DEP/RadGrad: Combining individualized learning plans and communities of practice to improve engagement, retention, and diversity in undergraduate computer science education

January 13, 2018

Contents

1 Introduction

To paraphrase Charles Dickens, it is the best of times and the worst of times for computer science (CS).

On the one hand, current CS research produces innovations with near-term societal impact with startling frequency, including: blockchain-based data storage, virtual currencies, autonomous vehicles, deep learning, quantum computing, and virtual/augmented reality. Some innovations can "cross the chasm" from research idea to mass market adoption within the length of a single undergraduate degree program [1]. Due in part to this velocity of innovation, the World Economic Forum estimates that more than 65% of current students will work in jobs that don't exist today [2]. According to the Computing Research Association, the number of undergraduate CS majors has tripled since 2006 and is expected to grow further [3].

On the other hand, U.S. high school students now rank near the bottom among 35 industrialized nations in math preparation [4]. Retention is poor: fewer than 40% of students who enter college with the intention of majoring in a STEM field such as CS actually complete the degree [5]. Diversity is actually decreasing: female participation in CS has declined to 18% from a peak of 37% in the mid-1980's [6]. In 2017, only one in five of those taking the Advanced Placement exam in computer science were underrepresented minorities [7]. Silicon Valley, the epicenter for CS innovation, suffers from a culture of entrenched and widespread sexual harrassment [8].

This combination of issues creates daunting challenges for undergradute computer science degree programs. In response to increased demand, programs have increased class sizes, instituted academic barriers to entry, and reduced some course offerings and faculty activities [3]. For example, class sizes approaching 1000 students in lower level courses now occur at UC Berkeley and other prominent computer science programs. Unfortunately, such programmatic responses can negatively impact on engagement, diversity and retention [9].

Some students seek alternatives to an undergraduate degree program to acquire CS skills, such as three to six month coding bootcamps [10]. However, such short-term educational programs cannot provide students with the analytical depth needed to engage with the leading edge of innovation.

In response to the velocity of innovation, some students turn away from their university's slow moving curriculum and toward online platforms such as Coursera, Udacity, and edX, which can quickly implement specializations in emergent disciplines such as big data, data science, and cybersecurity. Unfortunately, students are typically left to their own devices to select appropriate, high quality "extracurricular curriculum".

Recent approaches to addressing diversity in CS include BRAID [11] and non-profit organizations such as Girls Who Code and Black Girls Code [12, 13]. These approaches show great promise for their target demographic, but do not necessarily impact on overall engagement and retention.

In essence, we believe that computer science provides incredible social and economic opportunities, but we must develop new and better ways to improve *engagement* (i.e. create wider interest in pursuing CS), *retention* (i.e. create mechanisms to improve the chance that students, once pursuing a CS undergraduate degree, will complete it), and *diversity* (i.e. create ways to improve engagement and retention for women and underrepresented minorities).

The fundamental idea in this proposal is to provide students, faculty, and advisors with an alternative perspective on the undergraduate degree program—which traditionally boils down to a single kind of activity (coursework) and a single metric for success (grade point average). Our alternative perspective is called the *Degree Experience*, and it gives first class status to both curricular activities (courses) and extracurricular activities (discipline-oriented events, activities, participation in organizations, etc.) To establish the first class status of extracurricular activities, the Degree Experience perspective replaces GPA as the single metric for success with a three part metric called *ICE* that assesses student development with respect to Innovation, Competency, and Experience. Each student's Degree Experience also includes a representation of their disciplinary interests and career goals that helps them assess the relevance of potential curricular and extracurricular activities. Finally, the Degree Experience perspective is voluntary. It complements but

Figure 1: Research Questions

1. What factors help or hinder adoption of DEP/RadGrad?
2. Does the successful adoption of DEP/RadGrad into a undergraduate STEM degree program make a positive impact on engagement, retention, and diversity? Why or why not?
3. Can we identify student demographic factors that correlate positively or negatively with adoption, engagement, and retention?
4. Does the project provide evidence to support the DEP/RadGrad Theory of Change?
5. What factors help or hinder adoption of DEP/RadGrad across disciplines and institutions?

does not change any existing undergraduate degree requirements of a university.

Over the past two years, we have developed this idea into a conceptual framework called *Degree Experience Plans* (DEP) and a supporting technology platform called *RadGrad*. The design of DEP/RadGrad is influenced by research on diversity and retention and two educational research theories: Individualized Learning Plans (ILP) and Communities of Practice (CoP). ILPs help students connect their current studies to their future career goals. CoP identifies the importance of practitioner networks for both formal and informal learning. Based upon this prior research, and our pilot use of DEP/RadGrad with undergraduate students, we hypothesize that student populations adopting the Degree Experience perspective will show increased levels of engagement, retention, and diversity. Figure 1 presents the specific research questions to be pursued in this study.

In this project, we propose to deploy DEP/RadGrad to three undergraduate programs and assess its impact on engagement, retention, and diversity through a series of empirical studies. The results of this project will provide new insight into the factors affecting engagement, retention, and diversity. It will also provide new insight into issues surrounding the adaptation of ILP and CoP to the college setting. It will make a scalable, tailorable open source framework available to others who wish to replicate or adapt our approach to other computer science settings. Finally, it will create a foundation for further research on engagement, retention, and diversity across other STEM disciplines.

2 Related Work

To better motivate the design of Degree Experience Plans and RadGrad, this section summarizes recent work regarding retention and diversity in computer science, Individual Learning Plans, and Communities of Practice.

2.1 Retention and diversity in CS

There is a national need for undergraduate computer science degree programs to improve both retention (the percentage of students entering CS programs who finish the degree) and diversity (the percentage of graduates who are female and/or from an underrepresented minority group). We need to improve retention because the projected demand for skills in computer science far exceeds current production [3]. We need to improve diversity because a more diverse STEM population improves tech innovation at large. For example, mixed-sex teams filed 40% more information and technology patents than all-male teams [14], and management diversity leads to a $42M increase in S&P value of firms [15].

While the need is clear, addressing it is complicated. For one thing, gender diversity in computer science has actually fallen in the last 20 years [6], with no well accepted explanation for its cause. For another, some diversity-related issues start in middle and high school: Black students are less likely than White students to have computer science courses in middle and high school, and female students are less likely than male students to be told they would be good at computer science [16]. There is some research that provides evidence for a way forward: a study by Google concludes that four factors primarily influence young womens' decision to pursue CS: (1) social encouragement (positive reinforcement of CS pursuits from family and peers); (2) self perception (an interest in problem solving and a belief that those skills can be translated to a successful career); (3) academic exposure (availability of curricular and extracurricular CS activities); and (4) career perception (view of CS as a career with diverse applications and a broad potential for positive societal impact) [6].

If high school students graduate and enter an undergraduate degree program in computer science, then retention becomes an issue. More than half of the students who start out in science or engineering switch to other majors or do not finish college at all [17]. Initiatives to improve retention, such as the Threads undergraduate curriculum at Georgia Tech, emphasize giving students more control over their degree plan, a better understanding of how their studies relate to their career interests, and an increased emphasis on the importance of extracurricular activities [18].

2.2 Communities of Practice

Communities of Practice (CoP) is a theory of learning first proposed in 1991 [19], more fully developed in 1998 [20], and extended to "landscapes of practice" in 2015 [21]. A loose definition of Communities of Practice is "groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly." More specifically, three characteristics distinguish a community of practice from other kinds of communities: (1) There is at least one domain of interest shared by all members; (2) members engage in joint activities and discussions, help each other, and share information; and (3) members are practitioners in the domain, not just people with shared interests, and thus develop a shared repertoire of resources.

Communities of Practice is not a novel approach to social structure or learning style. In reality, it is a descriptive term originally arising from ethnographic studies of traditional apprenticeships upon recognition that the apprentice does not only learn from the master, but also from the surrounding community of journeymen and fellow apprentices. Communities of practice have been shown to exist in business, government, professional associations, and development projects, and exhibit many characteristics of systems in general: emergent structure, complex relationships, self-organization, and dynamic boundaries.

It might be assumed that university departments would naturally give rise to communities of practice, as departments by definition involve groups of people with common interests. Unfortunately, traditional curriculum and learning procedures can work against the creation and maintenance of communities of practice, as learning occurs through solitary, individual efforts, group structures rarely persist beyond a semester, structures are fixed, relationships are simple, and organizations have an apriori structure with static boundaries. For undergraduates in particular, the "practitioner" relationship is generally missing, as professors are practicing academics, not industry professsionals. (In contrast, the "practitioner" relationship, and thus a community of practice, can potentially develop between professors and Ph.D. students who also intend to pursue a career in academia.)

2.3 Individual Learning Plans

The Individual Learning Plan (ILP) is a tool developed for use in middle and high school to help students better align their academic activities with their post-high school goals, such as college, the military, or other post-secondary training. ILPs are mandatory in approximately 30 states as a mechanism to improve college and career readiness. They typically include an academic planner, career explorer, personality and learning style assessments, a resume builder, and an action plan. Students develop and manage their ILP in consultation with their teacher and career counselor throughout middle and high school.

Research on the effectiveness of ILP is still ongoing. However, initial results from focus groups and surveys indicate that providing access to ILP helps students to perceive school as more meaningful and useful, motivates them to pursue more rigorous in-school and out-of-school learning opportunities, and improves college and career readiness outcomes. [22].

2.4 Degree and Career Planning Technologies

(Related work on technologies for ILP, CoP, and Degree Planning will go here.)

3 DEP and RadGrad

3.1 Degree Experience Plans

Degree Experience Plans (DEPs) are a new conceptual representation for an undergraduate's degree program experience. They combine findings from research on diversity and retention, ILP, and Communities of Practice with the goal of improving engagement, diversity, retention, and post-graduation success. The DEP representation is not specific to computer science or even to STEM disciplines, though our current experience with it is limited to computer science. A Degree Experience Plan consists of the following seven entities:

Interests represent a set of discipline-specific topics relevant to the degree experience. In computer science, examples of Interests might include "blockchain", "big data", and "Java".

Career Goals represent professional outcomes that a student can pursue through the degree experience. In computer science, example Career Goals include "Data Scientist", "Augmented Reality Engineer", and "Security Analyst".

Courses represent the curricular activities associated with the undergraduate academic unit. For the University of Hawaii computer science department, approximately 40 courses are represented.

Opportunities represent extra-curricular activities that help a student progress toward one or more Career Goals and/or learn more about a specific Interest. The set of Opportunities available for a student to add to their DEP are "curated" by faculty members to ensure quality and relevance. Example opportunities include: a local Hackathon, a summer internship at a local high tech company, or participation in a faculty member's research project. Opportunities can also include online courses available through platforms like Coursera or edX, if the faculty have reviewed the offering and found it be useful and appropriate for their students. Opportunities are tagged if they represent an experience that involves a community of practice. Approximately half of the 70 Opportunities in the current RadGrad deployment for the UHM computer science department involve communities of practice.

Degree Plan comprises the set of Courses and Opportunities that a student has completed previously, is currently taking, or plans to complete in upcoming semesters. By explicitly representing and planning out curricular and extra-curricular activities, DEPs provide a more wholistic view of the student's disciplinary "experience", not just their classroom activities.

ICE is a three part measure to track both progress and success within the degree program. ICE is an acronym for its constituent measures, which are named Innovation, Competency, and Experience. To be a well-prepared computer science graduate according to RadGrad, students must earn 100 points in each of the three categories by the end of their degree program. ICE points are earned each semester by completing Courses and Opportunities. Typically, a student earns Competency points for completing Courses, and Innovation and Experience points for completing Opportunities. RadGrad admins are responsible for assigning the number of points earned for a given Course or Opportunity. For example, in the RadGrad deployment for the UHM computer science department, a student earns 6 points for a B in a Course, and 10 points for an A. For Opportunities, students earn 15 Innovation points for a weekend hackathon, and up to 25 points for a summer internship.

Levels respond to the need we identified during our pilot studies for RadGrad participation to have a physical manifestation. Students want to know who else is using the system, and what progress they have made so far, without having to login to the system. After several rounds of design, we decided on the use of laptop stickers with a custom RadGrad design, with a color scheme representing a six stage progression from zero ICE points to 100 points in all three categories. Pilot studies with approximately 50 students have shown that RadGrad levels and their physical manifestation as laptop stickers are an appealing and useful way to students to form communities around their degree experience plans. Students immediately put the sticker on their laptop and told us that they find it interesting to see who else has one and what color it is. Each student's level is also displayed in their navbar to the left of their ICE points once they login to RadGrad. The sample student illustrated in Figure 1 is at Level 3 (Green).

3.2 RadGrad

RadGrad is an open source application that implements the Degree Experience Plan conceptual framework along with user management and instrumentation to support evaluation. Development of RadGrad began in 2015 with the development of paper mockups and usability tests by and for computer science majors. RadGrad is now a functional web-based application using the Meteor framework, implemented in approximately 30,000 lines of Javascript and 7,000 lines of HTML. Figure 1 illustrates a hypothetical user's home page after they login to the system.

RadGrad implements the conceptual framework of Degree Experience Plans and provides backend database services, not just for the DEP entities, but also for users in five roles: students, faculty, advisors, mentors, and admins. More significantly, the two year RadGrad design process has resulted in a user interface that users find easy to understand and manipulate.

A phased roll-out of RadGrad to 50 students began in Fall, 2017. We introduce students to RadGrad through a 25 minute training session where we explain the motivation for and concepts behind Degree Experience Plans and how they are manifested in RadGrad. Students create an initial DEP, discover their level, and at the conclusion of the session obtain the laptop sticker associated with their current level. This initial roll-out has helped us refine our training approach and verified that the system is functional and ready for broader dissemination.

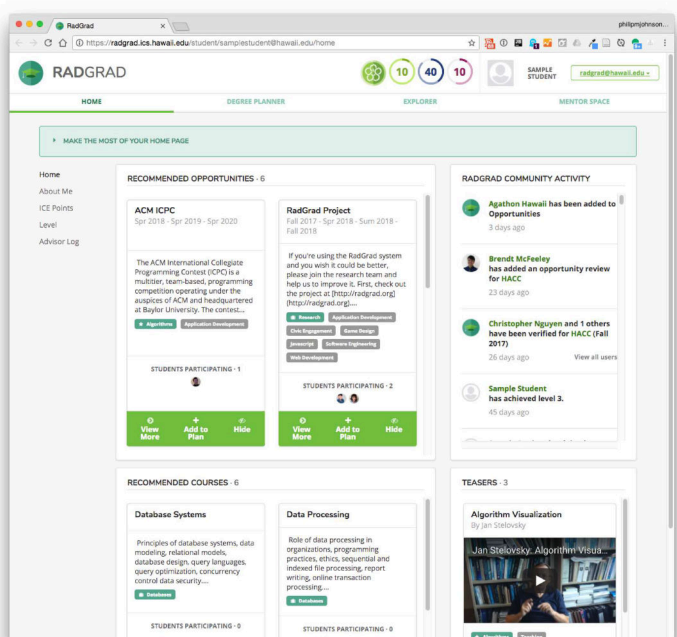

Figure 1: *RadGrad sample page*

We designed RadGrad to integrate into the department advising process. The UHM computer science department expects students to meet with an advisor once a semester to discuss their upcoming course selection and progress toward graduation. RadGrad provides features to simplify that aspect of advising in order to enable discussion of the larger degree experience: the student's interests and career goals and planning for complementary extracurricular activities over the remaining semesters of their program. Advisors and faculty use RadGrad to verify student participation in extracurricular activities, earning them the specified ICE points for that activity. Coursework is automatically verified in RadGrad by accessing institutional database records. We expect students to consult RadGrad consistently but infrequently: just once or twice per semester to update their Degree Experience Plan with new or different interests and career goals, update their future plans for curricular and extracurricular activities, obtain verification for activities they participated in previously, and receive their next Level laptop sticker if achieved. Pilot studies indicate that the chance to demonstrate progress through the program in the form of a higher level laptop sticker is an incentive to use the system at the start of each semester when verifications can earn additional ICE points.

3.3 DEP/RadGrad Theory of Change

Figure 2 illustrates our theory for how RadGrad usage can lead to increased engagement, retention, and diversity. First, we note that some students begin their undergraduate degree program already certain of their interest in pursuing a career in computer science and already knowledgeable about how to integrate computer science into their extracurricular activities. Those students are already engaged and not a retention risk. From our pilot studies, those students are enthusiastic RadGrad users, but mainly because RadGrad provides them with a reward (high ICE points and Level) for things they were doing anyway. Our theory of change is not intended for them.

Instead, our theory of change is oriented toward those students who are trying out computer science, but not yet sure if it is for them, which is more often the case for females and underrepresented minorities. Those users, upon starting with RadGrad, can use it to learn about the spectrum of disciplinary interests and career goals that they can prepare for through curricular and extracurricular activities. For new students, our theory suggests that RadGrad can help them explore interests and career goals though a "low investment" activities, such as joining clubs or attending one day or weekend events like hackathons. This exploratory phase can help students discover appealing career goals, disciplinary interests, and like-minded communities of practice, which the literature suggests will improve engagement and retention for females and underrepresented minorities. As they discover the interests and career goals that appeal to them, RadGrad can help those students become aware of extracurricular activites that require a greater investment of time and energy, such as research projects and internships. Our theory proposes that students who transition to high investment extracurricular activities are very likely to complete the undergraduate degree program.

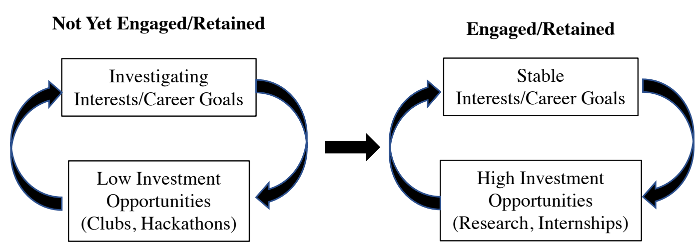

Figure 2: *DEP/RadGrad Theory of Change*

4 Project Plan

Figure 3 illustrates our project plan. Our plan is based on a project start date of September, 2018 with a duration of three years, and six major task categories. Tasks are organized around the Fall and Spring semesters. We will not explicitly collect data during summer semester, although RadGrad users can record their summer discipline-related activities (courses, internships, etc.) into their Degree Experience Plan which will then be picked up during the subsequent assessments. The next section describes the tasks and indicates the personnel who are primarily responsible for them.

UHM Deployment (Johnson, Moore): In this task, we will complete the deployment of RadGrad into the UH Manoa computer science department. Deployment means making sure that all computer science students are made aware of the existence of RadGrad, and that any student who desires to use the system receives the 25 minute training. We will use departmental mailing lists to publicize RadGrad. We will conduct trainings in the second semester introductory programming course to ensure that all computer science students currently in their second semester receive RadGrad training and develop an initial Degree Experience Plan. For higher level students, we will provide voluntary training sessions in the afternoons and evenings. We expect at least 50% of computer science students to have received training and developed their initial Degree Experience Plan by the end of this semester.

UHM Baseline Assessent (Johnson, Paek): At the beginning of Spring 2019 semester, we will conduct a "baseline assessment", in which we will gather data to assess adoption (A), diversity (D), engagement (E), and retention (R). We call this a baseline assessment because the impact of RadGrad at that point will be minimal, and because it will provide values that we can compare against in future assessments. (Section 5 defines A, D, E, and R in more detail.)

	Fall 18	Spring 19	Fall 19	Spring 20	Fall 20	Spring 21
UHM Deployment						
UHM Baseline Assessment		A, D, E, R				
UHM Registration		*	*	*	*	*
UHM Impact Assessment			A, D, E, R	A, D, E, R	A, D, E, R	A, D, E, R
UHM Interview Assessment		*		*		*
External Deployment						

Figure 3: *Project Plan.*

UHM Registration (Johnson, Paek, Leong): This task occurs for the first time at the beginning of Spring 2019 and repeats at the beginning of each subsequent semester for the remainder of the study period. In this task, we use 25 minutes of one lab period associated with our second semester introductory programming course to introduce students to RadGrad and help them develop their initial Degree Experience Plan. Except for a small number of transfer students, every computer science major takes this course, which means that after a few semesters of UHM Registration, almost all computer science students will know RadGrad and have developed an initial Degree Experience Plan. For the few remaining students who are transfers or missed the training, we will provide voluntary training sessions each semester.

UHM Impact Assessment (Johnson, Paek, Leong): Starting in Fall 2019, we will conduct an "impact assessment" each semester, which will gather data to assess current levels of adoption (A), diversity (D), engagement (E), and retention (R), then compare against previous semesters and the Baseline Assessment.

UHM Interview Assessment (Johnson, Paek, Leong): At the end of each academic year in the Spring, we will perform in-person or over-the-phone interviews of 20 graduating students, 20 students who have left the computer science program, and 20 graduates from the prior year who are now working or in graduate school. At least half of the students in each of these groups will be women or underrepresented minorities. These conversations will be recorded for subsequent analysis. In these interviews, we will obtain qualitative data regarding student perceptions of the causal factors underlying engagement, diversity, and retention.

External Deployment (Johnson, Moore): In the last year of the project, we will deploy DEP/RadGrad to the Department of Electrical Engineering program in Computer Engineering, and to the Department of Computer Science at San Jose State University. This task involves initializing the system with courses, degree program requirements, and extracurricular activities appropriate to each student community. It also involves training faculty and advisors on the use of the system and how to conduct the 25 minute introductory session with students and the development of their initial Degree Experience Plan. This task enable us to assess the readiness of DEP/RadGrad to support institutional transformation (i.e. within the University of Hawaii) as well as community transformation (i.e. in computer science departments at different institutions).

5 Research Design

This section presents the experimental design we will use to investigate the research questions presented in Figure 1.

Our design is a longitudinal, blocked, within subjects, observational study involving collection of both qualitative and quantitative data from the following three sources: (1) STAR, a UH institutional system that can identify the students enrollment in departmental classes each semester along with the self-identification of these students as female and/or Native Hawaiians; (2) RadGrad instrumentation, which will reveal the frequency and type of usage of RadGrad by each student, and (3) Interview data, providing insights into student attitudes not available through STAR or RadGrad.

5.1 Degree experience data model (DEDM)

Our research design is built upon an approach to collecting data about students and their use of RadGrad which we call the *degree experience data model* (DEDM). We will build the DEDM by taking snapshots of our student population each Fall and Spring semester over the course of the project period. The model allows us to evaluate both the impact over time of DEP/RadGrad on the subject population as a whole, on population slices of interest, and on specific individuals.

Subject population: Our subject population for each snapshot consists of all students enrolled in courses intended for majors during that semester. For example, this consists of 53 courses for the UH computer science department: two 100-level courses, six 200-level courses, eleven 300-level courses, and thirty four 400-level courses. Currently, approximately 150 students are enrolled in 100-level courses each semester, dropping to approximately 75 per semester in 400-level courses. The members of the subject population will change each semester as new students enter the program and old students either graduate or abandon the program.

Female, Native Hawaiian: For each student, we will access institutional records to determine whether they self-identify as female and/or (pure or part) Native Hawaiian. Our experimental design focuses on Native Hawaiians as a means to begin understanding the impact of our intervention on underrepresented minorities. The DEDM indicates a female subject with *F+* and a Native Hawaiian with *NH+*. Non-females and non-Native Hawaiians are represented with *F-* and *NH-*.

GradeLevel: We will assign each student in a semester snapshot to one of the following GradeLevels based upon the number of semesters in which they have taken CS courses: *Year1* (one or two semesters); *Year2* (three or four semesters); *Year3* (five or six semesters); *Year4* (seven or eight semesters); or *Year5* (nine or more semesters). For the UH computer science department, the number of students in each Grade Level generally parallel the course levels: approximately 150 Year1s, dropping to 75 Year4s and Year5s. Most students completing the degree program move through at least the first three GradeLevels.

RadGrad active: We will categorize each student in a semester snapshot as either an active user of RadGrad (*RadGrad+*) or not active (*RadGrad-*). To be classified as *RadGrad+* during a semester snapshot, the student must have: (a) logged into RadGrad at least once, and (b) changed their Degree Experience Plan (for example, by updating their set of Interests, Career Goals, or Opportunities). All other students will be classified as *RadGrad-*, indicating no modification of their DEP during that semester. While we intend for all students in a department to receive RadGrad training and develop an initial Degree Experience Plan, any further use of DEP/RadGrad is still voluntary and optional.

DEP Evolution: RadGrad's instrumentation allows us to see how interests, career goals, planned activities, and ICE points all evolve over time.

Figure 4 illustrates aspects of the DEDM over a hypothetical sequence of four semesters. Each semester, we collect a snapshot of data about the subject population, as illustrated by the four boxes labelled Fall 18, Spring 19, Sum 19, and Fall 19. The height of the boxes indicates the population total for that semester; for example, there were slightly more than 400 students in Fall 18, increasing to almost 500 students by Fall 19. Each box representing a semester is divided into five segments labelled 1 to 5. These segments represent the number of the subjects classified as *Year1*, *Year2*, *Year3*, *Year4*, and *Year5*. The figure illustrates that in this hypothetical example, retention is an issue: *Year1* students are always the largest numerical group within a snapshot, and the number of students drops significantly with increasing GradeLevel.

Within each GradeLevel segment of each snapshot, the figure contains is a vertical bar with a "+" on one side and a "-" on the other. This vertical bar partitions the segment into the students who are *RadGrad+* vs. *RadGrad-*. The figure illustrates a situation in which *Year1*s are always mostly *RadGrad-*, and *Year5*s are always mostly *RadGrad+*, though the precise division varies from semester to semester. Note that we could also partition GradeLevel by *F+*, *F-* and *NH+*, *NH-* to see how gender and underrepresented minority representation changes with GradeLevel.

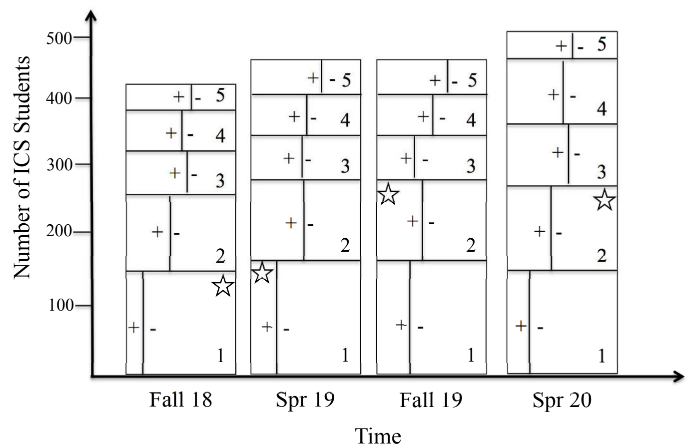

Figure 4: *Degree experience data model*

Finally, the star icon illustrates that we can track individual students as they change demographically over the course of their degree program. The hypothetical student represented by the star icons is a *Year1 RadGrad-* in Fall 18, a *year1 RadGrad+* in Spr 19, a *Year2 RadGrad+* in Sum 19, and a *Year2 RadGrad-* in Fall 19. Although it is not represented in the diagram, the DEDM will also indicate whether that student is *F+* or *F-* and *NH+* or *NH-*.

5.2 Assessing adoption

The DEDM provides a straightforward way the measure RadGrad adoption: it is simply the percentage of *RadGrad+* students. Depending upon our analytic needs, we can compute adoption on a per semester basis over the entire subject population, by GradeLevel, aggregated over one or more semesters, or based on a slice the population that is *F+* and/or *NH+* students.

While we expect adoption to be high based upon our pilot studies, we do not expect universal adoption. If the overall adoption rate per semester is approximately 75%, and the overall subject population per semester is at least 400, then there will be at least 100 *RadGrad-* students per semester that we can use as a proxy for a "control" group to compare against the *RadGrad+* "treatment" group. (Section 5.9 discusses limitations with this approach.)

5.3 Assessing engagement

According to [23], engagement measures the extent to which student are participating in educational practices that are strongly associated with high levels of learning and personal development. The North American National Survey of Student Engagement provides a generic instrument for assessing student engagement, though some researchers raise concerns with the application of these and other generic tools to computer science [24].

Based on the literature, we believe that ICE scores provide a valid, if not superior, alternative measure of engagement for undergraduates in computer science programs because they provide a fine-grained measure of verified student involvement with faculty-curated educational activities, both curricular and extracurricular. (Section 5.9 discusses limitations with this approach.)

To measure the impact of DEP/RadGrad on engagement, we will compare the average number of ICE points per *RadGrad+* student in a given GradeLevel over the first two semesters of the project period to the subsequent four semesters of the project period. We interpret ICE points during the first semesters as a form of "pre-test" measure of engagement, i.e. these scores measure student engagement before RadGrad became a part of the subject population's undergraduate experience. In the final four semesters, we will be able to measure engagement over time for students for whom RadGrad has been a part of their degree experience since their second semester. We exclude *RadGrad-* students in order to increase the internal validity of ICE as a measure of engagement.

If DEP/RadGrad has a postive impact on engagement, then we predict that the average number of ICE points per *RadGrad+* student per GradeLevel will be significantly higher in the final four semesters compared to the first two semesters. For example, we predict that the average number of ICE points earned by *Year3 RadGrad+* students in Fall 2020 will be significantly more than the average number of ICE points earned by *Year3 RadGrad+* students in Fall 2018. Note that impact will attenuate if more and more students achieve 100 points for each of the three categories, which DEP/RadGrad defines as the goal state for undergraduate preparation.

5.4 Assessing retention

The literature suggests that increased disciplinary engagement leads to increased retention. So, if DEP/ RadGrad increases engagement (as assessed above), then we can predict that DEP/RadGrad will make a positive impact on retention.

In this study, we measure retention as follows: for any given semester, a student is considered "retained" if they increase at least one GradeLevel or graduate within two semesters. We can use this definition to calculate an overall "retention rate" for the entire population on a semester-by-semester basis, which is the percentage of retained students in the subject population for that semester. We can also calculate retention rates for subpopulations, such as the retention rate for *FirstYear* students in Spr 18. Note that because we require data from the following two semesters to calculate retention for any given semester, we will only be able to calculate retention for the first four semesters of this six semester project.

Given our definition of retention rate, we predict that the retention rate for *RadGrad+* students will be higher than for *RadGrad-* students in any given semester, regardless of GradeLevel. However, we also expect this difference to be greater at lower GradeLevels, since *Year4* and *Year5* students are close to graduation and highly motivated to finish, so there is less attrition at this level.

It is possible that improved retention among *RadGrad+* students will have a ripple effect onto *RadGrad-* students (i.e. a rising tide lifts all boats). To see if that phenomena occurs, we will compare retention rates among *RadGrad-* students for a given GradeLevel over time. A ripple effect will manifest itself by an improving retention rate over time: for example, the retention rate for *Year1 RadGrad-* students in Spring 2020 will be higher than the rate for *Year1 RadGrad-* students in Spring 2018.

We must stress that our results regarding retention rates must regarded as preliminary at the end of the project period because they will be based on only four semesters of data. That said, the project will have put in place evaluation mechanisms that will enable us to provide useful conclusions regarding this question within a year or two after the project period.

5.5 Assessing diversity

If DEP/RadGrad has a positive impact on diversity, then if RadGrad is adopted by *F+* and *NH+*, then we predict that the percentage of *F+* and *NH+* students in a semester snapshot at the end of the project period will be significantly higher than the percentage of *F+* and *NH+* students in a semester snapshot at the beginning of the project period.

Because diversity is related to engagement and retention, we can assess these measures for *F+* and *NH+* students by doing the analyses as described above, but restricting ourselves to these subsets of our subject population.

Based upon the literature, we expect to observe high levels of adoption, plus positive changes in engagement and retention for females and Native Hawaiians over the course of the project period.

5.6 Assessing the DEP/RadGrad Theory of Change

Our theory of change suggests that RadGrad can improve engagement, retention, and diversity by providing a two-step pathway. An initial exploratory phase allows students to learn about different interests and activities with low investment extracurricular activities. RadGrad can then help students commit to the program with higher investment activities like research projects and internships that relate to their specific interests and career goals.

The DEDM records the evolution of interests, career goals, and planned and actual courses and opportunities for students. This information will enable us to see if the DEP/RadGrad Theory of Change manifests itself, and if so, whether it is associated with females and underrepresented minorities.

5.7 Investigating causality through interview data

The above assessments of engagement, retention, and diversity will provide evidence only of correlation; for example, we introduced DEP/RadGrad, and we saw changes in engagement. Because we are not randomly assigning subjects to control and treatment groups, the above analyses cannot support statistically-based evidence of a causal linkage.

To provide evidence for a more causal link between DEP/RadGrad and engagement, diversity and retention, we will conduct a series of interviews with students. We will sample students from all of our demographic categories: *RadGrad+*, *RadGrad-*, *F+*, *F-*, *Year1* through *Year5*, and both retained and unretained students. During this interview, we will ask them to evaluate their experience in the degree program, to list the experiences that were most useful to them, the obstacles they encountered, and what they would recommend about the program to future computer science students. To incentivize participation, we will offer a gift card. The interviews will be taped for subsequent coding and analysis.

In these interviews, we will not explicitly ask about RadGrad participation so that subjects are not biased toward it in their answers. Instead, we will analyze the responses for evidence that the student found DEP/RadGrad to be helpful, or (for *RadGrad-* students) that DEP/RadGrad would have proven useful to overcoming their obstacles. For example, if a student mentions that they found ICE points and Levels to be an important part of their undergraduate degree experience, then that provides causal evidence for the utility of RadGrad for that student. If a student recommends that future students use RadGrad, or engage in activities promoted by RadGrad, then that also provides causal evidence for the utility of RadGrad.

5.8 Addressing the research questions

This research design is intended to address all of the research questions in Figure 1. Our interview data will provide data on the factors that help or hinder adoption of DEP/RadGrad. Our measurements will enable us to assess whether or not successful adoption of DEP/RadGrad makes a positive impact on engagement, retention, and diversity. We can create submeasurements by various student demographics (year in the program, male or female, underrepresented minority or not) in order to determine if there are differences in engagement, diversity, and retention based upon demographic. Finally, our pilot programs in the third year of the project will provide evidence regarding the factors that help or hinder adoption of DEP/RadGrad in other disciplines and institutions.

5.9 Threats to validity

One threat to this experimental design is extremely high or extremely low adoption. If adoption is so high that there are almost no *RadGrad-* students, then there is no group to compare to *RadGrad+*. Conversely, if adoption is so low that there are no *RadGrad+* users, then the entire experimental design falls apart. Based upon our pilot studies, in which students reacted quite positively to RadGrad, we are hopeful that adoption will be high but not total.

A second threat is the presence of a large number of engaged students who are *RadGrad-*. These are students who are academic high achievers and who participate in extracurricular disciplinary activities, but who do not want to represent their degree experience in RadGrad. Again, our pilot studies provide contrary evidence: the subset of students we approached who evidenced high engagement were among the most enthusiastic about using RadGrad.

A related threat is that our decision to use ICE scores to measure engagement means our engagement data cannot be compared to engagement data gathered using generic surveys such as NSSE. A supplemental outcome of this project will be a study in which we administer the NSSE survey to a random selection of students, then compare this data to their ICE scores to see if a correlation exists.

As already noted, our design does not randomly assign students to the *RadGrad+* and *RadGrad-* groups, which weakens the interpretation of our empirical results. We nonetheless believe our design is in the best interests of the student population, and believe that over longer time and with replication, high quality understanding of the strengths and weaknesses of DEP/RadGrad will emerge.

Finally, the computer science department associated with this study is not static. New initiatives, teaching approaches, and faculty are likely to be introduced over the course of the project. Any of these could also have a significant positive impact on engagement, retention, and diversity (and, to be honest, we hope that they will.) As such changes appear, we will use interview data to gather evidence as to their potential impact, and they will be incorporated into the interpretation of the results from this project.

6 Broader Impacts

A first and foremost broader impact of this study is the generation of empirical insight into an approach for increasing retention and diversity across STEM disciplines, not just computer science.

A second broader impact is the development of production-ready open source technology for use across STEM disciplines to support planning of both curricular and extracurricular activities, and their relationships to interests and career goals. RadGrad has a modular structure and extensive documentation to support tailoring and enhancement as new use cases arise.

A third broader impact is the data that deployment of DEP/RadGrad makes available about the undergraduate STEM degree experience. Current research on retention and diversity suffers from a lack of detailed understanding of the disciplinary activities that students undertake, as well as how their interests and career goals form and change over time. DEP/RadGrad provides this data for researchers, while providing value to the students who provide it.

A fourth broader impact is the generation of new insights into Individualized Learning Plans and Communities of Practice in the context of the undergraduate degree program.

7 Project Sustainability

(Plans to sustain the project beyond the study period will go here.)

8 Results from prior NSF support

P. Johnson, *Human centered information integration for the Smart Grid*, NSF Grant IIS-1017126, 8/15/10 to 7/31/14, $413,467. *Intellectual Merit:* Insight into: the inadequacy of baseline data for energy competition research, experimental studies for assessing energy behaviors, energy competitions incorporating educational activities. *Broader Impacts:* two open source systems, WattDepot and Makahiki; data regarding energy education and gamification techniques. Selected publications: [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35].

References

[1] Rita Gunther McGrath. The Pace of Technology Adoption is Speeding Up, November 2013.

[2] Till Leopold. The future of jobs. Technical report, World Economic Forum, 2016.

[3] Tracy Camp. Generation CS: Computer Science Undergraduate Enrollments Surge Since 2006. Technical report, Computing Research Association, 2017.

[4] Jill Barshay. U.S. now ranks near the bottom among 35 industrialized nations in math, December 2016.

[5] Catherine Fry. Achieving systemic change: A sourcebook for advancing and funding undergraduate STEM education. Technical report, Association of American Colleges and Universities, 2014.

[6] Hai Hong and Abby Bouchon. Women who choose computer science: what really matters. Technical report, Google, Inc., 2014.

[7] Hadi Partovi. Code.org's plan for diversity in K-12 computer science. Technical report, Code.org, 2017.

[8] Maya Kosoff. Silicon Valley's sexual-harassment crisis keeps getting worse. *Vanity Fair*, September 2017.

[9] Engineering National Academies of Sciences. *Assessing and Responding to the Growth of Computer Science Undergraduate Enrollments*. The National Academies Press, October 2017. DOI: 10.17226/24926.

[10] Kyle Thayer and Andrew Ko. Barriers faced by coding bootcamp students. In *Proceedings of the ACM International Computing Education Research Conference*, 2017.

[11] Linda Sax. BRAID: A Diversity Program, 2017.

[12] Reshma Saujani. Girls Who Code, 2017.

[13] Kimberly Bryant. Black Girls Code, 2017.

[14] Catherine Ashcraft and Anthony Breitzman. Who invents it? Women's participation in information technology patenting. Technical report, National Center for Women and Information Technology, 2012.

[15] Cristian Dezso and David Ross. Girl Power: Female participation in top management and firm performance. Technical report, University of Maryland, 2007.

[16] Google Inc. Diversity Gaps in Computer Science: Exploring the Underrepresentation of Girls, Blacks and Hispanics. Technical report, Google, Inc., 2016.

[17] Nancy Kober. *Reaching Students: What Research Says About Effective Instruction in Undergraduate Science and Engineering*. National Academies Press, January 2015. DOI: 10.17226/18687.

[18] Jerri Barrett. Expanding the Pipeline: Key Learnings on Retaining Underrepresented Minorities and Students with Disabilities in Computer Science, November 2017.

[19] Jean Lave and Etienne Wenger. *Situated Learning: Legitimate peripheral participation*. Cambridge University Press, 1991.

[20] Etienne Wenger. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, 1998. Google-Books-ID: Jb8mAAAAQBAJ.

[21] Etienne Wenger, Mark O'Creevy, Steven Hutchinson, Chris Kubiak, and Beverly Wenger-Traynor. *Learning in landscapes of practice*. Routledge, 2004.

[22] Scott Solberg, Joan Wills, Kimether Redmon, and Laura Skaff. Use of Individualized Learning Plans: A Promising Practice for Driving College and Career Efforts. Technical report, National Collaborative on Workforce and Stability, Washington, DC, 2014.

[23] McCormick. National Survey of Student Engagement. Technical report, National Survey of Student Engagement, 2017.

[24] Jane Sinclair, Matthew Butler, Michael Morgan, and Sara Kalvala. Measures of Student Engagement in Computer Science. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, 2015.

[25] Robert S. Brewer and Philip M. Johnson. WattDepot: An open source software ecosystem for enterprise-scale energy data collection, storage, analysis, and visualization. In *Proceedings of the First International Conference on Smart Grid Communications*, pages 91–95, Gaithersburg, MD, October 2010.

[26] Robert S. Brewer, George E. Lee, and Philip M. Johnson. The Kukui Cup: a dorm energy competition focused on sustainable behavior change and energy literacy. In *Proceedings of the 44th Hawaii International Conference on System Sciences*, pages 1–10, January 2011.

[27] Robert S. Brewer. *Fostering Sustained Energy Behavior Change And Increasing Energy Literacy In A Student Housing Energy Challenge*. PhD thesis, University of Hawaii, Department of Information and Computer Sciences, March 2013.

[28] Robert S. Brewer, George E. Lee, Yongwen Xu, Caterina Desiato, Michelle Katchuck, and Philip M. Johnson. Lights Off. Game On. The Kukui Cup: A dorm energy competition. In *Proceedings of the CHI 2011 Workshop on Gamification*, Vancouver, Canada, May 2011.

[29] Robert S. Brewer. The Kukui Cup: Shaping everyday energy use via a dorm energy competition. In *Proceedings of the CHI 2011 Workshop on Everyday Practice and Sustainable HCI*, Vancouver, Canada, May 2011.

[30] Philip M. Johnson, Yongwen Xu, Robert S. Brewer, Carleton A. Moore, George E. Lee, and Andrea Connell. Makahiki+WattDepot: An open source software stack for next generation energy research and education. In *Proceedings of the 2013 Conference on Information and Communication Technologies for Sustainability (ICT4S)*, February 2013.

[31] George E. Lee, Yongwen Xu, Robert S. Brewer, and Philip M. Johnson. Makahiki: An open source game engine for energy education and conservation. Technical Report CSDL-11-07, Department of Information and Computer Sciences, University of Hawaii, Honolulu, Hawaii 96822, January 2012.

[32] Robert S. Brewer, Yongwen Xu, George E. Lee, Michelle Katchuck, Carleton A. Moore, and Philip M. Johnson. Energy feedback for smart grid consumers: Lessons learned from the Kukui Cup. In *Proceedings of Energy 2013*, pages 120–126, March 2013.

[33] Robert S. Brewer. Three shifts for sustainable HCI: Scalable, sticky, and multidisciplinary. In *Proceedings of the CHI 2014 Workshop "What have we learned? A SIGCHI HCI & Sustainability community workshop"*, Toronto, Canada, April 2014.

[34] Robert S. Brewer, Yongwen Xu, George E. Lee, Michelle Katchuck, Carleton A. Moore, and Philip M. Johnson. Three principles for the design of energy feedback visualizations. *International Journal On Advances in Intelligent Systems*, 3 & 4(6):188–198, 2013.

[35] Yongwen Xu, Philip M. Johnson, Carleton A. Moore, Robert S. Brewer, and Jordan Takayama. SGSEAM: Assessing serious game frameworks from a stakeholder experience perspective. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications (Gamification 2013)*, October 2013.