# Paradise Unplugged: Identifying Barriers for Female Participation on Stack Overflow

Denae Ford, Justin Smith
NC State University
Raleigh, NC, USA
{dford3,jssmit11}@ncsu.edu

Philip J. Guo
UC San Diego
La Jolla, CA, USA
pg@ucsd.edu

Chris Parnin
NC State University
Raleigh, NC, USA
cjparnin@ncsu.edu

## ABSTRACT

It is no secret that females engage less in programming fields than males. However, in online communities, such as Stack Overflow, this gender gap is even more extreme: only 5.8% of contributors are female. In this paper, we use a mixed-methods approach to identify contribution barriers females face in online communities. Through 22 semi-structured interviews with a spectrum of female users ranging from non-contributors to a top 100 ranked user of all time, we identified 14 barriers preventing them from contributing to Stack Overflow. We then conducted a survey with 1470 female and male developers to confirm which barriers are gender related or general problems for everyone. Females ranked five barriers significantly higher than males. A few of these include doubts in the level of expertise needed to contribute, feeling overwhelmed when competing with a large number of users, and limited awareness of site features. Still, there were other barriers that equally impacted all Stack Overflow users or affected particular groups, such as industry programmers. Finally, we describe several implications that may encourage increased participation in the Stack Overflow community across genders and other demographics.

## Keywords

Online Communities; Barriers; Females in Computing; Social Q&A

## CCS Concepts

• **Human-centered computing** → *Empirical studies in collaborative and social computing;*

## 1. INTRODUCTION

"Stack Overflow[1] is a programmer's paradise", according to programming blogs and their commentary [1, 2]. Programmers can go to this popular question and answer (Q&A) site and get quick answers to their questions on a range of topics from how to implement a simple sorting algorithm to settling disputes about mobile application security. Programmers who join the Stack Overflow community cite several benefits, including the ability to learn more about programming while helping others [3]. However, not everyone participates equally on Stack Overflow. A recent survey conducted by the site found that only 5.8% of the site's active contributors identified as a female [3], which is in contrast with the 20% of Stack Overflow female visitors [4]. Stack Overflow also acknowledges their dissatisfaction with low participation rates, *"Our internal stats suggest the imbalance is not quite as severe as the survey results would make it seem, but there's no doubt everyone who codes needs to be more proactive welcoming women into the field"* [3].

Females are virtually absent from online programming communities, even though they comprise about 20% of the software engineering field [5]. For example, David et al. found that females make less than 5% of all open source contributions [6]. There have been several theories to explain these low participation rates. Often females do not feel welcomed in these online communities [7] and overall are unfamiliar with their culture and expectations [8, 9], hindering their usage. In addition, some females face the challenge of maintaining a family and a work-life balance [10]; carrying more of the burden than males [11]. Time constraints associated with these social pressures can limit females availability to contribute online. However, excluding females from participating can be harmful, as increasing diversity in teams correlates with increasing team productivity [12].

What might lead female programmers to not use Stack Overflow? Consider Asha, a programmer persona who uses Stack Overflow, but has never contributed before. She encounters an unfamiliar exception in her code and searches the internet to find a solution. She selects the first hit that shows up: Stack Overflow. Asha then briefly scans the page to compare the question that is posted to the problem she is having. The question appears to be similar to her own so she then continues to the answers section. She first reads the top rated answer and then tries the solution provided. Unfortunately, it does not quite work for her, but with some tweaking she finds out how to resolve her issue with a better solution than the original answer. She goes back to Stack Overflow and wants to post a new answer that would help others, but for some reason unknown to us, something stops her and she never gives her improved answer.

We want to understand what barriers are preventing females, like Asha, from contributing and actively using Stack

[1]http://stackoverflow.com



*FSE'16, November 13-19, 2016, Seattle, WA, USA*



DOI: http://dx.doi.org/10.1145/2950290.2950331

Overflow. However, barriers can apply to males as well. There is evidence that the barriers are affecting the long-term viability and health of the Stack Overflow platform. Newcomers face many barriers: 90% of accepted answers provided by new users are self-answers. Higher rates of unanswered question are occurring [13]. The community's growth is limited and controlled by a gamified core of elite users: only 5% of the users answer 60% of the questions [14]. However, even gamified mechanics fail as most users stop doing actions once they earn a badge for doing that action [15].

To discover what barriers Stack Overflow users encounter, we use a mixed-methods approach. We conducted semi-structured interviews with 22 female developers in order to understand what prevents them from actively participating on the site. Our focus was on females because we wanted to understand the obstacles they face and identify possible solutions. We interviewed female developers from a wide range of experience and levels of using Stack Overflow (from lurkers to a top 100 user). From these interviews, we identified 14 barriers based on common experiences of participants. To validate and understand how these barriers might differently affect both female and male users, we sent a survey to software developers, receiving responses from 134 females and 1336 males. From the survey, we identified which barriers females and males face and which ones are gender-specific. To our knowledge, this is the first study of this magnitude to compare female and male barriers encountered in an online programming community.

The main contributions are: (i) 14 barriers arranged in three groups that describe why females do not actively participate on Stack Overflow, (ii) general survey results that explain female and male differences and experience-based factors related to the barriers, and (iii) implications that might help create an environment on Stack Overflow that encourages female activity and contributions.

## 2. BACKGROUND & RELATED WORK

**Women Who Code.** The number of women in computing is low and much research has been proposed to get this number up [16, 17]. With this strong push for female programmers, there are many resources to encourage their retention in computing and have them feel more confident in their abilities. Cuny and Aspray have outlined a list of recommendations as to how keep women in computing including mentorship and broadening institutional culture [16]. Having these recruitment and retention techniques is one way to increase female visibility. Another is having resources females feel comfortable using as they deem necessary [18]. In an effort to relate these disparities, there is research to support the experiences of females in classroom [19, 20] and corporate settings [21]. However, these do not prescribe solutions for online community-based resources.

Persistent problems continue to exist. Margolis and Fisher found that a persistent perception of a geek mythology culture in technology discourages women more than men and promotes expectations of male success and continual questioning of women's abilities [22]. Women start at a disadvantage to their male peers: women who code are nearly twice as likely to have less than 2 years programming experience than their male counterparts [3]. Although their years of experience is low, this value can be interpreted as a small gain for more female interest in programming. Encouraging the few female software engineers we have is important to get them in position to be role models for someone else [16]. Following the example of the aformentioned studies, we propose designs for females to participate more in these communities and take advantage of their piece of "programmer's paradise" [1].

**Lurkers and Non-Adopters.** Researchers have studied low participation in online communities outside of the technology field [23, 24, 25]. For example, Lampe et al. studied non-adopters of Facebook and found a divide that emerged between light users and heavy users of the community [26]. A common category of light user, colloquially referred to as a lurker, is one who is able to witness the interactions of community but is overall not engaged with a community. Lurkers exist in many types of communities, such as discussion lists [27], file-sharing tools [28], and bulletin board systems [29, 30].

According to Nonnecke et al. [31] people in communities lurk because: (1) there may be a mismatch between expectations from the community and the lurker, (2) lurkers are learning and getting familiar with the community without interjecting, and (3) lurkers already feel like they are members of the community and don't see a value in increasing their presence amongst other contributors. We investigate many of these factors of lurking in our work.

**Barriers to Online Contribution.** Research identifies multiple barriers for contributing to online communities. Online communities, such as open source projects, are prone to conflict [32], which can be discouraging to users. For new users in an open source programming community, Steinmacher et al. [8] identified social barriers such as documentation issues, technical hurdles, and cultural differences that can hold newcomers back. Ridings et al. found that psychological barriers such as trust in the community can restrict contributions to online as well [33]. In examining barriers in online knowledge sharing, Hew at al. found that a lack of time and comfort with expertise are major obstacles to contribution [34].

Though these barriers can restrict online activity for some, not all forms of activity are impeded by them. Vasilescu et al. found that males participate longer than females in Stack Overflow activity [35]. Though females posted more questions, both genders received the same amount of answers [4]. The relatively "unhealthy" [4] community that is generated from this activity causes females to disengage sooner than males. In our work, we dive deeper into the online contribution experience to explain why this disengagement occurs.

## 3. RESEARCH QUESTIONS

Vasilescu et al.'s study [35] hypothesizes why females do not participate in online communities, such as disgust with sexism and the lack of gender neutral language. In our research, we want to know if females identify with these community and communication barriers or if they are facing other types of barriers. Further, we want to know which barriers may be gender-related, though gender may not be the only influence on barriers. Finally, we want to better understand the relationship of other factors, such as lurking on the site and experience, with the perception of barriers. In summary, the three research questions we want to answer through this study are:

- *RQ1*: What **barriers** do females face on Stack Overflow?
- *RQ2*: How do barriers vary by **gender**?

**Table 1: Interview participant demographics. A * denotes the highly ranked user.**

| P# | Occupation | Exp. (Years) | Usage Classification |
|---|---|---|---|
| P1 | Undergrad Student | – | Lurker |
| P2 | Undergrad Student | – | Lurker |
| P3 | Grad Student | – | Active |
| P4 | Grad Student | – | Active |
| P5 | Undergrad Student | – | Active |
| P6 | Grad Student/Industry | 2 | Active |
| P7 | Grad Student | – | Lurker |
| P8 | Grad Student | – | Lurker |
| P9 | Undergrad Student | – | Lurker |
| P10 | Undergrad Student | – | Lurker |
| P11 | Grad Student | – | Lurker |
| P12 | Industry | 1 | Lurker |
| P13 | Industry | 7 | Active |
| P14 | Industry | 4 | Lurker |
| P15 | Industry | 1 | Lurker |
| P16 | Industry | 10 | Lurker |
| P17 | Undergrad Student | – | Active |
| P18 | Undergrad Student | – | Lurker |
| P19 | Industry | 10 | Active |
| P20 | Industry | – | Active |
| P21 | Industry | 10 | Active |
| P22* | Industry | 10 | Active |

- *RQ3*: How do the rating of barriers vary by **other factors**, such as site usage and experience?

# 4. METHODOLOGY

This study included two forms of data collection: semi-structured interviews with female users, including an interview with a top 100 user of Stack Overflow, and survey responses received from female and male developers. All interview scripts and survey forms are available online.[2] Our overall methology is similar to the work of Begel and Zimmermann [36].

## 4.1 Interviews

As the objective of this work is to find out what females consider as barriers to participating on Stack Overflow, it makes sense to start by asking females (and not males) about their experiences. We interviewed female developers to understand how they use Stack Overflow. We recruited female developers from a variety of usage levels (no accounts, occasional contributor, and active contributor) and experience levels (students and professionals). In addition, we recruited one of the top ranked active female users. Identifying a female user who is ranked as a top user presents the opportunity to find out what encourages her to use the site and identify strategies to overcome barriers. In interviewing this person, the determining factor that kept her coming back can be compared to those from the general interviews.

### 4.1.1 Interview Script

To create our interview script, we seeded questions based on potential factors listed in Vasilescu et al.'s work [35], and from reasons listed in a Stack Exchange post titled "Why do you post to Stack Overflow?".[3] We also asked questions about how developers used various features of the site, potential motivations for contribution, and possible interventions.

Subject areas discussed during the interview include:

- Personal usage of the site
- How people communicate on the site
- Interests in gamification of the site
- Personal incentives to using the site
- Response to potential scenarios on the site
- Modifications to the site that may increase usage

### 4.1.2 Distribution

To recruit participants for the usage interviews we sent emails to a "women in computer science" mailing list and recruited females the authors know in computing. The recruitment email asked participants to fill out a preliminary survey asking for their employment status, years of industry experience, whether they used Stack Overflow for programming tasks, and if they had an account on the site. We received 25 responses from the recruitment questionnaire. We required respondents to select a time slot to interview as part of the recruitment questionnaire. We interviewed the 21 participants who attended their scheduled time slot. Interview participants received no compensation for their participation.

We then contacted a high ranked female user on Stack Overflow. This user was ranked in the top 100 users of all time listing.[4] We confirmed the gender of this user with another social networking site linked to her Stack Overflow user page. The 22 participants are shown in Table 1. The participant demographics consisted of nine professional software developers working in industry, twelve students, and one who identified as both.

### 4.1.3 Procedure

Prior to beginning each interview, the participant was sent a consent form to remind them that their personal identity will remain anonymous and that audio and notes will be recorded throughout the interview. For consistency, we conducted interviews with the same interviewer for 30 to 45 minutes. We conducted interviews in a private room where the participant had the option of meeting there or on a private video call.

The high ranked user's interview was conducted after the general user interviews. We scheduled an online video for the interview through email correspondence. With the high ranked user, we discussed several themes that arose from the other user interviews and focused on how her experiences compared and contrasted with other users.

### 4.1.4 Analysis

We first transcribed the audio recordings for each interview. We then performed three phases of analysis on the interview transcripts. In the first, exploratory, phase we jointly identified themes within the transcripts. The themes we identified in this phase of analysis include statements participants made while describing themselves, statements

[2] http://go.ncsu.edu/StackOverflowBarriers

[3] http://meta.stackexchange.com/questions/18888/why-do-you-post-to-stack-overflow

[4] http://stackexchange.com/leagues/1/alltime/stackoverflow

describing a participant's experiences while programming, and statements describing barriers deterring them from using Stack Overflow. For the purposes of this paper, we focus on the latter theme and refer to these statements as *barriers*.

For the second phase of analysis, we divided all the transcripts divided among three of the authors. Each transcript was examined by two authors. Each author initially coded their assigned transcripts independently, marking statements they identified as barriers. To ensure all the investigators agreed on which statements expressed barriers, we jointly reviewed exemplary statements and revised our codes. In total, we coded 327 statements as barriers.

In our final phase of analysis we grouped together similar statements and labeled each grouping as a distinct barrier. To do so, each author reviewed the barrier statements in their assigned transcripts. Iteratively, whenever an author encountered a barrier statement that did not fit into one of the existing groupings, the other authors reviewed that barrier and created a new label. After completing this process, we filtered out barrier statements that did not meet both of the following criteria.

**A barrier was identified if and only if**:

- Two investigators independently found that barrier in a transcript.
- At least two participants described that barrier in their interview.

This criteria is consistent with other studies [37]. The 14 resulting barriers are described in Section 5 and summarized in Table 2.

## 4.2 General Survey

We constructed a survey with the barriers identified from the interview participants.

### 4.2.1 Design

The survey consisted of questions regarding the ratings for barriers and demographics. We included all 14 barriers with a statement that further described each one in the survey. The survey presented all barriers in a random order. Survey participants were asked to rate on a 5-point Likert scale from strongly agree to strongly disagree how much the barrier stopped them from contributing to Stack Overflow. Participants also had the opportunity to write in a barrier that was not already listed.

Demographic questions included participants' level of experience on Stack Overflow. Participants had the option of selecting all that applied of: "Lurker (I use the site to find answers without contributing)", "I have a Stack Overflow account", "I post answers to questions", "I post new questions", and "I vote on responses". Participants also had the opportunity to describe their usage in an open-response. We asked participants to fill in their employment status; multiple answers were allowed. The only required demographic question was gender: Female, Male, or Other where participants could write in their gender.

### 4.2.2 Distribution

We distributed the survey to the general developer population. We sent targeted emails, posted to programming forums, contacted large corporations, and posted in computer science Facebook groups. Survey participants received no compensation for their participation.

### 4.2.3 Analysis

We received data from 1470 partcipants: 134 females and 1336 males. With the ratings received for each barrier the data was segmented across different populations (including employment status, and Stack Overflow usage). We used the ratings received to derive the collective ranking of barriers per population.

# 5. RESULTING BARRIERS

To answer RQ1, we identified 14 barriers by jointly tagging the transcribed participant interviews. These barriers are described by participants as reasons they did not contribute to the Stack Overflow community. We are not suggesting that only females would be affected by these barriers, or that the barriers are in some direct way about their gender. However, we are suggesting that barriers to participating in the Stack Overflow community do exist. The barriers are grouped into 3 categories: Muddy Lens Perspective, Impersonal Interactions, and On-Ramp Roadblocks. To the right of each barrier name is the number of participants that acknowledged it. Each barrier decribed includes a quote from atleast one corresponding participant's interview. Table 2 summarizes and reports on the number of participants who identified with each barrier.

## 5.1 Muddy Lens Perspective

Interview participants expressed having a "muddy" or unclear interpretation of how Stack Overflow works. This section describes how particular perceptions and expectations (whether justified or not) acted as barriers to contribute.

### 5.1.1 Awareness of Site Features (11)

Stack Overflow provides many features beyond threads for asking and answering programming questions. For example, users can earn reputation and badges, upvote & downvote answers, post bounties, and personalize their profiles. Many of these features are designed to encourage users to interact with the site.

To understand how participants valued the different features, the interview script included scenarios that asked how a participant would use a feature in a hypothetical situation. For example, we would ask if they might answer more questions if they could gain a badge or edit an incorrect answer. We were surprised that many participants were completely unaware of most of the features we asked about. After learning more about the features, some participants felt more interested in participating in the community. For example, P8 stated, *"No one has told me that creating an account would help a lot. You get some kind of perks by joining. I have not [heard] of anything like that, but had I then, I definitely would have created an account."*

A lack of awareness of potential usage features is a common issue for tools with many features, such as Eclipse [38]. One effective strategy for raising awareness occurs from serendipitous observation of other peers using a new feature [38]. However, for female users there are not many other females or users they can identify with as peers on Stack Overflow. In this case, who will the community recommendations come from?

### 5.1.2 Nothing Left to Answer (10)

Interview participants expressed interest in contributing to the site, but they had trouble finding questions avail-

**Table 2: Summary of barriers and number of interview participants who faced them**

| Group | Barrier | Participant Count | Description |
|---|---|---|---|
| Muddy Lens Perspective | Awareness of Site Features | 11 | I feel I am simply unaware of and have not explored the more advanced features of the site. |
| | Nothing Left to Answer | 10 | I feel all the easy questions have already been answered, leaving only hard questions. |
| | Fear of Contributing to Clutter | 9 | I feel my question might just be a duplicate or unimportant question, so I refrain from posting. |
| | No "Good-Answer" Guarantee | 7 | When posting a question, I fear not getting a good answer. |
| | Perception of Slacking | 4 | I feel that I should not be spending time answering questions on Stack Overflow for my own personal benefit. |
| Impersonal Interactions | Fear of Negative Feedback | 18 | I fear my posts being harshly criticized by users on the site. |
| | Stranger Discomfort | 9 | I feel uncomfortable interacting with and relying on help from strangers online. |
| | Intimidating Community Size | 9 | I feel intimidated by the large community of users. I instead prefer connecting with a smaller and more intimate group. |
| | Posting is Hard, Friends are Easy | 6 | I feel the process of posting questions is too cumbersome compared to other resources such as asking friends for help. |
| On-Ramp Roadblocks | Abstraction Process | 20 | I feel my problems require too many dependencies or proprietary aspects for me to abstract away before having something I can ask to a general audience. |
| | Time Constraints | 17 | I feel making contributions on Stack Overflow requires more time than I have. |
| | Qualifications | 13 | I feel my expertise or answers would not be of any help to anyone else. |
| | Onboarding Hoops | 9 | I feel figuring out the unspoken social etiquette and community standards is too much work. |
| | Research Pressure | 9 | I feel discouraged by the amount of work I have to do to prove that I'm not asking a duplicated question. |

able to answer. The two types of questions found are: (1) those they can answer, but have already been answered and (2) those that are too hard to answer. Between these two options participants expressed that they did not find the opportunity to contribute to the community and lost interest in posting. P2 described her experience searching for questions to answer, *"For a while I'd just try to find questions that I could answer... but eventually, it gets to the point where you're like eh, I'm pretty useless, because all the questions are super hard and all the easy questions have already been answered."*

### 5.1.3 Fear of Contributing to Clutter (9)

Stack Overflow implements mechanisms that discourage users from posting duplicate questions. When a user encounters a question they suspect of being a duplicate, they might typically comment, "This is a duplicate. See *the other* answer". Participants acknowledged that they do not want to make the site any more confusing for other users by adding to the clutter of duplicates. P20 specifically mentioned this as one of her reasons for not contributing, *"I didn't want to add to a bunch of duplicate stuff that wasn't useful... I didn't want to contribute to that issue."*

For some participants, this apprehension prevented them from posting at all. These findings support work done by Preece et al. [30] when studying lurker behaviors.

Similarly, clutter also appears on the site in the form of irrelevant conversations. The conversations can be distracting from the final answer and make users dig through the treasures of the site to find the golden answer they are looking for. P12 described her hesitation to add to conversations, *"I feel like if I don't know why it's wrong, I'm not contributing to the discussion. I'm just adding noise."*

### 5.1.4 No "Good-Answer" Guarantee (7)

Not every question on Stack Overflow gets answered. Participants, like P7, worried that after spending time carefully crafting a question — no one would respond with an answer: *"That's part of the apprehension... that I'm going to post it and I'll never hear back."*

Participants are justified in feeling anxious that the site may not guarantee quality responses: for newcomers, 90% of their questions are answered by themselves [13]. Even when Stack Overflow users respond with answers, participants, like P14, noted their answer quality can vary greatly: *"I think part of [the reason I never signed up] is I've seen so many bad answers on there, like wrong answers."*

### 5.1.5 Perception of Slacking (4)

Stack Overflow is online all the time; developers can contribute during their free time, or while they are at work. Participants with industry positions expressed a hesitation to contribute to the site while at work. They explained that

others perceive posting while on the job as slacking, even though learning and helping others on Stack Overflow might be considered a form of professional development. One professional who described this perception was P21, *"I just don't feel comfortable doing it at work. You're deviating from your actual development tasks. And when the timeline is so tight, I try to get in and out and back to what I'm supposed to be doing."*

## 5.2 Impersonal Interactions

Interview participants mentioned the lack of personal connections and uncomfortable atmosphere they encountered on Stack Overflow. This sections describes how the impersonal interactions became a barrier for their usage.

### 5.2.1 Fear of Negative Feedback (18)

When engaging with peers over the internet, there is always the possibility of coming in contact with internet bullies [32]. These people seem to have no filter when responding to posts online. Some participants perceived the blunt responses of these individuals as rude and argumentative. For example, P1 described the responses to us, *"Have you seen some of the [responses on there]? [They] will just like brutally destroy their answers."*

As opposed to joining, participants, like P5, would rather disengage and question how they can fit in the community, *"It's hard enough to ask for help, then to ask for help and get rude help. You are kind of like, 'never mind'."*

### 5.2.2 Stranger Discomfort (9)

Participants perceived the style of communication on Stack Overflow as blunt and impersonal. Participants identified the lack of females and familiar people as a reason why they felt uncomfortable on Stack Overflow. For example, P20 mentions how the dialogue on the site reminded her of a boy's club, *"I've definitely seen some comments that's not offensive exactly but it feels like I'm walking into a boy's club. You just get that vibe, how they talk."* P20 goes on to mention not feeling welcomed on the site and therefore did not want to engage, *"It doesn't make me feel especially welcome so it doesn't like encourage me to want to post more questions myself."*

P5 also acknowledges "bro humor" and colloquial references in answers as they were geared towards more of a male audience and not her, *"I feel like it's very jokey, but it's in a bro humor type way. The type of things, not to be stereotypical, that guys find funny. And so they usually, in a way, end up objectifying women. Then it makes it weird, because I guess it's funny, but not really, because this affects my life for real."*

P7 reiterated a similar sentiment of discomfort with posting on the site, *"I tend to save the question-asking with people I [know and] feel more comfortable with."*

In other online communities, getting acclimated to the culture as a newcomer [8] can be a difficult barrier to overcome. This is especially true when a group of strangers lack diversity and are not open to the opinion of others. In summary, the lack of personal connections on the site can discourage females from engaging.

### 5.2.3 Intimidating Community Size (9)

The fact that Stack Overflow is such a widely utilized community was stated to be both a benefit and a drawback for participants. The site is large enough that it has a variety of detailed responses. However, the site is so large that it is hard to identify with the entire community. Participants, like P14, acknowledged that if there were sub-communities of people they actually knew within Stack Overflow they would be more comfortable using the site: *"I enjoy being part of a community as long as they are kind of small. When it becomes kind of a sea of people [it feels] daunting or intimidating."*

### 5.2.4 Posting is Hard, Friends are Easy (6)

Many participants acknowledged that Stack Overflow was a great tool. However, when it came to the opportunity to post to Stack Overflow, there are other painless options still on the table. Participants navigated a hierarchy of increasingly painful options by first going to a friend for help before even considering the most painful option of posting. The interesting finding in this hierarchy is that though asking a friend is usually a last resort, it is seen as a more viable option than constructing and posting a question on Stack Overflow. P2 explained why she preferred asking friends for help, *"I ended up asking other people instead. Other people who could at least explain what is going on."*

## 5.3 On-Ramp Roadblocks

In the process of posting to Stack Overflow, participants encountered obstacles that undermined their interest in posting. This section highlights those roadblocks as barriers to usage.

### 5.3.1 Abstraction Process (20)

Participants had difficulty asking questions about their code on Stack Overflow. There were two basic problems. Legally, software developers cannot post proprietary code, because they may divulge company secrets. Second, even if they could post all their code, the community may not understand it, because it is so specific to their application, or it would be too long or detailed for anyone to understand. As a result, in order for participants to even ask a question, they first needed to strip a question of all proprietary or unnecessary code. Sometimes this process could be tedious and too much of a burden, and for this reason, they avoided posting questions in general. For example, P11 described the difficulties she encountered with the abstraction process, *"Here's some code. You probably can't run it, because there's like 20 dependencies on it, but just look at it for me and tell me what you think I can do different. For general problems, that's one thing. But for more specific problems, it's a little more difficult to use Stack Overflow"*

### 5.3.2 Time Constraints (17)

In addition to getting familiar with the site, participants expressed a lack of time to interact with the site. There is not enough time to devote to voluntary programming contributions on Stack Overflow beyond the work day. Some participants mentioned other hobbies and interests outside of work they feel are more deserving of their free time; programming was not one of them. For instance, P12 preferred to spend her free time on other activities, *"Actually, I think I would enjoy answering questions on those more if I found one that were more related to my personal interests. Because I strongly associate computer science with work. So it's not something I'd choose to do in my free time."*

### 5.3.3 Qualifications (13)

Some interview participants acknowledged they do not feel they have the expertise to post to Stack Overflow. For example, P1 stated, *"I don't feel like my expertise [is enough] for me to actually post an answer that would be of any help to anyone else."* These participants are not confident in their abilities to interact in the community to help others. In addition, they do not feel like they are qualified to give valuable answers to others. This lack of qualifications and confidence has decreased their interest in contributing.

### 5.3.4 Onboarding Hoops (9)

When joining a new community there is always a process of establishing norms. As frustrating for users as it may be, it is a right of passage that has to take place [39]. A community may have many unspoken expectations on how members conduct themselves. Some participants acknowledged interests in using the site, but were not clear on how. A recent data mining study by Honsel et al. demonstrates that the new users violate site rules more than old users [40]. Not having the proper guidance has discouraged many participants from actively using Stack Overflow. Some participants, P20 for example, are unclear of the norms and rather than asking forgiveness for violating them, they would rather remain reserved, *"I feel like everyone else already knows what it is. And [I] want to stay away from that extra work to figure out how to use it. Just figuring out what the etiquette is, all the little social things too that are kind of unspoken"*

### 5.3.5 Research Pressure (9)

There is a level of research expected to be done prior to posting a question in the community. During the process of crafting a questions, the user is posed with a list of questions that seem similar to the question that user is constructing. This is the site's way of asking, "Have you done your research?" These pressures are reinforced here where it is common for the user to enforce a culture of doing homework before asking a question. Questioning the abilities of users has discouraged some participants, like P4, from posting questions on the site: *"I think there's only one case where I was close to posting a question, but then it said do your research, this question might already be there."*

## 6. RANKING AND RATING BARRIERS

Although we identified barriers through interviews, we want to identify which barriers can have a strong impact across genders and other groups. By distributing a survey, we can analyze the prevalence of these barriers in a broader population.

The third and fourth authors analyzed the free-form responses from the survey, by independently labeling the response with our 14 barriers or other category. In the analysis, we did not identify a new barrier. With the survey data, we answer our remaining research questions related to the ranking, differences in gender, and influences of other factors on barriers.

## 6.1 Gender Comparison

To answer RQ2, we performed a statistical analysis to identify contribution barriers across genders and understand how the populations differed in their response. We performed a two-tailed Wilcoxon rank-sum test on the ratings given to each barrier to compare populations. Table 3 demonsrates that 5 out of 14 barriers had a statistically significant difference between females and males. To be clear, we are not suggesting that only females are affected by these barriers, or that these barriers are primarily due to gender, but rather that 5 barriers were seen as significantly more problematic by females than by males. The barriers we identified cannot conclusively occur from gender differences alone. All barriers with $\alpha = .05$; $\alpha = .0012$ after Bonferroni correction [41] have been highlighted in green in table 3. This correction was derived by dividing the original alpha value of .05 by the 3 comparisons conducted on each of the 14 barriers. The three comparisons included gender, usage status, and employment status. In table 3, 5 is used to indicate strongly agree and 1 for strongly disagree. The columns labled F and M indicate the mode/median for females and males respectively. For example a value of 3.5 indicates a response between neutral and agree. The column labeled ES indicates the effect size which was calculated by taking the absolute value of subtracting the mean of the male distribution from the female distribution. The last column indicates the likert distribution for females from strongly agree to strongly disagree. The table is order from greatest to least agreement(combining strongly agree and agree ratings) for females. A diverging stacked bar chart comparing the distribution of barriers for females and males is also available online.[5]

**Table 3: Statistically significant differences in how females and males rated barriers.**

| Barrier | $p$-value | F | M | ES | F Likert |
|---|---|---|---|---|---|
| Fear Neg. Feedback | .004 | 4/4 | 5/3 | 0.4 |  |
| Fear Contrib. Clutter | .013 | 4/4 | 4/3 | 0.3 |  |
| Time Constraints | .497 | 4/4 | 4/4 | 0.1 |  |
| Nothing Left | .797 | 4/4 | 4/4 | 0.0 |  |
| Awareness | <.001 | 4/4 | 1/3 | 0.5 |  |
| Research Pressures | .374 | 4/4 | 4/4 | 0.1 |  |
| Posting is Hard | .059 | 4/3.5 | 4/3 | 0.2 |  |
| Abstraction Process | .729 | 4/3 | 4/3 | 0.0 |  |
| Qualifications | <.001 | 4/3 | 1/3 | 0.4 |  |
| Onboarding Hoops | .062 | 4/3 | 5/4 | 0.2 |  |
| No "Good-Answer" | .239 | 3/3 | 2/2 | 0.1 |  |
| Intim. Comm. Size | <.001 | 4/3 | 1/2 | 0.5 |  |
| Stranger Discomfort | <.001 | 1/2 | 1/1 | 0.4 |  |
| Percept. Slacking | .001 | 2/3 | 1/3 | 0.4 |  |

## 6.2 Other Factors

We understand that confounding factors, other than gender, may also affect Stack Overflow participation. For example, online interactions and programming experience could play a role. To answer RQ3, we review a few of these factors using a statistical analysis to compare how barrier ratings varied across Stack Overflow usage and employment status.

### 6.2.1 Usage Status

This analysis compared the participants who reported having a Stack Overflow account *(n=1003)* to those who did not *(n=467)*. We observed statistically significant differences (Wilcoxon rank-sum test with Bonferroni correction

[5] http://go.ncsu.edu/StackOverflowBarriers

$\alpha = .0012$) between account holders and non-account holders ratings of seven barriers. In all seven cases, non-account holders were more likely to identify with a barrier. Specifically, non-account holders were more likely to identify with the following barriers (listed in order of smallest p-value to greatest): *Awareness of Features*; *Stranger Discomfort*; *Intimidating Community Size*; *Fear of Contributing to the Clutter*; *Posting is Hard, Friends are Easy*; *Qualifications*; and *Onboarding Hoops*. These barriers suggest places where effort can be focused to encourage lurkers to jon.

#### 6.2.2 Employment Status

We also compared survey responses from those who identified as full time developers *(n = 1003)* to those who did not identify as such *(n = 467)*. We observed statistically significant differences (Wilcoxon rank-sum test with Bonferroni correction $\alpha = .0012$) between full time developer and non-full time developer ratings of two barriers. We found that fulltime developers were more likely to identify with *Time Constraints*. On the other hand, the non-full time developer group were more likely to identify with *Qualifications*.

## 7. DISCUSSION & FINDINGS

This section outlines observations and interventions for making Stack Overflow a more inclusive resources for females, as well as the entire development community. We are not suggesting that these experiences are unique to females, but have been outlined from their perspective and can help define a direction for the community.

### 7.1 The Female Experience

#### 7.1.1 Receiving and Giving Criticism

The design and community of Stack Overflow fosters a culture of criticism, which limits who is willing to contribute. The criticism does not always come from other users, but can arise in the design of the site imposing undesired criticism on user actions. For example, the process of writing a new question to Stack Overflow invokes the barrier of *Research Pressure*, by the site prompting the user to make sure they aren't posting a duplicate question and to first check if their answers have been already been posted. This is reinforced by users on the site, who may criticize the question for being a duplicate of another (whether justified or not). Even males rated *Research Pressure*, which characterizes the barriers presented when they want to post a question, as their highest ranked barrier.

Female users pointed out during interviews that they were reluctant to offer criticism, even when the interaction was indirect. For example, when participants were asked whether or not they would critique or edit an incorrect response posted on the site by another user, many were reluctant to do so because too many critiques already exist on the site. Others pointed out there were plenty of other users that can make the edit — an online version of the bystander effect.

Finally, a sense of dissatisfaction with the culture of criticism on the site strongly resonated with interviewees and survey respondents. This is mostly reflected by *Fear of Negative Feedback*: out of the 22 participants interviewed, 18 identified with this barrier as shown in Table 2. One participant spoke on being prepared to receive this negative feedback: *"If you know you're going to be questioned when you are right and wrong, you are going to make sure that you are 100% right. [If] you gonna question me, I'm going to be ready!"*. This comment references the combative type of dialogue that exists on the site. Even if males are not directly negatively criticizing females, seeing them do it to other males creates an unwelcoming atmosphere, as participants stated in interviews and we observed in survey responses. Overall, unwanted criticism can be demeaning and discouraging to those who have finally mustered up enough courage to engage with the community.

**Observation 1 -** The design and community of Stack Overflow fosters a culture of criticism, which limits who is willing to contribute.

#### 7.1.2 Qualified Personnel

According to interview responses, *Qualifications* ranked fourth for the most common barriers that arose. In short, even in the interviews with active contributors, there is still a concern of not being qualified enough to actively contribute to the site. From this study, we found a perception perpetuated through participants regarding the accomplishments and prior knowledge that users should have before contributing. One participant said, *"...you need to know more than you NEED to know to use Stack Overflow."*. These self-doubts as identified in the interviews have prevented participants from answering questions.

Participants also expressed how they have attempted and are interested in contributing, but have not found an opportunity to post on the site. One participant during the interview mentioned, *"all the questions are super hard and all the easy questions have been answered"*. Participants are under the interpretation that there is *Nothing Left to Answer* on the site that they could add value to.

**Observation 2 -** Participants want to help but do not feel like they can based on difficulty in finding opportunities to post.

#### 7.1.3 The Language Divide

There are two facets of language that occur on Stack Overflow that divide the community. One facet of language on Stack Overflow are complex questions that often go unanswered [13]. These complex questions come from the difficulty users have with the *Abstraction Process*. Interview participants identified this barrier as the most common encountered with 20 out of the 22 participants in Table 2. In addition, survey ratings identified this process as a barrier that both females and males agree (including strongly agree and agree) exist, 44.78% and 42.44% respectively. These findings suggest that this complex language is not simply a gender influenced barrier, but one that can affect all contributors to Stack Overflow. We need to dissolve this complex technial language divide [8] that exists on Stack Overflow to welcome those it discourages.

The other facet of language existing on Stack Overflow is the "boy's club" as mentioned by an interview participant. The current atmosphere on Stack Overflow is not inclusive and often times disrespectful to others who are not males. One survey participant acknowledged that though they identify as a female, they decided to go on Stack Overflow and portray themself as a male. This is under the assmption that males are more likely to be treated with respect and blend better into the community on the site. If Stack Overflow was a true "programmer's paradise", it would be a place where contributors should feel comfortable to be themselves

and not have to worry about trying to blend into the community. One suggestion proposed during an interview to encourage a comfortable environment is to use more gender inclusive language.

**Observation 3 -** There are two types of language that contribute to the unwelcoming atmosphere on Stack Overflow: complex questions and idle "boy's club" chatter.

### 7.1.4 Accounts Are Not Enough

Many survey respondents($\frac{1003}{1470}$) indicated that they had an account. However, about 5% of accounts, despite being active, are ever signed in when visiting Stack Overflow [3]. In addition, one-day flies, those who have posted a maximum of one question for the life of their account, make up 47% of the user base [13]. In our interviews with users, we were surprised by how many features, such as voting, badges, bounties, and moderators, participants are not simply aware of or knew how they operated. Once we asked interview participants about potential benefits to using site features, some participants agreed that they would use the site as a more active account holder. The basic description of some features intrigued participants to learn more about Stack Overflow and its benefits.

In our survey, *Awareness of Features* ($p$ = 5.3e-06) especially affected female users. The design of the site reinforces this answer searching process: (1) enter the question in a search engine, (2) Stack Overflow is the top rated answer, (3) read question and if correct, (4) read answer and if correct, (5) copy and paste code, then vote up if answer is very helpful, and (6) return to previous programming tasks until next roadblock. This process does not encourage users to explore the site. Given the multiple barriers related to onboarding and joining the community, users rarely get the opportunity to seek out the benefits to using the site beyond solving the their quick programming problem.

**Observation 4 -** Females and males both have accounts, however awareness of the features that go along with the account is a factor in not contributing.

## 7.2 Design Implications

### 7.2.1 Asking and Answering

**Five Minute v.s. Thirty Minute Questions.** P22, our most active interviewee, viewed answering Stack Overflow questions as something she can do as a relaxing activity instead of breaking out a phone to play on for 5 minutes. She classified answering questions into 2 categories: an intellectual challenge and relaxing social endeavor. This participants goes on to say, *"Both are very interesting...but my current emotional state will definitely affect which of them I'm more interested in at any given time."* We want to be able to address both interests which will provide options for a varied range of users. However, it is a bit of a task trying to identify which questions will take 5 minutes to answer versus 30 minutes manually. In fact, many participants reported having difficulty navigating and find an unanswered question they could respond to.

Research is needed to predict the difficulty of an asked question and how long it would take to provide an answer. This concept is similar to how Medium,[6] a popular blogging site, instruments and predicts how long it takes for someone to read an article. Incorporating this feature into Stack Overflow posts can increase the number of users that answer questions on the site. Unanswered questions could be ranked and rewarded by perceived difficulty in answering them. The algorithm for determining difficulty of these unanswered questions can include the number of views, time of original post (already provided by Stack Overflow), and background of people who have viewed but have not answered the post. Other features to include are the length of code snippets in posts and whether or not one exist in the post. Researchers can use feedback from the predicted time and actual time to answer in order to refine the algorithm. This would allow people to distinguish between questions that would take 5 minutes from those that might take 30 minutes to answer and being rewarded accordingly. This concept model of tailoring the activity of each user [42] tackles both the *Time Constraint* and *Qualifications* barrier.

**Implication 1 -** Research is needed to create and sustain a ranking algorithm of questions response time per user's skill and question difficulty. This features would encourage a wider range of users with varied availability.

**Quality Questions.** Instead of discouraging users, as the *Research Pressure* barrier acknowledges, the site can educate users on how to improve the likelihood of getting questions answered. One common barrier participants faced was the frustration to remove very specific details from questions during the *Abstraction Process*.

An algorithm can be used to help rephrase questions in a way that more users may be interested in responding. There are already algorithms on Stack Overflow for detecting whether the code in a posting is not formatted, but further refinement is needed. For example, a question with a title that explains the general concept that the question poster is trying to overcome will be easier to find through search results as it will contain buzz words that will be noticed by other users. We propose a new question framing algorithm to determine suggestions that include looking into questions that have high votes, the many types of answers a question receives, and the type of dialogue created through comments. Suggestions given to users can include how to phrase questions and remove specific details of the coding scenario.

Framing questions and answers appropriately shows that user contributions can be well received by the community [13] and help generate successful dialogue through comments [43]. As users have more answered questions, they will be able to see the value in their contributions of asking high quality questions and more inclined to post these questions to the site.

**Implication 2 -** Instead of discouraging users from posting questions, enhance the posting process by automatically providing feedback on the quality of the question in terms of how fast and how likely it will be answered.

### 7.2.2 Helping Hands

**Sub-communities.** One interview participant specifically stated that she would go out of her way to answer a question for another female user if she knew she was a female. Smaller groups on Stack Overflow will encourage personal connections participants said Stack Overflow lacked, while reducing the *Intimidating Community Size* barrier. In these sub-communities users may feel more comfortable disclosing profile pictures and conversations could be tailored towards

[6] https://medium.com/

the specific group interests. Within these sub-communities females can feel more comfortable in their niches [44] and encouraged to ask the questions they may be hesitant to ask on the full site. We want to create an inclusive community that encourages people to use the resource as well as contribute to it.

Another interview participant mentioned that they would explicitly cater to answering female questions more than males. She also goes on to mention how she identifies more with comparing personal progress to other females, *"At every point where you share experiences with another person who is so like you, but not you. It's like a control experiment and you can see how [you are] doing."* Leaderboards that show rankings of just female users may allow other females to more comfortably compare themselves to their own peers as opposed to the entire site. Some interview participants go on to mention a distinct difference in the opportunities for females and males by saying, *"They are just not even on the same race track."* There are organizations that acknowledge that there will be some difference between the support that is needed for different groups of developers. For the same reasons there are support groups offered for underrepresented groups in other programming settings, such as Hackbright Academy[7] and Black Girls Code[8], Stack Overflow should incorporate resource groups for those interested. It may be worthwhile to consider an inclusive space for females only.

**Implication 3 -** Sub-communities should be an available option to users who are interested in creating a more approachable group to interact with.

**Mentorship.** Currently, there are 18 moderators whose goal is to keep the 10 million question community[9] clean for the 4.6 million users on the site. Unfortunately, this ratio of moderators to questions to users does not scale. The community is not in a state "clean" enough for new users to navigate on their own. An interview participant mentioned the idea of having badges that are geared explicitly toward learning and helping other programmers. Participants would be interested in explicitly being informed they are helping others. This may be as simple as renaming badges to be attached to the title of a mentor. Having a mentor, a more experienced user, encourage their mentee, a less experienced user, throughout the process makes participants feel like they are not alone on the site [8]. This creates a better sense of belonging to the Stack Overflow community.

Many participants mentioned in both the interviews and surveys that there are many features and site etiquette that they are unaware of. The mentors can provide mentees with tutorials of the site and educate their mentee on Stack Overflow culture, expectations from the community, and intricate features as interested. By eliminating the *Awareness of Features* barrier, Stack Overflow potential contributors can be more prepared to engage in the community. Interview participants acknowledged that making the site feel like users are interacting with real people would encourage them to use the site. With this approach, the tone will be adjusted and move towards having similar qualities of actually asking a friend for help with a programming question.

[7] https://hackbrightacademy.com/
[8] http://www.blackgirlscode.com/
[9] http://stackoverflow.com/10m

Participants consistently mentioned that the overwhelming community size on Stack Overflow discouraged them from using the site. A mentorship program may increase new-user retention by combating this barrier.

**Implication 4 -** Retention of one-day flies and other contributors may increase if per user mentorship program is incorporated.

# 8. THREATS TO VALIDITY

**External Threats.** Although we interviewed females who are account holders of Stack Overflow, having more participants who are highly active on the site may have provided more experiences to review. This would have identified stronger differences between the different active users allowing these findings to apply to a wider population. In addition, having the opportunity to speak with multiple top female users provides the opportunity to compare their initial experiences on the site to their current usage.

**Internal Threats.** In receiving interview responses from just a female audience, we run the risk of not being able to compare the responses to males to identify if these experiences are common across genders. We explicitly chose to study females on Stack Overflow because many prior works have studied the site for the general population, not identifying the unique experiences that occur for females. However, to confirm that barriers exist and are not limited by genders, we incorporate a survey of both females and males.

**Construct Threats.** Through the survey many participants identified as having accounts. If we would have asked more open ended questions regarding account usage, we would have more insight as to incentives for using the site. The opportunity for gender swapping via the survey is also possible. One survey participant mentioned that their Stack Overflow account lists them as a male, though the participant identifies as a female. We did not explicitly ask interview participants if they identify as a different gender online than offline. However, in asking participants for their gender identity and providing a free response, we acknowledge that gender is more than a simple binary construct.

# 9. CONCLUSION

"Community is everything" at Stack Overflow [3]. Since 2008, the success of Stack Overflow has been based on those who post and answer questions. However, the bridge to paradise is burning and it is hard for female users to contribute and plug into the community. We used a novel approach of interviewing females to identify barriers hindering them from contributing and then confirmed the 14 barriers with survey responses from both females and males. We found 5 barriers females strongly identify with and discuss literature to support these findings. To encourage the growth of this widely used community across genders, we provide design implications to mitigate barriers of usage. Although we identified several promising interventions that can help overcome barriers, we encourage their implementation to study community effects.

# 10. ACKNOWLEDGEMENTS

We would like to thank reviewers for their valued feedback on this work. This material is based upon work supported by the National Science Foundation under grant number 1318323.

## 11. REFERENCES


[1] J. Hanlon. (2013) Five years ago, stack overflow launched. then, a miracle occurred. [Online]. Available: http://blog.stackoverflow.com/2013/09/five-years-ago-stack-overflow-launched-then-a-miracle-occurred/

[2] J. Atwood. (2008) Introducing stackoverflow.com. [Online]. Available: http://blog.codinghorror.com/introducing-stackoverflow-com/

[3] StackOverflow. (2015) 2015 stackoverflow developer survey. [Online]. Available: http://stackoverflow.com/research/developer-survey-2015

[4] B. Vasilescu, A. Capiluppi, and A. Serebrenik, "Gender, representation and online participation: A quantitative study," *Interacting with Computers*, p. iwt047, 2013.

[5] U. S. D. of Labor, *Labor Force Statistics from the Current Population Survey*, 2016. [Online]. Available: http://www.bls.gov/cps/cpsaat11.htm

[6] P. A. David and J. S. Shapiro, "Community-based production of open-source software: What do we know about the developers who participate?" *Information Economics and Policy*, vol. 20, no. 4, pp. 364–398, 2008.

[7] H.-A. Truong, G. Williams, J. Clark, and A. Couey, *Gender issues in online communications*. na, 1993.

[8] I. Steinmacher, T. U. Conte, M. Gerosa, and D. Redmiles, "Social barriers faced by newcomers placing their first contribution in open source software projects," in *Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing*, 2015, pp. 1–13.

[9] N. Abilio Oliveira, "Culture-aware q&a environments," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015, pp. 101–104.

[10] A.-M. Slaughter, "Why women still can't have it all," *Atlantic Monthly*, 2012.

[11] J. Rosener, "How women lead," *Harvard business review*, vol. 68, no. 6, pp. 119–125, 1990.

[12] B. Vasilescu, D. Posnett, B. Ray, M. G. J. van den Brand, A. Serebrenik, P. Devanbu, and V. Filkov, "Gender and tenure diversity in GitHub teams," in *CHI Conference on Human Factors in Computing Systems*, ser. CHI. ACM, 2015.

[13] R. Slag, M. de Waard, and A. Bacchelli, "One-day flies on stackoverflow: Why the vast majority of stackoverflow users only posts once," in *2015 12th Working Conference on Mining Software Repositories*. IEEE, 2015, pp. 458–461.

[14] C. Parnin, C. Treude, L. Grammel, and M.-A. Storey, "Crowd documentation: Exploring the coverage and the dynamics of api discussions on stack overflow," *Georgia Institute of Technology, Tech. Rep*, 2012.

[15] S. Grant and B. Betts, "Encouraging user behaviour with achievements: An empirical study," in *Proceedings of the 10th Working Conference on Mining Software Repositories*, ser. MSR '13. Piscataway, NJ, USA: IEEE Press, 2013, pp. 65–68.

[16] J. Cuny and W. Aspray, "Recruitment and retention of women graduate students in computer science and engineering: results of a workshop organized by the computing research association," *ACM SIGCSE Bulletin*, vol. 34, no. 2, pp. 168–174, 2002.

[17] E. S. Roberts, M. Kassianidou, and L. Irani, "Encouraging women in computer science," *ACM SIGCSE Bulletin*, vol. 34, no. 2, pp. 84–88, 2002.

[18] V. R. Lee, M. Recker, and T. Sumner, "Variable appropriation of an online resource discovery and sharing tool," 2013.

[19] D. Grigar, "Over the line, online, gender lines: E-mail and women in the classroom," *Feminist cyberscapes: Mapping gendered academic spaces*, pp. 257–281, 1999.

[20] R. M. Hall and B. R. Sandler, "The classroom climate: A chilly one for women?." 1982.

[21] C. Avin, B. Keller, Z. Lotker, C. Mathieu, D. Peleg, and Y.-A. Pignolet, "Homophily and the glass ceiling effect in social networks," in *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, ser. ITCS '15. ACM, 2015, pp. 41–50.

[22] J. Margolis and A. Fisher, *Unlocking the clubhouse: Women in computing*. MIT press, 2003.

[23] P. Bailey, N. Craswell, and D. Hawking, "Dark matter on the web," in *Poster Proceedings, 9th World-Wide Web Conference*, 2000, p. 2.

[24] L. Giles and S. Lawrence, "Accessibility and distribution of information on the web," *Nature*, vol. 400, no. 6740, pp. 107–109, 1999.

[25] A. Menking and I. Erickson, "The heart work of wikipedia: Gendered, emotional labor in the world's largest online encyclopedia," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 207–210.

[26] C. Lampe, J. Vitak, and N. Ellison, "Users and nonusers: Interactions between levels of adoption and social capital," in *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, ser. CSCW '13. ACM, 2013, pp. 809–820.

[27] B. Nonnecke and J. Preece, "Lurker demographics: Counting the silent," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2000, pp. 73–80.

[28] M. Muller, N. S. Shami, D. R. Millen, and J. Feinberg, "We are all lurkers: consuming behaviors among authors and readers in an enterprise file-sharing service," in *Proceedings of the 16th ACM international conference on Supporting group work*. ACM, 2010, pp. 201–210.

[29] M. Takahashi, M. Fujimoto, and N. Yamasaki, "The active lurker: Influence of an in-house online community on its outside environment," in *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work*, ser. GROUP '03. ACM, 2003, pp. 1–10.

[30] J. Preece, B. Nonnecke, and D. Andrews, "The top five reasons for lurking: improving community experiences for everyone," *Computers in human behavior*, vol. 20, no. 2, pp. 201–223, 2004.

[31] B. Nonnecke, D. Andrews, and J. Preece, "Non-public and public online community participation: Needs, attitudes and behavior," *Electronic Commerce Research*, vol. 6, no. 1, pp. 7–20, 2006.

[32] A. Filippova and H. Cho, "Mudslinging and manners: Unpacking conflict in free and open source software,"

in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ser. CSCW '15. ACM, 2015, pp. 1393–1403.

[33] C. Ridings, D. Gefen, and B. Arinze, "Psychological barriers: Lurker and poster motivation and behavior in online communities," *Communications of the Association for Information Systems*, vol. 18, no. 1, p. 16, 2006.

[34] K. F. Hew and N. Hara, "Empirical study of motivators and barriers of teacher online knowledge sharing," *Educational Technology Research and Development*, vol. 55, no. 6, pp. 573–595, 2007.

[35] B. Vasilescu, A. Capiluppi, and A. Serebrenik, "Gender, representation and online participation: A quantitative study of stackoverflow," in *Social Informatics (SocialInformatics), 2012 International Conference on*. IEEE, 2012, pp. 332–338.

[36] A. Begel and T. Zimmermann, "Analyze this! 145 questions for data scientists in software engineering," in *Proceedings of the 36th International Conference on Software Engineering*, ser. ICSE 2014. ACM, 2014, pp. 12–23.

[37] T. Roehm, R. Tiarks, R. Koschke, and W. Maalej, "How do professional developers comprehend software?" in *Proceedings of the 34th International Conference on Software Engineering*, ser. ICSE '12. IEEE Press, 2012, pp. 255–265.

[38] E. Murphy-Hill and G. C. Murphy, "Peer interaction effectively, yet infrequently, enables programmers to discover new tools," in *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. ACM, 2011, pp. 405–414.

[39] D. Ford and C. Parnin, "Exploring causes of frustration for software developers," in *8th International Workshop on Cooperative and Human Aspects of Software Engineering*. IEEE/ACM, 2015, pp. 115–116.

[40] V. Honsel, S. Herbold, and J. Grabowski, "Intuition vs. truth: Evaluation of common myths about stackoverflow posts," in *2015 12th Working Conference on Mining Software Repositories*. IEEE, 2015, pp. 438–431.

[41] Y. H. Yoav Benjamini, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[42] B. Johnson, R. Pandita, E. Murphy-Hill, and S. Heckman, "Bespoke tools: adapted to the concepts developers know," in *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing. IEEE*, 2015.

[43] F. Calefato, F. Lanubile, M. C. Marasciulo, and N. Novielli, "Mining successful answers in stack overflow," in *2015 12th Working Conference on Mining Software Repositories*. IEEE, 2015, pp. 430–433.

[44] Z. Sahib and J. Vassileva, "Designing to attract participation in a niche community for women in science & engineering," in *Computational Science and Engineering, 2009. CSE'09. International Conference on*, vol. 4. IEEE, 2009, pp. 909–914.