Computer Science is "Hard": Uncovering Cultural Identities Within Introductory Computing Courses

Submitted by Brittany Ann Kos, Doctoral Student, ATLAS
Advised by Dr. Diane Sieber, Associate Dean for Education, College of Engineering
Collaborating with Dr. Dirk Grunwald, Professor, Computer Science Department

1. PROJECT OVERVIEW

1.1 Background

The purpose of this study is to understand how cultural norms permeate introductory computing courses and recognize how students adopt or reject these identities in their academic careers. This study will focus on 2 introductory computing classes that serve three major degree programs: 1. B.S. in Technology, Arts, and Media (TAM-BS), 2. B.S. in Computer Science (CS-BS), and 3. B.A. in Computer Science (CS-BA). Each of these degree programs attracts a different type of student, but also have significant overlap in their degree requirements with one another. The TAM-BS and CS-BS degrees share many of the same humanities, science, and math requirements, and the CS-BS and CS-BA degrees share the same core Computer Science (CS) classes. All three of these programs share similar first years, since all students are required to take CS1: Introductory Programming (CSCI 1300), CS2: Data Structures (CSCI 2270), and Calculus 1 and 2. These introductory programming classes (CS1 and CS2) act as "gateway" classes. Students are required to take these early in their academic careers, since they are often prerequisites in more difficult or upper-division courses. This study will focus on these two classes and the impact they have on student's academic careers.

Within the introductory classes we will be concentrating on how the class develops its cultural norms and practices. In this study will use Goodnow's definition of cultural practices: "practices are actions that are repeated, shared with others in a social group, and invested with normative expectations and with meanings or significances that go beyond the immediate goals of the action" (1995). A social group can be as small as two individuals or as large as the university. In this study, we will be primarily focused on two types of social groups: the classroom setting and the department or major program; each of these settings will be viewed from the perspective of a student going through the program. As the students are exposed to these courses and departments, they will be forming personal identities that align or break from the culture that has been set up. These identities are tied directly with learning and cultural norms (Stevens et al, 2008). These identities will be formed early in the student's academic carers and can harm or assist the student's trajectory through each degree program.

1.2 Goals and Objectives

Ultimately this research has three goals; first, to discover how cultural norms develop, persist, and become enforced within these social settings, second, to uncover ways students adopt or reject these identities, and finally, how this process transforms the student's trajectory in their degree program. Ideally this investigation will cover a student's course through the introductory computing classes and a semester beyond. In the computing community, this will be a novel way to examine why students choose to stay in the field or leave within their first year. This study will break from traditional computer science education by incorporating theory, methods, and research from the learning sciences and sociology. This multidisciplinary approach will bring new perspectives to the academic community, highlight undiscovered cultural practices within the courses, inform teaching practices and departmental structure, and give students the tools to navigate these courses.

Research Questions:

1. How are cultural norms within CS1 and CS2 constructed and enforced by members of the community?
2. How are these cultural norms adopted or rejected by students in these courses?
3. How do these cultural norms affect student trajectories in these courses and within their degree program?

2. BACKGROUND

2.1 The Growing Computing Field and its Lack of Diversity

Computer Science and Information Technology are some of the fastest growing fields in America right now. Forbes reported that from 2013-2014, there was 14% growth in the computer systems and related services field (Biery, 2014). Computing jobs also account for 10% of the jobs listed by the U.S. News' report *The 100 Best Jobs* (2015). Despite these positive outlooks from the industry, few women and minorities enroll in computing fields. Nationally, women only hold 25% of the computing and mathematics jobs and people of Hispanic, African-, or Native-American descent only hold 10% of the engineering and science jobs (NSF, 2014). There is a large discrepancy between the gender and ethnicity of our national population and those few who earn a computing degree and stay within the field. This leads to a severe misrepresentation of the number of people creating and developing technology.

The problem of diversity in Computer Science has been addressed by many researchers in the last few decades. Some of the most notable work has been conducted by Margolis and her colleagues, their work has concluded that there are many factors that influence women and minorities to pursue a career in computing (2003, 2008, Fisher 1997). Some of these factors include early exposure, confidence gaps, perceived and actual ability, purpose, and attachment. These are all factors that persist in college and continue to influence students in their introductory computing courses. The previous research suggests that there is never one factor that affects student's desire to continue in the computing fields, but it is a combination of societal and cultural influences that persuade students to or away from these fields. The research we will be conducting looks at the holistic view of a student's experience within the introductory courses. We will be focused on connecting these different factors and determining the weight of their influence rather than isolating each element from its context.

The Computer Science community has made great strides in determining factors that deter females from continuing with computer science classes; however, there is a significantly smaller body of work focusing on the racial and ethnic divide that is also present in computing fields. In the last six years, the ACM Special Interest Group in Computer Science Education (SIGCSE) has published a number of papers focused on diversity, recruitment, and retention. In that time frame, more than half of the publications focused on recruiting women, more than a quarter were focused on minority women, but less than 15% investigated recruitment and retention issues in minority students (ACM, 2010, 2011, 2012, 2013, 2014, 2015).

2.4 Project Motivation

Due to the scattered nature of the current state of research on student recruitment and retention, we propose a paradigm shift in the focus of computing educational research. This project does not ask why *students* leave computing programs, but ask what *introductory classes* are doing to indicate that students should leave computing programs. This is a subtle, but novel shift from most research being conducted in computing education today. We will be focusing on the cultural agency the introductory courses and different departments exert on their students. This frames the problem of diversity at the departmental level instead of at the student level. This reframing has 2 benefits: 1. students can not longer be held accountable for leaving computing because they do not fit within the enforced departmental cultural norms and 2. the department will be in a position to change student recruitment and retention.

3. METHODOLOGY

3.1 Conceptual Frameworks

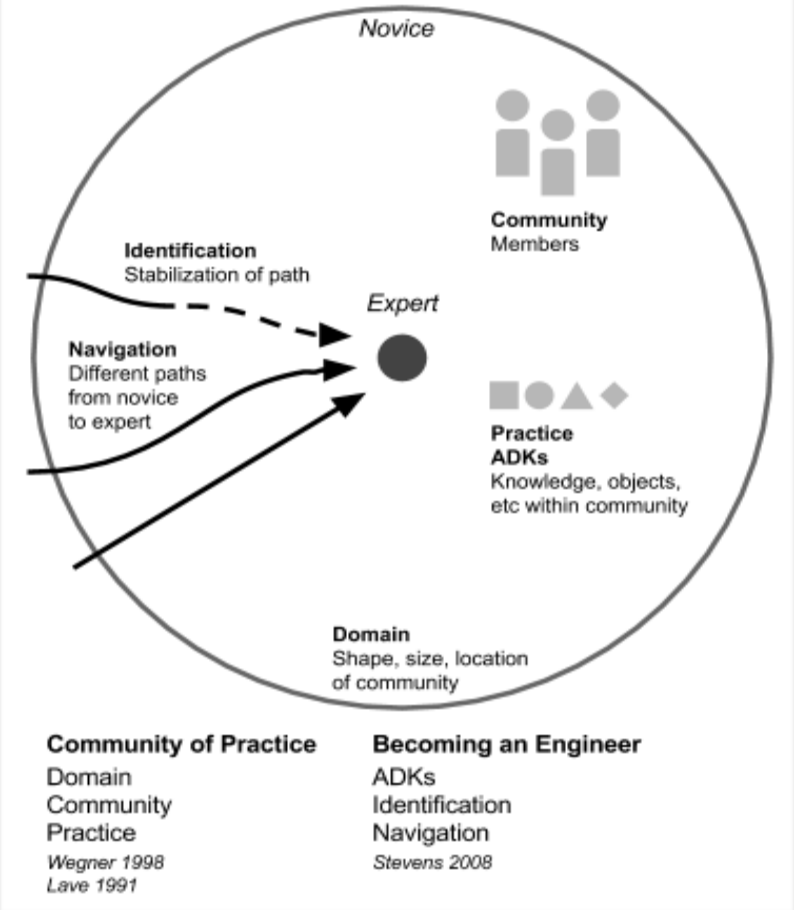
This study has three intersecting conceptual frameworks: Communities of Practice, Becoming an Engineer, and Mediated Discourse Analysis. These frameworks will be employed to gain a full understanding of a student's trajectory through the introductory classes of the three degree programs.

The idea of a Communities of Practice was first formed by Lave and Wenger (1991) and then reconceptualized by Wenger et al (1998). These communities can be defined as "groups of people who whate a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis" (1998). Communities of practice are structured by three elements: the domain, of the common identity; the community, or the social relationships; and finally the practice, or the framework, tools, and objects community

members use. In this study, we will be viewing the introductory courses and each degree program as a separate community of practice. Members can belong to multiple communities, but each community has it own set of goals, cultural norms, and expectations.

"Becoming an Engineer" is an analytical framework that was proposed by Stevens et al (2008). It involves three dimensions that track how a student "becomes" an engineer over time: accountable disciplinary knowledge (ADKs), identification, and navigation. We will be re-appropriating this framework for computing students, but using the same framework. This framework will be used in tandem with the communities of practice framework. "Becoming an Engineer" will inform how a student moves through a community and navigates their path from in a community from novice to expert. Figure 1 shows a representation of how we are visualizing communities of practice and how students "become" engineers.

Figure 1: Theoretical Framework



Finally, Mediated Discourse Analysis (MDA) is described as "a theory about social action with a specific focus on discourse as a kind of social action as well as upon discourse as a component of social action" (Scollon, 2001a). MDA provides the analysis methods and tools we will be using in our study. MDA uses mediated action, or community members using Discourse in the social world, as its unit of analysis. This allows us to focus our research on human *action*, not just discourse, or text. This also affords us to examine the relationship between discourse and social action in its correct context (Norris, 2005).

3.2 Research Design

This study will be conducted in 4 phases. The design of this study is an adaptation of *Methods of Critical Discourse Analysis* (Scollon, 2001b).

Phase 1: General ethnographic observation

This phase will be conducted in the classroom and study spaces each program provides. The purpose will be to identify which mediated means the students are using. We will focus on these mediated means as mechanisms that help students adopt or reject the program identities.

Phase 2: Specific ethnographic observations and informal interviews

This phase will begin once specific mediated means from Phase 1 have been identified. There will be two parts to this phase. The first part of this phase will be informal interviews with the students, TAs, instructors and other key personnel of these courses. The interviews will serve to substantiate what we've observed from phase 1. The second part will be continued ethnographic observation. We will study the scene the mediated means are taking place in, look at the events surrounding the means, and focus on the actions students are taking when participating in these means.

Phase 3: Specific semi-interviews

This phase will further investigate the mediated means that were found in phase 2 and confirmed in phase 3. We will be conducting semi-formal interviews with the participants about the scenes, events, and actions surrounding the means.

Phase 4: Focus groups

This will be the final stage of the study and will consist of focus groups, where we will hold additional semi-formal interviews with the students. These interviews will serve to validate, instead of investigate, the research findings.

3.3 Analysis Process

These phases will be qualitative in nature and will either be ethnographic observations, informal interviews, semi-formal interviews, and focus groups. All of the data will be collected through notes and recordings. The recordings will be continually transcribed throughout the study to keep an updated record of the data collected. The data will be continuously coded and analyzed to find trends within the scenes, events, and actions. This analyzed data will provide the groundwork for the following phases.

For each phase, we will focus on the factors that seem to be affecting student identity within these programs. After identifying the factors that help or hinder the adoption of a specific identity, we will focus on the mediational means that influence those factors. The interviews and focus groups will keep narrowing in on those identified mediational means to determine what specific means affect student identity in what way.

4. TIMELINE

4.1 Preliminary Findings

In the spring of 2015, we have been in the process of collecting data on two different introductory CS classes. This data was collected before the start of the TAM-BS degree, so the classes are from two different departments (CS and TAM). We have 16 hours of lecture recordings and 20 hours of observation notes. We are planning on doing data evaluation this summer on this data set.

We do, however, have a preliminary analysis from this data set that scrutinizes the way material is presented in these two classes. We have discovered a large discrepancy in the way students, professors, and teaching assistants refer to the course material. The CS classes acknowledge the difficulty of the material being covered and have normalized this type of discourse in the classroom. These classes have common phrases, such as "Computer science is hard" or "This isn't intuitive", which are usually referring to CS in general or the class overall. Alternatively, the TAM instructors almost never indicate how difficult the material might be, but instead focus on how hard the students will have to work to succeed. When students talk about their hardships, they refer to specific lessons or projects they are doing, not computing in general. We will be focusing on this difference more in depth is summer, but view these findings as a positive indicator that the frameworks and data collection methods we have chosen are valuable.

4.2 Project Timeline

We propose the timeline in Table 1. This is an aggressive timeline, but allows us to collect all the data from our research design in the time allotted.

Table 1: Proposed Project Timeline

	Summer 2015	Fall 2015	Winter 2015-16	Spring 2016	Summer 2016
Data Collection		Data collection from CS1 lecture Data collection from CS2 lecture Informal interviews		Data collection from CS1 lecture Data collection from CS2 lecture Semi-formal interviews	

		Send out survey		Focus Groups	
Analysis	Transcribe lectures from S15 Code data from S15	Build model cultural model for CS1 and CS2 Transcribe informal interviews	Transcribe lectures from F15-CS1 and F15-CS2 Code data from F15-CS1 and F15-CS2	Transcribe semi-formal interviews Transcribe focus groups	Transcribe lectures from S16-CS1 and S16-CS2 Code data from S16-CS1 and S16-CS2
Reports		Report S16 findings to SIGCSE '16	Report F15 findings to ICER '16	Kos does comprehensive exam with data	Report findings to SIGCSE '17

5. OUTCOMES

5.1 Benefits for CS and ATLAS

The undergraduate CS and ATLAS departments enroll hundreds of students each year and are rapidly growing every year. After the addition of the CS-BA degree in the fall of 2014, enrollment in the CS department nearly doubled and had another spike in enrollment this year (UCB). Currently, the TAM minor program enrolls over 600 students a semester and the addition of the TAM-BS degree the fall of 2015 is expected to draw in over 100 students initially. These students will be flocking to the introductory CS classes, when the department is still trying to recover from the addition of the CS-BA degree. Instructors and curriculum developers are still unprepared for this massive influx of students and face tremendous risk of losing students in these introductory classes, especially marginalized students.

This research will work to discover why students leave the program in this tumultuous time. Instructors will be able to use our findings directly in their classrooms and begin to counteract the attrition problem immediately and within the context of their large and diverse classrooms. If instructors are able to use this research in their classrooms, this can have a direct impact on the college's retention rates.

5.2 Benefits for CU

The benefits this study offers CU include offering the Discipline-Based Educational Research (DBER) community multidisciplinary theory methods, and research. Since this research has a foundation in engineering research, it also offers additional research that supports the methods and theory used in the engineering education community. This work can also become cross-disciplinary and inform other departments and degree programs. Finally, this work will be a paradigm shift in computing education, and will enrich CU Boulder's reputation for engineering and computing education research.

5.2 Personal Benefits

This study will represent the largest and most significant work I have contributed to. This project will become the basis for my dissertation and will provide the platform for future work I do in this area. The research collected and analyzed in this study will structure and inform my comprehensive exam in the spring of 2016. With support from the CGA and collaborators from computer science, I will be able to devote all of my time on the data collection and analysis that this study needs. Finally, this work will help me start my career as a computing education and discipline-based researcher. The CGA will give me the opportunity to conduct multidisciplinary research in a traditional discipline while incorporating learning science research.

6. REFERENCES

ACM (2010). In *SIGCSE '10 The 41th ACM Technical Symposium on Computer Science Education*. Milwaukee, WI, USA: ACM Press.

ACM (2011). In *SIGCSE '11 The 42th ACM Technical Symposium on Computer Science Education*. Dallas, TX, USA: ACM Press.

ACM (2012). In *SIGCSE '12 The 43th ACM Technical Symposium on Computer Science Education*. Raleigh, NC, MO, USA: ACM Press.

ACM (2013). In *SIGCSE '13 The 44th ACM Technical Symposium on Computer Science Education*. Denver, CO, USA: ACM Press.

ACM (2014). In *SIGCSE '14 The 45th ACM Technical Symposium on Computer Science Education*. Atlanta, GA, USA: ACM Press.

ACM (2015). In *SIGCSE '15 The 46th ACM Technical Symposium on Computer Science Education*. Kansas City, MO, USA: ACM Press.

Biery, M. E. The Fastest-Growing Industries Over The Last Year, 2014. Retrieved April 10, 2015, from SageWorks: Forbes.com: http://www.forbes.com/sites/sageworks/2014/04/28/the-fastest-growing-industries-over-the-last-year

Goodnow, J. J., Miller, P. J., & Kessel, F. (Eds.). (1995). *Cultural Practices as Contexts for Development*. Jossey-Bass Publishers.

Norris, S., & Jones, R. H. (2005). Discourse in Action: Introducing Mediated Discourse Analysis. Routledge.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Margolis, J., Estrella, R., Goode, J., Jellison Holme, J., & Nao, K. (2008). *Stuck in the Shallow End*. The MIT Press.

Margolis, J., & Fisher, A. (2003). *Unlocking the Clubhouse: Women in Computing*. MIT Press.

NSF et al. 2014. 2014 Science & Engineering Indicators.

UCB. Planning, Budget, and Analysis, 2014. Retrieved April 10, 2015, from University of Colorado Boulder: http://www.colorado.edu/pba/records/

Scollon, R. (2001a). *Mediated Discourse*. Routledge.

Scollon, R. (2001b). Methods of Critical Discourse Analysis. Routledge.

Stevens, R., Kevin, O., Garrison, L., Jocuns, A., & Amos, D. M. (2008). Becoming an Engineer: Toward a Three Dimensional View of Engineering Learning. *Journal of Engineering Education*, (July).

U.S. News and World Report. The 100 Best Jobs. Retrieved April 10, 2015, from U.S. News and World Report: http://money.usnews.com/careers/best-jobs/rankings/the-100-best-jobs

Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating Communities of Practice*. Harvard Business School Press.