

Teaching Professional Morality & Ethics to Undergraduate Computer Science Students through Cognitive Apprenticeships & Case Studies: Experiences in CS-HU 130 ‘Foundational Values’

Don Winiiecki, EdD, PhD
Boise State University
College of Engineering
Boise, Idaho USA
dwiniiecki@boisestate.edu

Noah Salzman, PhD
Boise State University
College of Engineering
Boise, Idaho USA
noahsalzman@boisestate.edu

Abstract—This proposal describes and details experience in guiding undergraduate computer science students to identify and address issues related to inclusion, diversity and social justice as they occur in computer science education and computer science professions. Specific details reported here arise from experience teaching a one-credit undergraduate course at Boise State University (CS-HU 130, Foundational Values).

Keywords— *Ethics, Professional Morality, Inclusion, Diversity, Social Justice, Computer Science, Higher Education*

I. INTRODUCTION

Even with only occasional attention on the popular press and academic research, one encounters a steady stream of stories describing how members of underrepresented groups in engineering fields have encountered personal and systemic instances of bias. These experiences are sometimes subtle, sometimes overt, and come from individuals and from the very systems in which people learn, work and interact with technologies [1]–[14].

This set of evidence makes one thing clear — attitudes (i.e., bias) that act to prefer some people and discount or dis-include others is a fact in the social world. Since computer science is part of the social world, its occurrence in CS is unsurprising. (It is also undeniable that this bias is officially excluded by the meritocratic fallacy in origin fables of computer science.)

With that in mind, the Computer Science department at Boise State University and its industry partners have committed to address this history of systemic bias in the field, across the undergraduate curriculum and into student internships in local industries. This effort has been substantively facilitated through an NSF RED (Revolutionizing Engineering & Computer Science Departments) grant that produces what we call the ‘Computer Science Professional’s Hatchery’ (aka CSP-Hatchery).^{1,2}

¹ ‘The Hatchery’ is named after a legacy industry in Idaho, USA, where Boise State University is located — the fish hatchery. Analogous to the purpose of fish hatcheries, the CSP-Hatchery is a planned ecosystem in which students are

This is manifested through an innovative curricular framework of one-credit courses that enables the BSU computer science department to rapidly develop and roll-out special-topics courses that address specific needs identified by local industry. Every undergraduate student is now required to take at least five of these one-credit courses, and in so doing become competent in particular technical and professional skills needed by their eventual employers. Targeted incentives are being used to facilitate faculty participation in this project.

The first of these one-credit courses (and the first Computer Science course required of all students) is CS-HU 130 ‘Foundational Values’. CS-HU 130 exposes students to case studies that document breaches in inclusion, diversity and social justice in both individual actions in computer science education and professional practice, and in the products of computer science work [3], for example, [8], [15]–[17].

For example, the experiences of Susan Fowler in her role as a ‘Site Reliability Engineer’ at Uber [8] highlights how one can assign responsibility for harassing behavior to individuals, but also to systemic failures in responsibility and accountability in Uber. Acting to remove the individual harasser would not also remedy the systemic problems in the workplace and could in fact allow new variants of the problem to arise in the future. This is not an engineering problem per se, but engineers acting as responsible agents in the social system can *engineer* hypothetically sustainable solutions to the problem.

Similarly, the failure of due diligence and responsibility by social institutions and computer science (machine learning) allowed for the occurrence of racial bias in criminal justice to be reified and ‘hardened’ through the use of the COMPAS software program intended to assign risk of recidivism among

prepared to thrive in the real world of professional computer science.

² As members of the ‘CSP-Hatchery’ project our expertise is not in computer science *as such*. Instead, we are a sociologist (Winiiecki) and an engineering education expert (Salzman). That said, we both also have long personal histories of involvement in engineering technology either vocationally or through academic study and credentialing.

individuals accused of crimes in the State of Florida [15]. Analyzing this case and identifying that the bias in machine learning reflected in COMPAS had a basis in unreliable data, and that a lack of awareness of basic social science facts allows computer science students to produce a very ‘engineering styled’ analysis of the problem and to prototype a path to sustainable solutions that go beyond computer science itself.

In accomplishing their analysis of these and other cases, students follow a rubric based on Rawls’ theory of justice [18], and models of organizational performance improvement [19]–[21], to propose systemic and sustainable changes in organizations that will help to reduce and remove influences that result in these breaches.

In addition, other courses in the curriculum are being developed and updated to include modules which connect traditionally ‘only technical’ content with socio-technical content that shows (a) how a ‘purely technical’ orientation to computer science actually produces an environment where undesirable bias can flourish, and (b) strategies and technical processes that act to mitigate and remove these underpinning factors from computer science and those affected by computer science.

In combination, these innovations aim to produce a learning ecosystem in which technical and social factors are always included in computer science learning and working communities. Through this process we aim to help produce a ‘new generation’ of computer scientists who understand the value in collaborating with social science experts and themselves becoming knowledgeable about the societal effects of computing. In turn, we aim to provide to current and emerging industries, entry-level professionals who are fluent in necessary technical skills and equipped with the moral sensitivity to identify and mitigate known and emerging risks and biases, all for the betterment of their organizations, and society and its members [22].

II. DISCUSSION

The core of CS-HU 130 ‘Foundational Values’ and the modules that incorporate a focus on inclusion, diversity and social justice into courses throughout the computer science curriculum is a framework for systematically analyzing case study material, and then proposing interventions that focus on individuals and on organizational forces that appear to be the principal causes of identified problems.

This framework is reflected in a worksheet used by students to analyze (i.e., take apart) the problems, and then synthesize (i.e., reassemble) systems of rules, incentives, and disincentives that address the identified problems in a way that may produce a more inclusive, diverse and just organizational whole. Note that we do not expect students to conclusively solve such problems in CS-HU 130, but rather to develop analytical skills and habits of mind that can be put to use in improving their learning and working environments as they grow.

This is accomplished through a variant of the ‘cognitive apprenticeship’ process, where the instructor leads a structured analysis and discussion of the case using the framework

students are expected to use [23]–[25] (the framework, an example case, and a worked-out example will be provided to attendees of this session). Students accomplish this activity four times in CS-HU 130. In the first instance, the instructor leads students through the entire framework, coaching students in the perspectives and process to be used. On the second, third, and fourth instances, the instructor begins this process but requires students to complete successively more of the framework in in-class laboratory-like sessions in teams of five. As in any laboratory class, the instructor consults with each student team, asking questions about their analyses, perspectives and processes, and guiding them to a coherent system of interventions involving rules, incentives and disincentives. The instructor intentionally questions groups about the unstated assumptions behind their analysis and their interventions in order to bring into explicit focus the basis of their lines of reasoning.

Across the three semesters this course has been offered, instructional process is being dynamically adjusted. One of these adjustments has produced a notable pattern of improvements in knowledge, comprehension, application, analysis and culminating synthesis³ of course content. This improvement occurs parallel with a straightforward application of the process of ‘fading’ – the gradual decrease of support offered to students as they demonstrate greater familiarity and fluency with the subject matter and its application in successively more complex problem sets. In this case, student teams demonstrate the ability to approach issues of bias, and loss of social justice as *problems associated with how social structures have been engineered* – thus as engineering problems – and in turn to approach the solution from the standpoint of human performance engineering [19]–[21]

Notably, reconfiguring of issues from the standpoint of social problems to engineering problems (albeit still from Rawls’ perspective) has also afforded a switch that appears to allow students skeptical of “liberal” conceptions of ‘social justice’ to participate and contribute in ways that suggest an important turn in both personal and professional responsibility. These often parallel episodes where a student is more easily able to see one’s self in a situation where bias would be experienced. This is consistent with what becomes possible when an individual adopts what Rawls called the ‘veil of ignorance’ and ‘original position’ [18].

It is not uncommon for students to rapidly ‘get the pattern’ and to introduce into class or E-mail the instructor with examples of stories from the popular press and their own lives that approximate the cases analyzed in class. This provides evidence of ‘near transfer’ that allows us to say that our approach is not only practicable in class situations but also affects students in their day to day live experiences. Anecdotally, some of these students are those who expressed the most resistance to inclusion of a class that explicitly incorporates an orientation to *inclusion, diversity, and social justice* in the catalog course description. In discussions with these individuals, they confirm that our approach to cognitive

³ Knowledge, Comprehension, Application, Analysis, and Synthesis constitute the first five levels of Bloom’s taxonomy of cognitive skills [26].

apprenticeship – and explicitly our focus on the (Rawlsian) notion that justice is itself a social good that does not require them to behave like (whatever is their image of) the “social justice warrior (scare quotes are intentional) – helps to break down their initial resistance to the content.

In order to assess the extent to which this transfer occurs outside of class and after CS-HU 130, we are accomplishing an ongoing set of interviews with students through their career at Boise State University. We find this transfer to be less common than we would hope, and we continue to develop new case examples and to share our strategy with other instructors, along with hints for how to build it into their courses.

Parallel with our developments we are encouraged to see that many incoming faculty members have begun to incorporate social-justice-oriented curricular innovations into courses. Regular conversations with faculty members allows us to learn more about their approaches and share them with students in CS-HU 130 with the message that this is an authentic concern of computer science. Some of our CS-HU 130 students have followed up with those faculty members and even joined their research teams with the stated aim of becoming part of that ‘new generation’ of computer science.

While we are still early in the development and rollout of this course and other aspects of the project it is apparent that – while planning and intentionality is a substantive element of the successes we have identified – other effects of our efforts are influenced by the social networks made possible by an organized college curriculum and an active faculty. With this in mind we are adding several additional parts of our research in order to map these social networks through a dedicated social networks analysis, and regular surveys of the student body to identify and assess the importance and impacts of emerging interests that create links between students and their experience in the social world as computer scientists.

Implications of the ‘CSP-Hatchery’ approach for helping students acquire positions of professional responsibility, morality and ethics will be discussed in light of the above.

REFERENCES

- [1] M. Bertrand and S. Mullainathan, “Are Emily and Greg More Employable Than Lakasha And Jamal? A Field Experiment on Labor Market Discrimination,” *Am. Econ. Rev.*, vol. 94, no. 4, pp. 991–1013, 2004.
- [2] S. Brayne, “Big Data Surveillance: The Case of Policing,” *Am. Sociol. Rev.*, vol. Prepublished August 29, 2017, 2017.
- [3] J. Buolamwini, “Gender Shades: Intersectional Phenotypic and Demographic Evaluation of Face Datasets and Gender Classifiers,” Master of Science, Massachusetts Institute of Technology, Cambridge, MA, 2017.
- [4] S. Ceci, W. Williams, and S. Barnett, “Women’s Underrepresentation in Science: Sociocultural and Biological Considerations,” *Psychol. Bull.*, vol. 135, no. 2, pp. 218–261, 2009.
- [5] S. Ceyer et al., “A Study on the Status of Women Faculty in Science at MIT,” Massachusetts Institute of Technology, Cambridge, MA, Manuscript, 1999.
- [6] J. Damore, “Why I was fired by Google,” *Wall Street Journal*, New York, 11-Aug-2017.
- [7] A. Etzioni and O. Etzioni, “Incorporating Ethics into Artificial Intelligence,” *J. Ethics*, Mar. 2017.
- [8] S. Fowler, “Reflecting on one very, very strange year at Uber.,” 19-Feb-2017. .
- [9] Google, “Machine Learning and Human Bias,” YouTube, 25-Aug-2017. [Online]. Available: <https://www.youtube.com/watch?v=59bMh59JQDo&feature=youtu.be&app=desktop>. [Accessed: 25-Aug-2017].
- [10] E. Huet, “Facebook’s Hiring Process Hinders Its Effort to Create a Diverse Workforce,” *Bloomberg*, 09-Jan-2017.
- [11] S. Lohr, “Facial Recognition is Accurate, if You’re a White Guy,” *New York Times*, New York, 09-Feb-2018.
- [12] G. Mann and C. O’Neil, “Hiring Algorithms are Not Neutral,” *Harvard Business Review*, 09-Dec-2016.
- [13] MIT Computer Science Female Graduate Students and Research Staff, “Barriers to Equality in Academia: Women in Computer Science at M.I.T.,” Massachusetts Institute of Technology, Cambridge, MA, Manuscript, 1983.
- [14] H. Natanson, “‘A Sort of Everyday Struggle’,” *The Harvard Crimson*, Harvard, MA, 20-Oct-2017.
- [15] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, “Machine Bias: There’s software used across the country to predict future criminals. And it’s biased against blacks.,” *ProPublica*, 23-May-2016. [Online]. Available: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. [Accessed: 24-May-2016].
- [16] J. Damore, “Google’s Ideological Echo Chamber: How bias clouds our thinking about diversity and inclusion.” 04-Aug-2017.
- [17] J. Young, “Holder’s recommended Uber diversity initiatives failed,” *San Francisco Chronicle*, San Francisco, 06-Jul-2017.
- [18] J. Rawls, *A Theory of Justice*, Revised edition. Cambridge, MA: Belknap Press of Harvard University Press, 1999.
- [19] S. Y. Chyung, “Human performance technology: From Taylor’s scientific management to Gilbert’s behavior engineering model,” *Perform. Improv.*, vol. 44, no. 1, pp. 23–28, 2005.
- [20] T. Gilbert, *Human Competence: Engineering Worthy Performance*. Washington, D.C.: Pfeiffer, 2007.
- [21] D. Langdon, *Aligning Performance: Improving People, Systems, and Organizations*. San Francisco: Jossey-Bass/Pfeiffer, 2000.
- [22] S. Vallor, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Cambridge: Oxford University Press, 2016.
- [23] J. Bergamin, “Being-in-the-flow: expert coping as beyond both thought and automaticity,” *Phenomenol. Cogn. Sci.*, vol. 16, pp. 403–424, 2017.
- [24] D. Schön, *Educating the Reflective Practitioner*. San Francisco: Jossey-Bass, 1987.
- [25] D. Schön, *The Reflective Practitioner*. New York: Basic Books, 1990.
- [26] B. Bloom, *Taxonomy of educational objectives; the classification of educational goals*. New York: Longmans, 1956.