2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #34351

Education and Evaluation for the NRT: Accounting for Numerous Requirements, Multiple Disciplines, and Small Cohorts

Dr. Nathalie Duval-Couetil, Purdue University at West Lafayette

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. She is responsible for the launch and development of the university's multidisciplinary undergraduate entrepreneurship program, which involves 2000 students from all majors per year. She has established entrepreneurship capstone, global entrepreneurship, and women and leadership courses and initiatives at both the undergraduate and graduate levels. Prior to her work in academia, Nathalie spent several years in the field of market research and business strategy consulting in Europe and the United States with Booz Allen and Hamilton and Data and Strategies Group. She received a BA from the University of Massachusetts at Amherst, an MBA from Babson College, and MS and PhD degrees from Purdue University.

Dr. Soohyun Yi, Texas Tech University

Soohyun Yi is an Assistant Professor of Educational Psychology at Texas Tech University. She received her Ph.D. in Educational Studies from Purdue University and M.A. in Psychological Measurement from Ewha Womans University. Her scholarly trajectory aims to improve education for underserved and underchallenged students with impactful research and evidence-based interventions. Longitudinal research methodology is the main area of her expertise, which has enabled her a) to investigate growth trajectories of motivation and career choices; b) to identify opportunity gaps within underserved groups; and c) to evaluate and improve educational interventions in STEM. With the expertise in quantitative research methodology, she is engaged in collaborative research with entrepreneurship education and other interdisciplinary programs.

©American Society for Engineering Education, 2021

Education and evaluation for the NRT:

Accounting for numerous requirements, multiple disciplines, and small cohorts

Education and evaluation for the NRT: Accounting for numerous requirements, multiple disciplines, and small cohorts

ABSTRACT

The purpose of this paper is to describe our approach to program evaluation for the National Science Foundation National (NSF) Research Traineeship (NRT) Program in Sustainable Food, Energy, and Water Systems at Purdue University. The NRT program is designed to educate and train the next generation of engineers, agronomists and scientists to meet local food, energy and water management needs with solar energy. Assessment and evaluation activities are important aspects of this work, but require significant attention to capture the range of activities undertaken by very small cohorts of interdisciplinary students and faculty. Our goal was to develop a "sustainable" evaluation plan given our observation that programs often begin with very ambitious assessment and data collection goals that diminish over time. This paper is a case study that describes the rationale for our assessment and evaluation choices, and select results from these activities.

INTRODUCTION

The National Science Foundation's Research Traineeship Program (NRT) supports university efforts to explore ways to equip master's and doctoral degree students with the skills, knowledge and competencies required to pursue STEM careers in academic, industry or government settings. The program focuses on interdisciplinary scientific areas that are considered a high national priority, and within them, education and training models that are considered innovative, evidence-based, and aligned with workforce and research needs. In doing so, the program emphasizes broad and diverse participation from underrepresented groups; seeks catalyzes institutional capacity building; and encourages strategic collaboration with industry, national research labs, and academic partners.

In NSF's portfolio of graduate training programs, that NRT Program took the place of the Integrative Graduate Education and Research Traineeship Program (IGERT). For many years, the IGERT Program also focused on the graduate education and preparing students to be leaders in science and engineering disciplines (Newswander & Borrego, 2021). Like the NRT, it emphasized collaborative research that transcended traditional disciplinary boundaries and was founded on the belief that diversity among participants contributed to their ability to solve "large and complex problems of significant scientific and societal importance at the national and international level" (NSF http://www.igert.org/public/about.html). From its inception in 1998, the IGERT program made 278 awards and funded approximately 6500 graduate students. The last call for proposals for the IGERT took place in 2013.

Assessment and evaluation activities, designed to measure impacts on training and career development, are important aspects of these graduate training programs, but require significant attention to capture the range of activities undertaken by very small cohorts of interdisciplinary students and faculty. Our goal was to develop "sustainable" evaluation activities given our observation that programs often begin with very ambitious assessment and data collection goals in year one that diminish over time, often due to the limited resources allocated to support these efforts. Further, tracking individual graduate student participation in various NRT program requirements can be labor intensive and is often done manually (e.g., combinations of disciplinary courses, interdisciplinary courses, professional development modules, and scholarly activities) and we sought to streamline these activities to reduce the administrative burden.

This paper is a case study that describes our approach, while also highlighting challenges associated with choosing evaluation methods that could be sustained over the grant period. The unique aspects of our plan were the creation of a professional development tracking “hub” within our university’s learning management system (LMS) and the development and dissemination of an annual survey to track various learning and career outcomes over time. These activities supplemented more traditional methods such as interdisciplinary course evaluations and annual interviews and focus groups conducted by an external evaluator. Below, we report on our rationale for, and some results of, these education and evaluation activities.

BACKGROUND

The NRT Program at Purdue University

The Purdue Sustainable Food Energy and Water System (SFEWS) NRT is a collaboration of agricultural scientists and practitioners, engineers, and economists, from several department and across two US academic institutions. External collaborators include two US national labs, three companies, and two international institutions. These individuals and entities were brought together to form an interdisciplinary traineeship program designed to develop graduate students into leaders able to find solutions related to sustaining food supplies, energy, and water (FEW) given that in two to three generations, the global population will grow to more than ten billion people. Since solar energy is the only locally available resource that can meet the entirety of humanity’s demands, the scientific goal of our NRT is to conduct solar energy research that will not compete with agricultural land.

This research is led by faculty from multiple disciplines including chemical engineering, electrical engineering, agronomy, agricultural and biological engineering, biological systems engineering, civil and environmental engineering, and agricultural economics. The interdisciplinary research being conducted explores novel photovoltaic arrangements, structures, and materials such that food, energy and water demands can be met simultaneously without any detriment to crop output. Further, given that these scientific solutions must be achieved in an economically viable way to be implemented and sustained, students are also exposed to courses focused on economic analysis and technology commercialization. To foster awareness of the value of diversity to interdisciplinary research and collaboration, they also participate in leadership and diversity training. The overarching goals of the NRT at the graduate student level are the following:

- Goal 1: Develop a program to equip students with an interdisciplinary mindset/skills
- Goal 2: Train students to generate SFEWS solutions
- Goal 3: Disseminate courses, modules, methods and tools for use among other institutions

The objectives are to instill in graduate students the skills necessary to be a team member, leader, mentor, communicator, teacher and entrepreneur. This is achieved through the large number and wide range of education, training and professional development components listed in Table 1.

Table 1. NRT Program Elements

Departmental requirements	“Core” master’s and doctoral requirements in disciplinary departments
New required interdisciplinary courses	Four courses developed or modified for the NRT: – Food and Energy Farms: Challenges to Sustainable Production on a Crowded Planet – System, Economics, and Supply Chain Analysis for Food, Energy and Water Production and Distribution – Discovery, Development and Translation in Food, Energy, and Water – Leadership and Diversity in Science and Technology
Elective courses within and across disciplinary departments	Students must take 2 courses that complement the NRT activities. Examples include: – Educational Methods in Engineering

	- Economic Benefit Cost Analysis - Intellectual Property for Engineers - Crop Physiology and Ecology
Professional development modules	30 professional development workshops offered through various entities at the university that align with progression through a doctoral program. Examples of the modules include: - Year 1: time management, success in graduate school - Year 2: responsible conduct of research, data management, teamwork, ethics, mentoring, oral presentations, writing conference abstracts - Year 3: writing academic papers, effective graphics for presentations, networking, responding to reviews, having difficult conversations - Year 4: dissertation and thesis writing strategies, preparing for a job search, writing a CV, interviewing, negotiating a job offer
Required research training modules	11 research modules are required. Examples include: - Create course modules (K-12, undergraduate) - Deliver classroom lectures on NRT topic - Define an undergraduate research project and mentor students - Present research at external conference - Interact with researchers at national labs, companies and at other universities - Take on role of lab safety manager and lab manager
Apply – Create – Translate (ACT) Modules	5 activities aligned with the research and goals of the program. Examples include: - Participate in STEM program for underrepresented groups - Participate in K-12 outreach activities - Technology commercialization activity through university programs or NSF I-Corps

To achieve these goals, graduate students are integrated into interdisciplinary research teams, where they actively and collaboratively work on important scientific and societal challenges in a rigorous manner. Our vision was to deliver this program in an educationally rigorous manner over the four to five-year period of a typical PhD student program, starting with broad exposure to concepts and fundamental skills, and culminating in the creation of novel solar energy solutions that are viable in real-world settings as conceptualized in Figure 1 (Anderson, Krathwohl, & Bloom, 2005).

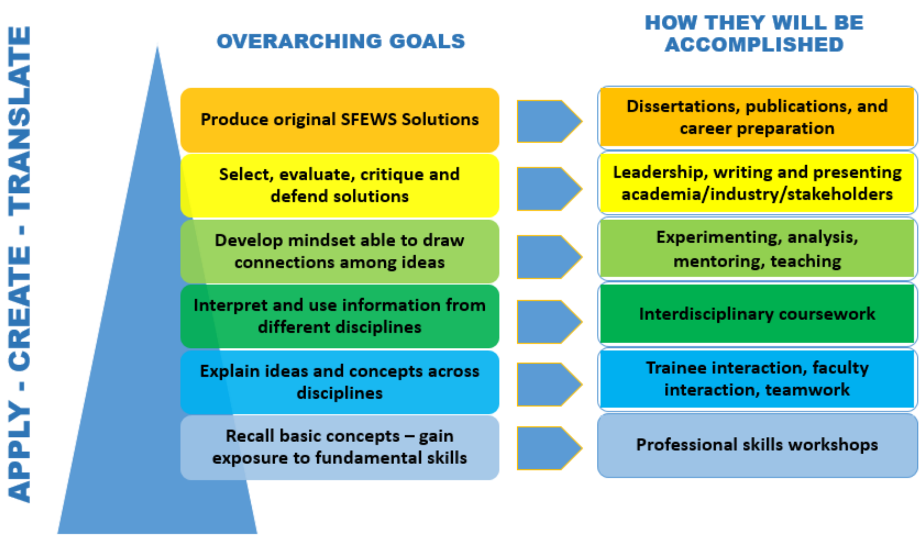


Figure 1. Aligning NRT Goals with Bloom's Taxonomy

Evaluation, Assessment and the NRT

Achieving the goal of solving large and complex problems of scientific and societal impact requires that graduate students and faculty involved in interdisciplinary research acquire the knowledge and competency that will facilitate this outcome. In addition to scientific outcomes, another goal of the NRT is to prepare students for careers within or outside of academia, given the large number of graduates who by choice or by necessity, will work in a variety of settings (St. Clair et al., 2017). Said another way, the objective of the NRT is to develop t-shaped professionals with deep disciplinary knowledge but also the traits that allow them to thrive in changing environments and a variety of settings.

There are many conference papers describing approaches to education and evaluation aspects of individual NRTs and IGERTs. Many use both qualitative and quantitative approaches to assessment administered by both internal and external evaluators. Some NRTs rely on the program coordinator for internal evaluation administration and/or graduate students and faculty. Evaluation approaches vary considerably and we found no evidence of programs sharing the same assessment instruments or protocols. A few examples are below. They describe evaluation from different viewpoints and we present them here to show examples of the diversity of methods employed, and some research outcomes and reflections.

- One paper described the use of specific assessment methods including competency rubrics, individual development plans, and ePortfolios for evaluation (Chang, Semma, Fowler, & Arroyave, 2021). The rubrics encompassed professional and technical skills including: 1) interdisciplinary knowledge generation, 2) collaboration, 3) conflict resolution, 4) oral communication, 5) written communication, 6) self-reflection, 7) ethics, 8) interdisciplinary research, 9) multidisciplinary skills, 10) disciplinary knowledge, 11) informatics, and 12) design. This paper only described the evaluation method and no results were presented.
- One NRT studied 12 participants in their 3rd, 4th and 5th year of graduate studies (Denton & Borrego) via semi-structured interviews of 10-40 minutes in length, focusing on the influence of the NRT over their career preparation and choices. Among participants, they found a lack of stigma around non-academic career paths, which was credited to the interaction of NRT students with non-academic entities through internships and capstone design projects outside of academia. Students were also satisfied with the formal and informal mentoring they received. However, they did not find the independent development plans, which were expected to support career development helpful. Instead, they found them onerous, challenging, and a “waste of time”.
- A study of an IGERT program surveyed nine students in a final cohort to identify when and why they felt or failed to feel motivated during the program and what curriculum modules were most valuable for their learning experiences (Wang et al., 2021). The authors conducted a Likert scale and open-ended question survey to collect students’ opinions and feedback about the program in the context of student motivation through autonomy, competence, relatedness and values. While the team collected a vast amount of data on very specific aspects of their program, they acknowledged that the sample size was too small to make decisive conclusions, and because they lacked a pre-test it was difficult to accurately capture the effect of the program.

From these examples, it is evident that there is no one standard or turnkey approach for NRT evaluation, at either a student, program or national level. It is also clear that it is challenging to create one given number of variables or constructs that are related to interdisciplinary education that could potentially be included.

There is one national study across IGERT sites titled "Essential Competencies for Interdisciplinary Graduate Training in IGERT" (Gamse, Espinosa, & Roy, 2013) sought to examine how grantee institutions prepared PhD Students to conduct interdisciplinary research. As part of this activity, the authors conducted a literature review in close consultation with NSF, to identify six core essential competencies that were considered essential to interdisciplinary research and collaboration. These included the ability to:

- Develop depth of knowledge in one discipline or field of study
- Recognize the strengths and weaknesses of multiple disciplines
- Apply the approaches and tools from multiple disciplines to address a research problem
- Work in a team with individuals trained in different disciplines
- Communicate research based in one discipline or field of study to academic researchers trained in different disciplines
- Communicate about interdisciplinary research to non-academic audiences (laypersons)

The authors then examined how IGERT projects developed trainees in these six areas and how these competencies were applicable to IGERT projects. To achieve this, the authors collected data from two key respondent groups: IGERT PIs and trainees from 40 programs initially funded in 2007 and 2008. This included interviews with 39 of the 40 PIS and from 431 trainees across the 40 programs. The study found that respondents considered most important the ability to communicate research in one discipline to researchers trained in another, and the capacity to work in a team setting with researchers from other disciplines. Understanding the contributions and limitations of different disciplines and the capacity to apply multidisciplinary approaches and tools were also considered essential.

This study also reported on the many evaluation approaches across IGERT institutions. PIs reported that the assessment of their trainees' capacity and progress occurred through external evaluators and individual faculty colleagues who supervised IGERT trainees. PIs noted that they tracked trainee publications presentations, post-graduation placements, and research collaborations. PI and trainee respondents identified several activities they believed contributed to trainees' ability to conduct interdisciplinary research, including interdisciplinary courses, lab and field experiences, and interdisciplinary team research projects. Trainees also felt outreach activities were the most helpful activity as a means to communicate research effectively to lay audiences. Of those surveyed, over 90 percent were confident in their ability to work in interdisciplinary teams and to communicate research in one discipline to researchers trained in others, and 71 percent were confident about applying approaches and tools from multiple disciplines to address a research problem.

Also in the study, one third of PIs reported uncertainty about how best to assess their trainees, or how best to measure the right things. Some issues cited were inherent challenges of assessing interdisciplinary skills. Some noted that no benchmark existed for the skills trainees had when they entered the program, so it was difficult to evaluate how much progress a trainee makes over time. Some also had difficult identifying measures able to assess skills such as leadership and teamwork, and finding time to fit in evaluation among the research and administrative demands that the IGERT grant required.

DESCRIPTION OF OUR EVALUATION APPROACH

Evaluation Challenges

In approaching the assessment for our NRT program, we were aware of complexity of program evaluation when many individuals and training components are involved. Also, our proximity to a prior IGERT grant experience through individuals on our NRT team, made us aware of potential administrative issues. These experiences, those described in the literature, and our own work in program evaluation (Duval-Couetil, 2013; Duval-Couetil, Reed-Rhoads, & Haghighi, 2010; Yi & Duval-Couetil, 2021) increased our awareness of the following evaluation challenges:

Evaluation Design and Measurement: Developing an individual's interdisciplinary research capacity is a broad concept encompassing knowledge, skills, abilities and even mindsets stemming from many experiences within and outside particular a particular research project, making it difficult to define and measure in a very objective way. While we found instruments that could measure aspects of it, there were few validated scales focused on the learning outcomes specific to the grant, and those that existed were too narrowly focused on specific topics or constructs, to use with our population of graduate students.

Over-promising and under-delivering: In a review of three engineering education related projects, including an IGERT, Trevisan (2004) found that all of the projects struggled to develop and support an appropriate evaluation design, that all over-promised in the grant proposal, and under-delivered. As described in the examples above, assessment methods can take many forms, and while ePortfolios may be a very valid way to capture and highlight interdisciplinarity and critical thinking skills, they require considerable additional work for graduate students, who have limited time available (Wheadon & Duval-Couetil, 2014). Therefore, they may perceive them as having limited value as a means to enhance their career development or their attractiveness to employers.

Small cohorts: At an institutional level, NRT cohorts are fairly small and diverse. Therefore, while it is possible to evaluate learning and satisfaction, these results are not generalizable to the larger graduate student population. And, while NRT offerings (courses, workshops) are often designed to be open to non-trainees (students whose stipends are not funded through the NRT), it is difficult and somewhat impractical to administer the same surveys and protocols to different groups given that they are also small and diverse. The value of the work by Gamse, Espinosa, & Roy (2013) was that it is based on a large population of respondents.

Timing: At the start of the program, a large enough number of students are recruited to represent a cohort. However, subsequently, students join and complete the program at different points in time, and not necessarily at the start of an academic year. This trickling in of students makes evaluation challenging as 1) it is difficult to compare students or student cohorts at one particular time/stage of a program, and 2) it is very difficult for the evaluator to track. Another challenge is that students are at different stages of their educational programs when they join (i.e., they are not necessarily first semester doctoral students, instead they may enter after already having been supported through another source of funding).

Tracking: Despite the relatively small number of students involved in the NRT, our experience with the IGERT taught us that ensuring that each trainee participated in each of the program components was a very administrative-intensive process, requiring lots of emailing back and forth with the program coordinator who used spreadsheets to keep track of professional development modules covered, conferences attended, and papers presented. It was imperative that this information be collected as it was required for annual reporting, but the process needed to be more efficient.

Resources for evaluation: While evaluation is required as part of the NRT, the amount typically allocated to evaluation and assessment activities, in practice, is fairly limited for on-site PI/co-PI involvement in startup educational activities, instrument development and evaluation protocols. Further, while external evaluators are required, at startup they are often far removed from the internal evaluators who have to get surveys and protocols up and running. As Trevisan (Trevisan, 2004) stated, during the budgeting process, tradeoffs are made among project components and that "evaluation tasks and activities were often thought of 'extra' duties, duties that could be accomplished on top of normal workload and project expectations"

(p. 9). About external evaluators, he says that they were sometimes brought into the evaluation well after the project had started and so evaluation activities were delayed until well into the project, "leaving evaluation support for the project in a tentative and fragile state" (p. 10).

Limited opportunities for scholarship: Given the small cohorts, diversity of backgrounds, and diversity activities involved in NRT training, there are few opportunities for scholarship for education researchers who often are part of the internal evaluation team. As stated, cohorts are small and diverse, limiting generalizability. Therefore, faculty are unable to devote significant time beyond what is required to evaluation research, particularly when other studies may be competing for their time. Further, PIs and co-PIs conducting the scientific research typically do not have expertise in education research, or time to participate in on-going evaluation activities.

Our Methodology

The overall purpose of the evaluation for the NRT is similar to the following overarching question as cited in (Denton & Borrego): *What experiences and program components may help engineering doctoral students increase their interest and preparation for a career in industry or government, and why are they effective?* However, much more granular evaluation and assessment activity must take place prior to evaluating students' overall experience as represented above. A number of assessment decisions must be made at the start of the grant to: 1) track student progress through the NRT program, and 2) meet NSF's annual reporting requirements. These assessment decisions often have to be made rather quickly, prior to program startup and enrollment, and it is best not to have to change the protocol over the period of the grant so as to be able to make comparisons across individuals and groups.

To our knowledge, there was no repository of instruments and surveys that have been used by prior NRT or IGERT programs, which could serve as a foundation for developing one's own. In lieu of that, we examined the literature on interdisciplinary research in an attempt to identify instruments able to measure student progress related to the areas noted above. In searching, we examined published instruments in related areas, identified skills perceived to be essential for the scientific workforce of the future (Begg et al., 2015); read about pitfalls and complexity in defining, fostering and assessing interdisciplinary skills (Holt et al., 2017); and, found an interdisciplinary collaboration framework by Carr (Figure 2), which represents how challenging it is to operationalize all of the constructs and variables associated with interdisciplinary research and career preparation in an efficient and sustainable evaluation program.

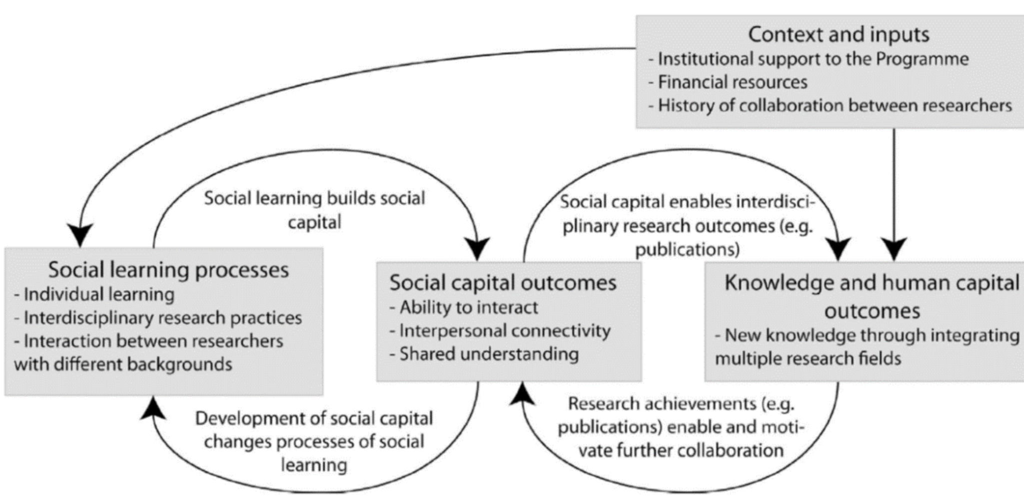


Figure 2. Conceptual framework for exploring the development of cross-disciplinary research. (Carr et al., 2008)

Given the challenges listed above, we chose three major thrusts for our internal assessment and evaluation efforts to be able to track progress and meet NSF's annual reporting requirements. These include: course evaluations for new courses developed; the development of a novel learning management tracking system for professional development and experiential (ACT) modules; and an annual progress survey for NRT Trainees. The results of these would be reviewed annually by our external evaluator, who would also conduct interviews and focus groups with NRT students and faculty, and conduct a periodic climate survey. A description of these various components is below.

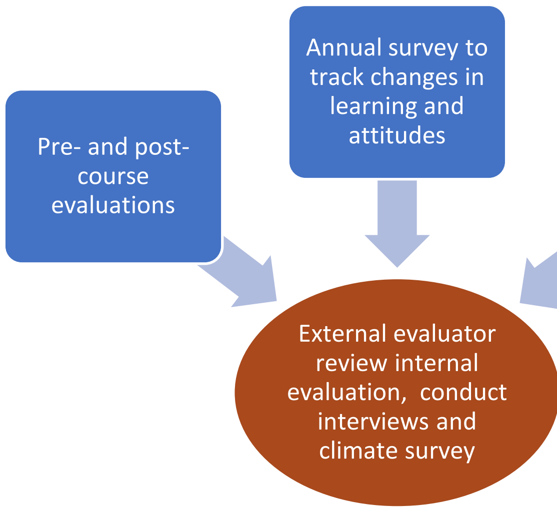


Figure 3. Main Program Evaluation Elements

NRT Course Evaluations: Course evaluations for new or modified interdisciplinary course for the NRT courses were fairly straightforward so we will not allocate much space to discussing them. These course evaluations were used in addition to university-administered evaluations. The survey consists of 4 Likert scale items including and two open-ended questions:

Likert scale (strongly disagree to strongly agree):

- The course enhanced my ability to analyze interactions between food, energy and water systems
- The course made me consider new technical solutions to food, energy and water challenges
- Working on multidisciplinary team projects enhanced my scientific knowledge
- Overall, the course stimulated my interest in interdisciplinary research

And two open ended questions:

- What are the strengths of the (example: SFEWS Systems and Economic Analysis for Food Energy & Water) course:
- How could the (example: SFEWS Systems and Economic Analysis for Food Energy & Water) course be improved?

Since the purpose of these courses was to also attract students from multiple disciplines who were not NRT trainees, questions able to distinguish non-NRT from NRT students were also embedded in the surveys. Typically, these evaluations were administered by getting program rosters from the NRT program coordinator and sending them out through a Qualtrics survey at the end of each semester.

Learning Management System for Professional Development Modules

As stated, tracking individual graduate student participation in and completion of various NRT program requirements can be labor intensive and is often compiled manually prior to the annual reporting deadline. This is particularly true for professional development and experiential modules given the number of them students are required to complete. These encompass workshops and seminars offered by departments and entities all over campus. It also includes academic and scholarly activities such as creating K-12 and undergraduate lesson plan, attendance at conferences, and papers published.

Since there is no unified "system" to track completion of requirements, this is typically done by having students report completion of various modules to the program coordinator. Traditionally, this occurred through email correspondence and noting data on spreadsheets to keep track of student progress. Therefore, there is a very high administrative burden associated with compiling this information, particularly when many reminders need to be sent to students as the annual NSF reporting deadline approaches.

To overcome these challenges, we created a professional development tracking "hub" within our university's learning management system (LMS). While this required a significant investment of time up front by the program coordinator, internal evaluator, and an instructional designer, we were confident there would be several benefits to taking this approach, including:

- One place to monitor student/cohort progress
- Progress could be tracked over multiple semesters
- Surveys could be integrated to assess learning and satisfaction with each module
- The responsibility would be on students to log their accomplishments
- Data collected would streamline annual evaluation and reporting activities
- Students from multiple institutions could potentially access the system

To create the LMS system, we first refined the modules that we had developed in the proposal into "Training Modules" (i.e., professional development) and "Apply/Create/Translate Modules" (i.e., experiential activities) (examples provided in Table 2). These modules provided the structure for the LMS "course" and associated assessments. Trainees were made students of the "course" and were responsible for logging professional development modules that were completed, most of which were not offered within the LMS but instead by units across the university. Here they could upload proof that a professional development module was completed, and at the same time provide a brief description of the experience, and complete a brief assessment (Figures 4 and 5). The LMS linked to Qualtrics which housed the survey. This system enabled the program coordinator and evaluator to see on one screen, in the gradebook, which students had completed which modules (Figure 6).

Table 2. *Training and Experiential Modules*

Training Modules	Apply/Create/Translate Modules
Training Module 1 – Intro to Graduate Studies Graduate school orientation Graduate student seminar Other training related to success in graduate school **Training Module 2 – Research and Ethics** Responsible conduct of research Data management Other training related to conducting research **OTHERS (not outlined in full here)** Training Module 3 - Lab Management Training Module 4 – Academic Writing Training Module 5 – Effective Communication Training Module 6 - Career Preparation	**Apply/Create/Translate Module 1 – Scholarship** Present research to SFEWS team or advisory board Present research at an external conference Collaborate on writing a grant proposal Write an academic paper Review a manuscript Other scholarly activities **Apply/Create/Translate Module 2 – Lab management** Manage a lab or specific activity within a lab Manage people in a lab Other lab-related activities or responsibilities **OTHERS (not outlined in full here)** Apply/Create/Translate Module 3 - Teaching and research mentoring Apply/Create/Translate Module 4 - Diversity Apply/Create/Translate Module 5 – Engaging with professional communities Apply/Create/Translate Module 6 – Intellectual property and technology commercialization

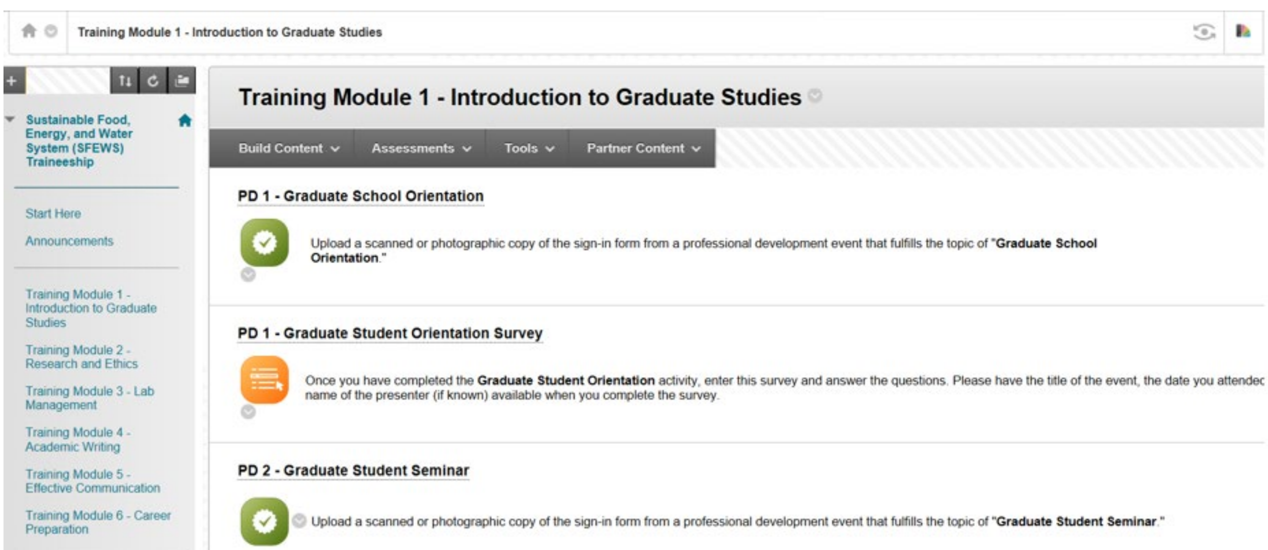


Figure 4. Example of modules as they appear in learning management system (LMS)

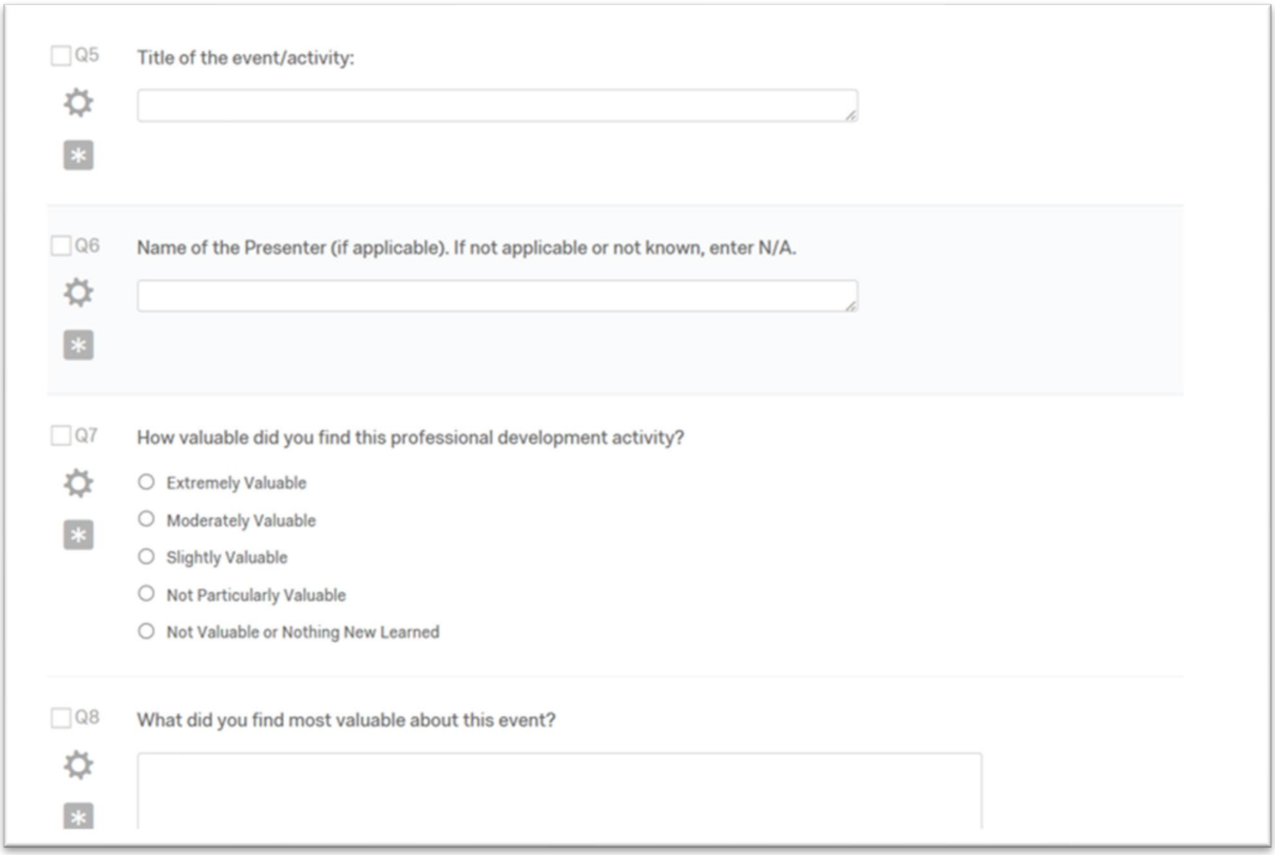
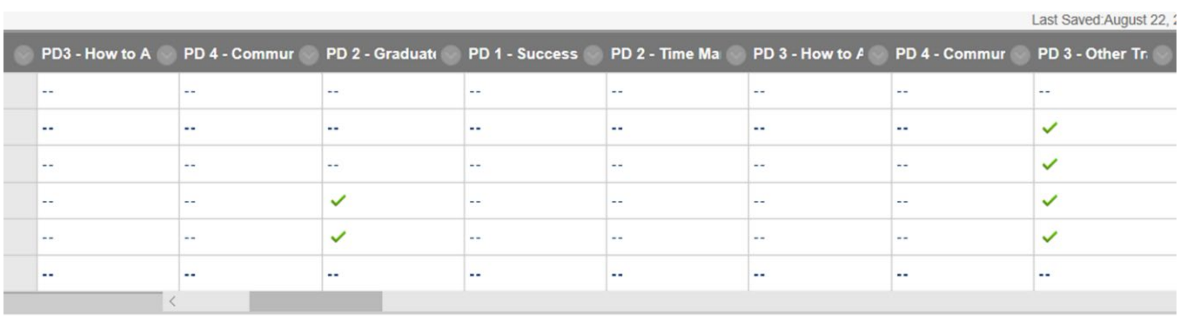


Figure 5. Example of the assessment associated with professional development modules.



Last Saved:August 22,

PD3 - How to A	PD 4 - Commur	PD 2 - Graduat	PD 1 - Success	PD 2 - Time Ma	PD 3 - How to A	PD 4 - Commur	PD 3 - Other Tr
--	--	--	--	--	--	--	--
--	--	--	--	--	--	--	✓
--	--	--	--	--	--	--	✓
--	--	✓	--	--	--	--	✓
--	--	✓	--	--	--	--	✓
--	--	--	--	--	--	--	--

Figure 6. Example of “gradebook” which shows what modules are completed.

<u>Annual Survey</u>

Typical educational evaluation research relies heavily on pre- and post-program surveys. These are convenient as they can be used to evaluate the impact of an instructional intervention by assessing students’ perceptions of changes in their knowledge and skills, personal attributes, or impact on their future aspirations or behaviors. They do so by rating themselves at the start (pre-survey) and at completion (post-survey) and the difference is attributed to the impact of the educational intervention. There are many problems with pre- and post-program surveys. Post-test scores can end up being lower than pre-test scores either because students at the beginning of a program didn’t know what they didn’t know and rated themselves higher than they did at the end of the learning experience.

We chose to do an annual survey that trainees would complete once per year because we envisioned additional challenges with administering pre- and post- surveys. Most importantly, we did so because trainees joined the program at different periods of their doctoral programs and at various times throughout the academic year. Thus, it would be very difficult to track when pre- and post- surveys should be administered. Further, it would be resource intensive to match pre- and post-survey data for small samples for diverse cohorts and periods. The annual survey could track individual students or cohorts over the grant period.

We were also concerned about over-surveying, given that we were also doing course evaluations and were tracking professional development. Response rates can be low when busy students are expected to complete multiple surveys. Another limitation was that we were also working across two institutions with very different resources and programmatic offerings. Finally, we were also aware of evaluation programs that are quite comprehensive (i.e., optimistic) at the start and get scaled back significantly due to a lack of resources or competing interests. Our goal was to create something that we could administer once per year and that could be sustained throughout the grant period. The question then became what to include in the survey to match the outcomes of the grant, and whether there were any validated instruments to help accomplish this.

Based on our literature review and past experiences in this area, we arrived at the following categories of items.

Table 2. *Annual Trainee Survey Categories*

Category/Construct	**# of items**
Background	5 items
Program satisfaction	7 items
Learning process (overall)	5 items
Learning process (solving FEW challenges)	1 open-ended
Career goal post-graduation	4 items
Change in career goals	2 items
Perceived mastery of, and interest in, professional development competencies	23 items
Interdisciplinary competence	6 items
Learning goals associate with the NRT	5 items
Demographic information	8 items

Annual Survey Results to Date

Below are select, preliminary results from the annual survey that demonstrate its ability to show the attitudes of the cohorts progressing through the program. As stated, the survey is administered once per year at the end of the spring semester, and therefore, we do not account for precisely when students entered the program, instead we are examining the trend as they progress through the program. These analyses demonstrate just how small the cohort dataset can be over a multi-year period. Despite our approach, it is challenging to get 100% participation each year, regardless of how many reminders trainees receive about how important the survey results are to the grant and NSF reporting activity.

The results below are based on five trainees in our first cohort who responded to the survey for three consecutive years. At the end of the grant period, we will be able to examine trends by cohort and by individual students to get a more precise view of the impact of the program over time.

Program Satisfaction. Participants rated how satisfied they were with the SFEWS program with a Likert scale (5 = strongly agree). The result is presented in Table 3. Participants strongly agreed that the interdisciplinary SFEWS program added value over a traditional disciplinary Ph.D. program (Year 1 *M* = 4.40 to Year 3 M=5.00). Participants' satisfaction with the unique interdisciplinary courses also increased (Year 1 *M* = 4.00 to Year 3 *M* = 4.67). Their satisfaction with the overall program experience (Year 1 *M* = 4.20 to Year 3 *M* = 4.33) fluctuated in year two but overall was high and slightly increased over time, which was the case for several program elements. Interaction with industry had the lowest satisfaction of all items.

Table 3. *Participant program satisfaction*

	Y1		Y2		Y3	
	M	*SD*	*M*	*SD*	*M*	*SD*
I am learning a great deal in the SFEWS program.	4.00	1.00	4.20	0.84	4.33	0.58
Organized in a way that helps me learn.	4.40	0.55	3.60	0.89	4.00	0.00
Developing my ability to apply theory to practice.	4.20	0.84	3.80	0.45	4.33	0.58
Additional value over a traditional disciplinary PhD.	4.40	0.55	4.60	0.55	5.00	0.00
Preparing me for my future career.	4.20	0.84	4.20	0.84	4.33	0.58
Overall program experience	4.20	0.45	3.80	0.45	4.33	0.58
Mentoring by my primary faculty advisor	4.60	0.55	4.40	0.55	4.33	0.58
Mentoring by other SFEWS-affiliated faculty	3.80	0.45	3.80	0.45	4.33	0.58
Interaction with other SFEWS trainees	4.40	0.55	4.00	0.71	5.00	0.00
Unique interdisciplinary courses offered through SFEWS	4.00	0.71	4.00	1.22	4.67	0.58
Professional development offered through SFEWS	3.40	0.55	3.80	0.84	4.00	0.00
Interaction with industry	3.00	1.22	2.80	0.84	3.33	1.15

Learning Outcomes. Participants' perceived self-efficacy for aspects of professional development was measured with a 5-point scale (Table 4). Overall, participants agreed that they were confident about research-related tasks and ethics (e.g., responsible conduct of research, lab safety, lab management, presentations); and their confidence overall increased slightly over time. Their perceived entrepreneurship-related self-efficacy (e.g., intellectual property protection, licensing and commercialization, entrepreneurship) was at high levels. Perceived competence for aspects of being an interdisciplinary researcher were also measured with a 5-point scale (Table 5), and were overall slightly higher overall than elements of professional development.

Table 4. *Professional Development Self-Efficacy*

	Y1		Y2		Y3	
	M	*SD*	*M*	*SD*	*M*	*SD*
Time Management	3.40	1.14	4.00	1.22	4.67	0.58
Dissertation preparation	4.40	0.89	4.20	0.84	4.33	0.58
Responsible conduct of research	4.00	0.71	3.80	0.84	4.00	1.00
Data management	4.00	0.71	4.20	0.84	4.33	0.58
Lab safety	3.80	1.30	4.40	0.89	4.33	0.58
Lab management	3.80	1.30	4.40	0.89	4.33	0.58
Writing academic papers	4.60	0.89	4.80	0.45	4.67	0.58
Grant writing	4.00	0.71	4.60	0.55	4.33	0.58
Reviewing academic papers	4.20	1.10	4.60	0.55	4.33	0.58
Creating presentations	4.20	0.84	4.40	0.89	4.00	1.00
Delivering oral presentations	4.40	0.89	4.20	0.84	4.00	1.00
Curriculum vitae presentation	3.60	0.55	4.60	0.55	4.33	0.58
Professional networking	3.80	0.84	4.60	0.55	4.33	0.58
Identifying mentors	3.80	1.30	4.60	0.55	4.33	0.58
Interviewing skills	4.20	0.84	4.80	0.45	4.33	0.58
Salary and job negotiations	4.20	0.84	4.60	0.55	4.33	0.58
Teaching university students	4.00	1.00	4.60	0.55	4.33	0.58
Mentoring university students	3.80	0.84	4.60	0.55	4.67	0.58
Curriculum development	3.60	1.14	4.00	1.22	4.33	0.58
Intellectual Property Protection	4.40	0.89	4.80	0.45	4.33	0.58
Licensing and Commercialization	4.40	0.89	4.20	1.10	4.33	0.58
Entrepreneurship	4.40	0.89	4.20	1.10	3.67	1.53

Table 5. *Perceived Interdisciplinary Competence*

	Y1		Y2		Y3	
	M	*SD*	*M*	*SD*	*M*	*SD*
Depth of knowledge in your primary discipline or field of study	3.80	0.45	4.40	0.55	4.00	0.00
Ability to recognize the strengths and weaknesses of multiple disciplines	4.00	0.71	3.80	0.84	3.67	0.58
Ability to apply the approaches and tools from multiple disciplines to address a research question	3.80	0.45	3.80	0.84	3.67	0.58
Ability to work in a team with individuals trained in different disciplines	4.20	0.45	4.60	0.55	4.67	0.58
Ability to communicate research based in one discipline or field of study to academic researchers trained in different disciplines	4.40	0.55	4.40	0.89	4.67	0.58
Ability to communicate about interdisciplinary research to non-academic audiences	4.40	0.55	4.20	1.10	4.33	0.58

Career Goals. Participants changed their career goals during the program. In Year 1, participants were most interested in faculty positions but that interest decreased over time (M = 3.60 to 3.33). Their interest in obtaining an industry research position increased over time (M = 3.40 to 4.33) and it was the highest among other options (e.g., a faculty position, a research position in a university, start-up) in Year

3. From the responses to an open-ended question, it seemed that participants got into the program with a solid career goal, primarily a faculty position, but they changed their mind as they became more aware of specific interests they had.

> *"Upon first entering the graduate program, I had complete desire to enter academia and become a professor. However, I am the type of person who likes to try new things, and after attending one seminar where they told us about how industry expects you to move around, I became a lot more interested in industry because tenure is not the main goal."*

> *"I became much more interested in pursuing policy as a larger career goal because I find looking at a problem at a high level is more interesting to me than just looking at purely technical problems."*

Table 6
Career Goal Changes

	Y1		Y2		Y3	
	M	*SD*	*M*	*SD*	*M*	*SD*
Obtain a faculty position	3.60	0.894	3.40	0.894	3.33	0.577
Obtain a research position in a university	3.00	0.707	3.60	0.894	3.00	0.000
Obtain a research position in industry	3.40	0.894	3.60	1.140	4.33	0.577
Commercialize my own technology or start a new venture	3.20	0.837	3.00	1.000	1.67	0.577

External Evaluation Activities

The data above provides an excellent foundation for interviews and focus groups conducted by our external evaluator, who has a limited but important role in our NRT program. She was involved in our startup efforts by sharing her experiences on other projects, and pointing us to survey instruments she had used in the past. She also reviewed and provided feedback on various drafts. Once we were out of startup mode, her activity consisted of conducting annual interviews with trainees and faculty, consulting with us during annual meetings, reviewing our internal evaluation data, and contributing to the NSF annual report. She also periodically administers a diversity/climate survey.

The nature of the findings of the external evaluator interviews with students were the following. It is clear that participants derived significant benefits from the program.

- Trainees indicated high levels of satisfaction with the program. Benefits most noted were: (1) learning about and experiencing practical interdisciplinary applications of their disciplinary studies, (2) contributing to a project with tremendous global, societal benefit, and (3) extending their knowledge and skill sets in ways that increased their future marketability.
- They spoke with great pride about the completion of the first set of solar array structures. They clearly recognized this as a significant accomplishment to which they had all contributed, while also emphasizing that it was only through collaboration that it had been realized.
- Some students noted that SFEWS course projects which had required them to pair with a peer from outside of their discipline had prepared them well for executing this real-world project.
- Benefits gained from cross-disciplinary collaboration included learning fundamental content and methods from other disciplines, building skills in cross-disciplinary communications, and having opportunities to integrate diverse sets of knowledge and skills to forge unique career paths.

- In addition to interdisciplinary learning through project research, interviewed trainees and associates reported benefit from interdisciplinary courses and from most professional development activities; of note were the tech commercialization and diversity courses.
- Trainees remained excited about the potential of SFEWS to contribute to resolving a significant global problem. A few interviewees indicated that they will continue to pursue FEWS research following their graduation. Their collective interest in applying their skills and knowledge to help people and society was an important inspiration for most trainees' involvement in SFEWS.
- Trainees frequently mentioned how SFEWS was preparing them well to pursue a wide range of career options outside of academia. Trainees reported that opportunities they were provided to communicate across disciplines and sectors (i.e., public, industry, academia), to work in teams, and to apply systems thinking had impacted their career preparedness and contributed to them thinking more broadly about career opportunities.
- Trainees felt the working climate of the program to be inclusive and supportive in ways that contributed to their productivity and success, but felt there were too few individuals of color and too few women in SFEWS communities.

Interviews were able to solicit many suggestions for improvement as well. Select examples are below.

- Trainees and associates spent a significant amount of time during their interviews discussing program communications. Most interviewees noted that faculty members, and the entire SFEWS team, would benefit from more opportunities to meet, discuss, and brainstorm for project problem solving.
- Trainees and associates believed that they had gained more from the SFEWS experience than had SFEWS faculty members because of their closer, collaborative working relationships with other trainees across disciplines.
- Several recommendations from trainees on the survey were related to communications, including holding more joint meetings, providing regular updates on project progress and challenges, and offering more opportunities for informal interactions that involve all team members.
- While findings of the internal and external evaluations indicate that trainees are satisfied with the mentoring they receive from their advisors, interviewees believe that their SFEWS experiences could be improved by more direct and open communication from their faculty advisors and all faculty with whom they collaborate for SFEWS.
- About half of interviewees commented on the need for new and different expertise on the SFEWS team, after having experienced the initial planning and building of the solar array structures. Expertise that would be useful to the project included structural engineering, landscaping, and commercial construction.

From the qualitative data above, it is clear that the external evaluator was able to gain insights from students and faculty on perceptions of and experiences with the NRT program. For example, she can understand more about the cohort experience, mentoring, communication and contributions to career development.

CONCLUSION

Putting an evaluation program into place and sustaining it to track and improve program outcomes requires considerable effort that is not always recognized. Even established, national programs such as the NSF NRT, have few resources and best practices from which to draw when proposing or executing and evaluation plan. Other issues are limited resources for evaluation activities, a hesitation to further burden busy students, and small and irregular cohorts, often spread across departments. Given these constraints, and our observation that evaluation plans can be over optimistic in year one, our goal was to create a sustainable evaluation program that would remain consistent over the grant period.

The plan we outlined above in this case study allows us to do so. The annual survey allows us to track individual students as well as cohorts regardless of when they join the program. It highlights areas that need to be probed by ourselves and our external evaluator. However, as stated above, there are many limitations with this system as well, given cohort sizes. First and foremost, when a trainee chooses not to respond to the survey we have gaps in our knowledge about cohort experiences. Second, when cohorts are small, fluctuations in the results (as we see above in Year 2) can be due to the responses of one individual, so examining these data in detail is essential. Ideally, it would be best to not rely on a survey and instead monitor information flowing continuously throughout the year that would enable us to pivot and make improvements in real time. However, this is very challenging to do in a systematic way since most of this real-time feedback occurs through anecdotal personal communication.

There are opportunities to innovate in program evaluation, particularly for large, multi-site initiatives. Among our contributions was developing a system housed in the University's LMS to track program requirements and collect a significant amount of assessment information related to trainee views on individual professional development components. While this has significant value, it isn't without challenges. Getting trainees to input information into the system is difficult and requires that the program coordinator to repeatedly remind them to do so. However, if students comply this does relieve some of the administrative burden on the program coordinator or evaluator. It is clear that trainees need to understand that reporting is an integral requirement of the program and that it is easier much easier for everyone if they log activity when it occurs, rather than wait until the end of the year or the annual reporting period.

There were several challenges with this system. First, the university changed LMS systems midway through the grant, which was very disruptive. There can also be glitches with making these work across institutions. Nevertheless, we feel that with some tweaking a system such as this can be an excellent way to collect a significant amount of education and assessment data, and streamline reporting activities. It is also potentially a way to disseminate standardized professional development modules across institutions, and specific programs such as the NRT. It should be noted that without a significant donation of time from our university's online learning unit, the development of the tracking system would not have been created since it was beyond the scope of the grant and no funding was allocated to the activity.

There are clearly opportunities for efficiencies within large national programs such as NRT which, in theory, have similar objectives. At startup, it would be helpful to have access to assessment surveys or interview protocols aligned with desired outcomes specific to the NRT program. Inevitably, without this, each NRT site is forced to reinvent the wheel by developing their own. This would also help us obtain larger sample sizes at a national level, which would lead to more generalizable findings, and allow for comparisons across institutions and demographic groups. Individual programs will always have unique aspects that should be assessed, but core objectives could be more similar across institutions. Startup of evaluation activities across NRT sites vary widely, thereby limiting best practices over all.

Assessment and evaluation activities, designed to measure impacts on training and career development, are important aspects of the NRT, and require significant attention to capture the range of activities undertaken by very small cohorts of interdisciplinary students and faculty. Our goal was to develop "sustainable" evaluation activities given our observation that programs often begin with very ambitious assessment and data collection goals in the first year, that diminish over time. The protocols we put in place were not overly ambitious, and have remained the same through the program period, which we consider a success.

REFERENCES

Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2005). *A taxonomy for learning, teaching, and assessing*: Longman.

Begg, M. D., Bennett, L. M., Cicutto, L., Gadlin, H., Moss, M., Tentler, J., & Schoenbaum, E. (2015). Graduate education for the future: new models and methods for the clinical and translational workforce. *8*(6), 787-792.

Chang, C.-N., Semma, B., Fowler, D. A., & Arroyave, R. (2021). *An Interdisciplinary Graduate Education Model for the Materials Engineering Field.* Paper presented at the 2017 ASEE Annual Conference & Exposition.

Denton, M. M., & Borrego, M. Non-Academic Career Pathways for Engineering Doctoral Students: An Evaluation of an NSF Research Traineeship Program.

Duval-Couetil, N. (2013). Assessing the impact of entrepreneurship education programs: Challenges and approaches. *Journal of Small Business Management, 51*, 14.

Duval-Couetil, N., Reed-Rhoads, T., & Haghighi, S. (2010). *Development of an Assessment Instrument to Examine Outcomes of Entrepreneurship Education on Engineering Students*. Paper presented at the 40th ASEE/IEEE Frontiers in Education Conference, Washington, DC.

Gamse, B. C., Espinosa, L. L., & Roy, R. (2013). Essential Competencies for Interdisciplinary Graduate Training in IGERT: Final Report. GS-10F-0086K. *Abt Associates*.

Holt, R. E., Woods, P. J., Ferreira, A. S. A., Bardarson, H., Bonanomi, S., Boonstra, W. J., . . . Holma, M. (2017). Avoiding pitfalls in interdisciplinary education. *74*(2), 121-129.

Newswander, L., & Borrego, M. (2021). *Igert Funding And The Institutionalization Of Interdisciplinary Graduate Education.* Paper presented at the 2009 Annual Conference & Exposition.

St. Clair, R., Hutto, T., MacBeth, C., Newstetter, W., McCarty, N. A., & Melkers, J. (2017). The "new normal": Adapting doctoral trainee career preparation for broad career paths in science. *12*(5), e0177035.

Trevisan, M. S. (2004). Evaluation Issues in the Renewal of Engineering Education: Lessons from NSF-Funded Projects. *9*, 1.

Wang, C., Clarkson, C. M., Andler, J., Handwerker, C. A., Korey, M., Frost, K. D., & Reeves, M. S. (2021). *Lessons Learned from the NSF IGERT Program: Cultivating Student Motivation in the Interdisciplinary and International Contexts.* Paper presented at the 2020 ASEE Virtual Annual Conference Content Access.

Wheadon, M., & Duval-Couetil, N. (2014). *Student Perspectives on Developing More Relevant Ph.D. Programs in STEM Disciplines through Professional Skills Training.* Paper presented at the ASEE Annual Conference, Indianapolis, IN.

Yi, S., & Duval-Couetil, N. Interdisciplinary Entrepreneurship Education: Exploring 10-Year Trends in Student Enrollment, Interest and Motivation. 2515127420979195.