/wwc/Handbooks

Wright, M., & Phan, P. (2018). The commercialization of science: From determinants to impact. *Academy of Management Perspectives*, *32*(1), 1–3.

Wright, M., Piva, E., Mosey, S., & Lockett, A. (2009). Academic entrepreneurship and business schools. *Journal of Technology Transfer*, *34*(6), 560–587. https://doi.org/10.1007/s10961-009-9128-0

Yi, S., & Duval-Couetil, N. (2021). Standards for evaluating impact in entrepreneurship education research: Using a descriptive validity framework to enhance methodological rigor and transparency. *Entrepreneurship Theory and Practice*, 10422587211018184.