

Engagement and Anonymity In Online Computer Science Course Forums

Mrinal Sharma*

University of California, San Diego
USA
m1sharma@ucsd.edu

Anshul Shah

University of California, San Diego
USA
ayshah@ucsd.edu

Emma Hogan

University of California, San Diego
USA
emhogan@ucsd.edu

Hayden McTavish*

University of California, San Diego
USA
hmctavish@ucsd.edu

Vardhan Agarwal

University of California, San Diego
USA
v7agarwa@ucsd.edu

Ismael Villegas Molina

University of California, San Diego
USA
isvilleg@ucsd.edu

Kristen Vaccaro

University of California, San Diego
USA
kvaccaro@ucsd.edu

Zimo Peng

University of California, San Diego
USA
z5peng@ucsd.edu

Caroline Sih

University of California, San Diego
USA
sih.caroline@gmail.com

Adalbert Gerald Soosai Raj

University of California, San Diego
USA
asoosairaj@ucsd.edu

Abstract

Online discussion boards, designed to facilitate learning from peers and instructors in an accessible space, are a vital part of course design, especially in large scale computer science classes. Previous work has shown that women in computer science tend to use anonymity more often than men on these boards, a trend not found in humanities, social science or business courses. In this work, we build on these findings using an intersectional lens, analyzing both gender and race/ethnicity. We find this combined analysis reveals differences in anonymity that are not apparent when examining gender alone. For example, we find a significantly greater difference in anonymity use between Hispanic men and women than would be expected from analyzing race/ethnicity and gender independently. We additionally analyze type of content (e.g., questions, answers), course, platform, and data source to characterize the many factors at play in measuring students' choice to participate anonymously. In doing so, we show that different approaches used in prior work for eliciting information on gender — whether using registrar data, a survey, or imputing gender based on name — changes how over 20% of students are classified, particularly affecting nonbinary students and Asian students. Understanding when students participate

anonymously can help educators and platform designers to make students' experience of online discussion boards more welcoming.

CCS Concepts

• **Social and professional topics** → **Computer science education**; • **Human-centered computing** → **Empirical studies in collaborative and social computing**.

Keywords

anonymity, course forums, inclusion

ACM Reference Format:

Mrinal Sharma*, Hayden McTavish*, Zimo Peng, Anshul Shah, Vardhan Agarwal, Caroline Sih, Emma Hogan, Ismael Villegas Molina, Adalbert Gerald Soosai Raj, and Kristen Vaccaro. 2023. Engagement and Anonymity In Online Computer Science Course Forums. In *Proceedings of the 2023 ACM Conference on International Computing Education Research V.1 (ICER '23 V1)*, August 7–11, 2023, Chicago, IL, USA. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3568813.3600121>

1 Introduction

Online course discussion boards, designed to make learning from peers and instructors easier and more accessible, are a vital part of course design, especially in large scale computer science classes. Many of these course forums give students the option of making their posts anonymous, which prior work has suggested encourages students with low confidence to participate [55, 68]. Previous work has shown that women in computer science tend to use anonymity more often than men on these boards [29, 59, 75], a trend not found in humanities, social science or business courses [59]. However, as with many efforts at broadening participation in computing [38], this prior work has primarily focused on gender.

* Both authors contributed equally to this research.

We introduce an intersectional analysis to understand engagement and use of anonymity in online course forums. *Intersectionality* is a theory that argues that when multiple axes of identity — including race, gender, and class — overlay, they often create experiences that differ from the individual parts. As originally introduced by Crenshaw, this framework focused on overlapping and interdependent systems of discrimination, recognizing that the harms experienced by Black women (as an intersectional identity) can be greater than those they might experience from gender or race alone [13]. More recently, intersectionality has been defined more inclusively, as a way of thinking about “*different kinds of difference*”, where complex and varied effects can result when differences intersect, and researchers should “*avoid conflating positionings, identities and values*” [80].

In this work, we analyze use of anonymity starting from traditional single-factor approaches, then building into an intersectional approach. First, we consider gender. We also introduce a new analysis, looking at race/ethnicity. Finally, we analyze both aspects of identity at the same time. In doing so, we identify impacts of gender and race/ethnicity *jointly* that would be not visible from analyzing either factor alone. This forms our first research question:

- RQ1** To what extent does anonymity usage rate for all posts differ based on student demographics?
- (a) based on student gender?
 - (b) based on student race/ethnicity?
 - (c) based on an intersection of gender and race?

To do so, we collect a data set of 7847 posts by 2216 students, 1578 of whom made at least one post. These posts are drawn from Piazza and Ed Discussion¹ forums for eight undergraduate computer science courses at a large, research-focused university over two terms during the 2021–2022 academic year. To understand student demographics, we collected data from two sources. For students in the first term, we collected demographic information from the university registrar, including gender, race, and ethnicity. For students in the second term, we used self-reported data on gender, race, and ethnicity, that students submitted via a pre-course survey.

In addition to studying anonymity use in general, we study the impact of several types of context: post types, platforms, and courses. First, we follow prior work [29, 75], and analyze differences in anonymity use for the different post types: questions, answers, follow ups, and notes. And because our data collection includes posts from two platforms and eight courses, we additionally study the effect of platform and course on anonymity use, once again using an intersectional approach. These form our second research question:

- RQ2** To what extent does anonymity usage rate differ based on intersectional demographics in different contexts?
- (a) for different post types, i.e., questions, answers, follow-ups, notes?
 - (b) for different courses?
 - (c) for different online discussion forums, i.e., Piazza, Ed Discussion?

We also bring together two related threads of prior work: the use of anonymity and decisions to lurk [54, 58]. Lurkers are community

members who read but do not post. While some have argued that listening is a powerful and necessary form of participation [12], lurkers have traditionally been viewed in similar ways to anonymous posters [41]. As a result, researchers have extensively studied what factors influence decisions to contribute, including both aspects of the individual and the environment [44, 51, 65]. Interestingly, prior work shows mixed results on whether gender impacts decisions to post [54, 58]. We again take an intersectional approach:

- RQ3** To what extent does lurking rate differ based on intersectional demographics?

Finally, prior efforts to study effects of gender on anonymity use have revealed a number of challenges around data quality that we attempt to address. In some cases, researchers have been forced to reduce gender to a binary (men/women) due to low counts of nonbinary students [29]. However, in others, researchers have used name-based gender predictions to predict binary gender [75], in ways that we argue are fundamentally flawed. To audit the performance of these name-based gender prediction tools, we use the same pipeline to generate name-based inferences for our data. By comparing these predictions with our other data sources (survey and registrar data), we can assess the impact on perceived gender, again using an intersectional approach, in our final research question:

- RQ4** To what extent do prediction models for imputing gender match self-reported data?

2 Theoretical Framework

This paper adopts theoretical framings from two areas: computing education and social computing.

2.1 Computing Education

From the computing education literature, we emphasize the role of course-focused discussion forums for learning communities. Learning communities are a long-standing movement in higher education, with significant impacts on academic achievement, retention rates, and curricular integration [35]. These communities build on foundations that learning is most successful as a social process — emphasizing cooperative and collaborative practices [15]. Generally these efforts aim to create opportunities for small groups of students to learn, often directing their own activities [8].

From the introduction of learning communities, it has been recognized that they can be either in-person or virtual [35]. As a result, researchers have designed methods to support the growth of learning communities using online technologies. For example, Scardamalia and Bereiter’s Knowledge Forum², prototyped as early as 1983, was motivated by a belief that students were a largely wasted resource and could engage more deeply through network technology [60]. In practice, this was wildly successful. The Knowledge Forum “*restructured the flow of information in the classroom, so that questions, ideas, criticisms, suggestions, and the like were contributed to a public space equally accessible to all*,” creating “*an emergent hypertext that represented the collective rather than only the individual knowledge*” [60]. In doing so, these virtual spaces collectively represent knowledge and engage students in active learning.

¹Previously known as EdStem

²Also known as CSILE (Computer Supported Intentional Learning Environments)

In addition, we argue online course discussion forums also support students as they become members of a community of practice. In the community of practice literature, learning takes place through informal interactions and social activity that moves newcomers towards expertise in a profession. Online course discussion forums enable students to understand norms and practices for online spaces that are integral parts of the computing profession, like Stack Overflow and Github. These online spaces have been extensively studied with respect to gender differences. On Stack Overflow, only 5.8% of contributors are women [22], women receive less recognition for their contributions [10, 40, 78], and women who post face hostility and harassment [9]. This hostility exists despite the fact that women’s contributions are more likely to be accepted — although this is only true when their gender is obscured [74]. We suggest that course discussion forums are an important space where students learn practices for online communication as computer scientists.

2.2 Social Computing

From the social computing literature, we identify the impact of anonymity on social interactions. The social computing field builds extensively on Goffman’s *Presentation of Self in Everyday Life* [25]. This seminal sociological theory develops the metaphor of the world as a stage — arguing that, whether intentional or habitual, people put on performances for those around them. Those performances aim to suggest a particular public image, that both achieves the individual’s goals while also aligning with others’ expectations. People do so by regulating and controlling information in a social interaction, a process called *impression management*. However, impression management can have costs, because it requires keeping one’s internal preferences, beliefs, or information secret.

Often these costs are framed in terms of the functional model of self-disclosure [6]. In this model, there are both risks and benefits to disclosing personal information. Benefits include social validation, self-expression, and releasing the stress of keeping one’s thoughts secret. Risks include increased vulnerability, loss of privacy, and negative impacts on self-presentation [2, 4, 18, 73]. In online spaces, this can happen frequently as users experience context collapse [16, 39], where the actual audience for a post is larger or different from their *imagined audience* [36, 39, 66].

In this landscape, anonymity can play an important role. In particular, anonymity offers a disinhibition effect [72], making people feel safer communicating. Early evidence of this was found in person, where the “strangers on a train” phenomenon encouraged people to disclose highly personal information to those they did not expect to encounter again [57]. Online this disinhibition effect has been found to have both positive and negative effects. Negative effects include an increased aggression and antisocial behaviors, for example, posting offensive content [7] or trolling [11, 26, 70]. However, anonymity can also have important positive effects, allowing people space to share information that might reflect poorly on them, for example parenting struggles [2], mental health issues [17], or experiences with abuse [4]. Providing anonymity allows users to experience the benefits of self disclosure, including self-expression and social support, without the costs to their self-presentation [34, 71, 76]. As a result, researchers have found that anonymous and pseudonymous accounts can help users share more authentically online [28, 73, 79].

In the case of an online course discussion forum, we could imagine a situation where a student debates whether to ask a question they think will reveal their lack of understanding about the course material: they must weigh the benefits (potential social validation, learning the answer to specific questions) with the risks (negative effects to their image). By providing anonymity, online course discussion forums, can leverage the disinhibition effect both by reducing risks to self-presentation and also reducing students’ inhibitions about messaging instructors.

3 Related Work

As the ratio of students to faculty increases due to burgeoning computer science enrollment, online forums such as Piazza have taken on a significant role connecting students with teaching assistants, instructors, and other students. With more than 5 million users worldwide, a significant number of students depend on Piazza to get help from course staff [49]. This is particularly true for computer science courses. Through personal communication, we learned that 34% of the courses using Piazza at our institution in the year we studied were computing courses, and about 25% of the total use of Ed Discussion comes from computing courses. There are an increasing number of spaces available for these online discussions, including tailored online discussion boards (e.g., Piazza, Ed Discussion), learning management systems (e.g., Canvas, Blackboards), and other social media (e.g., Discord, Slack). As the use of these systems increases, it becomes more important to understand the impact of them on classrooms.

3.1 Anonymity in Classrooms

As technology becomes more important in classrooms, a number of researchers have studied whether and how to support anonymity inside the classroom. For example, many researchers have looked at classroom response systems (also referred to as electronic response systems), such as the iClicker, that can allow anonymous answering within the classroom. Researchers have found that anonymity increases engagement [24] and participation in in-class discussions [45]. Others have found anonymity useful for in-class peer assessments, to reduce negative social effects [77]. Some have sought to understand the factors that influence use of anonymity in classrooms. Latham et al. suggested a preference for anonymity typically indicates “*the introvert, the less willing to show self-performance, and the feeling of distance from others*” [33]. In a comparison between online and face-to-face settings, Miyazoe et al. found that although all students appreciated pseudonyms, women increased participation when anonymous [43].

3.2 Anonymity in Classroom Discussion Boards

Researchers have more narrowly explored use of anonymity on classroom discussion boards. Past studies suggest that students are more comfortable posting anonymously, and anonymity can increase student willingness to participate [24, 55]. In this way, online forums can present an opportunity for students who might not otherwise ask questions and discuss key concepts. In computer science classes, anonymity appears to play an especially important role in women’s use of discussion boards. Several studies suggest that women use anonymity more often than their men peers [29, 59, 75]. Thus, having anonymity features as an affordance of educational

social media could be associated with challenges in computer science education surrounding sense of belonging and confidence of women [59, 75]. This pattern of anonymity usage, in turn, may shape students’ experience of an online discussion board. In particular, women posting anonymously more often has the potential to cause student perceptions that most posts, including anonymous ones, are made by men[29].

Although there is a clear consensus from prior work around anonymity rates, there is less consistency in results for lurking behaviors. A census study of Piazza in 2015 found that women provided more answers than men [59]. However, one recent study of computer science classes found no statistically significant difference in the answering rates of men and women [75] and recent work on graduate students identified no effect of gender on lurking behavior [58].

In this work, we extend this gender-focused analysis with an intersectional lens. Recent work has noted the importance of intersectional analyses in engineering [37]. We begin by conducting an intersectional analysis of anonymity use by race/ethnicity and gender. Prior work has examined some factors of context (in particular, distinguishing behavior in questions vs. answers), we extend this by examining the effect of post type, course, and platform. We also audit a gender imputation pipeline that has been used in prior work [75, 78], identifying flaws in this approach.

4 Methods

In this section we describe our data collection and preparation methods, alongside information about the institutional context. We received approval from our IRB for this data collection and analysis.

4.1 Data Collection

4.1.1 Discussion Forums Our dataset consisted of 7847 posts on the Piazza and Ed Discussion forums for eight undergraduate computer science courses at a large, public, research university located in North America. We requested discussion board data from instructors of every four-unit computing course considered a core requirement for the CS major at the university during one term; of these fourteen courses, we received permission from six different courses. We supplemented this data with an additional two courses in the following term, taught by one of the authors. All data was collected during the 2021–2022 academic year. Our data for each course included for every post: the posting student’s email, anonymity status for the post, and post type (question, answer, note or followup). Notes operate like announcements, sharing important information or general insights. Follow ups are opportunities for expanded discussion about a similar topic, once a question has been answered. These might provide additional information, engage in discussions, or offer solutions related to the initial note or question. In total 5 courses used Piazza, and 3 courses used Ed Discussion. After the processing described in Sections 4.2.2 and 4.2.1, we use and analyze 7133 posts total. Table 1 describes the number of posts for each post type and platform.

4.1.2 Students & Demographic Data We collected information for 2216 students, 1578 of whom made at least one post (Table 2). For students in the first term, we collected demographic information from the university registrar, including gender, race, and ethnicity.

	Piazza		Ed Discussion	
Questions	1813	47.8%	1555	46.5%
Answers	650	17.2%	205	6.1%
Follow ups	1281	33.8%	1557	46.5%
Notes	44	1.2%	26	0.8%
Total	3790		3343	

Table 1: Total number of posts (and percentages) for each post type for the two platforms in our data set.

For students in the second term, we had access to self-reported data on gender, race, and ethnicity, that students submitted via a pre-course survey. Approximately 367 students enrolled in courses included in the dataset both terms; for these students we have both registrar and self-reported survey data.

For Piazza, we also could not analyze demographic information for two groups of users: those who posted with full anonymity (anonymous not just to students, but to instructors as well), or those who posted and then unenrolled from the course before we collected our data. This constituted 320 posts (3.5%).

4.1.3 Gender Data The university registrar collects student demographics during the application process. The university’s data collection process provides a number of options: *Male*, *Female*, *Trans Male / Trans Man*, *Trans Female / Trans Woman*, *Genderqueer / Gender Non-Conforming*, *Decline to State*, or a self-description option. The registrar mapped these to four categories, which we refer to as *Men* (*Male* and *Trans Male / Trans Man*), *Women* (*Female* and *Trans Female / Trans Woman*), *Unknown* (*Decline to State*), and *Non-binary Students* (all other options). In the survey, students were asked to report their pronouns, given a set of options (he/him/his, she/her/hers, they/them/theirs, Prefer not to say, or a short text input), and asked to select one or more of the options. These were collapsed into four categories (*Men* for students who wrote *only* he/him, *Women* for students who wrote *only* she/her, *Nonbinary Students* for students who wrote any other pronouns (e.g., she/they), and unknown for students who did not provide pronouns.

Prior work has several key limitations related to how gender data is collected and reported. Existing studies largely report gender as a binary, excluding nonbinary individuals [29, 59, 75]. Moreover, many cases of prior work do not provide clarifying information on how gender is assessed. To our knowledge, Piazza only allows users to check an optional box indicating that they are women³ as a means of indicating their gender. Past work using Piazza data either provides no information about how this limitation was addressed in data collection [59] or uses software tools such as genderComputer⁴ and gender-guesser⁵ to impute a binary gender from students’ names [75]. In our work, we measure the differences between gender data if generated with these predictions and the survey and registrar data we collect, to provide an estimate of the accuracy of this approach.

³Specifically, the Piazza profile interface, accessed March 2022, had a single checkbox labelled “female” which students could mark or leave blank, with no other gender categories. The interface has more recently been updated to include “non-binary”.

⁴<https://github.com/tue-mdse/genderComputer>

⁵<https://pypi.org/project/gender-guesser/>

Gender	Black		Asian		Hispanic		White		Total	
	P	E	P	E	P	E	P	E	P	E
Women	1	9	190	207	23	69	26	58	240	343
Men	8	15	438	291	53	43	75	62	575	412
Nonbinary Students	0	1	4	15	2	8	0	6	6	30
Total	9	25	632	513	78	120	101	126	821	785

Table 2: Demographics of students included in dataset. Counts are separated for Piazza (P) and Ed Discussion (E).

4.1.4 Race/Ethnicity Data The university also collects race/ethnicity information during the application process. Students can select one or more checkboxes indicating their race/ethnicity. The survey contained a similar multiple-selection checkbox menu, with an added option “Other” that allowed students to write in their own description. Table 3 shows the race/ethnicity options in each dataset.

While the registrar allows students to respond with multiple races or ethnicities, any multiracial response is collapsed to a single value. According to university documents, the single value is the earliest race/ethnicity on the list (in Table 3) selected by the student.

African American/Black
Chicano/Latino
American Indian/Alaska Native
Native Hawaiian/Pacific Islander
Asian/Asian American
White

Table 3: Race/ethnicity options provided by the registrar and survey.

To ensure compatibility, we applied the same rules to the data from the survey, collapsing multiracial responses to a single value. This allowed us to run consistent aggregated analysis across all classes, using both data sources. There are clear limitations to not having data on multiracial student experiences, which we hope future researchers will investigate with datasets that include this information.

Three race/ ethnicities were removed from our analysis, due to small counts. These were Native Hawaiian/Pacific Islander students (2 in the dataset, 0 who made public posts) and American Indian/Alaska Native students (4 in the dataset, 2 who made public posts). In addition, the data contained one category (International) without documentation. The 14 students (11 of whom posted publicly) with this race/ethnicity label were removed from the analysis.

4.2 Data Processing

4.2.1 Removing Instructor Posts Ed Discussion data indicates for each post whether it was made by an instructor/TA or a student. For Piazza, this information is only included for answers. To remove instructor posts from the data, we tagged every email associated with an instructor answer, added in the emails of the instructors who taught the course, and removed any post associated with any of those emails. 11369 posts were removed in this way from the Piazza dataset. We also removed 5376 posts from the Ed Discussion dataset, using the indicator on each post about whether the poster was an instructor.

4.2.2 Removing Private Posts Some of the data we received included private posts that students had made visible only to instructors. We removed this data from our analysis of anonymity rates, though we did use it when studying lurking behavior. The data from Ed Discussion provided a simple way to differentiate between private and public posts, so we simply filtered out all posts in the private category. Piazza provides no way to differentiate between private and public posts from the datasets we had access to. Instead, we used view count, which corresponds to the number of unique users who view the post. We removed any post whose view count did not exceed the number of instructor/TA accounts in the course. 4266 and 3081 posts were removed from the Piazza and Ed Discussion data, respectively.

4.3 Institutional Context

No courses gave credit for participation on online discussion forums, except for Introductory CS B which tracked top contributors, whose grades would be rounded up if close enough to a grade boundary. In addition to receiving help via a discussion board, students in all courses could attend student teaching assistant and instructor office hours, as well as view recordings of lectures. All courses gave students the option to be anonymous to their classmates, but only two courses (Introductory CS C and Introductory Lab) had the option to post anonymously to instructors as well.

In accordance with COVID-19 requirements, all courses in the first term (that is, all courses except Introductory Lab and Introductory CS A) were offered remotely for the first 5 weeks. Afterwards, some of these switched to either in-person or hybrid modalities. Several courses were co-taught by two instructors, who offered different modalities. All modalities are listed in Table 4. In this table, modalities are described after the 5th week.

Most courses included some form of inclusion practices (e.g., peer instruction, active learning), while Introductory CS A offered some

Course	# Students	Modality
Introductory CS A	433	Remote, In-person
Introductory CS B	155	Remote, Hybrid*
Introductory CS C	116	Hybrid*
Introductory Lab	195	Remote, In-person
Basic Data Structures	245	Remote, In-person*
Software Engineering	110	In-person*
Systems Programming	193	Remote, In-person*
Advanced Data Structures	159	Remote*

Table 4: Course enrollment and modality.

unique activities (e.g., inclusion minutes, live coding with errors). Nevertheless, the Computer Science & Engineering department can have a competitive environment. Our university had a 24% acceptance rate overall in 2022. At the time of data collection, it had an extremely selective lottery system process for students who were not directly admitted to the major. To be eligible for the lottery, students faced a GPA cutoff as well as required introductory courses, which could create a high-stakes environment for those seeking to transfer.

All courses except for Introductory CS A and Introductory CS C require at least one prerequisite course on this list. Introductory CS C expects students to have had prior exposure to computing, such as a high school AP course; Introductory CS A does not.

5 Analysis

In this work, we provide an intersectional analysis of use of anonymity, to build on prior work that looked only at gender. Similar to prior work, we encounter challenges where the number of students with some identities, particularly intersectional identities (e.g., Black women), are too small to report on. Following the reporting standards for our institution⁶, any group with fewer than five students is omitted. This becomes more common when the analysis focuses on narrower slices of the data (e.g., only looking at answering behavior).

5.1 Representing Anonymity Choices

Some students post much more frequently than others. To avoid over representing behavior by students who participate frequently, we measure and analyze anonymity decisions on a per-student basis. In addition, individual students post both anonymously and by name. To capture these choices, we capture how likely each student is to post anonymously, that is, if someone posts, what fraction of their posts are anonymous?

Therefore, for each demographic category, we report the *mean anonymity rate* for a given demographic group. The mean anonymity rate is the ratio of anonymous posts to all posts, calculated for each student, and averaged over all students in a demographic category. This captures the likelihood of a student in a given category to choose to use anonymity for a post. For example, a demographic category comprised of 3 students—one student who made 2 anonymous posts out of 10 total posts, another student who made 3 anonymous posts out of 3 total posts, and another who made 0 anonymous posts out of 2 total posts—would have a mean anonymity rate of $\frac{\frac{2}{10} + \frac{3}{3} + \frac{0}{2}}{3} = 0.4$. Students who do not post are omitted from this calculation.

5.2 Independent Analysis of Gender And Race/Ethnicity

For **RQ1**, we begin by analyzing gender and race/ethnicity independently. One way analyses (gender alone and race/ethnicity alone) are conducted using a series of pairwise z-tests of proportions [61]. For example, in a gender analysis, we compute three comparisons: *men-women*, *men-nonbinary students*, and *women-nonbinary students*. To control for these multiple comparisons, we use a Holm-Bonferroni correction, correcting for 3 tests for gender and 6 tests

for race/ethnicity. We report z-score, p-value and effect size (Cohen’s d) for all statistically significant results after applying this correction. In some statistical contexts, it is standard to run an initial test, and only to run pairwise, post-hoc comparisons if the initial comparison is statistically significant. In our context, we chose not to do this since the most standard tool for such an approach, analysis of variance (ANOVA), does not perform well for proportion data that varies from 0–1, such as an anonymity rate. We instead run pairwise comparisons. This approach has precedent in computing education when working with categorical variables [31].

5.3 Intersectional Analysis

For **RQ1(c)** and **RQ2**, we conduct analyses intersectionally. Following the most common approach in intersectional analysis, we use regression models [3]. In particular, we use a beta regression [14], which is a popular model when the dependent variable takes the form of fractions or percentages. This allows us to control for any effects of race/ethnicity and gender separately, and identify cases where the intersection has a significant effect above and beyond that. Following convention, all comparisons are made relative to a “baseline” group, the largest group. In our analysis, this is the category Asian for race/ethnicity and the category men for gender. When running beta regression with categorical interaction effects, it is possible to encounter issues when counts for some categories are 0 or extremely small. To handle this, if our optimizer encountered an error we remove nonbinary students from our beta regression model (for some races/ethnicities, nonbinary students have extremely small counts), then if there is still an issue we remove black students from the model. We never needed to remove any further categories, and we only remove either of these categories when we could not get a result from beta regression without doing so.

5.4 Lurking Analysis

For **RQ3**, to understand the effect of gender on lurking behavior, we conduct a chi square test of independence. We do so for three analyses (gender, race/ethnicity, and the intersection). We do not use a regression model, because lurking behavior is a categorical response with three levels: no posts, private posts, and public posts.

An important note about this analysis is that we conduct it only on Piazza data. While Piazza includes a file of all students (including

	Total	Women	Men	Nonbinary Students
All Posts	52.8%	52.8%	53.7%	28.5%

(a) Anonymity rates for all posts by gender

	Total	Asian	Black	Hispanic	White
All Posts	52.8%	57.0%	31.9%	45.7%	41.3%

(b) Anonymity rates for all posts by race/ethnicity

Table 5: Differences in anonymity rate for all posts. Analysis of (a) gender and (b) race/ethnicity independently.

⁶<https://diversity.ucsd.edu/accountability/undergrad.html>

those who do not post), Ed Discussion does not. There do not appear to be any significant platform design choices (as in the anonymity of followups) that impact lurking behavior, so we suspect the Ed Discussion analyses would have a similar proportion of lurkers.

5.5 Gender Data Source Analysis

In our final analysis for **RQ4**, we directly compare different data sources for gender. We have three sources of gender information: self-reported survey data, registrar data, and predicted gender based on poster name. The last of these has been used in prior work for estimating gender, when other gender information is unavailable [75, 78]. To make the name-based gender predictions, we follow prior work [75], using two software tools *genderComputer* and *gender-guesser*. If *genderComputer* is unable to predict a binary gender, *gender-guesser* is used. When neither tool predicts a binary gender, it is reported as unknown.

We calculate confusion matrices for these three data sources. For classes from the first term we studied, we compare registrar data to the automated predictions. For classes in the second term we studied, we compare survey data to the automated predictions. And for the subset of students who enrolled in classes we studied in both the first and second term, we compare registrar to survey data. We report statistics of accuracy for all three comparisons.

6 Results

6.1 RQ1: Use of Anonymity For All Posts

6.1.1 Analysis By Gender Table 5a shows the anonymity rate for all posts analyzed by gender. We found that men and women post anonymously much more than their nonbinary classmates. The pairwise *z* test for women and nonbinary students proved to be statistically significant ($z=3.24$, $p<0.01$, $d=0.56$) as was the test for men and nonbinary students ($z=3.38$, $p<0.001$, $d=0.57$). However, women’s and men’s anonymity usage rates are similar ($z=-0.39$, $p=0.70$, $d=-0.02$).

6.1.2 Analysis By Race/Ethnicity We also analyze anonymity rate for all posts by race/ethnicity alone (Table 5b). We find that Asian students post anonymously much more than their White, Hispanic, and Black counterparts with *z* scores, *p* values, and effect sizes ($z=4.94$, $p<0.001$, $d=0.36$), ($z=3.36$, $p<0.001$, $d=0.26$), and ($z=3.31$, $p<0.001$, $d=0.58$), respectively. Pairwise comparisons between other race/ethnicity groups do not reveal statistically significant differences.

6.1.3 Intersectional Analysis Taking an intersectional lens, we analyze the impact of gender and race/ethnicity jointly. Figure 1 shows the mean anonymity rate for each intersectional identity group, while Table 6 shows the results of the beta regression. In Figure 1, white boxes indicate a group where fewer than five students posted.

As in the independent analyses, we find no effect of gender for men and women but do find effects for nonbinary students. For race/ethnicity, we once again find that White students are less likely than Asian students to use anonymity, but the reduced use of anonymity for Black and Hispanic students is no longer statistically significant. Importantly, we find there are significant effects *beyond* what we would expect from analyzing race/ethnicity and gender independently. In particular, Hispanic women have a much lower

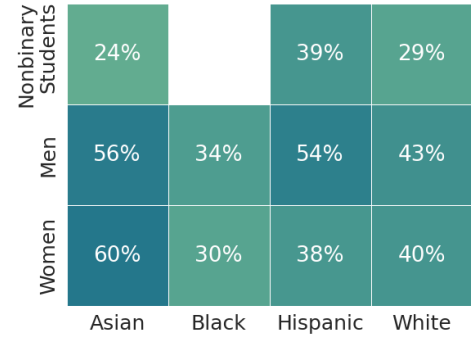


Figure 1: Intersectional anonymity rate for all posts

	β	Error	z value	$\Pr(> z)$
(Intercept)	0.15	0.05	2.964	0.003
<i>Independent Factors</i>				
Gender (W)	0.07	0.08	0.830	0.407
Gender (NB)	-0.82	0.30	-2.731	0.006
Race/Eth. (Black)	-0.53	0.28	-1.894	0.058
Race/Eth. (Hispanic)	-0.07	0.15	-0.500	0.617
Race/Eth. (White)	-0.35	0.12	-2.771	0.006
<i>Intersectional Factors</i>				
Gender (W) & R/Eth. (B)	-0.20	0.50	-0.389	0.697
Gender (W) & R/Eth. (H)	-0.48	0.21	-2.280	0.022
Gender (N) & R/Eth. (H)	0.40	0.53	0.756	0.450
Gender (W) & R/Eth. (W)	-0.17	0.20	-0.826	0.409
Gender (N) & R/Eth. (W)	0.44	0.62	0.708	0.479

Table 6: Beta regression output for anonymity rate for all posts

anonymity rate (38%) than we would expect, particularly given that Hispanic men use anonymity frequently (54%). We investigate possible reasons for this effect in the analysis for **RQ2(b)** and **(c)** and discuss in the Discussion.

Key Findings: Women and men’s anonymity use was similar, but nonbinary students were much less likely to use anonymity. Asian students used anonymity much more than their White, Hispanic and Black counterparts. Intersectional analysis revealed that Hispanic women have a much lower anonymity usage rate than Hispanic men.

6.2 RQ2(a) Use of Anonymity By Post Type

Online course discussion forums allow several types of posts: questions, answers, notes, and followups. Although we find there are too few notes to analyze (70 notes, which is less than 1% of the dataset), we conduct an intersectional analysis on anonymity use for the other three post types. Table 10a shows the anonymity rate for different post types by gender, Table 10b shows the anonymity rate for different post types by race/ethnicity. Figures 2, 3, and 4 show the intersectional anonymity rates for questions, answers, and follow ups, paired with tables that show the beta regression results for each analysis (Table 7, 8, and 9).

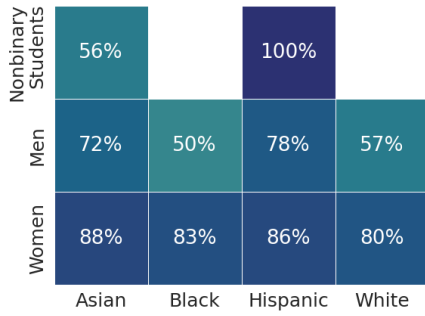


Figure 2: Intersectional anonymity rate for questions

	β	Error	z value	Pr(> z)
(Intercept)	0.60	0.06	10.390	< 0.001
<i>Independent Factors</i>				
Gender (women)	0.35	0.09	3.774	< 0.001
Race/Eth. (Black)	-0.55	0.37	-1.489	0.136
Race/Eth. (Hispanic)	0.13	0.17	0.759	0.448
Race/Eth. (White)	-0.41	0.15	-2.745	0.006
<i>Intersectional Factors</i>				
Gender (W) & Race/Eth. (B)	0.45	0.68	0.668	0.504
Gender (W) & Race/Eth. (H)	-0.14	0.26	-0.532	0.594
Gender (W) & Race/Eth. (W)	0.25	0.25	1.010	0.312

Table 7: Beta regression output for anonymity rate for questions

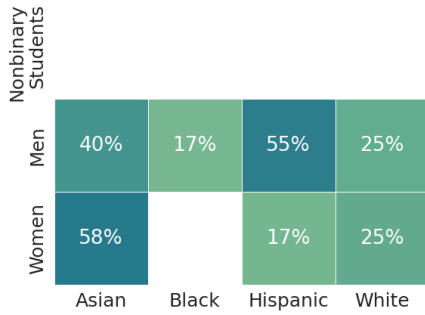


Figure 3: Intersectional anonymity rate for answers

	β	Error	z value	Pr(> z)
(Intercept)	-0.21	0.10	-2.117	0.034
<i>Independent Factors</i>				
Gender (women)	0.40	0.20	1.946	0.052
Race/Eth. (Hispanic)	0.36	0.38	0.953	0.341
Race/Eth. (White)	-0.35	0.25	-1.415	0.158
<i>Intersectional Factors</i>				
Gender (W) & Race/Eth. (H)	-1.22	0.54	-2.232	0.026
Gender (W) & Race/Eth. (W)	-0.39	0.49	-0.803	0.422

Table 8: Beta regression output for anonymity rate for answers

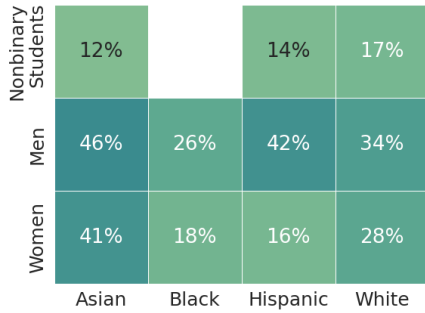


Figure 4: Intersectional anonymity rate for follow ups

	β	Error	z value	Pr(> z)
(Intercept)	-0.11	0.06	-1.748	0.080
<i>Independent Factors</i>				
Gender (women)	-0.12	0.10	-1.210	0.226
Gender (nonbinary students)	-0.71	0.33	-2.182	0.029
Race/Eth. (Black)	-0.45	0.30	-1.488	0.137
Race/Eth. (Hispanic)	-0.09	0.17	-0.556	0.578
Race/Eth. (White)	-0.28	0.15	-1.850	0.064
<i>Intersectional Factors</i>				
Gender (W) & Race/Eth. (B)	-0.04	0.51	-0.071	0.944
Gender (W) & Race/Eth. (H)	-0.45	0.24	-1.891	0.059
Gender (N) & Race/Eth. (H)	0.10	0.60	0.169	0.866
Gender (W) & Race/Eth. (W)	-0.01	0.23	-0.051	0.959
Gender (N) & Race/Eth. (W)	0.37	0.68	0.548	0.584

Table 9: Beta regression output for anonymity rate for follow ups

6.2.1 Questions Our analysis of questions revealed significant effects for gender (women) and race/ethnicity (White), where women are more likely to use anonymity and White students are less

likely to use anonymity (Table 7). None of the intersectional factors were statistically significant. Although Hispanic nonbinary students were *much* more likely to post anonymously, this effect did not quite achieve statistical significance.

Content	Total	Women	Men	Nonbinary Students
Question	75.6%	86.3%	69.8%	75.8%
Answer	39.6%	46.7%	38.2%	—
Followup	38.9%	34.0%	43.3%	13.1%
Note	40.3%	56.5%	28.9%	—
All Posts	52.8%	52.8%	53.7%	28.5%

(a) Anonymity rates by post type and gender

Content	Total	Asian	Black	Hispanic	White
Question	75.6%	76.8%	58.9%	82.4%	65.4%
Answer	39.6%	44.0%	—	33.8%	24.1%
Followup	38.9%	43.4%	22.9%	28.0%	31.3%
Note	40.3%	42.0%	—	—	—
All Posts	52.8%	57.0%	31.9%	45.7%	41.3%

(b) Anonymity rates by post type and race/ethnicity

Table 10: Differences in anonymity rate by post type. Analysis of (a) gender and (b) race/ethnicity independently.

6.2.2 Answers Our analysis of answers found that none of the independent factors were statistically significant. Though the p-value is above 0.05, indicating a lack of a significant difference, we saw a relatively large coefficient for the gender term. Once again, however, we found an intersectional effect for race/ethnicity and gender. As in the case of overall posts, Hispanic women are much less likely to post answers anonymously (Table 8) than would be expected from analyzing race/ethnicity or gender alone.

6.2.3 Follow ups Our analysis of follow ups found only one statistically significant effect: nonbinary students are less likely to use anonymity than other students in follow up posts. No other differences are statistically significant. We did, however, observe relatively larger correlations with Black students and Hispanic women students, though not statistically significant.

6.3 RQ2(b) Use of Anonymity by Course

Our data was collected from eight courses in the core requirements of the Computer Science major. Our next analysis focused on the impact of course. Table 11 shows the anonymity rate for all students in each of the courses. We note a striking result: the anonymity rate was *much* lower in the first introductory course (Introductory CS A)

Platform	Course	Anonymity
Ed Discussion	Introductory CS A	18.9%
	Systems Programming	49.2%
	Advanced Data Structures	63.3%
Piazza	Introductory CS B	73.1%
	Introductory CS C	66.1%
	Introductory Lab	66.7%
	Basic Data Structures	69.2%
	Software Engineering	73.2%

Table 11: Anonymity rate by platform and course.

relative to all other courses. In this first course, the anonymity rate was only 18.9%, compared with other courses that range from 49.2%–73.2%. This effect was statistically significant after Holm-Bonferroni post-hoc corrections, with ($z=-21.12$, $p<0.001$, $d=-1.19$). Table 11 is categorized first by the platform used, but within that roughly according to when students take it within their studies. As Table 11 shows, anonymity rates are not only higher in advanced courses, they are higher in the other introductory courses (Introductory CS B, Introductory CS C, and Introductory Lab) as well. There appears to be something quite unique about the experience of Introductory CS A. Drawing from the syllabus and student makeup of this course, we discuss possible reasons for this in the Discussion. However, this unusual course affects the overall analysis as 27.0% of the students in our dataset are enrolled in Introductory CS A.

Given the strong effect of this one course, Introductory CS A, on anonymity use, we examine how race/ethnicity and gender identity affect anonymity for students in this class and outside of it. Figure 6 and Table 12 show the relationship between gender, race/ethnicity, and anonymity rate for students in Introductory CS A. Across almost all categories, the anonymity rate is similarly low; no term in the beta regression is statistically significant (the higher rate of anonymity for Hispanic nonbinary students is not statistically significant because of a small sample size).

By contrast, Figure 7 and Table 13 show the relationship between gender, race/ethnicity, and anonymity rate for students in all other classes. Here there are clear, statistically significant effects. Women use anonymity more often than men. And White students use anonymity less often than Asian students. While the results are not statistically significant, Figure 7 also shows that Black men have anonymity rates below 50%, whereas most other categories of students have substantially higher anonymity rates. We note that several of the trends from the previous research questions no longer appear when Introductory CS A is removed. Most notably, the lower anonymity rate for Hispanic women noted in RQ 1 is no longer statistically significant. Instead, we suggest the finding from RQ1 is a result of a large portion of the Hispanic women in our dataset enrolling in Introductory CS A and being less likely to use anonymity, as a function of that course’s design. Further, without students from Introductory CS A, there are too few non-binary students to report results. We discuss both of these in the Discussion.

6.4 RQ2(c) Use of Anonymity By Platform

Finally, we examine whether there are any differences by platform. As shown in Table 11, all three Ed Discussion courses have lower anonymity rates (18.9%–63.3%) than the Piazza courses (66.1%–73.2%). This effect is statistically significant with ($z=16.78$, $p<0.001$, $d=0.84$). Unfortunately, because each class uses a single platform, we cannot statistically distinguish between effects of platform and of course.

Nevertheless, two researchers investigated the platform affordances, to see whether this difference could be explained by platform design choices. We do identify one important difference. On

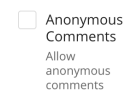


Figure 5: Ed Discussion anonymity checkbox

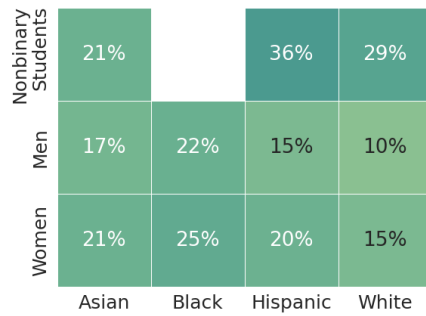


Figure 6: Intersectional anonymity rate for Introductory CS A

	β	Error	z value	Pr(> z)
(Intercept)	-1.26	0.13	-9.948	<0.001
<i>Independent Factors</i>				
Gender (women)	0.08	0.15	0.533	0.594
Gender (nonbinary students)	0.06	0.32	0.174	0.862
Race/Eth. (Hispanic)	-0.10	0.26	-0.390	0.696
Race/Eth. (White)	-0.22	0.28	-0.817	0.414
<i>Intersectional Factors</i>				
Gender (W) & Race/Eth. (H)	0.07	0.31	0.233	0.816
Gender (N) & Race/Eth. (H)	0.60	0.58	1.034	0.301
Gender (W) & Race/Eth. (W)	0.05	0.34	0.150	0.880
Gender (N) & Race/Eth. (W)	0.49	0.63	0.771	0.441

Table 12: Beta regression output for anonymity rate for Introductory CS A

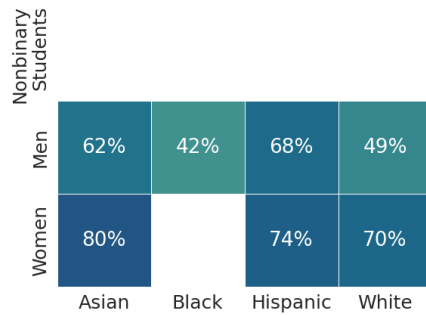


Figure 7: Intersectional anonymity rate for all courses except Introductory CS A

	β	Error	z value	Pr(> z)
(Intercept)	0.32	0.05	5.790	2.37e-09
<i>Independent Factors</i>				
Gender (women)	0.44	0.10	4.585	< 0.001
Race/Eth. (Black)	-0.50	0.36	-1.385	0.166
Race/Eth. (Hispanic)	0.15	0.16	0.887	0.375
Race/Eth. (White)	-0.35	0.13	-2.638	0.008
<i>Intersectional Factors</i>				
Gender (W) & Race/Eth. (H)	-0.24	0.30	-0.823	0.410
Gender (W) & Race/Eth. (W)	0.12	0.26	0.459	0.646

Table 13: Beta regression output for anonymity rate for all courses except Introductory CS A

Piazza, the decisions whether to let students post anonymously is made on a per-course basis. On Ed Discussion, in contrast, there are some post types where the original poster needs to indicate *for each post* whether anonymous comments are allowed (Figure 5). However, this is only true for notes, which is the least common post type, comprising less than 1% of our dataset. For this reason, we suspect that the platform differences are primarily based on courses, but future work could collect data for similar or identical courses using different platforms to explore this question more.

Key Findings: While there were differences in anonymity use based on post type, the dominant effect was based on course. One large, introductory course, Introductory CS A, had a much lower anonymity rate than all other courses. And the demographic makeup of this course explains the intersectional result noted in RQ1.

6.5 RQ3 Lurking

Understanding anonymity use is important, but so is understanding students who do not participate in discussions at all. These students, often referred to as *lurkers* in social computing contexts, only read without posting themselves. Table 14 shows the number of students who used discussion boards but never post or only send private messages to instructors. More than half of students never post publicly (56.9%). On Piazza 34.9% never post at all, and

an additional 22.0% only send direct messages to instructors. This pattern of behaviour is similar for women (31.3%, 24.9%, 43.8%) and for men (35.8%, 20.9%, 43.3%), where these values represent the frequency of people who do not post, post privately, and post publicly. The Piazza dataset only contains 15 nonbinary students, but the pattern does appear to be similar. A chi square test of independence for posting behaviour does not find a significant difference between women and men ($\chi^2 = 6.24$, $p = 0.18$). Lurking behavior is also similar across race/ethnicity. A chi square test of independence fails to reject the null hypothesis that race/ethnicity and posting behaviour are independent ($\chi^2 = 6.62$, $p = 0.36$). One exception is Black students, (50.0%, 15.4%, 34.6%), who are least likely to post, but also the smallest group of students, so the effect is not statistically significant. And finally, we find no statistical significance in our intersectionality analysis. Conditioning on each category of race/ethnicity, chi square tests of gender and posting behaviour do not reject the null hypothesis of independence ($p = 0.74, 0.21, 0.42, 0.44$), and conditioning on each category of gender, race/ethnicity and posting behaviour do not reject the null hypothesis of independence ($p = 0.52, 0.49, 0.41$). Overall, we find no statistically significant differences from analyzing the choice to post or not, as it relates to race/ethnicity and gender identity. That is, while we see differences in *how* students participate (anonymously

Students who never post

Gender	Black		Asian		Hispanic		White		Total	
Women	< 5		135	31.1%	18	29.5%	16	30.8%	179	31.3%
Men	10	47.6%	347	35.2%	55	39.0%	71	37.2%	499	35.8%
Nonbinary Students	0		< 5		< 5		< 5		6	
Total	13	50.0%	483	33.9%	73	36.4%	88	36.2%	732	34.9%

Students who only send private messages

Gender	Black		Asian		Hispanic		White		Total	
Women	< 5		109	25.1%	20	32.8%	10	19.2%	142	24.9%
Men	< 5		199	20.2%	33	23.4%	45	23.6%	291	20.9%
Nonbinary Students	0		< 5		0		0		< 5	
Total	< 5		310	21.9%	53	25.7%	56	22.8%	463	22.0%

Table 14: Intersectional lurking behavior

or not), we do not see differences in choosing whether to participate at all. This validates a core motivation to provide anonymity.

Key Findings: We identify no significant differences in lurking behavior by gender or race/ethnicity.

6.6 RQ4 Impact of Gender Data Source

Table 15c shows the confusion matrices for the three data sources for gender. First, the gender imputation process fundamentally prevents predicting nonbinary gender, so 100% of nonbinary students have their gender predicted incorrectly. Beyond this, the name-based prediction tools make large numbers of errors, compared to both the registrar and survey data. The most common errors were being unable to identify the gender of the student, for approximately 500 students. The model had a 25.9% error rate (8.7% misclassified, 17.2% unable to identify) when comparing with the survey results and a 23.9% error rate (6.9% misclassified, 17.0% unable to identify) when compared to the registrar results. Among students for whom we had both registrar and survey data, the registrar data closely corresponded to survey data, with an error rate of just 2.7%. This clearly demonstrates that for our data, name-based predictions would have introduced large numbers of errors into gender analysis relative to registrar and survey data (which remain consistent with each other).

Again, an intersectional lens reveals even more information. In Table 16, we highlight how the error rate from name-based predictions differed by race/ethnicity. For these percentages, we combine our registrar and survey data, since these sources have high agreement with each other. Relative to students of most other races/ethnicities, Asian students were more than twice as likely to be misclassified (31.6%), echoing problems identified in prior work [64]. Black students had the second highest rate of error (23.5%), also greater than the rates for White or Hispanic students.

Finally, we note one striking difference between the survey and registrar data: students were more than ten times as likely to indicate pronouns other than exactly one of ‘she/her’ or ‘he/him’ on the survey (> 3%) as to check the nonbinary category in the registrar data (0.3%). Survey data, asking about pronouns, seems to be capturing information that is lost in registrar data. We consider survey data more accurate than registrar data because 1) it is more recent (in the same term vs. at the time of application to the university) and 2) students may have concerns about sharing information with the administration (c.f., [21]) that they do not feel about individual instructors or classes.

Key Findings: Name-based gender prediction introduces a large number of inaccuracies, which are significantly worse for some identities: Black students, Asian students, and/or nonbinary students.

		Survey			
		M	W	NB	U
Prediction	M	358	25	9	<5
	W	22	301	22	<5
	NB	-	-	-	-
	U	87	63	<5	<5

(a) Survey vs. Name-based Prediction
Error Rate = 25.9%

		Registrar			
		M	W	NB	U
Prediction	M	1090	40	<5	36
	W	95	417	<5	6
	NB	-	-	-	-
	U	245	102	0	<5

(b) Registrar vs. Name-based Prediction
Error Rate = 23.9%

		Survey			
		M	W	NB	U
Registrar	M	253	0	<5	2
	W	0	100	<5	2
	NB	0	0	0	0
	U	5	<5	0	0

(c) Survey vs. Registrar data
Error Rate = 2.7%

Table 15: Impact of Gender Data Source. Confusion matrices show matching and unmatching data from different sources. In general diagonal entries are correct (e.g., prediction correctly labels gender) while off-diagonal entries are incorrect. Results are presented for men (M), women (W), nonbinary students (NB), and unknown gender (U).

Race/ Ethnicity	Total Error	Misclassified	Failure to Classify
Black	23.5%	2.0 %	21.6%
Asian	31.6%	7.8%	23.8%
Hispanic	8.4%	4.5%	3.9%
White	11.0%	7.1%	3.9%
Unknown	9.3%	4.2%	5.1%
Total	24.7%	7.0%	17.8%

Table 16: Error rate by race/ethnicity

7 Discussion

7.1 Successful Interventions to Increase Inclusion

Prior work has found an important effect of gender on anonymity use in course discussion boards. We find parts of these prior results that we confirm, but also important differences. For example, we find that women use anonymity significantly more often, but only after excluding one of our largest introductory courses, Introductory CS A. This aligns with findings from prior work that found that women tend to have a higher anonymity usage rate than men [29, 59, 75]. But at the same time, this reveals the significant impact of classroom norms on decisions around use of anonymity.

There are several potential explanations for this improvement in women’s experiences in Introductory CS A and willingness to post by name. One possible reason might be that this course contains the most non-majors. Relatedly, this course is only one in the curriculum that expects *no* prior knowledge. Given this, students might have fewer concerns about self-presentation risks when asking questions. However, while the proportion is greatest in this course, all of the lower division courses in our department are open to and frequently taken by non-majors. But the other courses in our data set don’t show this effect.

Instead, through an investigation of the syllabus and course policies, we suggest that this effect is primarily due to active choices in the course design to increase inclusion and sense of belonging in this course. The instructor for this course has adopted a number of best practices suggested in the literature [19, 42, 52, 56]. For example, the first five minutes of each class are used as an *inclusion minute*, which cover topics like emphasizing the importance of women in the history of computing [62] and encouraging students to reflect on their potential for growth rather than where their knowledge starts [47]. In addition, through a number of practices, including live programming *with errors*, the instructor normalizes the struggle that students may experience when they learn programming for the first time. Though we do not directly evaluate the effectiveness of these interventions, we are encouraged by the dramatic difference we see in these courses. We believe our work can motivate future experimental studies to investigate the impact of said interventions on large-scale classroom inclusion.

7.2 Intersectional Impacts on Anonymity Use

Researchers within the computer science education community have recognized the importance of intersectional approaches for a variety of topics [27, 32, 50, 53, 63, 69], however, this work is the

first to take this approach looking at anonymity and lurking behaviors on course discussion forums. In extending beyond prior work on gender, we find a strong impact of race/ethnicity on anonymity use. While prior work suggested that minority groups in computing would be most likely to use anonymity, we actually find it is the largest group of students — Asian students — who are most likely to use anonymity. This introduces an interesting point of reflection into work on anonymity use, where a group is both a minority in the United States while a majority within Computer Science. This is a case that may also highlight the need to integrate additional identities; including social class, for example, might capture important differences within this community to explain this effect.

We also find intersectional effects. For example, there are very strong intersectional effects for Hispanic students, a finding that would not be clear looking at the two factors, gender and race/ethnicity, independently. Through our further investigations of context, including courses, we argue this finding can be attributed to the fact that there are many more Hispanic women in Introductory CS A than there are in the other courses in our dataset. In fact, there are many more Hispanic men outside of Introductory CS A. This is a prime illustration of Simpson’s paradox [67], and a motivation to be as comprehensive in recording identity and contextual information as possible, as argued in prior work in the computing education community [46].

Indeed, the vast majority of Hispanic women, Black women, and nonbinary students in our data set are enrolled in Introductory CS A. This raises a potential concern that these students find an unwelcoming environment and leave the computer science curriculum early. While our results suggest success in building a supportive environment in Introductory CS A, they may indicate the need to expand similar practices into the upper division.

7.3 Lurking Compared with Posting Anonymously

When examining lurking behaviour, we found no statistically significant differences based on gender or race/ethnicity. This stands in contrast to the differences we do find with respect to anonymity rate. That is, while we see differences in *how* students participate (anonymously or not), we do not see differences in choosing whether to participate at all. These results are consistent with many of the reasons given for allowing anonymity on discussion boards: disparities are confined to the way students participate, but no group of students is less likely to participate at all.

This data also highlights a possible drawback in the decision to allow anonymity on discussion boards. Despite the diversity in publicly posting students, this diversity is invisible because of the anonymity option. As suggested by prior work, this might reinforce the impression that women are using the platform less than they actually are [29]. Keeping all this in mind, it is clear that anonymity is used quite often given the option. Our work echoes findings from prior work that students’ propensity to engage with in class questions increases with the degree of anonymity provided [24]. Particularly with the rise of remote or hybrid options (particularly since the onset of the COVID-19 pandemic), thinking about these anonymous options is important. For example, platforms like Zoom and Microsoft Teams might integrate anonymous posting in their

chats.⁷ At the same time, to address risks of artificial invisibility, platforms might test partial revelation of identity. For example, chats might provide summary statistics about the race/ethnicity or gender of students who contribute or optional identifiers (like Memoji⁸) for posts. However, any design explorations in this space must also address the harsh reality of women being taken less seriously than men in many computing tasks [74].

7.4 Unreliability of Gender Imputation:

We conclude with a critique of gender imputation in computing education research. In a recent effort focused on roles for computing in social change, researchers have argued that “*technical experts have a unique contribution to make in illustrating the inherent limits of those approaches*” [1]. We make an argument for such a rebuttal here. We find that using the standard automated gender prediction pipeline misclassifies or removes 23.9% of students from the study, when compared with self-reported data. These tools have profound limitations for reliably identifying students’ gender identity. This is particularly true when investigated with an intersectional lens. While the models perform better for White and Hispanic students, we note alarming error rates for Asian students (30.1%), Black students (23.5%), and nonbinary students (100%). While it may be possible to improve accuracy for some students with better training data, the inaccuracy for nonbinary students seems to be a fundamental limitation of the tool’s premise.

Although this is not the focus of our work, some prior work makes use of gender imputation when investigating external perceptions of gender based on username, rather than an individual’s gender identity [10]. In this case, it is important to keep in mind that these instruments are, to our knowledge, not validated as measuring perceptions of gender. Questions around the validity of this instrument — such as whose perceptions these models best encode — are crucial to answer if using this tool as a proxy for human perception.

We ask researchers to consider the implications of using a tool which misclassifies, or fails to classify, those whose gender identity does not match expectations associated with their name. Researchers should consider how their methodology could contribute to erasure of groups whose identity differs from what this tool detects [48]. If, instead of the registrar and survey data we used in our study, we had used this tool, such a choice could have contributed to missed findings about Asian, Black, and especially nonbinary students. We further note that some of the inherent biases in this tool, most evident for nonbinary students who cannot exist in this labelling system, may also affect other students, including transgender men and women. Transgender students may be more likely to have names in administrative documents that do not correspond to their gender identity [20]. More fundamentally, there are limitations to any automated gender prediction as pertains to trans students. As noted by researcher Os Keyes, “*a trans-inclusive system for non-consensually defining someone’s gender is a contradiction in terms*” [30]. Researchers should be cautious and critical of tools for automatic gender detection.

⁷Currently this is only offered in Zoom webinars.

⁸<https://support.apple.com/en-us/HT208986>

8 Limitations

Instructors shared data for only 8 out of 14 requested courses. Consequently, there may be volunteer bias in the data (e.g., more inclusive practices in these courses). And as we note in Section 4.1, our methods for collecting information on gender, race, and ethnicity have limitations. The registrar data to which we had access allows students to respond with multiple races or ethnicities, but unfortunately collapses those responses to a single value. This means that our analysis has clear limitations for multiracial student experiences [23]. In addition, the survey only asked students for their pronouns. Gender identity and pronouns do not necessarily map in a single way for every student.

Finally, our institutional context may limit generalization: the large scale, competitive admissions, and research focus likely impact students’ experiences. For example, prior work has shown that the communication patterns established between the course staff and students as well as informal hierarchies created among students can impact the classroom environment [5]. Our institutions’ large class sizes and course staff (instructors, TAs, tutors, etc.) may play a role in online environments as well. Finally, our data was collected at a time where courses were continuing to adjust to the COVID-19 pandemic. Among other changes, this has led to an increased use of social media (e.g., Discord), which may impact students’ conversation dynamics on course discussion forums.

9 Future Work

We are excited by the results that we have found — including evidence that efforts for inclusion can have real and meaningful impacts on students’ willingness to engage publicly and by name on course discussion forums. However, there are a number of future directions to pursue. First, future work should investigate *why* students make their decisions to post anonymously or publicly. Prior work has suggested that the trend of women preferring to post anonymously could be related to lower confidence, but researchers have not tested this directly. Either by collecting quantitative confidence scores through surveys, or by supplementing this largely quantitative research area with qualitative interview data, we might build a better understanding of the factors in students’ decisions. The functional model of self disclosure suggests there are likely both risks and benefits that may be specific to this context and could be learned through qualitative work. In addition, there may be further contextual information that could be analyzed using the posts themselves. Further, research might investigate how and when students intentionally select discussion forum handles that do not reveal gender. Finally, we believe this work (along with others on anonymity [24, 45]) suggests productive directions for increasing engagement *within* classrooms, for example, by building and studying tools that support anonymity within live Q&A features used during class.

References

- [1] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G Robinson. 2020. Roles for computing in social change. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 252–260.
- [2] Tawfiq Ammari, Sarita Schoenebeck, and Daniel Romero. 2019. Self-declared throwaway accounts on Reddit: How platform affordances and shared norms enable parenting disclosure and support. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–30.

- [3] Social Justice Analysis. 2022. *Using intersectionality to understand structural inequality in Scotland: evidence synthesis*. Technical Report. Scottish Government.
- [4] Nazanin Andalibi, Oliver L Haimson, Munmun De Choudhury, and Andrea Forte. 2016. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 3906–3918.
- [5] Lecia Jane Barker, Kathy Garvin-Doxas, and Michele Jackson. 2002. Defensive Climate in the Computer Science Classroom. In *Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education (SIGCSE '02)*. Association for Computing Machinery, New York, NY, USA, 43–47. <https://doi.org/10.1145/563340.563354>
- [6] Natalya N Bazarova and Yoon Hyung Choi. 2014. Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of Communication* 64, 4 (2014), 635–657.
- [7] Michael Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Greg Vargas. 2011. 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community. In *Proceedings of the international AAAI conference on web and social media*, Vol. 5. 50–57.
- [8] Katerine Bielaczyc and Allan Collins. 2009. Learning communities in classrooms: A reconceptualization of educational practice. *Instructional design theories and models 2* (2009), 269–291.
- [9] Sian Brooke. 2019. “Condescending, Rude, Assholes”: Framing gender and hostility on Stack Overflow. In *Proceedings of the Third Workshop on Abusive Language Online*. 172–180.
- [10] SJ Brooke. 2021. Trouble in programmer’s paradise: gender-biases in sharing and recognising technical knowledge on Stack Overflow. *Information, Communication & Society* 24, 14 (2021), 2091–2112.
- [11] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2017. Anyone can become a troll: Causes of trolling behavior in online discussions. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 1217–1230.
- [12] Kate Crawford. 2011. Listening, not lurking: The neglected form of participation. *Cultures of participation* (2011), 63–74.
- [13] Kimberlé Crenshaw. 1989. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *u. Chi. Legal f.* (1989), 139.
- [14] Francisco Cribari-Neto and Achim Zeileis. 2010. Beta Regression in R. *Journal of Statistical Software* 34, 2 (2010), 1–24. <https://doi.org/10.18637/jss.v034.i02>
- [15] K Patricia Cross. 1998. Why learning communities? Why now? *About campus* 3, 3 (1998), 4–11.
- [16] Jenny L Davis and Nathan Jurgenson. 2014. Context collapse: Theorizing context collisions and collisions. *Information, communication & society* 17, 4 (2014), 476–485.
- [17] Munmun De Choudhury and Sushovan De. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Proceedings of the international AAAI conference on web and social media*, Vol. 8. 71–80.
- [18] Michael A DeVito, Ashley Marie Walker, and Jeremy Birnholtz. 2018. “Too Gay for Facebook” Presenting LGBTQ+ Identity Throughout the Personal Social Media Ecosystem. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–23.
- [19] Inclusion Diversity and Access (DI&A) Initiative. 2023. *Inclusion Icebreakers*. Technical Report. Colorado School of Mines.
- [20] Corinne Dorsey. 2022. Transgender students fear being ‘deadnamed’ at college. *Washington Post* (2022).
- [21] Brendan Farrington. 2023. DeSantis seeks details on transgender university students. *AP News* (2023).
- [22] 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.
- [23] Karly S Ford, Ashley N Patterson, and Marc P Johnston-Guerrero. 2021. Monoracial normativity in university websites: Systematic erasure and selective reclassification of multiracial students. *Journal of Diversity in Higher Education* 14, 2 (2021), 252.
- [24] Mark Freeman, Paul Blayney, and Paul Ginns. 2006. Anonymity and in class learning: The case for electronic response systems. *Australasian Journal of Educational Technology* 22 (Nov. 2006), 568–580. <https://doi.org/10.14742/ajet.1286>
- [25] Erving Goffman et al. 1959. *The Presentation of Self in Everyday Life*. Doubleday.
- [26] Cheng Guo and Kelly Caine. 2021. Anonymity, user engagement, quality, and trolling on Q&A sites. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.
- [27] Apriel K Hodari, Maria Ong, Lily T Ko, and Rachel R Kachchaf. 2014. New enactments of mentoring and activism: US women of color in computing education and careers. In *Proceedings of the tenth annual conference on International computing education research*. 83–90.
- [28] Xiaoyun Huang and Jessica Vitak. 2022. “ Finsta gets all my bad pictures”: Instagram Users’ Self-Presentation Across Finsta and Rinsta Accounts. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–25.
- [29] David A. Joyner, Lily Bernstein, Ian Bolger, Maria-Isabelle Dittamo, Stephanie Gorham, and Rachel Hudson. 2022. Anonymity: A Double-Edged Sword for Gender Equity in a CS1 Forum?. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2022)*. Association for Computing Machinery, New York, NY, USA, 766–772. <https://doi.org/10.1145/3478431.3499289>
- [30] Os Keyes. 2018. The misgendering machines: Trans/HCI implications of automatic gender recognition. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–22.
- [31] Sophia Krause-Levy, William G. Griswold, Leo Porter, and Christine Alvarado. 2021. The Relationship Between Sense of Belonging and Student Outcomes in CS1 and Beyond. In *Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021)*. Association for Computing Machinery, New York, NY, USA, 29–41. <https://doi.org/10.1145/3446871.3469748>
- [32] Sophia Krause-Levy, William G Griswold, Leo Porter, and Christine Alvarado. 2021. The relationship between sense of belonging and student outcomes in CS1 and beyond. In *Proceedings of the 17th ACM Conference on International Computing Education Research*. 29–41.
- [33] Alyson Latham and N Sharon Hill. 2014. Preference for anonymous classroom participation: Linking student characteristics and reactions to electronic response systems. *Journal of Management Education* 38, 2 (2014), 192–215.
- [34] Alex Leavitt. 2015. “ This is a Throwaway Account” Temporary Technical Identities and Perceptions of Anonymity in a Massive Online Community. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 317–327.
- [35] Oscar T Lenning and Larry H Ebberts. 1999. *The Powerful Potential of Learning Communities: Improving Education for the Future*. ASHE-ERIC Higher Education Report, Vol. 26, No. 6. ERIC.
- [36] Eden Litt. 2012. Knock, knock. Who’s there? The imagined audience. *Journal of broadcasting & electronic media* 56, 3 (2012), 330–345.
- [37] Stephanie Lunn, Leila Zahedi, Monique Ross, and Matthew Ohland. 2021. Exploration of Intersectionality and Computer Science Demographics: Understanding the Historical Context of Shifts in Participation. *ACM Trans. Comput. Educ.* 21, 2, Article 10 (mar 2021), 30 pages. <https://doi.org/10.1145/3445985>
- [38] Jane Margolis. 2017. *Stuck in the Shallow End: Education, Race, and Computing*. MIT press.
- [39] Alice E Marwick and Danah Boyd. 2011. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society* 13, 1 (2011), 114–133.
- [40] Anna May, Johannes Wachs, and Anikó Hannák. 2019. Gender differences in participation and reward on Stack Overflow. *Empirical Software Engineering* 24 (2019), 1997–2019.
- [41] Cath Mazuro and Namrata Rao. 2011. Online discussion forums in higher education: Is lurking working. *International Journal for Cross-Disciplinary Subjects in Education* 2, 2 (2011), 364–371.
- [42] Mark Miller. Unknown. *Framing the ‘Inclusive Minute’*. Technical Report. Lawrence Livermore National Laboratory.
- [43] Terumi Miyazoe and Terry Anderson. 2011. Anonymity in blended learning: Who would you like to be? *Journal of Educational Technology & Society* 14, 2 (2011), 175–187.
- [44] Michael Muller. 2012. Lurking as personal trait or situational disposition: lurking and contributing in enterprise social media. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 253–256.
- [45] Asako Ohno, Takahiro Yamasaki, and Kin-ichiroh Tokiwa. 2013. A discussion on introducing half-anonymity and gamification to improve students’ motivation and engagement in classroom lectures. In *2013 IEEE Region 10 Humanitarian Technology Conference*. IEEE, 215–220.
- [46] Alannah Oleson, Benjamin Xie, Jean Salac, Jayne Everson, F Megumi Kivuva, and Amy J Ko. 2022. A Decade of Demographics in Computing Education Research: A Critical Review of Trends in Collection, Reporting, and Use. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1*. 323–343.
- [47] John Ousterhout. “A little bit of slope makes up for a lot of y-intercept”. <https://gist.github.com/gtallen1187/e83ed02eac6cc8d7e185>.
- [48] Dylan Paré. 2021. Queering computing and computing education. (2021).
- [49] Piazza. New Piazza Features and Product Updates. <https://medium.com/piazza/new-piazza-features-and-product-updates-2fe3c05a6854>
- [50] 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 (ITICSE ’20)*. Association for Computing Machinery, New York, NY, USA, 219–224. <https://doi.org/10.1145/3341525.3387360>
- [51] Jennifer Preece and Ben Shneiderman. 2009. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS transactions on human-computer interaction* 1, 1 (2009), 13–32.
- [52] Adalbert Gerald Soosai Raj, Jignesh M Patel, Richard Halverson, and Erica Rosenfeld Halverson. 2018. Role of live-coding in learning introductory programming.

In *Proceedings of the 18th kali calling international conference on computing education research*. 1–8.

- [53] Yolanda A. Rankin, Jakita O. Thomas, and Sheena Erete. 2021. Black Women Speak: Examining Power, Privilege, and Identity in CS Education. *ACM Trans. Comput. Educ.* 21, 4, Article 26 (oct 2021), 31 pages. <https://doi.org/10.1145/3451344>
- [54] Pei-Luen Patrick Rau, Qin Gao, and Yinan Ding. 2008. Relationship between the level of intimacy and lurking in online social network services. *Computers in human behavior* 24, 6 (2008), 2757–2770.
- [55] Lynne D. Roberts and Camilla J Rajah-Kanagasabai. 2013. "I'd be so much more comfortable posting anonymously": Identified versus anonymous participation in student discussion boards. *Australasian Journal of Educational Technology* 29, 5 (Nov. 2013), 612–625. <https://doi.org/10.14742/ajet.452>
- [56] Marc J Rubin. 2013. The effectiveness of live-coding to teach introductory programming. In *Proceeding of the 44th ACM technical symposium on Computer science education*. 651–656.
- [57] Zick Rubin. 1975. Disclosing oneself to a stranger: Reciprocity and its limits. *Journal of Experimental Social Psychology* 11, 3 (1975), 233–260.
- [58] Isabel Ruthotto, Quintin Kreth, Jillian Stevens, Clare Trively, and Julia Melkers. 2020. Lurking and participation in the virtual classroom: The effects of gender, race, and age among graduate students in computer science. *Computers & Education* 151 (2020), 103854.
- [59] Pooja Sankar, Jessica Gilmartin, and Melissa Sobel. 2015. An examination of belongingness and confidence among female computer science students. *ACM SIGCAS Computers and Society* 45, 2 (July 2015), 7–10. <https://doi.org/10.1145/2809957.2809960>
- [60] Marlene Scardamalia and Carl Bereiter. 2006. Knowledge building. *The Cambridge* (2006).
- [61] Randall E. Schumacker. 2023. *Learning Statistics Using R*. SAGE Publications, Inc., 55 City Road, London, Chapter 12. <https://doi.org/10.4135/9781506300160>
- [62] Oscar Schwartz. Untold History of AI: Invisible Women Programmed America's First Electronic Computer.
- [63] Kristin A Searle and Yasmin B Kafai. 2015. Boys' Needlework: Understanding Gendered and Indigenous Perspectives on Computing and Crafting with Electronic Textiles.. In *ICER*. 31–39.
- [64] Paul Sebo. 2021. How accurate are gender detection tools in predicting the gender for Chinese names? A study with 20,000 given names in pinyin format. *Journal of the Medical Library Association* 110, 2 (2021). <https://doi.org/10.5195/jmla.2022.1289>
- [65] Joseph Seering, Jessica Hammer, Geoff Kaufman, and Diyi Yang. 2020. Proximate social factors in first-time contribution to online communities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [66] Bryan Semaan, Heather Faucett, Scott Robertson, Misa Maruyama, and Sara Douglas. 2015. Navigating imagined audiences: Motivations for participating in the online public sphere. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1158–1169.
- [67] Edward H. Simpson. 1951. The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 13, 2 (1951), 238–241.
- [68] Melissa Sobel, Jessica Gilmartin, and Pooja Sankar. 2016. Class size and confidence levels among female STEM students [Impact]. *IEEE Technology and Society Magazine* 35, 1 (2016), 23–26.
- [69] Amber Solomon, DeKita Moon, Armisha L. Roberts, and Juan E. Gilbert. 2018. Not Just Black and Not Just a Woman: Black Women Belonging in Computing. In *2018 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–5. <https://doi.org/10.1109/RESPECT.2018.8491700>
- [70] Erika M Sparby. 2017. Digital social media and aggression: Memetic rhetoric in 4chan's collective identity. *Computers and Composition* 45 (2017), 85–97.
- [71] Frederic Stutzman and Woodrow Hartzog. 2012. Boundary regulation in social media. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 769–778.
- [72] John Suler. 2004. The online disinhibition effect. *Cyberpsychology & behavior* 7, 3 (2004), 321–326.
- [73] Lee Taber and Steve Whittaker. 2020. "On Finsta, I can say 'Hail Satan'": Being Authentic but Disagreeable on Instagram. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–14.
- [74] Josh Terrell, Andrew Kofink, Justin Middleton, Clarissa Rainear, Emerson Murphy-Hill, Chris Parnin, and Jon Stallings. 2017. Gender differences and bias in open source: Pull request acceptance of women versus men. *PeerJ Computer Science* 3 (2017), e111.
- [75] 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*. ACM, Virtual Event USA, 438–444. <https://doi.org/10.1145/3408877.3432395>
- [76] Emily Van der Nagel and Jordan Frith. 2015. Anonymity, pseudonymity, and the agency of online identity: Examining the social practices of r/Gonewild. *First Monday* (2015).
- [77] Ellen Vanderhoven, Annelies Raes, Hannelore Montrieux, Tijs Rotsaert, and Tammy Schellens. 2015. What if pupils can assess their peers anonymously? A quasi-experimental study. *Computers & Education* 81 (2015), 123–132.
- [78] Bogdan Vasilescu, Andrea Capiluppi, and Alexander Serebrenik. 2012. Gender, representation and online participation: A quantitative study of stackoverflow. In *2012 International Conference on Social Informatics*. IEEE, 332–338.
- [79] Sijia Xiao, Danaë Metaxa, Joon Sung Park, Karrie Karahalios, and Niloufar Salehi. 2020. Random, messy, funny, raw: Finstas as intimate reconfigurations of social media. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
- [80] Nira Yuval-Davis. 2006. Intersectionality and Feminist Politics. *European journal of women's studies* 13, 3 (2006), 193–209.