Changing the Online Climate via the Online Students: Effects of Three Curricular Interventions on Online CS Students' Inclusivity

Lara Letaw
Oregon State University
Corvallis, Oregon, USA
letawl@oregonstate.edu

Rosalinda Garcia
Oregon State University
Corvallis, Oregon, USA
garciros@oregonstate.edu

Heather Garcia
Oregon State University
Corvallis, Oregon, USA
Heather.Garcia@oregonstate.edu

Christopher Perdriau
Oregon State University
Corvallis, Oregon, USA
perdriac@oregonstate.edu

Margaret Burnett
Oregon State University
Corvallis, Oregon, USA
burnett@oregonstate.edu

ABSTRACT

Motivation: Although CS Education researchers and practitioners have found ways to improve CS classroom inclusivity, few researchers have considered inclusivity of online CS education. We are interested in two such improvements in online CS education—besides being inclusive to each other, online CS students also need to be able to create inclusive technology.
Objectives: We have begun developing a new approach that we term "embedded inclusive design" to address both of these goals. The essence of the approach is to integrate elements of inclusive design education into *mainstream* CS coursework. This paper presents three curricular interventions we have developed in this approach and empirically investigates their efficacy in online CS post-baccalaureate education. Our research questions were: How do these three curricular interventions affect (RQ1) the climate among online CS students and (RQ2) how online CS students honor the diversity of their users in the tech they create?
Method: To answer these research questions, we implemented the curricular interventions in four asynchronous online CS classes across two CS courses within Oregon State University's Ecampus and conducted an action research study to investigate the impacts.
Results: Online CS students who experienced these interventions reported feeling more included in the major than they had before, reported positive impacts on their team dynamics, increased their interest in accommodating diverse users, and created more inclusive technology designs than they had before.
Discussion: These results provide encouraging evidence that embedding elements of inclusive design into mainstream CS coursework, via the interventions presented here, can increase both online CS students' inclusivity toward one another and the inclusivity of the technology these future CS practitioners create.

ICER 2021, August 16–19, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8326-4/21/08…$15.00
https://doi.org/10.1145/3446871.3469742

CCS CONCEPTS

• **Human-centered computing → Interaction design process and methods**; **Heuristic evaluations**; • **Applied computing → E-learning**; • **Social and professional topics → Gender**.

KEYWORDS

GenderMag, diversity, inclusion

ACM Reference Format:
Lara Letaw, Rosalinda Garcia, Heather Garcia, Christopher Perdriau, and Margaret Burnett. 2021. Changing the Online Climate via the Online Students: Effects of Three Curricular Interventions on Online CS Students' Inclusivity. In *Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021), August 16–19, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3446871.3469742

1 INTRODUCTION

Although bits and pieces of computing curricula continually change as CS technologies advance, the CS major course sequence has changed little in decades: it begins with a few introductory computing courses followed by a sequence of data structures, algorithms, theory, 1-2 programming language fundamentals and/or implementation courses, 1-2 operating systems courses, a collection of software engineering courses, and some electives [2]. These courses reinforce and build upon one another: programming language courses use some theory and data structures taught earlier, operating systems courses use data structures and computer architecture concepts, and so on.

But courses that address how the software students build can affect the people who use it, such as CS ethics, HCI, or usability courses, are essentially sidelined, with little connection to the rest of the CS education that students receive [65]. The message comes through to CS students loud and clear: concepts in ethics, society, and humans, are unimportant to CS professions. In response to growing evidence of such problems, CS education researchers have called for not only an increase in teaching ethics and social consequences in CS, but also increasing coverage of such topics in mainstream CS courses [20, 24–26, 44, 65].

One effect that has been called out less often is that CS education shows direct evidence of producing CS professionals who are unable to *create inclusive technology*. For example, in one recent study of post-secondary computing faculty, 49% of the faculty reported having seen their students struggling to prevent or even recognize how their biases affected the software they were designing [60]. In the same survey, 54% of the faculty also reported having seen students in their courses finding it difficult to design for diverse people's abilities and usage styles [60]. As Putnam summarizes, as long as CS education continues to sideline the concept of designing inclusive software, it perpetuates "the cycle of ignorance among ... developers <that> maintains the status quo of exclusion and marginalization" [64].

We are developing a new approach to mainstreaming inclusivity in CS education, in which we integrate portions of an inclusive design method called GenderMag into a variety of *mainstream* CS courses. We call our approach "embedded inclusive design". In essence, the approach incrementally embeds bits of inclusive design into the work students already do. The goal is to increase not only the inclusivity of CS education itself, but also students' attitudes toward product inclusivity of the software they are creating. Because students "do equity" as part of their mainstream CS work, we hypothesize that the approach will produce students who (1) are more inclusive to each other and (2) can create more equitable software than their predecessors.

In this paper, we focus on inclusivity in asynchronous, online, post-secondary CS education. We present three curricular interventions of the embedded inclusive design approach, and evaluate their effectiveness in four online CS classes, as a first step toward our embedded inclusive design vision. Three were separate offerings of a third-year (junior-level) database course (CS-DB) and one was an offering of a third-year (junior-level) software engineering course (CS-SE). Both are required courses for the online CS major. Neither course typically includes inclusive design.

The setting was our online Ecampus post-baccalaureate CS program, taken by people who previously earned a non-CS baccalaureate degree and now are taking CS courses to add on a baccalaureate degree in CS. Each "class" is entirely asynchronous—there are no synchronous class meetings, and people from different locations and timezones around the world can, within limits, set their own schedules. However, students in a given class commit to starting and ending the course on traditional term-calendar boundaries, to completing the assignments by certain deadlines and, in some classes, need to work (asynchronously) with other students in that class on a team. Instructors and TAs are permanently assigned for the duration of a term, and they answer individual questions (via email or discussion platforms), provide timely feedback on assignments, run discussion forums, and so on. Classes tend to be large—in our investigation, class sizes were 218, 150, 213, and 226 students. A total of 64 of these students, plus 11 more from a baseline course for a total of 75 participants, opted in to allowing their work to be used for research purposes.

Within this educational setting, we conducted an Action Research investigation. Action Research is a type of longitudinal field study that involves engaging with a community to address some problem and through this problem solving to develop scholarly knowledge [34]. As per Action Research's longitudinal focus, our involvement spanned months. Specifically, we had consistent involvement over 9 months (three terms) with four faculty members at Oregon State University. We structured our investigation around the following research questions:

- RQ1: How do these curricular interventions affect the climate among online CS students?
- RQ2: How do these curricular interventions affect online CS students' respect for users' diversity, and their ability to create more inclusive technology for these users?

2 BACKGROUND: GENDERMAG'S FACETS AND PERSONAS

Our approach leverages the GenderMag method's components and foundations. GenderMag [14] is an evidence-based method for avoiding, finding, and fixing inclusivity "bugs" in software. The process aspect of the method is a specialized cognitive walkthrough, but here we describe only its facets and personas, since those are the portions that our curricular interventions leveraged.

GenderMag's cognitive styles (cognitive "facets" in GenderMag) form the core of the GenderMag method. Each facet captures different individuals' diverse cognitive approaches by defining an evidence-based range of possible values. The five facets capture diversity of *motivations* for using tech; *information processing style*; *computer self-efficacy*; *learning style* (by process or by tinkering); and *attitude toward risk*. (These facets will be detailed further later in Section 4.) GenderMag defines "inclusivity bugs" as omissions of a technology product to support these five facets' full ranges of values. For example, technology features that support risk-tolerant users but present barriers to risk-averse users have inclusivity bugs.

Such barriers are cognitive style inclusivity bugs because they disproportionately impact people with particular cognitive styles. They are also gender-inclusivity bugs because the facets capture (statistical) gender differences in how people problem-solve [4, 14, 17, 18, 75, 80].

GenderMag uses three personas to bring the facets to life: Abi (Abigail/Abishek), Pat (Patricia/Patrick), and Tim (Timara/Timothy). Abi's and Tim's values for each of these facets lie at opposite ends of the spectrum, and Pat has values within. The Abi persona represents facet values whose proportions disproportionately skew towards women, Tim represents facet values that disproportionately skew towards men, and Pat provides a third set of values [14]. The interventions we describe in this paper include snippets of these three personas, shown later with our curricular interventions (Section 4).

Empirical studies have found GenderMag to be effective at identifying inclusivity bugs and at pointing toward effective fixes [12, 14, 22, 35, 61, 69, 80]. However, in the realm of CS education, the only work relating to GenderMag is Oleson et al.'s Action Research investigation into how to teach GenderMag in face-to-face university CS classes [59]. No prior work has investigated incorporating aspects of GenderMag into online CS courses, or the effects of doing so on the inclusivity climate of any CS course.

3 RELATED WORK

Many researchers have reported issues with inclusivity in online CS education's climate. For example, one study reports that instructors are more likely to respond to forum posts by White male students [5]. On one popular online discussion question-and-answer site, Stack Overflow, women tended to ask fewer questions, answer fewer questions, have lower reputation scores than men, and experience barriers to participation such as feeling intimidated and being unable to identify other participants of their gender [28, 79]. Although researchers investigating the Piazza online gathering site for students have reported better experiences by women than in Stack Overflow, a recent study of over 2500 Piazza users reported that Piazza women still feel the need to keep their identities and genders anonymous, and when they did not, to be less likely than men to receive answers to their questions by members of the same gender [77].

Further, Phirangee and Malec identified three "othering" themes—professional, academic, and ethnic—which are experienced by women in online learning. Despite differences in these three themes, each type of othering resulted in students feeling excluded from their online learning communities [62]. Dym et al. likewise reported that LGBTQ+ programmers in online commmunities expect few women and LGBTQ+ individuals to become CS students without additional support or encouragment because, in these participants' experiences, the field exhibits little diversity and a heterosexist climate [23]. Dym et al.'s results are not unique; similar results have been reported by other investigations of the experiences of individuals with queer gender and/or sexual identities [53, 55, 73].

Non-inclusive academic climates can affect students' performance and completion in these education environments. Kizilcec and Halawa found that women were less persistent with lectures and assessments in online courses. They also found that feelings of social belonging are influenced by success in the classroom, which can negatively impact women when they have higher attrition rates. Similarly, in an online conversion program helping individuals change career paths to computer science, women were much less likely to finish the program than men [38].

Both online and in-person CS education research have investigated factors, including gender differences, that contribute to students' feelings of exclusion. For example, Pournaghshband and Medel point out that much of society embraces a widely accepted "fallacious archetype" of what a successful CS student looks like: a young, White male with at least mid-level socioeconomic status [63]. Kuttal et al. reported differences in women's and men's experiences when completing a remote pair-programming assignment [46]. One of these differences was that their communication and gender awareness differed significantly—women relied more on non-verbal communication, which is difficult in an online setting. Women also preferred co-located pair programming whereas men were comfortable with a remote setting. Several have reported women in in-person CS classes to have less passion about technology per se but more passion about "computing with a purpose", and lower confidence in their computing abilities [3, 8, 21, 52]. Low confidence can become even lower when students compare themselves to others, such as in [38] where women students reported that they constantly compared themselves to more experienced students and became less confident when they saw experienced students struggle. Gender differences in confidence have in turn been linked to gender differences in communication in CS classes; for example, Alvarado's study found that women were less comfortable than men were when communicating with their instructor [3].

A significant body of work has investigated increasing recruitment and/or retention across genders in *in-person* CS education, and these works have brought about improvements in both recruitment and retention. Among the especially well-known practices are: pair programming (e.g., [82, 84]), meaningful or socially relevant assignments (e.g., [7, 11, 51]), and leveling the playing field with mechanisms like having everyone start with a language new to all or eschewing programming entirely to instead focus on problem solving (e.g., [43]). Some of these practices, such as giving socially relevant assignments or changing the language used in the course, are not reliant on synchronous or in-person presence in classes and thus can transfer directly to asynchronous online CS classes. However, all of these practices tend to be unidirectional—they aim to make more students feel includ<u>ed</u>, but do not generally aim to improve students' inclus<u>ivity</u> toward others.

Closest to our own research is work on using universal design to improve inclusivity for disabled stakeholders in in-person CS classes. For example, the AccessComputing project created a web development course that integrates accessibility and universal design into its curriculum [1]. To increase feelings of inclusion by both women and students with disabilities, Blaser et al. have proposed including universal design principles in engineering courses in order to prepare future engineers better as well as improving representations of disabled users and engineers [9]. This research rests on prior investigations of what diverse students value in their courses and jobs, reporting that women in engineering often value contribution to society more than men do, which suggests that women may be drawn to inclusive and universal design (e.g., [32]). Similarly, Izzo et al. have found that teaching universal design in college courses in order to include people with disabilities helps both students and instructors to improve accessibility, awareness, and instructional flexibility [39]. Putnam et al. [64] and Waller et al. [81] have both experimented with integrating accessibility concepts across multiple face-to-face courses in the major, treating accessibility as an integral part of design and development. Others have investigated including stakeholders with a disability (e.g., a wheelchair-bound user) in design/evaluation team sessions [50, 70]. Our approach applies many of the Putnam and Waller recommendations to our project, and also leverages elements of the stakeholders strategy, but our "stakeholders" are research-based personas instead of actual people. Most important, our education setting is asynchronous, online courses rather than face-to-face courses.

Thus, although there is extensive work on CS education's lack of inclusivity, there are a limited number of previous studies that investigate how to *improve* inclusivity, and even fewer in *online* CS courses. The common themes in the small body of existing work on improving online CS education's inclusivity are ways that the instructors, prerequisites, or course advertising can help.

For example, work from Kizilcec's lab found that women in online learning tend to enroll in courses that are taught by female instructors and are less rigorous [42]. They noted that the preference for less rigor and fewer prerequisites may be due to a preference for meeting all requirements and expectations for success in the course, and pointed out that these preferences may be mitigated by clearly communicating expectations. Kizilcec's lab also found that having two instructors, one man and one woman, helped retain

Course	When	#	Activity	Who involved	Intervention(s)
CS-DB	Week 7	1	Exploration	Individual	InclusivityFacets, InclusivityHeuristics
	Week 7	2	Extra Credit Assignment (reflection / application)	Individual	InclusivityFacets, InclusivityHeuristics
CS-SE	Weeks 1+2	3	Exploration	Individual	InclusivityFacets, InclusivityHeuristics
	Weeks 1+2	4	HW1 (facet reflection)	Individual	InclusivityFacets
	Weeks 1+2	5	HW1 (design / evaluate / revise)	Individual	InclusivityHeuristics, InclusivityDesign
	Weeks 1+2	6	Learning Quiz	Individual	InclusivityFacets, InclusivityHeuristics
	Weeks 1+2	7	Team Facet Discussion	Team	InclusivityFacets
	Weeks 5+6	8	HW3 (integration with others' designs)	Team	InclusivityDesign
	Weeks 5+6	9	Peer Heuristic Evaluation, HW3 (design revisions)	Classmates, Team	InclusivityHeuristics, InclusivityDesign
	Weeks 9+10	10	HW5 (climate and users reflection)	Individual	InclusivityFacets, InclusivityHeuristics
	Weeks 9+10	11	Course Feedback (extra credit cognitive styles reflection)	Individual	InclusivityFacets, InclusivityHeuristics

Table 1: Summary of curricular interventions. CS-DB students experienced two interventions (InclusivityFacets and InclusivityHeuristics) implementations through an extra credit assignment and exploration during one week of the course. CS-SE students experienced all three interventions (added InclusivityDesign) through different implementations, spanning the entire course. All activities are available in our Supplemental Document [47]. Activity# serves as an ID; we refer back to these throughout this paper.

women in online computer programming courses but having only a woman instructor was met with negative reactions from some of the women in the course [41]. Work from that lab also found that adding gender-inclusive elements to course presentation increases women's enrollment in STEM courses [40]. For example, Cheryan et al. found that classroom decoration impacts women's interest and success in computer science—even in virtual classrooms [19]. In particular, having more neutral elements such as nature pictures is better for women than having elements that are perceived as masculine or stereotypical for computer scientists (e.g., action movie posters). All of these studies show ways that incorporating gender-inclusive elements can help women to feel comfortable in online computer science. However, these studies focus mainly on course advertising or presentation, not on how curriculum changes themselves can improve both feelings of inclusion and acts of inclusion, by students in online CS education. That is the gap this paper aims to fill.

4 THREE CURRICULAR INTERVENTIONS

We are working on an emerging approach we term "Embedded Inclusive Design". The essence is to embed elements of inclusive software design into the curricula of *mainstream* CS courses.

Toward this end, we have developed three curricular interventions for asynchronous online CS education. *InclusivityFacets:* The first curricular intervention is a set of activities to enable students to learn the GenderMag cognitive styles (termed "facets" in GenderMag literature). *InclusivityHeuristics:* The second is a set of activities to enable students to learn the GenderMag Heuristics, which are based on these cognitive styles, and to use the heuristics to evaluate technology. *InclusivityDesign:* The third is a set of activities to enable students to improve the inclusivity of the technology they create, using the GenderMag Heuristics. All interventions included mechanisms to assess student learning of inclusive design concepts from these interventions. Table 1 enumerates each activity and the interventions to which they contributed.

We hypothesize that, because inclusive design will be integrated with what students are learning as part of their major, these curricular interventions will impact online CS education in these primary ways: it will positively impact online CS students' feelings of belonging in the major (investigated in RQ1); it will positively impact online CS students' inclusiveness toward other students in the major (also investigated in RQ1); it will positively impact online CS students' attitudes toward diverse users of software products (investigated in RQ2); and it will improve the inclusivity of the software the students create (investigated in RQ2).

4.1 The Approach's Theory Foundations

Our interventions build upon three foundations. The first two are the Community of Inquiry (CoI) framework [29] and Quality Matters (QM) standards, which are popular models on distance and online learning [67, 68] for creating high-quality course designs to meaningfully engage students. The third foundation is the Pedagogic Content Knowledge (PCK) investigation into teaching GenderMag and related inclusive design concepts.

The Community of Inquiry (CoI) model [29] is for incorporating curriculum into online courses to motivate and engage students. The CoI framework guides students to construct meaning through three different presences: (1) social presence, which focuses on interactions with peers, (2) cognitive presence, which focuses on interactions with the learning materials, and (3) teaching presence, which involves interactions with teachers and instructional staff. According to Fiock [27], CoI is one of the most widely used models for building community in online environments. From Fiock's CoI practices for online course design, we used the following practices.

To foster cognitive presence, we integrated a cognitive styles discussion, where students were asked to reflect on their facet values

(Activity7 in Table 1, which summarizes the intervention implementations for each course). Reflection is a key aspect of cognitive presence [29], so we built multiple opportunities for reflection into the homework assignments (Activities2,4,7,10,11). A third addition that aimed to strengthen cognitive presence included an exploration module (Activities1,3), where students could actively engage with learning materials.

To foster teaching presence in our interventions, we added instructional team cognitive profiles in Activity3, in which the instructor and TAs described their own facet values and identified as an Abi, Pat, and/or Tim. We also increased teaching presence through an instructor video introducing the interventions, regularly collected feedback about the interventions from students, and regularly provided feedback to the students on their work.

To foster social presence in our interventions, we provided multiple discussion opportunities (Activity7 and Piazza, which allowed students to post questions that would be answered by the teaching team or other students) and used collaborative project-based group work (Activities7-9).

Our second foundation is the Quality Matters (QM) standards. QM standards are research-based guidelines for quality course design [67, 68]. A QM rubric is used to evaluate course quality/presence of eight components; some we harnessed for our interventions are discussed below.

For example, to address QM's Standard 1 on providing course introduction materials, we added an introduction video to the first CS-SE learning module and we incorporated the cognitive styles team discussion (Activity7), which gave students an opportunity to introduce themselves to their peers. To address QM Standard 4 on instructional materials, we added the GenderMag learning exploration, connected the exploration to the stated learning outcomes and other activities in the course, clearly explained the connection between the exploration to current research in the field, and modeled academic integrity by citing sources (Activities1,3). Last, to address QM Standard 5 on providing learner activities and interactions, we designed project-based collaborative group work to support active learning and social interaction (Activities7-9).

Finally, our curricular changes build upon research into applicable pedagogic content knowledge (PCK). Originally introduced by Shulman [71], PCK is the intersection of pedagogical knowledge (background in effective teaching techniques and practices) and content knowledge (background in the subject being taught) that enables faculty to teach particular content. PCK is not general: PCK is specific to the topic at hand (e.g., photosynthesis, quadratic equations) and to the audience [78]. Thus, our curricular changes build upon the foundations, resources, and results from Oleson et al.'s investigation into PCK that enables faculty to teach inclusive design skills using GenderMag [59].

The most useful of Oleson et al.'s PCK elements for our work have been PCK2 (Credibility), PCK4 (Concretization), PCK5 (Modeling), PCK6 (Theory of Mind), PCK9 & PCK10 (on Stereotyping), and PCK11 (Resistance). For example, for PCK2-Credibility, we emphasized the research foundations and provided references to make clear that students could check the evidence basis for themselves (in Activities1,3). Another example is PCK4-Concretization, in which the TAs "concretized" what the facets looked like in themselves (Activities1,3) (shown later in Figure 4), and the students actively followed suit in their own reflections (Activities2,4,7,10,11). The TAs doing so is also a use of PCK5-Modeling, in which they modeled correct usage. Still another example is PCK10-Stereotyping, which recommends having students perform an inclusive design process themselves to reduce stereotyping; this is precisely what we did in CS-SE.

4.2 Curricular Materials

Table 1 enumerates the activities our interventions used, and here we describe the curricular materials supporting those activities.

Exploration (Activities1,3): Following recommendations from Fiock's CoI work [27], we created an interactive exploration for the GenderMag Heuristics. We devised the GenderMag Heuristics to fix gender inclusiveness issues in software, deriving them from the evidence-based GenderMag facets [13–15, 36]. The GenderMag exploration covered InclusivityFacets and InclusivityHeuristics in multiple ways. For InclusivityFacets, the GenderMag personas were presented [14] to help students see from the perspective of others (PCK6-TheoryofMind [59]). The heuristics themselves show how facet values may affect software usage (supports PCK4-Concretization [59]) and how to design GUI elements for diverse users. Additionally, embedded quiz questions (Figure 1) gave students immediate feedback (following CoI recommendation in [27]). Figure 2 shows an excerpt from the heuristics and an example of applying it. CS-DB offered the GenderMag exploration (Activity1) or an exploration for the general usability Nielsen's Heuristics [30, 58] for extra-credit, but CS-SE included only the GenderMag exploration (Activity3) as a required part of the course. The versions of the heuristics used for our interventions are in the Supplemental Document [47].

During CS-DB and before CS-SE, we iterated on the exploration. We added the Pat (middle-spectrum) persona to help emphasize that individuals often reflect a mixture of Abi and Tim facet values. Additionally, we renamed GenderMag Heuristics to Cognitive Style Heuristics (CSH) to communicate that GenderMag is about cognitive diversity but retained gender discussion in the exploration. We also replaced the persona documents with abbreviated versions (Figure 3) to emphasize the InclusivityFacets (personas

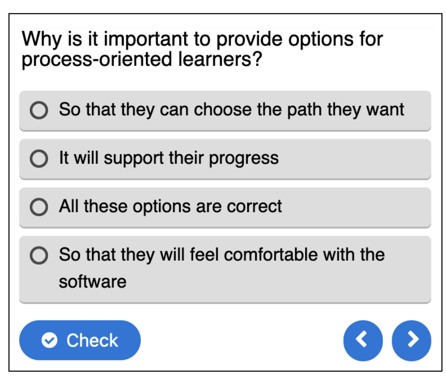


Figure 1: Quiz widget from cognitive styles exploration. Provided to CS-DB and CS-SE students as a low-stakes way to check their understanding of the cognitive styles content.

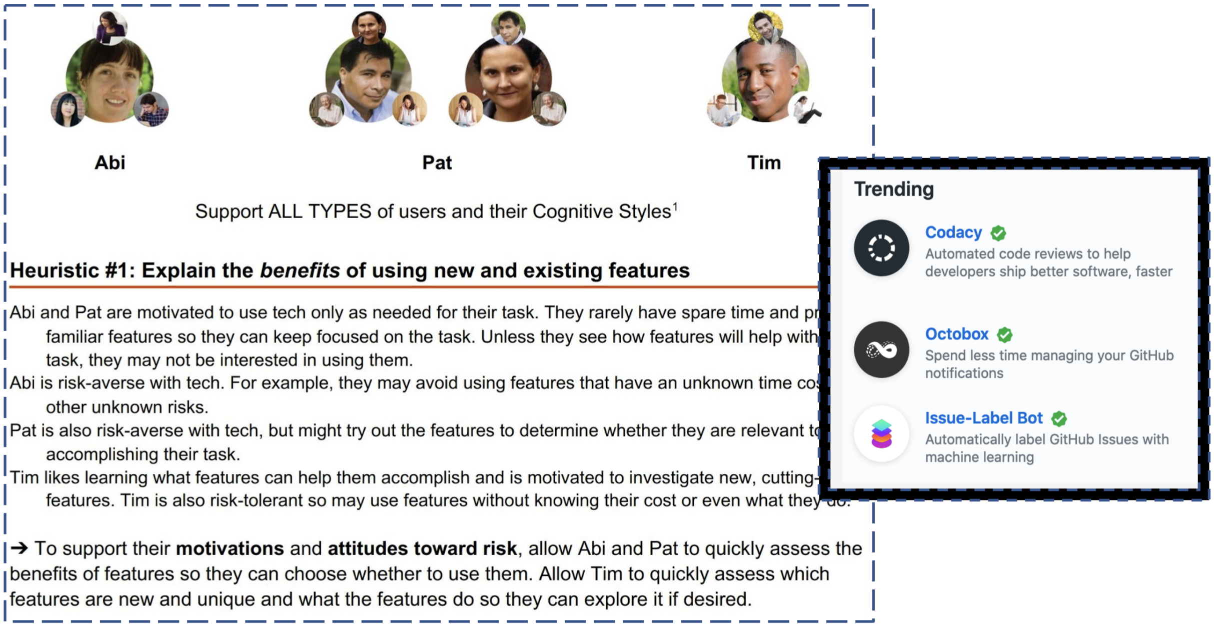


Figure 2: One of eight Cognitive Style Heuristics. (Left): Each heuristic has a short summarizing title, explanation of how it supports each persona, and a guideline for applying to software design. Current version is shown. (All versions are in the Supplemental Document [47].) (Lower Right): Example of applying Heuristic #1 by briefly explaining benefits of each feature.

Figure 3: GenderMag persona snippets as presented to CS-SE students as part of learning exploration (Activity3). Adapted from the full GenderMag personas [14]. Multiple races/genders/ages represented to help students see they can be like any persona.

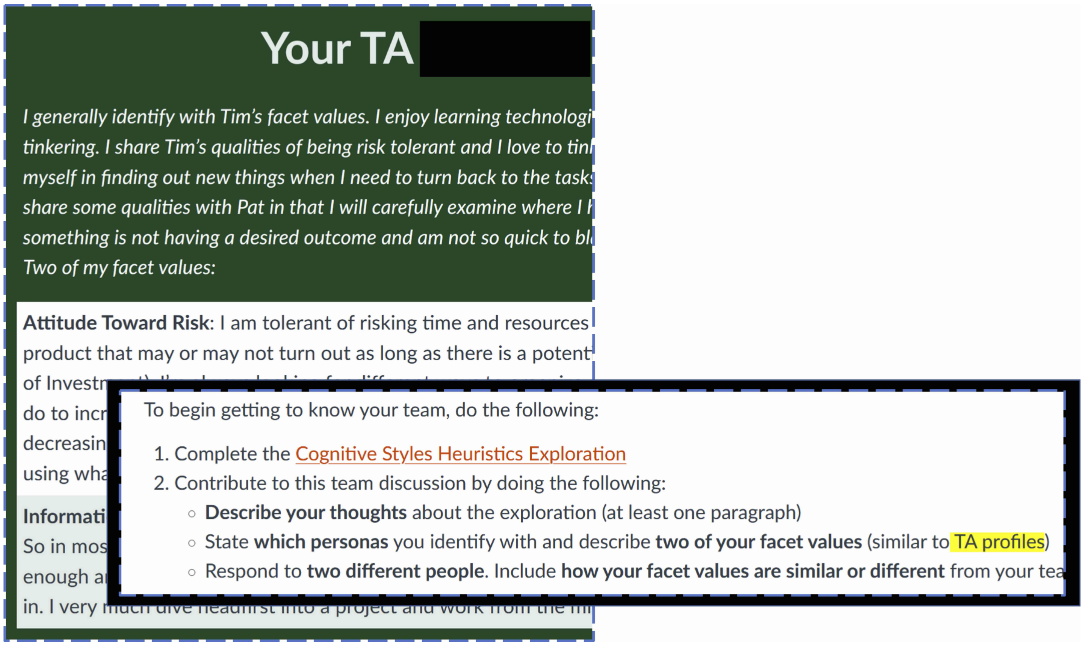


Figure 4: (Left): Cognitive style profile of a CS-SE teaching assistant. Includes TA's persona and facet value identifications. Instructor and four TAs provided profiles like these to help students learn the GenderMag course material, give students examples of talking about cognitive styles, and to help students see the diversity within instructional team. (Right): Team discussion assignment from CS-SE. Students had first two weeks of course to discuss their persona identification(s), facet values, and similarities/differences. (See Supplemental Document [47] for reusable versions.)

are not needed for a heuristic evaluation, which is meant to be a "discount" usability evaluation technique [57]). CS-SE used this revised version of the CSH exploration. Additionally, the instructor and teaching assistants added profiles of their cognitive styles (Figure 4 (Right)), which both communicated their own diversity and provided examples to guide students on later assignments.

Extra-Credit Assignment (Activity2): For extra-credit, CS-DB students could complete a 400-word reflection discussing thoughts on the exploration and giving an example of how they would apply what they learned. The structure of the assignment was defined by a CS-DB instructor, who had used the structure for other extra-credit assignments.

Homework 1 Facet Reflection (Activity4): CS-SE students completed an individual reflection to identify their own facet values. Students were asked to self-identify each of their five facet values, compare their cognitive styles to those of Abi and Tim, and identify how cognitive styles may impact their software usage.

Homework 1 Design Process (Activity5): As part of the same homework, students completed and evaluated a paper prototype. Students were not required to make changes but could lose points if they did not adequately justify how their design reflected three heuristics: (#2) Explain what existing features do, and why they are useful, (#3) Let people gather as much information as they want, and no more than they want, and (#4) Keep familiar features available.

Learning Quiz (Activity6): Students in CE-SE took a learning quiz about applications and aspects of the CSH. They were able to take the quiz unlimited times.

Team Facet Discussion (Activity7): As part of a group reflection, students discussed the exploration and shared facets with their term-long team, as in Figure 4 (Right).

Homework 3 Design Integration (Activity8): Activity5 led into a group assignment where students coordinated with their team to combine prototypes (each a different feature) into a GUI prototype for an entire application.

Peer Heuristic Evaluation, Homework 3 Design Revision (Activity9): Students then posted their work for critique by classmates outside the team, evaluated another team's work, and received feedback with which they revised their GUI design. Reviews included evaluations for 8 of the CSH as well as a suggestion and something they liked.

Homework 5 Climate and Users Reflection (Activity10): During Weeks 9+10, students individually reflected on the curriculum. They were asked how it had affected interactions with their teammates and how they viewed users.

Extra-Credit Course Feedback (Activity11): For extra-credit at the end of the term, students could leave feedback further reflecting on the cognitive styles content. Students were asked about its positive / negative effects on them and why the content might / might not be useful in the future.

The curricular materials are available in the Supplemental Document [47].

5 METHODS

We investigated our research questions using these three curricular interventions in four online classes spanning nine months (three CS-DB classes, one CS-SE class), via Action Research. Action Research is an iterative and longitudinal form of field research where the goals of solving a problem and obtaining scholarly knowledge about that problem are intertwined. One implication is that the division between researchers and participants is also blurred—researchers can act as participants and participants can act as researchers [34, 48, 74]. In our investigation, the four classes were taught by four different faculty members. All four *experienced* the interventions from a faculty perspective (participant role) and *conducted* the interventions (researcher role), and two also *contributed* to the interventions by creating/improving them (researcher role). Another implication is the treatment is not "fixed," but rather is iteratively improved in response to data arriving over time. To achieve rigor, Action Research often makes extensive use of triangulation to affirm credibility and validity of results, providing conclusions only when multiple data sources/perspectives/methods produce the same findings. We return to our triangulation mechanisms in Section 7.

Our participants were students in these classes, and our data came from their work products (Table 1) and questionnaires. From an education perspective, each assignment was required of all students (or available to all students, if extra-credit). From a research perspective, students who did the assignment could opt in to having their work used in our research, for which they were compensated with a $10 Amazon eGift Card. As per IRB and university policies, the instructors/TAs involved in the research did not know which students' work had been opted-in until after final course grades were turned in.

5.1 Participants and Tasks

The 75 students who opted to allow their work to be used in this study were all CS students, taking classes that are part of the online post-baccalaureate (post-bacc) CS degree program described in Section 1. Online CS post-bacc students have a median age of 30. Thus, many students participating had significant work, life, and educational experience in various fields. Additionally, the online program has almost three times as many men (74%) as women (25%) but this gender gap is much better than the on-campus CS program, which has only 15% women. About 50% are White. These demographics are consistent with our participants'. In terms of race, 50% of participants in CS-DB and 41% of CS-SE participants (excluding baseline participants from previous terms) self-identified as White. The other participants included Asian, Hispanic/Latino, Middle Eastern, and African American students. In terms of self-identified gender, as shown in Table 2 we had 66% men, 31% women, and 3% Agender or FTM, excluding baseline participants.

For each class implementing the interventions (i.e., all except CS-SE baseline class), students were asked to learn about the GenderMag Heuristics (or Nielsen's Heuristics) and complete assignments relating to the material as described in Section 4.

	Men	Women	Agender	FTM	Unk.	Total
CS-DB	20	8	1	1	0	30
CS-SE	22	12	0	0	11	45
Combined	42	20	1	1	11	75

Table 2: Participant count by self-identified gender for each course. Participants self-identified their gender in an open-ended questionnaire response. The unknown genders are of participants from our baseline section of CS-SE, who did not learn the Cognitive Style Heuristics.

5.2 Data and Analysis

We collected data on the effects of the GenderMag curriculum through curricular interventions as described in Section 4, a post-questionnaire for CS-DB, and pre- and post-questionnaires for CS-SE. The questionnaires, adapted from the NCWIT Student Experience of the Major survey [56], asked students about their perception of the curriculum, feelings of inclusion, and perceptions of the CS major. These can be found in the Supplemental Document [47]. In CS-SE, the questionnaire data were mostly quantitative . We also collected students' course feedback every two weeks. Data collected from CS-DB also includes data from students who completed the extra-credit assignment on Nielsen's Heuristics. With this information, we can look not only at the impact of the GenderMag curriculum, but also compare these results to those of the Nielsen's participants. Of special interest to this paper are the design work products we collected from CS-SE students. As a basis of comparison, we were able to collect design work products from students who took CS-SE in a pre-intervention term as a baseline of design comparison.

With these past assignment submissions, we had GenderMag experts judge student designs that used GenderMag Heuristics and those that did not. Four experts, ranging from 9 months to 3 years of experience, were presented with a total of 12 designs: 8 that used GenderMag Heuristics and 4 that did not. We redacted each design to remove potential hints about the design process so experts would not know whether it was changed using the GenderMag Heuristics. The experts then judged the designs by marking if they were better/worse/same for each two values of the GenderMag facets, or if they were not sure. Three experts judged each design and, if 2+ experts agreed, the design was considered improved for the facet value.

Our quantitative analyses use descriptive statistics only. Inferential statistics would not be appropriate in this investigation because no "controls" were in place to cleanly isolate variables, and the number of students opting into the research from any one class offering was very small (discussed further in Section 7.3).. To allow numeric summaries of qualitative data, we qualitatively coded as follows. First, we segmented the qualitative data (open ended questionnaire responses and written student assignments) so each response was a segment. Next, as our work crosses through several areas of research (education, online education, HCI, CS, social sciences…), we followed Hsieh & Shannon's [37] conventional content analysis approach, which is appropriate for describing phenomena that cannot be well-encompassed by existing research, to develop our codeset. Working bottom-up, categories of phenomena

	n	Abi	Tim	Pat
Women	12	42%	8%	50%
Men	22	18%	41%	41%
Overall	34	27%	29%	44%

Table 3: Persona self-identifications by gender for the CS-SE students. Women skewed closer to Abi, and men skewed toward Tim. Almost half the participants were Pat's, and the other half was almost evenly divided between Abi's and Tim's. Abi: Participant self-identified with >=3 of Abi's 5 facet values. Tim: >=3 of Tim's facet values. Pat: Everyone else. The frequencies of specific facet values with which they identified are shown in Table 4.

emerge from the data (we used affinity diagramming [31] for this process). Two researchers independently coded small portions of the data to check agreement. We achieved >=80% agreement on >=20% of the data for each dataset (Jaccard method) [72]. Given this level of consensus, one coder coded the remainder of the data. All code sets can be found in the Supplemental Document [47].

6 RESULTS

To investigate how our new approach affected different diversity outcomes, we analyzed students' coursework (Activities1-10 in Table 1) and questionnaire responses to investigate their feelings of inclusion, inclusivity attitudes toward others, team processes, and the inclusivity of the technology they created, beginning with RQ1.

6.1 RQ1: Climate Effects

6.1.1 Learning About Me. The first climate aspect we consider is how students felt about themselves. Recall that CS-DB and CS-SE students both did InclusivityFacets activities (Activities2,4). Within CS-SE Activity4, students identified their cognitive facet values.

Who did CS-SE students think they were? Their reports are summarized in Table 3. For purposes of our analysis, we labeled participants according to their self-reported facet values: if they self-reported >=3 facets like Abi's, we labeled them Abi; >=3 facets like Tim's, we labeled Tim; and any other combination we labeled Pat. As the table shows, the number of Abi's and Tim's were about the same, and Pat's were the most common.

Women skewed closer to Abi's than Tim's and men skewed closer to Tim's than Abi's. Pat's were very popular with both of these genders. (Though given a free-form textbox to enter any gender description, all CS-SE participants reported only man or woman as their gender identifications.) Women and men's opposite skews are consistent with prior research (e.g., [14, 80]).

The facet value reflection portion of Activity4 was followed by the prompt "How did identifying your facet values affect your understanding of how you use software?" A similar reflection, CS-DB students' Activity2, included the prompt "what were your thoughts while going through the [GenderMag] exploration?"

Several students in both courses responded with insights derived via the vocabulary of their GenderMag exploration (Activities1,3). For these students, the vocabulary offered ways to *name* their own problem-solving styles, *recognize* when these styles led them astray, and *appreciate* these styles when they led to success.

Facet	Facet Value	Frequency
Attitude toward risk	Averse (Abi)	16
Information processing	Selective (Tim)	16
Learning style	Tinkering (Tim)	15
Information processing	Comprehensive (Abi)	14
Motivations	Task (Abi)	14
Computer self-efficacy	High (Tim)	13
Attitude toward risk	Tolerant (Tim)	11
Learning style	Process (Abi)	8
Motivations	Tech Interest (Tim)	7
Computer self-efficacy	Low (Abi)	2

Table 4: How many CS-SE students self-identified as having each GenderMag facet value. Part of Activity4 (Table 1). Students self-identified with Abi and Tim facet values about as often; they were cognitively diverse. Students identifying as having both values of a facet are not counted here.

> P30018-SE-Abi: *"Identifying my facet values was tremendously helpful [for articulating what had] been abstract... I feel much more confident."*

> P0810202022-DB: *"[I might use cognitive styles knowledge to] protect those like me from going down the rabbit hole too deep at the detriment of other tasks."*

> P30097-SE-Tim: *"Identifying my facet values helped me [understand which features of technology] are most helpful [for my learning...] the most successful I have been... was when I just jumped right in... <Facet: Tech learning style (tinkering)>"*

One student also appreciated the GenderMag content for its applicability to a wide range of gender identities.

> P0514202109-DB: *"I'm trans and I noticed the gender inclusion right away"*

By the time CS-SE took place, we had expanded our data collection to include questions about what made students feel included/excluded in the course (post-questionnaire in Supplemental Document [47]). The results were very strong: 82% of students reported that the InclusivityFacets activity of learning cognitive styles (Activity3) made them feel included (Table 5). In fact, "very included" was the most common response to this question—for all persona-types and genders.

6.1.2 "Safety in Numbers" Self-Validation. In CS-SE, the next InclusivityFacets activity was a within-team discussion about facet values and persona identifications (Activity7) (described in Section 4). This activity produced a "safety in numbers" form of self-validation.

Students quickly found commonalities with teammates. For example, although only about 27% of participants were Abi's, 77% identified as having at least one Abi facet value. Thus, mostly-Abi's found overlap even with mostly-Tim students.

> P30097-SE-Tim: *"I consider myself to be mostly 'tech-literate,' but the first time I was introduced to macOS for work I was lost. <Facet: Computer self-efficacy>"*

	n	Very included	Somewhat included	Neither	Somewhat excluded	Very excluded	Included	Excluded
Abi	9	56%	22%	22%	0%	0%	78%	0%
Tim	10	70%	20%	10%	0%	0%	90%	0%
Pat	15	53%	27%	20%	0%	0%	80%	0%
Women	12	58%	25%	17%	0%	0%	83%	0%
Men	22	59%	23%	18%	0%	0%	82%	0%
Overall	34	59%	23%	18%	0%	0%	82%	0%

Table 5: Effects of LEARNING COGNITIVE STYLES on CS-SE participants feeling included/excluded. Right: Almost everyone felt included and nobody felt excluded. Left: Results leaned more toward very included than somewhat included. Tim's felt especially included. Maximum row values within section are highlighted.

These discoveries seemed to provide a form of self-validation in knowing they weren't the only one with their styles of problem-solving.

> P30018-SE-Abi: *"... excited to work on a team with a fellow Abi... [anticipating] running through the whole process start to finish. <Facets: Information processing, Tech-learning style>"*

> P33731-SE-Abi: *"... our [facet values] are nearly identical! Like yourself I can be very timid about new programs and apps when I don't fully understand what all is going on. <Facets: Computer self-efficacy, Risk>"*

The "safety in numbers" sentiment came out strongly when multiple students realized none of their team felt expert at CS.

> P33731-SE-Abi: *"... [most of us] decided we are a Pat and have some Tim tendencies, [and all realized we are] not nearly experts. <Facet: Computer self-efficacy>"*

> P30018-SE-Abi: *"I have been taking CS classes for 3 years now and... felt like an imposter because I wasn't a tinkerer. These cognitive styles point out the benefits of my caution as well as validate them to myself and amongst my peers. <Facets: Computer self-efficacy, Tech-learning style>"*

6.1.3 Exploring Each Others' Differences. Besides finding commonalities, students also explored facet value differences within their teams. For example, P30683-SE, who self-identified as having three Tim facet values and two Abi facet values, chose to point out their Abi learning style.

> P30683-SE-Tim: *"Wow! It's crazy to read about your willingness to use LaTeX... I really hate tinkering. <Facet: Tech learning style>"*

P37307-SE (an Abi) agreed, saying that spending time tinkering might dis-serve their employer; they would instead contact tech support.

> P37307-SE-Abi: *"[tinkering with it would be] wasting a company's money...<Facets: Motivations, Tech-learning style>"*

Some of these discussions of differences led to discussions of coveting one anothers' facet values, and how they aspired to change. Interestingly, their aspirational facet values were sometimes direct opposites.

> P31766-SE-Pat wanting to become more Abi-like: *"...working on being more comprehensive...<Facet: Information processing>"*

> P35173-SE-Tim with Abi's Information processing style, wanting to become even more Tim-like: *"[My comprehensive style is] probably often a detriment <Facet: Information processing>"*

Soon, students began to express understanding of teammates' diverse collection of facet values, and what those might mean for their teamwork.

> P32624-SE-Pat: *"My teammates' information processing style and learning style were a 180 degree pivot [from mine]. <Facet: Information processing>"*

> P33842-SE-Pat: *"[I previously] assumed most students were tech-savvy and enjoyed technology [but now realize that] an interest in CS doesn't suddenly make one a Tim. <Facets: Computer self-efficacy, Motivations>"*

From these discussions, strong "includedness" results emerged. As Table 6's leftmost section (Bottom Row) shows, our implementations of the InclusivityFacets intervention (Activities3,4,6,7,10,11) made 88% of students feel included, and made nobody feel excluded. Disaggregating into Abi-like, Tim-like, and Pat-like CS-SE students (Top Rows), and by gender (Middle Rows) shows very high results for every persona-type and every gender. In summary, the InclusivityFacets intervention helped everyone, but Abi and women especially seemed to benefit.

Note the rightmost section, regarding simply *learning* about cognitive styles (e.g., in Activity3) (as opposed to discussing them, e.g., in Activity7). As that section shows, for Tim's, Pat's, and men, learning alone was nearly as effective as discussing them. However, for Abi's and women, discussing cognitive styles in addition to learning them had a pronounced positive effect.

6.1.4 Teamwork: Toward "I'm OK, You're OK" [33]. Team members' understanding of one another's facet values began to affect their teams' processes. For example, some used their team's "facet inventory" to decide which teammates to consult (e.g., for Activity9).

	n	Included (Discuss)	Excluded (Discuss)	Very inclu.	Some inclu.	Neither	Some exclu.	Very exclu.	Included (Learn)	Excluded (Learn)
Abi	9	100%	0%	67%	33%	0%	0%	0%	78%	0%
Tim	10	90%	0%	40%	50%	10%	0%	0%	90%	0%
Pat	15	80%	0%	27%	53%	20%	0%	0%	80%	0%
Women	12	92%	0%	50%	42%	8%	0%	0%	83%	0%
Men	22	86%	0%	36%	50%	14%	0%	0%	82%	0%
Overall	34	88%	0%	41%	47%	12%	0%	0%	82%	0%

Table 6: Effects of COGNITIVE STYLES on CS-SE participants' feelings of inclusion. Students' post-questionnaire responses after the term showed almost everyone felt included by *discussing* cognitive styles with teammates, but Abi and women especially seemed to benefit. (Recall that only two genders participated in the CS-SE investigation.) Left section: Summary, where "Included" = Somewhat/very included, and "Excluded" = Somewhat/very excluded. Middle section: Detailed breakout for discussing cognitive styles. Right section: Summary of students' responses for simply *learning* cognitive styles, as per Table 5. Maximum row values within section are highlighted .

	n	Team			Non-team		
		Included	Neither	Excluded	Included	Neither	Excluded
Abi	9	89%	11%	0%	45%	33%	22%
Tim	10	90%	0%	10%	40%	60%	0%
Pat	15	80%	7%	13%	27%	46%	27%
Women	12	66%	17%	17%	17%	50%	33%
Men	22	95%	0%	5%	45%	45%	10%
Overall	34	85%	6%	9%	35%	47%	18%

Table 7: Effects of team/class interactions on CS-SE students feeling included/excluded. Students' responses to post-questionnaire showed that students' feelings of inclusion were very high in interactions with teammates, and far more so than with their other classmates. Left: Effects of TEAM interactions. Right: Effects of NON-TEAM CLASSMATE interactions. Maximum row values within section are highlighted .

> P33075-SE-Pat: *"It was helpful… knowing how each other approaches using… software because we… had someone who could look at a feature and give a different perspective."*

> P37307-SE-Abi: *"extremely glad we did the Cognitive Styles Exploration early...[If we were] looking to improve [the team's] project with respect to a Tim-type user, we already had an understanding of which teammates could relate to that user the most"*

Over the course of the term, teams' understanding of their diverse facets affected both how they were treated by their teammates and their feelings of being accepted by their teammates. In essence, students began to understand that their teammates' facet value differences did not equate to differences in ability.

> P36170-SE-Pat: *"[My teammates understand that we] tend to work differently [and thus we were] less demanding on each other."*

> P30683-SE-Tim: *"[I'm] resistant [to] new technologies...[but my team understands that I'm not] being lazy...some people just aren't really tinkerers."*

They also began seeing value in their differences.

> P33075-SE-Pat: *"[The cognitive styles curriculum] was helpful in the group… we had each cognitive style reflected… This resulted in an overall more usable web app than if it were developed by… people all of the same cognitive style"*

6.1.5 Climate by the Numbers. Using the post-questionnaire, we collected CS-SE students' assessment of their feelings of being included by their teams. As Table 7 shows, for every persona and every gender, CS-SE students overwhelmingly reported feeling included by their teammates—with whom students did cognitive styles sharing via Activity7. These numbers are in stark contrast to their feelings of being included by their non-team classmates—with whom students were *not* assigned cognitive styles sharing—which did not reach even 50% for any persona-type or any gender. Women felt the least included, but their inclusion rate among their teammates was still a dramatic increase over their inclusion rate by other classmates. These data, in combination with Table 6, suggest that the InclusivityFacets intervention was directly responsible for these effects.

	n	Represented?		Belong?		Likely to complete?		Cognitive styles increased interest?
		Pre	Post	Pre	Post	Pre	Post	
CS-DB	8							0.59
CS-SE	34	0.67	0.68	0.81	0.88	0.91	0.99	0.75

Table 8: CS major/minor climate. (Left section:) Feelings of representation, belonging, and likelihood to complete CS major/minor increased for CS-SE participants between the beginning and end of course (not measured for CS-DB). (Right section:) Both CS-DB participants and CS-SE participants were more interested in the CS major/minor after using GenderMag Heuristics. (Note: For CS-DB, 8 of 30 participating students used GenderMag Heuristics; the others used Nielsen's). Maximum row values within section are highlighted. *Units*: Position on Likert scale from 0 (least) to 1 (most).

Course	Source	% of students mentioning user diversity
CS-DB	Extra Credit Assignment (Activity2)	88%
CS-DB	Post-Questionnaire	63%
CS-DB	Course total	100%
CS-SE	Team Facet Discussion (Activity7)	91%
CS-SE	Feedback Assignments (including Activity11)	74%
CS-SE	HW5 (Activity10)	74%
CS-SE	Post-Questionnaire	68%
CS-SE	Course total	100%

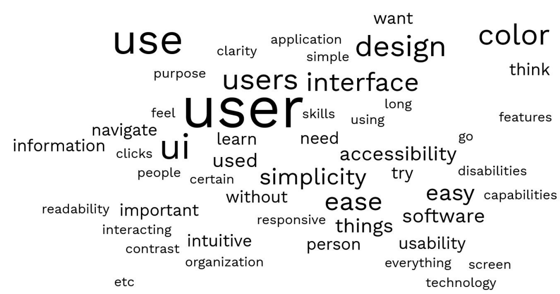
Table 9: CS-DB and CS-SE students whose work products mentioned user diversity. From all data sources throughout term. Bottom row for each course computes union of students counted in that course's rows above. By the end of the term, 100% of the participating students in all three CS-DB classes and in CS-SE had brought up this point. Activities refer to Table 1.

The climate "bottom line" was that these implementations of the InclusivityFacets intervention increased students' interest and/or feeling of belonging in the CS major/minor (Table 8) .

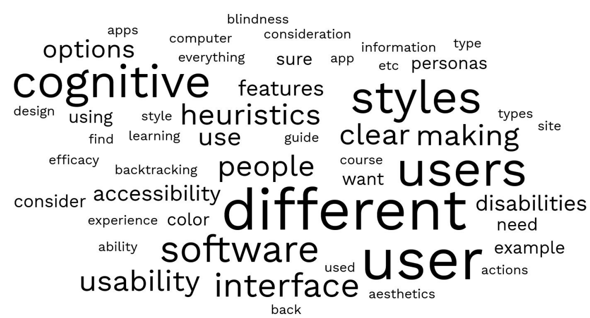
6.2 RQ2: Honoring User Diversity, Creating Inclusive Tech

The results about climate are encouraging, but are only part of the goals of our interventions. In addition, we hoped the three interventions together would change students' attitudes toward their ultimate users, so they would not only recognize and honor their users' diversity, but also gain skills to create more inclusive technology.

After experiencing the cognitive styles and heuristics curriculum, both the CS-DB and CS-SE students began to talk about the need to support diverse users. As Table 9 shows, *every one* of the students in both courses who experienced the cognitive styles and heuristics curriculum remarked in one way or another on this point. An interesting contrast to the CS-DB and CS-SE students in Table 9 is with the CS-DB students not in that table, i.e., the 22 CS-DB students whose Activity2 assignment covered Nielsen's Heuristics instead of GenderMag Heuristics. About two-thirds of the 22 Nielsen's CS-DB students did not talk about user diversity at all. Instead, these students spoke of accommodating a generic "the user". Further, most



(a) Before learning GenderMag Heuristics



(b) After learning GenderMag Heuristics

Figure 5: GUI design priorities of CS-SE students at the beginning of the course (Top) versus the end (Bottom). Students went from talking about "color" (7 counts), "ease" (6) and "simplicity" (5), to "different" (13), "styles" (10), and "cognitive" (10).

of the few CS-DB Nielsen's students who did talk about diversity mentioned only diverse levels of experience. The following quotes detail a few of these differences, and Figure 5 shows a comparison between how CS-SE participants spoke about users before vs. after the curricular interventions.

> P0512200012-DB-Nielsen (about different levels of experience): *"...tailoring a program to ...beginners [and also] advanced users..."*

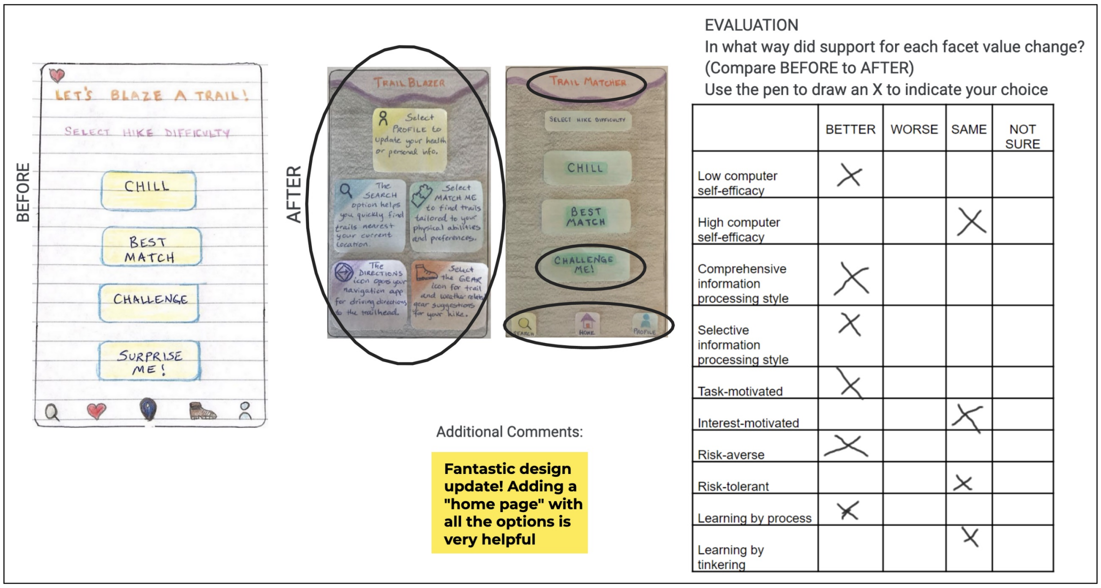


Figure 6: GenderMag expert GUI design scoring example. Before (Left) and after (Middle) versions of participant (P37283-SE) GUI feature design evaluated by a GenderMag expert for facet value support. Experts circled changes they noticed, then judged the design as being better, worse, or the same in its support for each facet value (Right). This "after" version was judged to be more inclusive on 6 of the 10 facet values.

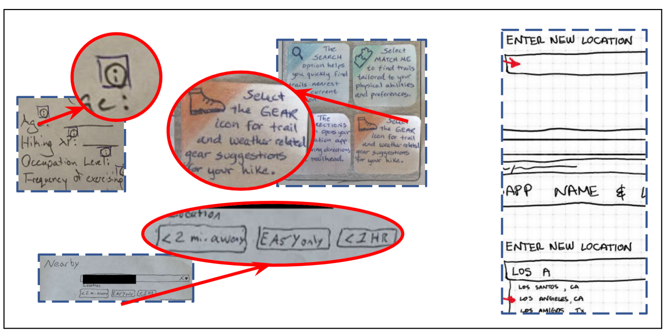


Figure 7: Student design changes. Inclusive design changes using GenderMag Heuristics (Left), versus a non-inclusive design change made using Nielsen's Heuristics (Right, bare search box with auto-complete requires tinkering).

> P0514201819-DB-GenderMag: *"…people with different problem-solving instincts and software-interaction behaviors"*
>
> P30097-SE-Tim: *"[helped me] think of users as not just one type... [and] remember that each user will use my app in different ways"*
>
> P30774-SE-Tim: *"[I used to categorize users as] haves v. have-nots in terms of…technology skills [but instead] it's all a massive gray area…Some people are risk-averse…and some will tinker for hours for the sake of tinkering. <Facets: Risk, Tech-learning style, Motivations>"*

To see whether and how the interventions had given students the means to bring their new understanding of users' diversity into the technology they created, we turned to our expert judges. (Recall from Section 5 that multiple experts' judgments were also cross-checked against each other.) Figure 6 shows how one expert judged P37283-SE's hiking app prototype, which the student improved by applying the GenderMag Heuristics. The expert complimented the student on these design changes, which scored improvements in six of the 10 facet values. For example, low-self-efficacy users are better supported by a new, informative home screen (Facet: Computer self-efficacy), and the formatting of the information on the new screen is well-suited to both comprehensive and selective information processors (Facet: Information processing style).

In contrast, students from the baseline offering of CS-SE used only standard heuristics, such as Nielsen's, and these designs sometimes brought inclusivity setbacks. For example, Figure 7 compares several features designed by students using the GenderMag Heuristics against a similar feature set designed by a student using Nielsen's. The expert judged the Nielsen's design as being non-inclusive due to its reliance on users wanting to tinker (Facet: Tech-learning style). However, students using the GenderMag Heuristics added information buttons, explanations, and filters, which support users with both comprehensive and selective information processing styles (Facet: Information processing style).

Finally, Figure 8 illustrates a straightforward set of changes that together greatly improved the prototype's inclusivity.

By the end of the term, many CS-DB and CS-SE students were expressing a sense of responsibility for technology problems diverse users might experience. For example, upon learning that low-self-efficacy users might blame themselves [14] when technology does

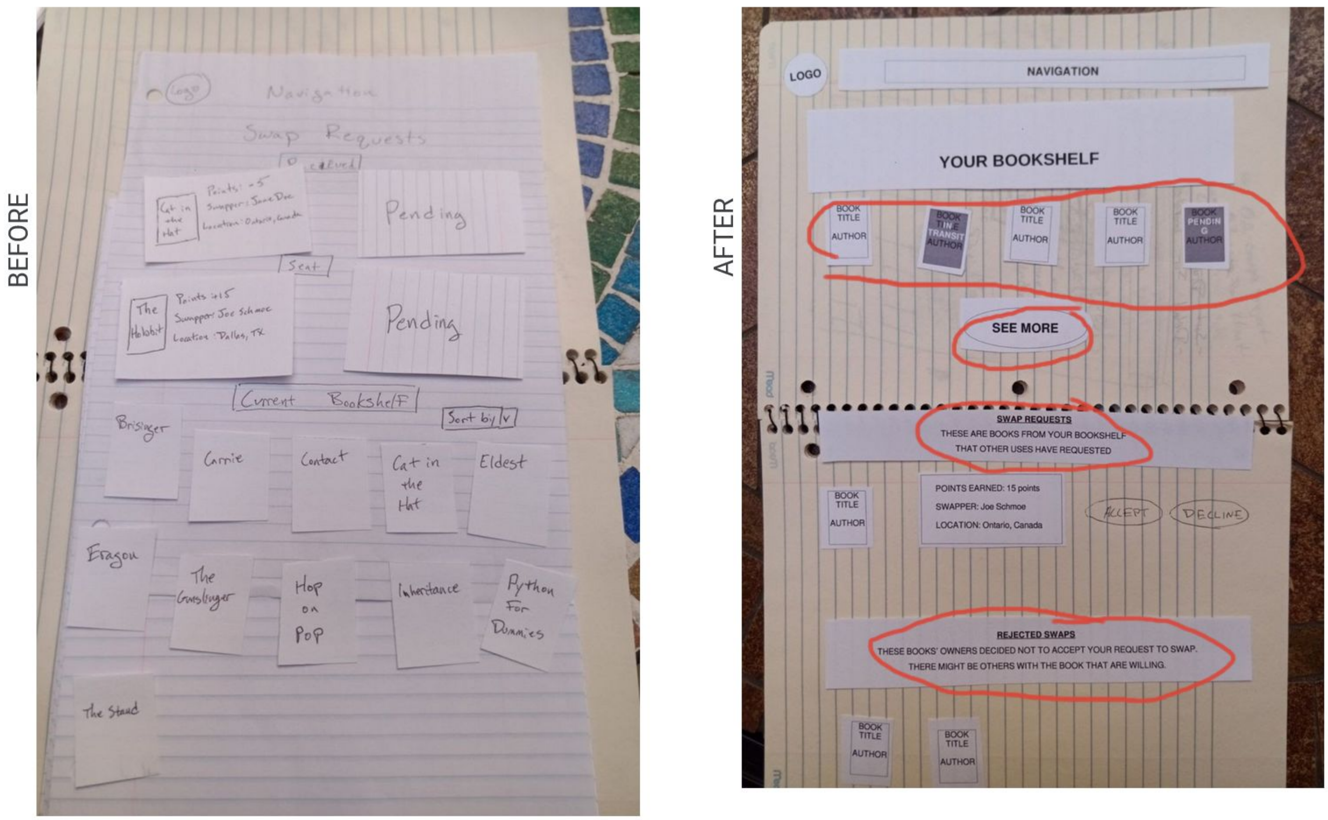


Figure 8: Example of inclusivity improvement. Before (Left) vs. after (Right) applying the GenderMag Heuristics. Brief subsection descriptions (Lower Half) added support for risk-averse, task-motivated users (by giving enough information for them to determine what the features in that subsection are for) and risk-tolerant, tech-interest motivated users (by saving them time on locating features most interesting to them).

not perform as expected (recall Figure 3), P30683-SE was vehement that the blame lay with the developer, not the user.

> P30683-SE-Tim: *"That's not right! I felt a sense of responsibility to users like these. <Facet: Computer self-efficacy>"*

> P37987-SE-Pat: *"if you do not fit [the user's] cognitive type, then you may not fully understand how they interpret [a feature]"*

> P36673-SE-Abi: *"I will now be more careful to incorporate features that allow you to undo actions."*

> P30774-SE-Tim: *"[when I see others] struggle with technology, especially in this context of the pandemic, I will view them with more compassion"*

7 DISCUSSION

7.1 Results Triangulation

Consistent with Action Research methods, we safeguarded the reliability of our results through extensive use of triangulation, which we enumerate in Table 10.

As the table shows, every result was cross-confirmed across multiple sources of evidence. Each major result occupies a column. The cells under the first two columns show the activities and/or questionnaires that showed evidence for each Research Question; the last three columns do so for each curricular intervention. Summing up the evidence from CS-DB and CS-SE, the total results for each research question are evidenced by 5 data sources, and for each intervention were evidenced by 4 to 6 data sources. In summary, evidence from multiple courses and data sources pointed to the same results for all research questions and interventions.

7.2 Perspectives on Gender, Feminism, and Responsible CS—and Lessons Learned

The impetus behind our approach is the goal of supporting pluralism in the technologies that CS students, as future CS practitioners, create. As Bardzell explained, supporting pluralism means creating technologies that "resist any single, totalizing, or universal point of view" [6], i.e., emphasizing attention to individual differences within genders rather than universality. The concept of pluralism is a central tenet of feminist HCI as per Bardzell [6]. Under differing vocabularies, the same concept is also widely espoused in

	Results by RQ		Results by Curricular Intervention		
	RQ1-climate	RQ2-users	InclusivityFacets	InclusivityHeuristics	InclusivityDesign
CS-DB	Q	Q	2,Q	2,Q	
CS-SE	10,11,Q,Q	5,10,Q,Q	4,7,10,11	5,9,10,11	5,9,XX

Table 10: Results triangulation. Row 1: Triangulation of results within the CS-DB course. (Note that InclusivityDesign is not applicable to CS-DB). Row 2: Triangulation of results within the CS-SE course. Each number refers to an activity number whose work product produced the result for the column header above it. Q refers to a questionnaire result instead of an activity result, and XX to GenderMag expert evaluations where 2+ experts agreed. (Left columns): Triangulation of results by Research Question. (Right columns): Triangulation of results by curricular intervention.

others' views on HCI feminism, queer feminism, and queer theory (e.g., [16, 49, 66, 76]; summarized in [10]). The concept of pluralism is also a central idea behind work in universal or inclusive design, suggesting that the approach presented here may also be applicable to diversity dimensions beyond gender diversity (e.g., disabilities, age, socioeconomic, via InclusiveMag-generated foundations [54]). All of these ideologies embrace the common theme of avoiding making technology that attempts to force diverse individuals to all problem-solve and work with technology in one and only way.

Like researchers in the "Responsible CS" and "critical CS" schools of thought (e.g., [20, 26, 45]), we view integrating such issues into modern CS education to be important and necessary. CS students, as future CS practitioners, cannot build a digital world that honors diversity and pluralism if CS education does not offer them skills to do so. The interventions we presented attempted to do so, and our results so far are encouraging.

Even touching upon gender equity issues can seem controversial, and three students explicitly shared their views of the inadvisability of our interventions. As Table 11 shows, one expressed discomfort with the inclusivity of the interventions themselves, three said that the interventions invoked gender stereotyping, and one implied that talking about gender at all was unnecessary. However, several others applauded the interventions, for reasons ranging from feeling from more included to being able to build better software to becoming a more responsible computing professional; Table 11 shows a few of the positive responses.

From these views, in combination with the interventions' foundations from the CoI Model, QM, and the relevant PCK (Section 4), we derived three lessons.

Integrate throughout the term: First, the interventions worked best when integrated throughout the course as in CS-SE, not just in a single assignment as in CS-DB. Not only did the integration enable CS-SE students to learn more skills for software creation, the term-long experience also produced very strong climate impacts (e.g., recall Table 8).

Incorporate in team discussions: Although too late for data gathering, another class of CS-SE using these interventions has just finished. In this post-study offering, team discussions were not part of the interventions, and the students did not seem to have gained as much comfort with the materials as a result.

Emphasize the research behind the interventions: Over time, we have become more explicit about the evidence behind these interventions, with several pointers and citations pointing the way, and this has been positively received. However, we have not emphasized the research investigating whether GenderMag encourages stereotyping. (In fact, it reduces stereotyping [36].) Future terms will bring this evidence forward, to see whether that helps allay stereotyping concerns.

7.3 Limitations

No empirical study is perfect. One reason is the inherent tradeoff among different types of validity [83]. Field studies, including Action Research studies, achieve real-world applicability, whereas controlled studies achieve isolation of variables. Our Action Research study therefore had many uncontrolled variables, such as multiple courses with multiple instructors in multiple terms.

This leads to limitations in generalizability. Our four classes were offerings of only two courses, and our educational setting was (1) post-bacc CS students, who are different than post-secondary CS students; and (2) asynchronous online courses, which are different from face-to-face courses. Another generalizability limitation is that fewer than 10% of students in these classes opted in out of >800 students[1] Thus, interpretations we made from our data might be different had we studied different students.

Limitations like these can only be addressed by additional empirical studies using a variety of empirical methods in a variety of educational settings. Given these limitations, we do not view our results as being generalizable beyond the particular context of our investigation, but rather as encouraging evidence of the potential of these interventions.

8 CONCLUSION

In this paper, we have presented three new curricular interventions for online CS courses. Our educational setting was two courses (four classes) over nine months in an asynchronous online CS education program for post-bacc students. We evaluated whether these interventions improved class climate, led students to honor their users' diversity, and led students to build more inclusive software for these users. Among our results were:

- RQ1 (climate): Students gained new acceptance of themselves, reported positive impacts on team dynamics, and felt more included in the major.
- RQ2: (users, tech): Students came to recognize and respect their users' diversity and were effective at designing more inclusively.

[1]Other instructors in this program have likewise experienced low opt-in rates for research projects. One reason could be that many post-bacc students have full-time jobs and an incentive of $10 may not be attractive enough for this population.

Student and view	Pro/Con	Why Pro	Why Con
Woman-DB: There are probably developers ... not using these heuristics [are] limiting the ...effectiveness of their software.	→	Better CS	
Man-SE: [when I see others] struggle with technology...I will view them with more compassion... I am happy...this class did more to diversify representation.	→	Responsible CS, Better at CS	
FTM (Transgender)-DB: I noticed the gender inclusion right away... I appreciate that the industry is finally addressing these issues... I want to see that reflected in my education... [It's] good to... increase your user base...	→	Responsible CS, Better at CS, Includes me	
Agender-DB: ...bring to mind the gender stereotype orientedness of the heuristics [but I'll] disassociate them from the idea of gender and just think about different user types... nonbinary-exclusive... irritated me.	→ ←	Better at CS	Stereotyping, Excludes me
<unknown gender>-SE: I feel like we have a tendency to tie gender to things unnecessarily... I wish we would stop fueling stereotypes.	←		Unnecessary, Stereotyping
<unknown gender>-SE: I think the cognitive styles content can be harmful...the concept may be used to stereotype...	←		Stereotyping

Table 11: Students' pro (→) vs. con (←) views on these interventions. We gathered these statements from our data (Section 5) and from anonymous course evaluations. Thus, these views represent those of >800 students in all course offerings, not just the opting-in students. This table shows all three CON comments from these >800 students, and a subset of the PRO comments.

Students' attitudes toward these interventions were generally quite positive, although a few raised issues that will require more improvement in our implementation of the interventions. Among their reasons for positive responses were the feeling that the interventions helped them to both create better software and to be a more responsible CS professional. Perhaps this is the most important outcome of all—accepting responsibility for the impacts one's software creation can have on diverse users.

ACKNOWLEDGMENTS

We thank the CS-SE and CS-DB students and instructors; Aishwarya Vellanki for her contributions to this research; Kyle Harms, Patricia Morreale, and Jodi Tims for helping to nurture and develop these ideas; and Benjamin Brewster for enthusiastically supporting implementation in the OSU Ecampus. This work was supported in part by the OSU Ecampus Research Fellows Program, Center for Inclusive Computing at Northeastern University, and NSF #1901031 and #2042324. Any opinions, findings and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the sponsors.

REFERENCES

[1] AccessComputing. 2015. *Web design & development I.* http://www.uw.edu/accesscomputing/webd2/

[2] ACM/IEEE Joint Task Force on Computing Curricula. 2013. Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. (20 December 2013). https://www.acm.org/education/curricula-recommendations Accessed: Feb. 2, 2020.

[3] Christine Alvarado, Yingjun Cao, and Mia Minnes. 2017. Gender Differences in Students' Behaviors in CS Classes throughout the CS Major. In *Proceedings of the 2017 acm sigcse technical symposium on computer science education.* 27–32.

[4] Manon Arcand and Jacques Nantel. 2012. Uncovering the nature of information processing of men and women online: The comparison of two models using the think-aloud method. *Journal of theoretical and applied electronic commerce research* 7, 2 (2012), 106–120.

[5] Rachel Baker, Thomas Dee, Brent Evan, and June John. 2018. *Bias in Online Classes: Evidence from a Field Experiment.* Technical Report. Stanford Center for Education Policy Analysis. https://cepa.stanford.edu/sites/default/files/wp18-03-201803.pdf

[6] Shaowen Bardzell. 2010. Feminist HCI: taking stock and outlining an agenda for design. In *Proceedings of the SIGCHI conference on human factors in computing systems.* 1301–1310.

[7] Lecia J. Barker, Charlie McDowell, and Kimberly Kalahar. 2009. Exploring Factors That Influence Computer Science Introductory Course Students to Persist in the Major. In *Proceedings of the 40th ACM Technical Symposium on Computer Science Education* (Chattanooga, TN, USA) *(SIGCSE '09).* Association for Computing Machinery, New York, NY, USA, 153–157. https://doi.org/10.1145/1508865.1508923

[8] Sylvia Beyer. 2014. Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education* 24, 2-3 (2014), 153–192.

[9] Brianna Blaser, Katherine M. Steele, and Sheryl Elaine Burgstahler. 2015. Including Universal Design in Engineering Courses to Attract Diverse Students. In *2015 ASEE Annual Conference & Exposition.* ASEE Conferences, Seattle, Washington. https://www.jee.org/24272.

[10] Samantha Breslin and Bimlesh Wadhwa. 2014. Exploring Nuanced Gender Perspectives within the HCI Community *(IndiaHCI '14).* Association for Computing Machinery, 45–54. https://doi.org/10.1145/2676702.2676709

[11] Michael Buckley, Helene Kershner, Kris Schindler, Carl Alphonce, and Jennifer Braswell. 2004. Benefits of using socially-relevant projects in computer science and engineering education. In *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education.* 482–486.

[12] M. Burnett, R. Counts, R. Lawrence, and H. Hanson. 2017. Gender HCl and microsoft: Highlights from a longitudinal study. In *2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC).* 139–143. https://doi.org/10.1109/VLHCC.2017.8103461

[13] Margaret Burnett, Anicia Peters, Charles Hill, and Noha Elarief. 2016. *Finding Gender-Inclusiveness Software Issues with GenderMag: A Field Investigation.* Association for Computing Machinery, New York, NY, USA, 2586–2598. https://doi.org/10.1145/2858036.2858274

[14] Margaret Burnett, Simone Stumpf, Jamie Macbeth, Stephann Makri, Laura Beckwith, Irwin Kwan, Anicia Peters, and William Jernigan. 2016. GenderMag: A Method for Evaluating Software's Gender Inclusiveness. *Interacting with Computers* 28, 6 (10 2016), 760–787. https://doi.org/10.1093/iwc/iwv046 arXiv:https://academic.oup.com/iwc/article-pdf/28/6/760/7919992/iwv046.pdf

[15] Margaret M. Burnett, Laura Beckwith, Susan Wiedenbeck, Scott D. Fleming, Jill Cao, Thomas H. Park, Valentina Grigoreanu, and Kyle Rector. 2011. Gender pluralism in problem-solving software. *Interacting with Computers* 23, 5 (07 2011), 450–460. https://doi.org/10.1016/j.intcom.2011.06.004 arXiv:https://academic.oup.com/iwc/article-pdf/23/5/450/1875429/iwc23-0450.pdf

[16] Judith Butler. 2011. *Gender trouble: Feminism and the subversion of identity.* routledge.

[17] Shuo Chang, Vikas Kumar, Eric Gilbert, and Loren G Terveen. 2014. Specialization, homophily, and gender in a social curation site: Findings from Pinterest. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing.* 674–686.

[18] Gary Charness and Uri Gneezy. 2012. Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization* 83, 1 (2012), 50–58.

[19] Sapna Cheryan, Andrew N. Meltzoff, and Saenam Kim. 2011. Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education* 57, 2 (2011), 1825–1835. https://doi.org/10.1016/j.compedu.2011.02.004

[20] Lena Cohen, Heila Precel, Harold Triedman, and Kathi Fisler. 2021. A New Model for Weaving Responsible Computing Into Courses Across the CS Curriculum. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 858–864.

[21] Mathilde Collain and Deborah Trytten. 2019. You don't have to be a white male that was learning how to program since he was five. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. 968–974.

[22] Sally Jo Cunningham, Annika Hinze, and David M. Nichols. 2016. Supporting Gender-Neutral Digital Library Creation: A Case Study Using the GenderMag Toolkit. In *Digital Libraries: Knowledge, Information, and Data in an Open Access Society*, Atsuyuki Morishima, Andreas Rauber, and Chern Li Liew (Eds.). Springer International Publishing, Cham, 45–50.

[23] Brianna Dym, Namita Pasupuleti, Cole Rockwood, and Casey Fiesler. 2021. " You don't do your hobby as a job": Stereotypes of Computational Labor and their Implications for CS Education. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 823–829.

[24] Rodrigo Ferreira and Moshe Y Vardi. 2021. Deep Tech Ethics: An Approach to Teaching Social Justice in Computer Science. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 1041–1047.

[25] Casey Fiesler, Mikhaila Friske, Natalie Garrett, Felix Muzny, Jessie J Smith, and Jason Zietz. 2021. Integrating Ethics into Introductory Programming Classes. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE'21). New York, NY, USA: ACM*.

[26] Casey Fiesler, Natalie Garrett, and Nathan Beard. 2020. What do We teach when We teach tech ethics? A syllabi analysis. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 289–295.

[27] Holly Fiock. 2020. Designing a community of inquiry in online courses. *The International Review of Research in Open and Distributed Learning* 21, 1 (2020), 135–153.

[28] Denae Ford, Justin Smith, Philip J Guo, and Chris Parnin. 2016. Paradise unplugged: Identifying barriers for female participation on stack overflow. In *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*. 846–857.

[29] D Randy Garrison, Terry Anderson, and Walter Archer. 1999. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The internet and higher education* 2, 2-3 (1999), 87–105.

[30] Emily Gonzalez-Holland, Daphne Whitmer, Larry Moralez, and Mustapha Mouloua. 2017. Examination of the use of Nielsen's 10 usability heuristics & outlooks for the future. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 61. SAGE Publications Sage CA: Los Angeles, CA, 1472–1475.

[31] Elizabeth Goodman, Mike Kuniavsky, and Andrea Moed. 2012. *Observing the User Experience, Second Edition: A Practitioner's Guide to User Research* (2nd ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[32] Jerilee Grandy. 1994. Gender and ethnic differences among science and engineering majors: Experiences, achievements, and expectations. *ETS Research Report Series* 1994, 1 (1994), i–63.

[33] Thomas A Harris. 1967. *I'm OK, you're OK.* Random House.

[34] Gillian R Hayes. 2014. Knowing by doing: action research as an approach to HCI. In *Ways of Knowing in HCI*. Springer, 49–68.

[35] C. Hilderbrand, C. Perdriau, L. Letaw, J. Emard, Z. Steine-Hanson, M. Burnett, and A. Sarma. 2020. Engineering Gender-Inclusivity into Software: Ten Teams' Tales from the Trenches. In *2020 IEEE/ACM 42nd International Conference on Software Engineering (ICSE)*. 433–444.

[36] Charles G. Hill, Maren Haag, Alannah Oleson, Chris Mendez, Nicola Marsden, Anita Sarma, and Margaret Burnett. 2017. *Gender-Inclusiveness Personas vs. Stereotyping: Can We Have It Both Ways?* Association for Computing Machinery, New York, NY, USA, 6658–6671. https://doi.org/10.1145/3025453.3025609

[37] Hsiu-Fang Hsieh and Sarah E Shannon. 2005. Three approaches to qualitative content analysis. *Qualitative health research* 15, 9 (2005), 1277–1288.

[38] Hui-Ching Kayla Hsu and Nasir Memon. 2021. Crossing the Bridge to STEM: Retaining Women Students in an Online CS Conversion Program. *ACM Transactions on Computing Education (TOCE)* 21, 2 (2021), 1–16.

[39] Margaretha Vreeburg Izzo and William M Bauer. 2015. Universal design for learning: enhancing achievement and employment of STEM students with disabilities. *Universal Access in the Information Society* 14, 1 (2015), 17–27.

[40] René Kizilcec and Andrew Saltarelli. 2019. Psychologically Inclusive Design: Cues Impact Women's Participation in STEM Education. In *Proceedings of the 2019 CHI Conference on human factors in computing systems (CHI '19)*. ACM, 1–10.

[41] Rene Kizilcec, Andrew Saltarelli, Petra Bonfert-Taylor, Michael Goudzwaard, Ella Hamonic, and Rémi Sharrock. 2020. Welcome to the Course: Early Social Cues Influence Women's Persistence in Computer Science. In *Proceedings of the 2020 CHI Conference on human factors in computing systems (CHI '20)*. ACM, 1–13.

[42] Rene F Kizilcec and Anna Kambhampaty. 2020. Identifying course characteristics associated with sociodemographic variation in enrollments across 159 online courses from 20 institutions. *PloS one* 15, 10 (2020), e0239766–e0239766.

[43] Maria Klawe. 2013. Increasing female participation in computing: The Harvey Mudd College story. *Computer* 46, 3 (2013), 56–58.

[44] Amy J Ko, Alannah Oleson, Neil Ryan, Yim Register, Benjamin Xie, Mina Tari, Matthew Davidson, Stefania Druga, and Dastyni Loksa. 2020. It is time for more critical CS education. *Commun. ACM* 63, 11 (2020), 31–33.

[45] Amy J Ko, Alannah Oleson, Neil Ryan, Yim Register, Benjamin Xie, Mina Tari, Matthew Davidson, Stefania Druga, and Dastyni Loksa. 2020. It is time for more critical CS education. *Commun. ACM* 63, 11 (2020), 31–33.

[46] Sandeep Kaur Kuttal, Kevin Gerstner, and Alexandra Bejarano. 2019. Remote Pair Programming in Online CS Education: Investigating through a Gender Lens. In *2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, 75–85.

[47] Lara Letaw, Rosalinda Garcia, Heather Garcia, Christopher Perdriau, and Margaret Burnett. 2021. Supplemental Document to "Changing the Online Climate via the Online Students: Effects of Three Curricular Interventions". https://doi.org/10.6084/m9.figshare.14294744

[48] Kurt Lewin. 1952. Group decision and social change. Readings in social psychology. *Newcombe and Hartley (Eds.), Henry Holt, New York* (1952).

[49] A. Light. 2011. HCI as heterodoxy: Technologies of identity and the queering of interaction with computers. *Interacting with Computers* 23, 5 (2011), 430–438. https://doi.org/10.1016/j.intcom.2011.02.002

[50] Stephanie Ludi, Matt Huenerfauth, Vicki Hanson, Nidhi Rajendra Palan, and Paula Garcia. 2018. Teaching inclusive thinking to undergraduate students in computing programs. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. 717–722.

[51] Jennifer Mankoff. 2006. Practical service learning issues in HCI. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. 201–206.

[52] Jane Margolis and Allan Fisher. 2002. *Unlocking the clubhouse: Women in computing.* MIT press.

[53] Allison Mattheis, Daniel Cruz-Ramírez De Arellano, and Jeremy B Yoder. 2019. A model of queer STEM identity in the workplace. *Journal of homosexuality* (2019).

[54] Christopher Mendez, Lara Letaw, Margaret Burnett, Simone Stumpf, Anita Sarma, and Claudia Hilderbrand. 2019. From GenderMag to InclusiveMag: An inclusive design meta-method. In *2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, 97–106.

[55] Ryan A Miller and Megan Downey. 2020. Examining the STEM Climate for Queer Students with Disabilities. *Journal of Postsecondary Education and Disability* 33, 2 (2020), 169–181.

[56] National Center for Women and Information Technology (NCWIT). 2015. *Survey-in-a-Box: Student Experience of the Major.* https://www.ncwit.org/resources/survey-box-student-experience-major-0

[57] Jakob Nielsen. 1994. *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier.* Academic Press, Inc., USA, 245–272.

[58] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 249–256.

[59] Alannah Oleson, Christopher Mendez, Zoe Steine-Hanson, Claudia Hilderbrand, Christopher Perdriau, Margaret Burnett, and Amy J. Ko. 2018. Pedagogical Content Knowledge for Teaching Inclusive Design. In *Proceedings of the 2018 ACM Conference on International Computing Education Research* (Espoo, Finland) *(ICER '18)*. Association for Computing Machinery, New York, NY, USA, 69–77. https://doi.org/10.1145/3230977.3230998

[60] Alannah Oleson, Meron Solomon, and Amy J Ko. 2020. Computing Students' Learning Difficulties in HCI Education. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[61] Susmita Hema Padala, Christopher John Mendez, Luiz Felipe Dias, Igor Steinmacher, Zoe Steine Hanson, Claudia Hilderbrand, Amber Horvath, Charles Hill, Logan Dale Simpson, Margaret Burnett, et al. 2020. How gender-biased tools shape newcomer experiences in OSS projects. *IEEE Transactions on Software Engineering* (2020).

[62] Krystle Phirangee and Alesia Malec. 2017. Othering in online learning: An examination of social presence, identity, and sense of community. *Distance Education* 38, 2 (2017), 160–172.

[63] Vahab Pournaghshband and Paola Medel. 2020. Promoting Diversity-Inclusive Computer Science Pedagogies: A Multidimensional Perspective. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*. 219–224.

[64] Cynthia Putnam, Maria Dahman, Emma Rose, Jinghui Cheng, and Glenn Bradford. 2016. Best practices for teaching accessibility in university classrooms: cultivating awareness, understanding, and appreciation for diverse users. *ACM Transactions on Accessible Computing (TACCESS)* 8, 4 (2016), 1–26.

[65] Inioluwa Deborah Raji, Morgan Klaus Scheuerman, and Razvan Amironesei. 2021. You Can't Sit With Us: Exclusionary Pedagogy in AI Ethics Education. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 515–525.

[66] Jennifer A Rode and Erika Shehan Poole. 2018. Putting the gender back in digital housekeeping. In *Proceedings of the 4th Conference on Gender & IT*. 79–90.

[67] Ayesha Sadaf, Florence Martin, and Lynn Ahlgrim-Delzell. 2019. Student Perceptions of the Impact of Quality Matters–Certified Online Courses on Their Learning and Engagement. *Online Learning* 23, 4 (2019), 214–233.

[68] Kay Shattuck, Whitney Alicia Zimmerman, and Deborah Adair. 2014. Continuous improvement of the QM rubric and review processes: Scholarship of integration and application. *Internet Learning* 3, 1 (2014), 5.

[69] Arun Shekhar and Nicola Marsden. 2018. Cognitive Walkthrough of a Learning Management System with Gendered Personas. In *Proceedings of the 4th Conference on Gender & IT* (Heilbronn, Germany) *(GenderIT '18)*. Association for Computing Machinery, New York, NY, USA, 191–198. https://doi.org/10.1145/3196839.3196869

[70] Kristen Shinohara, Cynthia L Bennett, and Jacob O Wobbrock. 2016. How designing for people with and without disabilities shapes student design thinking. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. 229–237.

[71] Lee Shulman. 1987. Knowledge and teaching: Foundations of the new reform. *Harvard educational review* 57, 1 (1987), 1–23.

[72] Steven E Stemler. 2004. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation* 9, 1 (2004), 4.

[73] Jane G Stout and Heather M Wright. 2016. Lesbian, gay, bisexual, transgender, and queer students' sense of belonging in computing: An Intersectional approach. *Computing in Science & Engineering* 18, 3 (2016), 24–30.

[74] Ernest T Stringer. 2007. *Action Research*. Sage.

[75] Simone Stumpf, Anicia Peters, Shaowen Bardzell, Margaret Burnett, Daniela Busse, Jessica Cauchard, and Elizabeth Churchill. 2020. Gender-inclusive HCI research and design: A conceptual review. *Foundations and Trends in Human–Computer Interaction* 13, 1 (2020), 1–69.

[76] Lucy Suchman. 2009. Agencies in technology design: Feminist reconfigurations. In *Proceedings of 5th European Symposium on Gender & ICT, Digital Cultures: Participation–Empowerment–Diversity*.

[77] Adrian Thinnyun, Ryan Lenfant, Raymond Pettit, and John R Hott. 2021. Gender and Engagement in CS Courses on Piazza. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 438–444.

[78] Jan H Van Driel, Nico Verloop, and Wobbe De Vos. 1998. Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching* 35, 6 (1998), 673–695.

[79] Bogdan Vasilescu, Andrea Capiluppi, and Alexander Serebrenik. 2014. Gender, representation and online participation: A quantitative study. *Interacting with Computers* 26, 5 (2014), 488–511.

[80] Mihaela Vorvoreanu, Lingyi Zhang, Yun-Han Huang, Claudia Hilderbrand, Zoe Steine-Hanson, and Margaret Burnett. 2019. From Gender Biases to Gender-Inclusive Design: An Empirical Investigation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300283

[81] Annalu Waller, Vicki L Hanson, and David Sloan. 2009. Including accessibility within and beyond undergraduate computing courses. In *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 155–162.

[82] Linda L Werner, Brian Hanks, and Charlie McDowell. 2004. Pair-programming helps female computer science students. *Journal on Educational Resources in Computing (JERIC)* 4, 1 (2004), 4–es.

[83] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. 2012. *Experimentation in software engineering*. Springer Science & Business Media.

[84] Kimberly Michelle Ying, Lydia G Pezzullo, Mohona Ahmed, Kassandra Crompton, Jeremiah Blanchard, and Kristy Elizabeth Boyer. 2019. In their own words: Gender differences in student perceptions of pair programming. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. 1053–1059.