Student Experiences with GitHub and Stack Overflow: An Exploratory Study

Trishala Bhasin  
Department of Computer Science  
University of Victoria  
Victoria, BC, Canada  
trishalabhasin@uvic.ca

Adam Murray  
Department of Computer Science  
University of Victoria  
Victoria, BC, Canada  
adammurray@uvic.ca

Margaret-Anne Storey  
Department of Computer Science  
University of Victoria  
Victoria, BC, Canada  
mstorey@uvic.ca

Abstract—Programmers who want to improve their skills in software development rely heavily on developer social platforms such as GitHub and Stack Overflow to enhance their learning. Stack Overflow provides answers to questions they have about languages or library skills they wish to acquire, while contributing to open-source projects hosted on sites like GitHub gives them valuable experience. Students also use these platforms during their education: most will rely heavily on Stack Overflow at some point in their schooling, while many can benefit from contributing to GitHub projects to build their expertise and professional portfolios. We already know from previous research that developers face barriers participating on these platforms, so we may expect that at least some students will experience similar barriers and possibly even bigger challenges. This paper describes a semi-structured interview study with university students to explore how they use the GitHub and Stack Overflow platforms. We identify the barriers they face and benefits they report from using these tools. We conclude with some preliminary recommendations on how to reduce the hurdles students may face with these and other developer social platforms, and suggest future work to mitigate these roadblocks.

I. INTRODUCTION

Developer social platforms, such as GitHub, Bitbucket, Stack Overflow and Reddit, provide a valuable resource for modern software developers [2]. For example, developers use GitHub and Bitbucket to host software projects and to make meaningful contributions to other projects, they use Stack Overflow and Reddit to enhance their own knowledge and skills, and to solve problems they may face while developing code. Developers benefit in other ways when they explicitly contribute to resources hosted on these open platforms (e.g., contributing by authoring or voting up/down a resource) as they gain visibility in the community due to the social and public nature of the platforms [21] [23]. Furthermore, their public contributions often comprise portfolios they can use when applying for or switching jobs [20].

There has been extensive research in recent years on how developers both contribute to and benefit from these platforms. As GitHub and Stack Overflow are the most widely used, they have been studied the most frequently. This research has also revealed many barriers (such as lack of confidence or time) that certain developer demographics experience when contributing to resources hosted on GitHub and Stack Overflow [11]. However, relatively little has been done to investigate how students use GitHub and Stack Overflow, or other popular developer platforms. Do students experience similar or even different benefits to more senior developers, and do they encounter similar or new barriers that may impede them from the benefits they could achieve?

We do know that students are an important part of these developer social platforms and that they participate in a variety of ways, including hosting their code online, contributing to other people’s code, and learning more about programming by finding answers to their questions and by helping others [3]. Platforms such as Reddit offer a medium for students to learn and grow through personalization, customization, and opportunities for networking and collaboration [4]. This was observed in a study by McLoughlin et al. that focused on the interactive aspect of today’s Internet called the modern Web 2.0. Additionally, many software development companies request a student’s GitHub profile as a part of their hiring process1. Hence, many students do strive to maintain a positive open community presence as it increases their chances of getting hired in industry.

What has been less studied are the barriers students face while using open platforms such as GitHub and Stack Overflow. We feel this is important to investigate because as educators we may be in a position to act on some of these barriers at an early stage of student development, and we may also be able to share insights with the designers of these and other developer social platforms.

In this research, we conducted an exploratory study and semi-structured interviews with students at different stages of their studies to investigate how they use and contribute to content posted on GitHub and Stack Overflow. We report in depth on the barriers they encounter and the benefits they perceive from using or contributing to these platforms as these have not been studied much by others. Many of the

1https://techbeacon.com/app-dev-testing/what-do-job-seeking-developers-need-their-github
barriers students encounter overlap the barriers reported in earlier research for developers, but we find that students face even more doubt in their abilities to contribute and lack the motivation to help others even though doing so could help them. Our research also points to recommendations that educators and students may follow to increase participation and meaningful contributions, and our findings may also be relevant to designers of developer social platforms.

In the next section of the paper, we first discuss related work that describes what we know about developer experiences using GitHub and Stack Overflow, and what we already know about student participation on these platforms. We then report on our study, findings, and recommendations.

II. BACKGROUND

Prior research on participation and interaction in developer social platforms has focused on different demographics of the participants in the communities formed around each of these tools. For example, some studies have looked at the factors that attract engagement [2], while others have investigated the participation of women in particular [3] [7]. For those who wish to participate, prior research has shown that the barriers can be intrinsic as well as extrinsic [8].

These (so far mostly exploratory) studies have surveyed small segments of the developer community, such as seasoned or more junior developers, university students, etc., that exist within the community, analyzing their involvement and providing suggestions for mitigating barriers to their participation [10]. Communities on social platforms like GitHub and Stack Overflow are heterogeneous, each platform is used quite differently, and the problems faced on each platform can also vary greatly. We discuss related work about these platforms below and also refer to related work later in the paper.

A. Developer Social Platforms and Their Communities

Developer social platforms such as GitHub, GitLab, and Bitbucket, and question-answer forums such as Stack Overflow are integral parts of collaborative software development [2]. They provide a space where people can actively showcase their programming projects to other people, seek collaborations on their work, ask questions, and resolve issues.

The formation of developer communities of practice begins when a developer decides to share the source code for their project in an open-source repository and the project gains attention from people in the form of active collaboration. One such case is that of the Linux operating system, one of the most successful software projects under development. The entire development of the Linux kernel happens through the contributions of the open-source developers who built an ecosystem around it [8]. The sharing of source code on these platforms is associated with the idea that there should be content available online so the community can use the information in any way they choose. The content being generated is heavily dependent on participation by people, and as Sun et al. [9] mentioned in their literature review on understanding lurkers2 in online communities, the majority of the code is generated by a minority of people accessing the data.

B. Student Participation in Open-Source Communities

Students are an important part of the communities that form around open social developer platforms [1] and they directly benefit from this participation. Students interact with developer social platforms during their degrees to seek help with their assignments or to collaboratively develop projects and learn new languages. Sometimes students get involved in internships where they have the opportunity to work on open-source projects3. The participatory aspect of developer social platforms is used by students to connect with and learn from each other [4] [22].

C. Barriers to Participation in Open Communities

Previous studies identify barriers to participation on communities that form around developer social platforms. Indeed, some research shows a 90-9-1 pattern of participation, where 90% of the participants don’t effectively participate in online conversations, 9% show a small amount of activity, and only 1% of the participants represent practically all of the online activity [6]. Studies of question-answer forums like Stack Overflow suggest that community newcomers face many barriers: 90% of accepted answers provided by new users are self-answers, and a lot of questions go unanswered [5]. The community’s growth is also limited by a core of elite users: only 5% of the users answer 60% of questions [3].

Barriers can be social or technical in nature. Social barriers can include discrimination around participation on the basis of stereotypes, or an unwelcoming environment for new users [6]. Technical barriers, which are more prevalent on code sharing platforms like GitHub, can include a lack of guidelines around how to make contributions, and a lack of documentation to help explain the code [11]. Mendez et al. found that there is always a population of people who are willing to participate but face barriers, and there are always people who are not willing to participate [11]. In Figure 1, we synthesize the barriers that are discussed in the existing literature and that we anticipate we may find in our study with students.

Fig. 1. The barriers can be divided into two sets: firstly, barriers that relate to a lack of interest [11], including a lack of passion [8] and altruism; and secondly, barriers that are faced when developers are interested in participating [11], including a lack of confidence, misconceptions about community culture [6] and support along with poor management of time.

---

2 People who consume content without contributing [6]

3 https://summerofcode.withgoogle.com/
III. METHODOLOGY

In this paper, we address the following research questions:

RQ1: How do students participate and contribute to open content hosted on GitHub and Stack Overflow?

RQ2: What are the barriers that students face in their participation on GitHub and Stack Overflow?

RQ3: For the students who participate, what benefits do they experience from their active participation on the developer social platforms under investigation?

To answer our questions, we used semi-structured interviews to gather insights on students’ personal experiences, attitudes, perceptions, and beliefs [12] [13]. We anticipated that interviews would help us better understand the social phenomena around using these platforms, and allow us to share meaning around the experiences and views of our student participants. We have provided some supplementary material online4.

We decided to focus our study on more senior undergraduate, graduate and newly graduated students as many of them would have had some co-op/internship experience or experiences in their courses in which they would have been introduced to these platforms. In addition, these students may be more inclined to use these platforms and to contribute to them. In order to be more inclusive and to collect results from students at different stages of their degree, we sent an invitation (as described below) to three cohorts:

- Cohort 1: Students in 3rd year or higher of an undergraduate Computer Science (BSc) or Software Engineering (BSEng) degree. This cohort spends most of their time working on courses.
- Cohort 2: Graduate students taking a Master’s (MSc) or Doctoral (PhD) degree in Computer Science. This cohort devotes time toward research and graduate projects.
- Cohort 3: Students on a paid co-op/internship or that recently graduated and are working for a software development company.

The invite was sent using mailing lists to all of the students registered in 3rd year and above of computer science and software engineering, graduate students in computer science, and new graduates from software engineering and computer science. We received 45 responses, of which we conducted interviews with 20 students based on their availability, job status, registration status as a full-time student, and willingness to share their in-depth experiences with the platforms. We recognize the response rate was low (as we emailed many students); however, as this was an exploratory study and many students or new graduates are busy, this low response rate was not surprising to us. Table I shows the demographics of the participants in our interviews.

A. Semi-Structured Interview Design

The interview questions were a mix of open-ended and closed-ended questions that centered on identifying the participant’s interactions with the developer social platforms under investigation, such as asking them about their recent use of or contributions to GitHub, or their recent participation on Stack Overflow. There were 20 questions related to GitHub participation and 12 questions related to Stack Overflow. The interviews were held via Skype and were voice recorded (with permission). Each interview lasted approximately one hour and was conducted by a member of our research team who was not in a position of power over the participants. The selected participants were asked to complete consent forms that were sent to them via email before the interview, and the interview transcripts were all kept anonymous. The interview study was approved by the Ethics department of our university.

B. Analysis of Semi-Structured Interviews

After conducting the interviews, we transcribed the recordings, and then read (in multiple rounds) the answers gathered in order to conduct a ‘thematic analysis’ [14] [15] of the transcribed interviews. Two members of our team independently coded the interviews and later held agreement sessions to note any overlap of the assigned codes. Most of the analysis was conducted using Atlas.ti5.

1) Generating codes: We began to identify similarities among the responses of various participants. The two researchers analyzed the interview transcriptions using the following steps:

- We processed the interview transcriptions by highlighting statements that could indicate a potential barrier or benefit. As an example of a barrier, a participant answered, “Yeah I would totally be more comfortable in a private repository, as in a huge open-source project there are so many people and contributors it is very hard to get your contribution here as there are so many people working on their things.” This shows a hesitation to participate when the projects are public facing.

---

4Supplementary material: https://anonymous.4open.science/r/e49e11c4-fa55-4b07-a29a-823655dcd4b9/

5A qualitative data analysis tool https://atlasti.com/
We highlighted individual responses that looked similar. For example, “If fixing the issues is not taking much of my time, I will fix them, otherwise I will not” and “The projects were so big that if I started fixing one or two issues then it would take several days, investing that much time is not possible for me at the moment” are similar to each other as they both reflect the time-consuming nature of the process of raising issues in open-source projects, and the participants do not contribute for this reason. So we marked them under the same code.

We iterated to refine our codes, using features in Atlas.ti to compare our codes and to notice differences. We shared our codes at each iteration. For the most part, we easily came to agreement as we were closely aligned in the codes we developed, differing slightly in the terminology we used initially. At this step, we also added examples to our agreed codes to capture our understanding of the codes for future rounds.

2) Clustering codes and identifying themes: We clustered the codes to reveal underlying patterns and to develop insights. We used a ‘bottom-up card sorting technique’ [10] for the barriers and benefits we identified. The clustering step was conducted separately for codes associated with GitHub and Stack Overflow. For clustering, we followed these steps:

- We observed similarities in some of the codes. For example, “I do not trust strangers' algorithms directly to use in my projects” and “I do not trust strangers to work on my projects” reflect some underlying trust issues, thus we placed them under the cluster associated with trust issues.
- To ensure our open card sorting was consistent across our research team, we conducted the process of clustering the data in a collaborative mode on Coggle.it, and we held regular agreement sessions that helped us come up with consistent clusters. We also had an additional researcher regularly review our codes and clusters.

The result of our clustering process is shown in Figure 2 and Figure 3 for the barriers we identified for GitHub and Stack Overflow.

### IV. FINDINGS

We present the findings for the three research questions we posed. We discuss the barriers (RQ2) in detail as this was the focus of our study to help derive recommendations on how to address barriers. We also confirm that students do experience the benefits we anticipated before our study.

A. **RQ1: How Students Use GitHub and Stack Overflow**

The majority of study participants used GitHub and Stack Overflow at some point during their degrees. As shown in Table II, 11 of 20 people we interviewed had an account on both GitHub or Stack Overflow. Of these, seven participants shared that they are active on both platforms. The table also shows if the participants indicated they “contributed” to public content hosted on these platforms for purposes outside of their school work or co-op jobs.

For GitHub, participants either contributed to private repositories as part of their co-op job or course work, or collaborated with teams of other students for coursework. For Stack Overflow, all participants reported using the platform to find help with their code and enhance their coding logic. However, the degree of interaction with GitHub and Stack Overflow varied across the participants. We also asked them if they “contributed” content to projects hosted on GitHub or contributed to questions/answers on Stack Overflow.

While most interactions consisted of school- or job-related work, some participants indicated contributions outside of these commitments to either open-source projects on GitHub or to asking or answering questions on Stack Overflow.

### TABLE II

**Participant Interactions with GitHub and Stack Overflow**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>P No.</th>
<th>Account Created</th>
<th>Contributed to Open Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>P2</td>
<td>GitHub</td>
<td>GitHub</td>
</tr>
<tr>
<td></td>
<td>P10</td>
<td>GitHub</td>
<td>GitHub</td>
</tr>
<tr>
<td></td>
<td>P12</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P17</td>
<td>Both</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>P20</td>
<td>GitHub</td>
<td>GitHub</td>
</tr>
<tr>
<td>C2</td>
<td>P1</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P5</td>
<td>Both</td>
<td>GitHub</td>
</tr>
<tr>
<td></td>
<td>P6</td>
<td>Both</td>
<td>GitHub</td>
</tr>
<tr>
<td></td>
<td>P7</td>
<td>GitHub</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>P8</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P9</td>
<td>GitHub</td>
<td>GitHub</td>
</tr>
<tr>
<td></td>
<td>P11</td>
<td>GitHub</td>
<td>GitHub</td>
</tr>
<tr>
<td></td>
<td>P19</td>
<td>Stack Overflow</td>
<td>Stack Overflow</td>
</tr>
<tr>
<td>C3</td>
<td>P13</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>P14</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P15</td>
<td>GitHub</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>P16</td>
<td>Both</td>
<td>Both</td>
</tr>
<tr>
<td></td>
<td>P18</td>
<td>Both</td>
<td>Stack Overflow</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>GitHub: 7</td>
<td>GitHub: 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stack Overflow: 1</td>
<td>Stack Overflow: 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Both: 11</td>
<td>Both: 7</td>
</tr>
</tbody>
</table>

B. **RQ2: Barriers Students Experienced with GitHub and Stack Overflow**

Research has shown that every public-facing community [16] has barriers to entry and participation. To help us answer RQ2, our interviews highlighted the specific barriers our participants experienced on GitHub and Stack Overflow. These barriers to participation were faced by all the students.

1) **Barriers observed with GitHub:** Table III contains barriers, the codes associated with the barriers, and some quotes corresponding to each barrier from the interviews associated with GitHub. Below, we provide explanation for each barrier that we identified.

**Lack of self-confidence.** Some participants believed that their programming skills were not as good as other people who performed similar activities on GitHub. Some participants indicated an inferiority complex with respect to the ability of others to understand and share their code. This barrier was also
<table>
<thead>
<tr>
<th>Barrier</th>
<th>Code</th>
<th>Example of Participant Responses from Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of self-confidence</td>
<td>I am inexperienced</td>
<td>“There are online communities dedicated to the code bases. I know it is being taken care of by people who are more experienced than me.” (P2)</td>
</tr>
<tr>
<td></td>
<td>I’m not sure if it is a bug or a gap in my understanding</td>
<td>“I never found a flaw I was sure about. I am never sure if it is bug or just my perspective.” (P20)</td>
</tr>
<tr>
<td></td>
<td>My projects are not good enough</td>
<td>“I am sure there is a better version of my code somewhere there, someone else can do better than me.” (P15)</td>
</tr>
<tr>
<td></td>
<td>I will not get attention in any big projects</td>
<td>“I am more comfortable collaborating privately as then my contribution matters. It is hard to get your contribution seen in bigger and more open projects.” (P3)</td>
</tr>
<tr>
<td></td>
<td>My projects are too small to get attention from people</td>
<td>“I posted my class work once but never got any stars or other activity so I figured I need to post something big and important.” (P5)</td>
</tr>
<tr>
<td></td>
<td>If my projects are good enough I post them online</td>
<td>“None of my existing work is good enough to be open.” (P15)</td>
</tr>
<tr>
<td>Desire to remain in comfort zone</td>
<td>I’m more comfortable collaborating within my known circle</td>
<td>“I don’t like to interact much, I hesitate a lot and it is in my nature. The thought of having to interact with so many people at the same time scares me.” (P2)</td>
</tr>
<tr>
<td></td>
<td>I have never sought collaboration online</td>
<td>“I don’t remember going to Twitter and publishing ‘I need collaborators for this project’ I usually ask people I know.” (P1)</td>
</tr>
<tr>
<td></td>
<td>I prefer to contribute to repositories of people I know</td>
<td>“It is hard to fill my knowledge gap for repositories of strangers if their documentation is not complete. It’s not like I can just ask them like I would if I knew them.” (P4)</td>
</tr>
<tr>
<td>Time related barriers</td>
<td>I do not have time to create PRs so I raise issues</td>
<td>“I participate in open-source but I don’t have much time so I usually report issues, and sometimes I would create a branch and start working on it and maybe create a PR. Then I stop following it.” (P1)</td>
</tr>
<tr>
<td></td>
<td>My school work keeps me occupied</td>
<td>“Usually during an academic term I have a heavy load of mid-terms and assignments and I cannot do anything else.” (P13)</td>
</tr>
<tr>
<td></td>
<td>I raise issues if it is not taking much of my time</td>
<td>“If the issue is small enough for me to explain, only then I raise it.” (P1)</td>
</tr>
<tr>
<td></td>
<td>I do not have time to look for projects to collaborate</td>
<td>“I’m busy with my research and collaborating in other people’s repositories requires time and effort that I cannot put in at the moment. Maybe after I graduate I might.” (P18)</td>
</tr>
<tr>
<td></td>
<td>I do not have time to raise issues</td>
<td>“I am always running on a tight schedule, I have never bothered to even point out any issue. After my Master’s thesis is done I will have less stress and more active time there.” (P6)</td>
</tr>
<tr>
<td>Expecting immediate benefits from participation</td>
<td>I do not see any direct benefit to participation in open-source</td>
<td>“It is not giving any immediate value in return as compared to some other development activities like competitive coding. It is not beneficial to me in short or long run.” (P9)</td>
</tr>
<tr>
<td></td>
<td>I am more interested in participation when my work gets attention</td>
<td>“When you don’t get recognition from other people, it gets boring. I have released many projects and when they don’t get any attention, I get bored and stop working on it.” (P11)</td>
</tr>
<tr>
<td>Lack of interest in participa-</td>
<td>Open participation is not my hobby</td>
<td>“Contributing to open-source has always been an extra step towards my career. It is not something I would do for fun as it is not one of my main hobbies.” (P9)</td>
</tr>
<tr>
<td>tion</td>
<td>I prefer to work on my own code and have no interest in contributing to the code of others</td>
<td>“I have never felt like working on other people’s projects on GitHub. I’d rather make my own projects and work on that.” (P11)</td>
</tr>
<tr>
<td>Not willing to make an effort</td>
<td>I participate in only those open-source projects that I am personally using</td>
<td>“I only fix bugs on projects if it affects me in some way like I need to use the code for my own project.” (P14)</td>
</tr>
<tr>
<td></td>
<td>I am not willing to recruit contributors for my projects</td>
<td>“Online collaborations require a certain way to communicate like reviewing PR, raising and resolving issues which is extra work that I have failed to keep up with in the past.” (P3)</td>
</tr>
<tr>
<td></td>
<td>If creating a PR is a lot of work, I just raise issues</td>
<td>“Making a PR means following community guidelines and then writing code, waiting for comments to edit it again. I’d rather just report bugs.” (P1)</td>
</tr>
<tr>
<td>Trust issues</td>
<td>I do not trust strangers to work on my projects</td>
<td>“I do not know if a stranger will be able to understand my project and there are chances that my work can be misinterpreted.” (P6)</td>
</tr>
<tr>
<td></td>
<td>I do not trust strangers’ algorithms to use in my projects</td>
<td>“With algorithms you have to understand math behind it and the theory behind it. I don’t know what kind of knowledge they have in my project domain so it is hard for me to open my projects for contribution.” (P10)</td>
</tr>
</tbody>
</table>

85
prevalent in other communities when we studied the literature about lurkers [6]. Lurkers who feared receiving judgement from the community about their work did not trust the quality of their own work, and consequently were uncomfortable and reluctant to post online.

**Desire to remain in comfort zone.** Some participants indicated being more comfortable when interacting with people within their known social circles. They noted that they avoid interacting with people outside their circles as they feel shy with outsiders. This may be due to the perception that within a known circle the participants can “be wrong”, a feeling they are uncomfortable sharing with outsiders. Some participants also displayed introversion as they were not comfortable interacting with outsiders even in the form of comments.

**Time-related barriers.** Some participants indicated an interest in participating in open-source communities but could not because of other commitments. Such participants were occupied with their school work. Some participants with more experience, such as the MSc and PhD students, indicated that due to time constraints they were unable to finish contributing to projects but they raise issues and add comments instead, which is less time consuming than creating a pull request. It is noteworthy that this does not appear to arise from a lack of interest, but rather that participants worried they lacked the time to ensure their code contributions were of sufficient quality.

**Expecting immediate benefits from participation.** Actively participating and collaborating with people can help grow one’s professional and technological skills. Through participation in online communities, one can enhance their coding skills, contribute across a diversity of projects, develop collaborations with others regardless of geographical boundaries, have an opportunity to learn from people with different expertise, and enhance collaboration skills. However, some participants expected an immediate return on investment from their contributions and failed to recognize the long-term benefits. This was also observed in a study conducted on Microsoft online communities [8]. The participants in the Microsoft study mainly perceived incentives to include situations where an employee completes tasks and in return is remunerated, or where a student works on class projects and in return receives grades.

**Lack of interest in participating.** Some students indicated a lack of interest in participating in open-source projects hosted on sites like GitHub. If they had free time, they would rather participate in other activities, such as hobbies. They see
the act of working on projects only as an activity that would add to their professional profiles and aid them in getting jobs. For the most part, participants who experienced this barrier did not intend to contribute to open-source projects.

**Not willing to make an effort.** The participants whose responses indicated this barrier lie somewhere between ‘time-related barriers’ and ‘lack of interest in participating’: they want to participate but are not interested in taking the extra steps needed to contribute. Some noted they would only contribute to software projects they use personally. These participants would occasionally make pull requests or leave comments on the repositories they already access as it required minimal extra effort. They did not always share their projects online and they preferred to collaborate with people they could meet in person as they felt communicating through issues and comments on repositories was too much work.

**Trust issues.** Some of the participants in our study do not share their code on these platforms (particularly algorithmic issues they encounter) as they do not trust the expertise of people they have never met. These students prefer discussing and developing their code with peers they trust. They also feel that the problems they are facing require qualifications, such as coding proficiency, that are hard to verify in people they have never collaborated with before.

2) **Barriers observed with Stack Overflow.** Some of the themes that emerged in this section are similar to those we observed with GitHub, however, there is a difference in the nature of the barriers in each of these communities. The barriers we identified on Stack Overflow, the codes associated with each of these barriers, and some quotes corresponding to each barrier are presented in Table IV.

**Stack Overflow is not an obvious choice for seeking help.** When some of our interviewees discover issues in their code, they type their problems into search engines and use whatever links appear on the first page of the search results. If a link to Stack Overflow is in the search results, then they will try this first. For these participants, Stack Overflow is one of the online sources that might contain an answer to their question at that moment, but not a place to routinely search for answers and make regular contributions.

**Looking for a quick answer.** Some participants indicated that they visit Stack Overflow to look for solutions to the problems they encounter in their code, but they never post anything there. They feel that it takes too long to post questions and then wait for responses from other members of the community. These participants prefer to read various resources and determine the solution themselves.

**Contributing is too much work.** Some interviewees were interested in contributing on Stack Overflow but felt that doing so was too much work due to community standards and the formatting required to post questions, even responses to questions need to follow formatting rules.

**Desire to remain in their comfort zone.** This barrier is the same as the GitHub barrier ‘Desire to remain in comfort zone’. Some students mentioned they are too shy to participate in such communities as they prefer not to share their questions and answers outside their known circle. Some try to avoid interactions with people they do not know.

**Poor expectations from the Stack Overflow community.** Some participants felt their issues tended to be specific to their projects, and that if they post questions on Stack Overflow, they would not make much sense to anyone else. This misalignment prevents these participants from posting questions at all. Other participants shared a hesitation toward seeking help with their problems as they believed they should do more research themselves. They felt that their questions may look *absurd* to the community. Hence, they avoided relying on Stack Overflow to get help with their code or even sharing a response they might have to another user’s question.

**Lack of self-confidence.** Some participants felt they were not skilled enough to answer others’ questions, and as a result, they refrained from participating. A similar theme appeared in response to contributions on GitHub.

**Unwilling to help others.** This barrier is similar to one of the reported GitHub barriers. Some students do not participate as they have no desire to contribute to Stack Overflow and give back to the community; they believe that reading and using answers is enough for them. They only see the responses they receive as a search result and not as a place where they could potentially interact with others by answering questions and/or asking questions of their own. This mindset is explained in another study that identifies differences between active users and lurkers [8].

C. **RQ3: Benefits to Participation**

We pulled examples from our interviews that show the benefits our participants experienced from their active participation on these two platforms. We then categorized the examples into codes which we later organized into clusters that represent a certain type of benefit. We describe these benefits below, noting which platforms they apply to (GH, SO, or both). Quotes from the interviews are included as examples, and we note which interviewee is responsible for each quote with (Px).

1) **Learning from the community:** Developer social platforms offer a space for people in the development field to share their code, collaborate with each other, and get help with their problems.

**Gaining experience in collaborative development (GH)**
Participants shared that as they contribute to repositories, they gain experience working with a team. This helped them grow the soft and technical skills needed for jobs where they were expected to collaborate in work. Some statements from the interviews that point to this benefit include: “People read each other’s code and one of the benefits is that you can actually learn how to code by reading code. Open-source projects bring an opportunity to read code from different people who use different techniques and learn from them. There’ve been times when I’ve tried to mimic other people’s style of coding for my own project.” (P3)

**I get to learn from other people’s work (GH, SO)** Some participants shared that whenever they were looking for ways to accomplish a certain task, they looked at other people’s code
on these platforms. One of the interviewees noted that “there is always an opportunity to learn a lot from other people, sometimes you see varied expertise, so you get chances to learn.” (P19)

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Code</th>
<th>Example of Participant Responses from Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stack Overflow is not an obvious choice for seeking help</td>
<td>I type my problem into a search engine</td>
<td>“I only go to SO if I have a very specific request that I cannot solve after say 5-7 days.” (P9)</td>
</tr>
<tr>
<td>Looking for a quick answer</td>
<td>Long waiting time to get answers on Stack Overflow</td>
<td>“I am looking for quick help and I know SO will take time to give an answer to any question I post there.” (P7)</td>
</tr>
<tr>
<td>Contributing is too much work</td>
<td>I do not ask questions as the task of asking questions is too much work</td>
<td>“I noticed that there was a kind of standard way of asking questions. If it is not that way then you get down-voted very fast, there have been times when my questions got closed because of not being appropriate. Over time I realized that asking a question is not enough, you also need to say what you have already tried and you need to give references and parts of the source code and different links to show people what you tried.” (P19)</td>
</tr>
<tr>
<td>Desire to remain in their comfort zone</td>
<td>I am comfortable asking questions from people I know</td>
<td>“I am afraid of the response I will get if I post a question. I was once being told that my question is absurd when I posted it on Stack Overflow so I stopped posting.” (P8)</td>
</tr>
<tr>
<td>Poor expectations from the Stack Overflow community</td>
<td>My problems are specific to my use case</td>
<td>“I don’t care enough to share my improvement of an answer.” (P11)</td>
</tr>
<tr>
<td>Lack of self-confidence</td>
<td>I’m inexperienced</td>
<td>“I feel shy while asking questions.” (P20)</td>
</tr>
<tr>
<td>My contribution is not that special</td>
<td>My answers are specific to my use case</td>
<td>“I don’t know, maybe I do not want to deal with interaction that follows the asking of a question there.” (P11)</td>
</tr>
<tr>
<td>Unwilling to help others</td>
<td>I don’t care about helping others with my answers</td>
<td>“I don’t care enough to share my improvement of an answer.” (P11)</td>
</tr>
</tbody>
</table>

There is always support available (GH, SO) Some participants stated they enjoy the development support that is always available whenever they get stuck on a problem. When asked about the benefits they experienced while developing projects collaboratively, one of our interviewees responded, “Huge community support is available readily without geographical boundaries, no centralization of power, sense of equality.” (P14)

2) Altruism: Some of the more experienced students that participated in our interviews stated they like to share their knowledge with everyone.

I like helping fellow coders by answering their questions (SO) Some participants stated they like to answer questions as it is their way of giving back to the community they use for help. They also felt that responding to questions takes little effort compared with the benefits an answer might provide to someone who needs help. For example: “If I know the solution, why not post it? I know how frustrating it is to find answers.” (P1)

I like it when others find my code useful (GH) We asked the participants who share public repositories of their coding projects how it feels getting recognized by other people. One noted, “There are two projects that I published recently, I noticed that students are using for their class assignments, they email me questions and report errors, I feel that it is cool that some people find it useful.” (P1)

I report bugs on projects as it makes the product better for everyone (GH) This was noted by participants who use open-source communities for their own work. They appear to not only use the code but also care about the code base as a whole: they usually report any bugs they spot so that the project can be improved. As one interviewee mentioned, “There is a framework that is relatively new, I’m using that and it is very good and useful. Every time I come across a bug, I report that as I feel it will make the product better for everyone.” (P4)

3) Increased confidence from displaying their skills online: Some participants experienced increased confidence when they shared their code online.

I like to improve my reputation (SO) Some participants liked to maintain their reputation on their Stack Overflow profile by collecting badges. They achieve these badges by actively participating on the site: they up-vote and down-vote answers, respond to questions, and help moderate the site. These activities make them feel confident in their skills and are referenced on their resumes to show prospective employers. As one interviewee noted, “I try to post an answer next to other answers telling them that well, other answers are correct as well but I came up with this one maybe you wanna take a look at this. It might receive up-votes that increases my reputation.” (P4)

I use my online repository as a portfolio (GH) Many participants reported using their GitHub account as an online development portfolio, keeping their portfolio up-to-date with their recent work and referencing this work on their resume. They actively browse other public repositories and contribute to them (with PRs or comments) to show they are regular members of the software development community. Some ex-
amplified by this include: “MyGitHub has a lot of projectsI worked on, and I mention my GitHub on my resume” (P2), “Once a company asked me for myGitHub page, and they didn’t select me as I didn’t have much going on there. Since then I make sure to stay active there.” (P15)

I feel good when people use my code (GH, SO) Some participants reported that they feel good when they are recognised on developer social platforms. This feeling affirmed their skills and helped motivate them to work harder. One of our interviewees mentioned, “If you get the attention, recognition, it is always nice. Makes me feel good about time I spend developing the project.” (P1)

V. DISCUSSION OF OUR FINDINGS AND RECOMMENDATIONS

We found from our study that all of the students we interviewed were aware of GitHub and Stack Overflow and used these platforms in some way, either in their courses or professionally. Most of our participants indicated they have accounts on these sites: almost all have a GitHub account, while over half have Stack Overflow accounts. However, not all students who have an account contribute, and not all contributions are evenly weighted. In future work, we would like to question how and to what degree students contribute; however, this was not a focus of the present study.

In terms of benefits, we found that students experience similar benefits to other developers as reported in the previous literature (after all, students are junior developers). However, some benefits may be of particular importance to students. As students have a greater need to rapidly learn new material, they can benefit from connecting with others, and the visibility of their contributions can be used to build their portfolios and boost their confidence.

The main focus of our study was to uncover barriers that students may face and to understand if these barriers were a significant impediment (or if they faced new barriers). In Section 2, we summarized the barriers from previous work. In our study, we revealed more nuanced barriers to students and in Figure 4 we map to the barriers previously reported. However, we see that lack of confidence starts early in student use of these tools and that building confidence quickly could help students benefit from contributing more to the open knowledge hosted on these platforms. We also learned that students often had low expectations about using Stack Overflow, which may indicate that more training could help students use this tool more effectively. Finally, many of the students lacked motivation to contribute and poor awareness of the benefits they may gain may be an issue in this regard.

Although this is an exploratory study, we suggest recommendations for students, educators and platform designers.

Recommendations for Students

- We recommend students invest time to learn how to use these platforms and to contribute more so they can build portfolios to improve future job prospects.
- We recommend students contribute to active and important open-source projects so that they can enhance their collaboration skills working with larger teams.
- We encourage students to ask questions and to post answers in Stack Overflow, as learning how to ask and answer questions are important skills, and that these platforms provide training in this regard.

Recommendations for Educators

- Peer parity has a positive effect on people’s motivation to participate in these platforms [7]. Educators could increase awareness that other students and even more senior developers also face challenges such as poor self-confidence in using social coding platforms, and that overcoming this barrier is worthwhile.
- More emphasis could be placed on using GitHub and Stack Overflow earlier in educational courses, with encouragement and even specific assignments on how to use them, and projects that involve making actual contributions that are later adopted in open-source projects or accepted as good questions or answers in Stack Overflow.

Recommendations for Platform Designers

- We suggest platform designers pay more attention to the barriers that students may face and actively involve them in the design and evaluation of their platforms.
- Offer greater support for learning how to use these tools and provide students with support when they get stuck.

VI. LIMITATIONS

Our exploratory study has limitations. We interviewed a relatively small sample of students. We also opted to do interviews over a larger survey so we could more deeply explore student experiences and barriers. We intentionally do not report counts or numbers as this study is intended to be qualitative in nature. Our insights could be used to design a future survey to quantify some of our findings.

The semi-structured nature of our questions has limitations as some bias may have occurred in the follow-up questions we asked. There may be other possible biases such as the order of questions asked and the specific wordings. However, we ran pilots to try to mitigate some of these biases. Only two coders were involved in the coding of the interviews, but a third reviewer inspected the codes and themes throughout the process, watching for potential biases and inconsistencies.

We recognize that the findings of our study may not generalize to other students or to other platforms.

During our recruitment phase, we explicitly mentioned our interest in understanding how students use GitHub and Stack Overflow, and shared our goal to identify barriers and benefits they may experience. This information may have influenced who participated in our study: in particular, we expected that students that do not use these tools much would not have participated. Our findings are also limited by the use of a single
method. However, we did find that the themes that emerged were consistent with the existing literature in terms of the overarching barriers and benefits we identified.

VII. FUTURE WORK AND CONCLUDING REMARKS

We would like to investigate how students acquire (or don’t) the motivation and skills to use these platforms over the course of their educational programs and how using them may be of benefit. In this small study, we were not able to investigate barriers that some genders or other demographics may face, but doing so is important future work. In closing, we propose that these open platforms could play a much greater role in educating future developers and that more research will help make this possible.

REFERENCES


