Social Barriers Faced by Newcomers Placing Their First Contribution in Open Source Software Projects

Igor Steinmacher
Department of Computing
Federal University of Technology – Paraná
igorfs@utfpr.edu.br

Tayana Uchôa Conte
Institute of Computing
Federal University of Amazonas
tayana@icomp.ufam.edu.br

David F. Redmiles
Department of Informatics
University of California, Irvine
redmiles@ics.uci.edu

Marco Aurélio Gerosa
Institute of Mathematics and Statistics
University of São Paulo
gerosa@ime.usp.br

ABSTRACT
Newcomers' seamless onboarding is important for online communities that depend upon leveraging the contribution of outsiders. Previous studies investigated aspects of the joining process and motivation in open collaboration communities, but few have focused on identifying and understanding the critical barriers newcomers face when placing their first contribution, a period that frequently leads to dropout. This is important for Open Source Software (OSS) projects, which receive contributions from many one-time contributors. Focusing on OSS, our study qualitatively analyzed social barriers that hindered newcomers' first contributions. We defined a conceptual model composed of 58 barriers including 13 social barriers. The barriers were identified from a qualitative data analysis considering different sources: a systematic literature review; open question responses gathered from OSS projects' contributors; students contributing to OSS projects; and semi-structured interviews with 36 developers from 14 different projects. This paper focuses on social barriers and its contributions include gathering empirical evidence of the barriers faced by newcomers, organizing and better understanding these barriers, surveying the literature from the perspective of the barriers, and identifying new potential research streams.

Author Keywords
Onboarding; barriers; joining; open source software; newcomers; social barriers; qualitative study; open collaboration; online communities

ACM Classification Keywords
D.2.9 [Management]: Programming teams, K.4.3 [Organizational Impacts]

General Terms
Human Factors

INTRODUCTION
Open collaboration communities leverage and depend on new contributors [23]. According to Qureshi and Fang [63], motivating, engaging, and retaining new contributors promotes a sustainable community. Kraut et al. [47] claim that newcomers are a source of innovation, new ideas, and work procedures.

According to Forte and Lampe [23], open collaboration relies on an online environment that presents a low barrier to entry. However, during the onboarding period, newcomers may be susceptible to several barriers such as expectation breakdowns, reception problems, setup misconfiguration, and learning curve. Each of these may have varying levels of importance to and impact on the overall joining process.

To better support newcomers' onboarding, barriers must be identified and understood. Focusing on open source software (OSS) development communities,, in which newcomers are usually left to learn on their own [67], the qualitative study described in this paper aimed to identify, understand, and organize the barriers faced by newcomers. Since OSS communities require developers with specific skills, and delivering a task to an OSS project is usually a long, multi-step process, some newcomers may lose motivation and even give up contributing if there are too many barriers to overcome during this process. As Karl Fogel [22] states, *"if a project doesn't make a good first impression", newcomers may wait a long time before giving it a second chance."*

Previous research related to newcomers' joining process examined the dynamics driving OSS contributors, mostly focusing on the motivations for becoming a member, roadmaps to becoming a core developer, or indicators of potential long-term commitment [30, 37, 68, 86, 88]. An understudied aspect of the OSS joining process is what happens during the period after a newcomer decides to participate and before their first code contribution is accepted and included in the shared repository. This period is particularly relevant to OSS projects, as many newcomers do not want to join or remain at the project, only to post a single contribution (e.g., a bug correction or a new feature). What happens in this period affects, for example, students in computer courses whose assignments include OSS project contribution, and professional developers who find a bug or wish to customize a particular software product. With a more in-depth understanding of the barriers, researchers and community can invest their efforts in building or improving tools and processes, ultimately gaining more contributions, such as '*drive-by commits*' [60], which are small changes made by developers who are only casually or briefly interested in a project and do not intend to have a prolonged engagement.

In this study, we qualitatively analyzed four different sources: (i) a systematic literature review (SLR) aimed at identifying the literature's empirically evidenced barriers; (ii) feedback from 9 students after they contributed to OSS projects; (iii) 24 open-question responses from 9 OSS

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CSCW '15, February 14–18, 2015, Vancouver, BC, Canada.
Copyright 2015 ACM xxx-x-xxxx-xxxx-x/xx/xx...$15.00.

projects; and (iv) semi-structured interviews with 36 developers from 14 different projects, including newcomers, dropouts, and experienced members. To analyze this data, we used Grounded Theory procedures [79]. We found 58 barriers, which we organized into 6 main categories comprising technical and social barriers. The 45 technical barriers primarily relate to running or developing environments' domain technologies and concepts, documentation, and source-code; the remaining 13 social barriers, which were similar across different projects and are the focus of this paper, relate to issues that arose from the relationship between the newcomers and the established community.

The contributions of this paper include: (i) providing detailed empirical evidence of the social barriers faced by newcomers to OSS projects, which was drawn from data obtained from 69 practitioners contributing to 19 different projects; (ii) relating this empirical evidence to results from 21 studies identified in a systematic literature review, as well as to related studies on open collaboration communities; (iii) organizing the barriers into a single model; and (iv) providing a discussion of the social barriers (and possible mitigation strategies) in the context of CSCW and related literature. We hope that OSS communities and researchers will take advantage of this paper to better understand the barriers in their context and design strategies to deal with them.

This paper is organized as follows: related work; data collection and analysis method, resulting conceptual model; some discussion; and conclusions.

RELATED WORK

Many barriers influence open communities newcomer behavior and permanence. In the following subsection, we present significant findings related to social barriers in several online communities, such as Wikipedia, Reddit, MovieLens, Facebook, and Q&A communities. Afterwards, we delve into the literature on OSS projects, this paper's focus.

Newcomers in Various Open Collaboration Communities

Perhaps due to its new contributors decline [27], Wikipedia has been the subject of several studies, including Zhu et al.'s [90], which identified that received feedback influences newcomers' engagement. In addition, Halfaker et al. [28] and Suh et al. [80] found that newcomers are negatively impacted by how many edit reverts they suffer. From a more positive angle, Choi et al. [13] found that welcome messages, technical assistance, and constructive criticism over time retarded the natural decline in newcomers' editing. Similarly, Faulkner et al. [21] found that modifying first time warnings prompted additional newcomer contribution. Offering a sentiment that seems to undergird each of these studies' implications, Halfaker et al. [27] concluded that *"Wikipedia has changed from the encyclopedia that anyone can edit to the encyclopedia that anyone who understands the norms, socializes himself or herself, dodges the impersonal wall of semi-automated rejection, and still wants to voluntarily contribute his or her time and energy can edit."*

Recent studies have also focused on other online communities, including social media, support, and discussion networks. Burke et al. [8] analyzed Facebook newcomers and found that when newcomers learn by example and receive proper feedback, one can more easily predict their future contributions. Wang et al. [84] investigated an online health support group, finding that participants exposed to both emotional and informational support demonstrated lower dropout rates. Lampe and Johnston [52] analysis of Slashdot newcomers revealed that new, un-moderated commenters were less likely to post a second comment. Similar conclusions were drawn by Joyce and Kraut [41], who analyzed Usenet group newcomers.

Newcomers to Open Source Software Projects

Developer motivation and project attractiveness are well-explored OSS research topics [55, 66, 71, 86], and many studies also investigate how newcomers join projects and become core project members [18, 19, 50, 54, 57]. Motivation and project attractiveness entice outsiders and bring them into the joining process. However, contribution barriers can impede developers from completing the joining process. These barriers comprise technical and social impediments, including steep learning curve, lack of community support, and difficulties finding out how to start. Furthermore, these barriers influence both developers who are willing to make a single contribution and those who are willing to progress and become a project member. Given their importance, remarkably little is known about these barriers.

Many papers in the literature focus on the joining process by defining stages and activities in the path to becoming core members or long-term contributors. For instance, the Onion Model [57, 86] was presented as a general layered structure to organize OSS project member roles as well as the process a developer needs to follow in order to contribute. Jergensen et al. [37] studied whether this model still held true in large project ecosystems, finding little evidence that individuals migrate from the edges of a project to the core through a gradual socialization process. Herraiz et al. [32] found that at least 52.5% of developers commit before sending a message to the mailing list, thereby contradicting the onion model. They also found two groups with clearly different joining patterns: volunteer and hired developers. Whereas volunteers tend to follow a systematic joining process, hired developers usually experience a quick integration. In addition to the Onion Model, other studies mapping the OSS project joining process include Von Krogh et al. [50], which proposed the concept of a joining script. Ducheneaut [18] similarly analyzed mailing list archives, offering an in-depth look at a successful newcomer's socialization history. Based on this individual, the author identified a set of socialization activities that contributed to his success.

Other parts of the literature focus on the forces of motivation and attractiveness that drive newcomers toward projects. Lakhani and Wolf [51], for example, found that extrinsic benefits (e.g.; better jobs, career advancement) primarily motivate new contributors, together with enjoyment, challenges derived from writing code, and improved programming skills. Hars and Ou [30] reported that internal motivation plays a role, but note that external factors, such as

building human capital and personal software solution needs, are more influential. Shah [71] distinguished between two different contributors: need-driven and hobbyists. Hertel et al. [33] identified seven factors as distinct motivational sources. Several other studies dealt with motivation in OSS [4, 16, 40, 42, 48, 59, 65]. Moreover, motivation is generally extensively studied in software engineering [2, 24, 25, 72, 87]. Attractiveness is also a force influenced by several characteristics, such as license [66], source code attributes [53, 55], and code base [12].

Regarding newcomers' retention and long-term contribution, Qureshi and Fang [63] quantitatively identified four distinct classes of newcomer retention behavior, considering their initial amount of interactions with core members and the growth of these interactions. Fang and Neufeld [19] qualitatively revealed that initial participation conditions do not effectively predict long-term participation, but that situated learning and identity-construction behaviors were positively linked to sustained participation. Other research revealed that retention is influenced by the familiarity with the project's coordination practices [68] and individual's attitude [88].

While the current literature focuses on motivation and forces that lead developers to the project's core, studies neglect those newcomers who do not envision a long-term engagement as well as those who want to place a single contribution. Counterexamples include Hannebaun et al. [29] and Steinmacher et al. [77], which explicitly focus on barriers that influence newcomers' first contribution. Some other studies proposed ways to help newcomers' first contributions. Wolff-Marting et al. [85], for example, proposed two patterns to support newcomers in overcoming contribution barriers. Also aiming to support newcomers, Cubranic et al. [14] presented Hipikat, a tool that builds a group memory comprising source code, email discussions, and bug trackers. The tool enables newcomers to request recommendations based on existent artifacts. With a similar goal in mind, Wang and Sarma [82] presented a tool to enable newcomers to identify bugs of interest and resources related to that bug, as well as to visually and interactively explore the bug's appropriate socio-technical dependencies. Canfora et al. [9] proposed and evaluated an approach to identifying and recommending mentors for open source project newcomers by mining data from mailing lists and source code versioning systems.

As opposed to the traditional focus on motivation, attractiveness, and retention in OSS projects, our study focused specifically on the initial contribution barriers. Other studies cite barriers that influence newcomers' overall experiences, however do not provide an in-depth understanding of the barriers, their relations, and their relevance in multiple projects. Thus, knowledge about contribution barriers is spread thin across the literature, and there is little exploration of the barriers faced by newcomers who ultimately stop contributing or do not become members.

Additionally, we observed that the majority of the related research on specific barrier effects in open collaboration communities (including OSS) predominantly relied on quantitative evidence. Those that qualitatively uncovered barriers [6, 18–20, 50, 81] missed research questions or objectives related to the identification of first contribution barriers faced by newcomers.

In this paper, we systematically review the OSS literature and discuss it in the context of the evidenced data from practitioners and from the literature of open collaboration communities. Our study aims to provide this neglected in-depth understanding of the initial contribution barriers and to add to the dominant quantitative-based research on newcomers' joining process.

METHOD

Our study aimed to identify and understand the barriers that hinder newcomers from posting their first contribution to an OSS project. We chose a qualitative approach [17] because this phenomena occurs in a complex, social environment, in which the context of its occurrence is important. Moreover, a qualitative view complements the existing literature, which relies mostly on quantitative evidence.

Our study relied on four different sources: a systematic literature review; feedback from students after having them attempt to contribute to OSS projects; responses to an open-ended question sent to OSS communities; and interviews with both newcomers and experienced OSS project members. Three models emerged from these studies, and we compiled these models to generate a broad model of barriers for newcomers to OSS.

Figure 1 depicts the method followed. The first data set was collected from students who contributed to OSS projects. From their feedback, we noticed that there were recurrently reported barriers. Motivated by the students' reports, we conducted a systematic literature review to identify the available research related to these barriers. From the literature review, we built a model reporting what was already published in OSS domain. To gather more understanding about the phenomena, and to check how the practice was related to the literature, we surveyed a small set of OSS developers. We analyzed the data from these practitioners alongside the data from the students to conceive a preliminary model of barriers. The results of this preliminary model were used as input for an in-depth investigation. This investigation was made by means of semi-structured interviews conducted with both experienced members and newcomers in order to identify and explore the barriers from different perspectives. By analyzing data from different sources and perspectives, we aimed for a broad understanding of the barriers.

In the following, we detail the methods used to conduct the systematic review and the practitioner interviews.

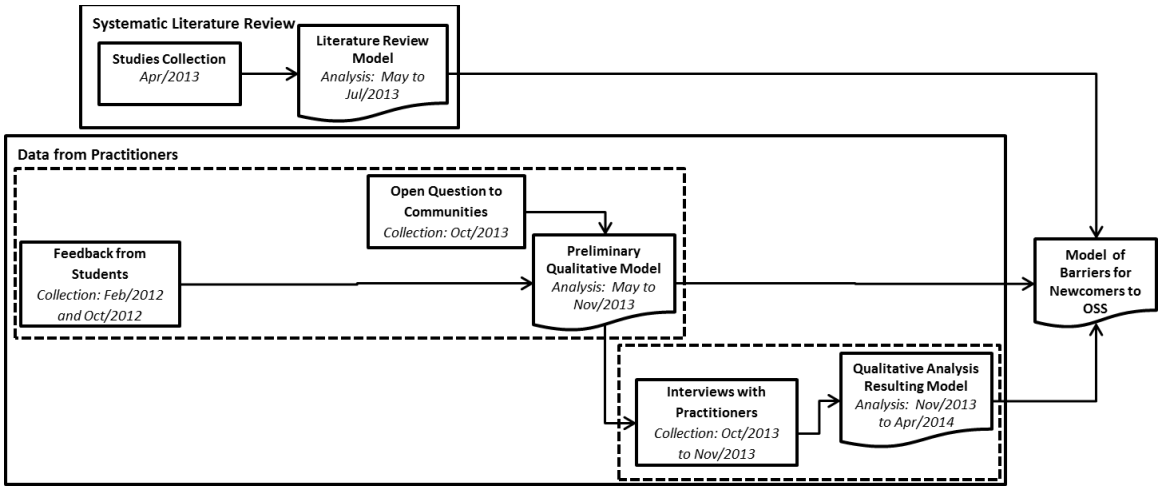

Figure 1: Research Method

Systematic Literature Review

We conducted a systematic literature review to identify the empirically evidenced and reported barriers [76]. The goal was to generate a list of the barriers newcomers encounter that can influence their first project contributions. We then aggregated the barriers evidenced by the different studies in a single model. Since we were interested in comparing the empirical data collected from the practitioners, we restricted the systematic review to the OSS domain.

This study has been undertaken as a systematic literature review (SLR) based upon guidelines established by Kitchenham et al. [43, 44]. The following question guided our SLR: *What are the barriers that influence newcomers' onboarding to OSS projects?*

After some iterative refining, the following query was used to retrieve the studies from the ACM, IEEE, Scopus, and Springer Link digital libraries. These libraries were selected because they index relevant venues for this study, support searches using Boolean expression, and provide access to complete texts. The search was performed in April 2013.

```
("OSS" OR "Open Source" OR "Free Software" OR FLOSS OR FOSS) AND (newcomer OR "joining process" OR newbie OR "new developer" OR "new member" OR "new contributor" OR novice OR beginner OR "potential participant" OR retention OR joiner OR onboarding OR "new committer")
```

We considered for selection papers that were available for download, written in English, dealt with newcomers' OSS project onboarding, presented empirical results, and were published in peer-reviewed journals or event proceedings. We excluded studies that analyzed only newcomers' motivation for joining a project. We focused on the issues newcomers face after deciding to contribute.

For each selected paper, we conducted snowball sampling [35], checking if the authors of the selected studies published other relevant studies not retrieved from the digital libraries. We checked their profiles in ACM, IEEE, DBLP, and personal homepages (when available).

After running the query on the digital libraries systems, our resulting sample comprised 291 candidate papers. For each paper, two independent researchers analyzed title, abstract, and keywords. In a consensus meeting, we agreed on 33 papers. We checked other papers published by the authors of these 33 candidate studies, finding 20 other candidate papers. After analyzing these papers, we selected nine relevant papers, resulting in 42 candidate papers. After further analysis, 21 papers were deemed relevant and were considered for data extraction. Next, we applied Grounded Theory procedures [79] (open and axial coding) to classify the barriers reported in the selected studies [76].

Data from practitioners

This subsection presents the method for the qualitative practitioner study. We gathered this data from: (i) students that contributed to OSS projects; (ii) answers to an open question sent to OSS projects; and (iii) semi-structured interviews with newcomers to and members of OSS projects.

The first source comprised feedback received from four PhD candidates and five undergraduate students who contributed to OSS projects as part of a course assignment. All the students were newcomers to the projects they were contributing to. The PhD candidates were all males, experienced developers, and 30 years old or older. The undergraduate students, including four males and one female ranging between 21 and 24 years old, were attending the last semester of the Internet Systems course and therefore were preparing to join the software development industry. The students received an assignment to contribute to an OSS project.

The students contributed to JabRef (2 graduate/1 undergraduate), LibreOffice (3 undergraduate), and Mozilla Firefox (3 graduate) projects. After the conclusion of the assignment, their feedback was collected by means of an open-ended questionnaire. The open questions enabled students to debrief and explain the problems they faced while trying to place their first code contribution. The data was collected at the end of the course (February 2012 for the graduate students and October 2012 for the undergraduate students).

The second data source comprised answers to a questionnaire sent to OSS project developers. The data was obtained from 24 answers to an open question sent to OSS project mailing

lists and forums. The messages were posted and the answers received during October 2013. We sent messages to nine projects from different domains: Apache OpenOffice, aTunes, Audacity, LibreOffice, emacs, FreePlane, jEdit, Mozilla Firefox, and OpenVPN. None of these projects delivers development frameworks or scaffolding technologies, which are generally more complex, demanding a higher degree of specific skills and knowledge. These characteristics could hide some possible barriers encountered by newcomers, since they could face complex problems related to the inherent project technology and domain.

The questionnaire delivered to the community members comprised two questions designed to profile the contributor (project and contribution time) as well as an open question: "*In your opinion, what are the main difficulties faced by newcomers when they want to start contributing to this project? (Consider technical and non-technical issues).*" We received 24 complete answers to the questionnaire, as detailed in Table 1.

Table 1. Project to which participants mainly contribute (left) and Period of contribution for questionnaire respondents (right)

Project	Count
LibreOffice	6
OpenOffice	3
aTunes	3
Mozilla Firefox	3
Audacity	2
jEdit	1
OpenVPN	1
FreePlane	1
Emacs	1
Did not inform	3

Time contributing to the project	Count
Less than 6 months	7
Between 6 months and 1 year	3
Between 1 year and 3 years	6
More than 3 years	8

The final data collection entailed semi-structured interviews with practitioners. Semi-structured interviews include a mixture of open-ended and specific questions, which are designed to elicit foreseen and unexpected information types [70]. We conducted interviews in order to complement the findings gathered from sources 1 and 2, thereby deepening and broadening our understanding of the newcomers' barriers. We recruited subjects belonging to four different groups: *Experienced Members* (project owners, managers, or developers that commit code directly to the software repository), *Successful Newcomers* (participants that started to contribute to the project less than one year before the interview), *Dropout Newcomers* (volunteers that tried to contribute to the project, but gave up), and *Onboarding Newcomers* (volunteers trying to place their first contribution).

The participants were recruited primarily through mailing list and forums from 15 different projects. In addition, we also directly invited different types of newcomers, identifying them by mining and following projects' mailing lists and issue trackers. Only adults older than 18 were eligible to participate in this study; but we made no distinction related to gender, nationality, or other identity categories. Participants were expected to have software development experience, primarily because we were interested in the barriers to contribute to a project and not to learn how to program. The interviews were conducted in English from October 2013 to March 2014.

We interviewed 36 participants from 14 different projects (Pardus, TextMate, zxing, Gephi, Hadoop, jEdit, Moodle, Integrade, Noosfero, Apache OpenOffice, cogroo, etherpad, JabRef, and LibreOffice), including 11 experienced members, 16 newcomers that succeeded, 6 dropout newcomers, and 3 newcomers that were still trying to place their first contributions.

Table 2 shows some profiling information of the students and interviewees. The participants received an ID, shown in the first column. The first character of the ID represents the profile of the participant: "S" for student, "E" for Experienced member, "N" for Successful newcomer, "D" for Dropout newcomer, and "O" for Onboarding Newcomer.

Table 2. Profile of the participants (H = Hours per week in OSS; F = First Project?; C = Country; Y = Years in the project)

ID	H	F	Project	C	Y	ID	H	F	Project	C	Y
E1	< 5	N	JabRef	FR	8	N13	10-20	N	LibreOffice	BR	1
E2	05-10	Y	Etherpad	DE	3	N14	05-10	Y	LibreOffice	BR	1
E3	10-20	N	JabRef	DE	3	N15	N/I	Y	Etherpad	FR	0
E4	05-10	N	jEdit	CA	10	N16	05-10	N	JabRef	DE	0
E5	05-10	N	LibreOffice	DE	15	D1	05-10	N	Hadoop	US	0
E6	> 20	N	LibreOffice	HU	10	D2	< 5	Y	Hadoop	IN	0
E7	> 20	N	Moodle	AU	5	D3	< 5	N	JabRef	DE	0
E8	> 20	N	Noosfero	BR	5	D4	< 5	Y	OpenOffice	BR	0
E9	> 20	N	Pardus	TR	8	D5	< 5	Y	LibreOffice	IN	0
E10	05-10	N	Cogroo	BR	5	D6	< 5	Y	OpenOffice	IN	0
E11	< 5	N	Noosfero	BR	7	O1	< 5	N	OpenOffice	IN	0
N1	< 5	Y	JabRef	DE	0	O2	10-20	Y	LibreOffice	CN	0
N2	< 5	Y	Gephi	BR	0	O3	< 5	Y	OpenOffice	GR	0
N3	05-10	Y	Gephi	IN	1	S1	N/I	N	Mozilla	BR	0
N4	05-10	Y	Moodle	IN	0	S2	N/I	Y	LibreOffice	BR	0
N5	< 5	Y	JabRef	DE	0	S3	N/I	Y	LibreOffice	BR	0
N6	< 5	Y	jEdit	US	0	S4	N/I	Y	Firefox	BR	0
N7	< 5	Y	TextMate	US	0	S5	N/I	Y	Jabref	BR	0
N8	> 20	Y	Zxing	GR	0	S6	N/I	Y	Firefox	BR	0
N9	< 5	Y	Cogroo	BR	0	S7	N/I	N	Jabref	BR	0
N10	< 5	Y	Integrade	BR	0	S8	N/I	N	Jabref	BR	0
N11	< 5	Y	Cogroo	BR	0	S9	N/I	Y	LibreOffice	BR	0
N12	N/I	N	Etherpad	UK	0						

We used a semi-structured format in which a script (interview guide) supported the interviewing process. We started with pilot interviews with five developers involved in Open Source Software Development to adjust the script. After that, we recruited the subjects and conducted the interviews. All interviews were conducted using text-based chat tools, like Google Talk, because the participants used this means of communication in their work, and because it facilitates data collection and interview scheduling.

Each interview was individually conducted and the data was saved on a local computer. Interviews began with a short explanation of the research, followed by some questions to profile the interviewees regarding their technical experience and main occupation. The questions in the interview script guided the interview, but were not necessarily directly asked.

Data Analysis

We qualitatively analyzed the data using procedures of Grounded Theory (GT) [79]. We selectively applied open coding, whereby concepts are identified and their properties and dimensions are discovered, and axial coding, whereby connections among codes are identified and grouped according to their properties to represent categories.

We split our analysis into two steps. The first (preliminary) step consisted of analyzing the data obtained from the students' feedback and the open questions sent to communities. In the second step, the codes and categories found in the preliminary study were used as seeds for the interview coding. During open coding, we assigned codes to sentences, paragraphs, or revisions. This procedure overlapped the axial coding, in which we identified connections between categories. We executed open and axial coding several times to refine the emerging codes and categories.

In the first step, open coding was conducted in parallel by three researchers. Each researcher independently quoted and coded the documents. After coding, the researchers discussed the quotes and codes until they came to a consensus for the entire document set. This was done to mitigate the bias caused by a single researcher and to reach a common understanding about the nomenclature and criteria. After the discussion, we began axial coding iterations, followed by discussions and changes in codes and categories. The result of this study was a Preliminary Qualitative Model of the barriers faced by newcomers to OSS.

For the second step, we analyzed the data obtained from the interviews. The analysis process was similar to the one applied in the first step, however, we used the codes and categories identified in the first step as seeds to the open coding. This time, a single researcher conducted open and axial coding and discussed with the other two the doubts and proposed new or merged categories.

After obtaining the model from the interview analysis, we iteratively reanalyzed the models obtained from all studies, relying on their respective data. The goal of this reanalysis was to combine the findings to create a single model accommodating all the barriers evidenced. Once again, we merged some barriers and reorganized the categories.

In the next section, we present our model. As our focus is on social barriers, we will only discuss the 13 barriers related to the relationship between newcomers and established open collaboration community.

RESULTING BARRIERS MODEL

The resulting model (Figure 2) aggregates all the barriers evidenced in the intermediate models. It comprises 58 barriers, organized in 6 categories and in several subcategories. In the following subsections, we focus on the 13 social barriers, supported by empirical evidence and related to the open collaboration communities literature.

Whenever we introduce a social barrier, we present the amount of times (different documents) that the barrier was identified per data source. The numbers represent, in order: number of studies from the SLR that evidenced the barrier (out of 21); number of students that reported the barrier in their feedback (out of 9); number of mentions in the open questions (out of 24); and number of interviewees that reported the barrier (out of 36). We also provided the number of projects in which the barriers were evidenced, considering only the data from practitioners.

Reception Issues

We identified four barriers related to reception issues, as presented in Table 3. This table presents the sources in which the barrier was evidenced and the number of times it was reported per source and per profile. Reception issues were evidenced in all sources, and were reported by both newcomers and experienced members.

Not receiving an answer (5-1-0-1, 1 project) was found to be a problem. During the feedback sessions, students reported that their forum post was never answered and they ended up working on an incorrect issue: "*They never answered our forum post. We spent a lot of effort in something that was already being done…*" [S5] In the interviews, an experienced member highlighted: "*In my opinion, the first [barrier] is not getting any reply*" [E1].

We also identified studies focusing on the reception barrier in our SLR [50, 73, 77, 78]. Analyzing the Freenet project, von Krogh et al. [50] found that "*… only 29 (10.5%) participants did not receive any reply to their initial posting and subsequently did not appear on the developer list again.*" Singh [73] reported a similar behavior: "*non returning newcomers can be attributed to not receiving a response…*"

CSCW research similarly studies this particular barrier, and the results reported are in accordance to those in the OSS literature and evidenced in our qualitative analysis. Joyce and Kraut [41] analyzed newcomers' posts to Usenet and found that newcomers who got a reply to their first posts were 12% more likely to post to the community again. Analyzing newcomers to Slashdot, Lampe and Johnston [52] evidenced similar results.

Compounding the lack of answers, we also found delayed answers (2-1-0-3, 3 projects) to be a contribution barrier affecting newcomers' motivation: "*[a problem was] a huge delay to receive an answer. It was necessary to send more than one email to receive an answer after a week. Demotivating. I was about to give up*" [S6] It also bothered a newcomer to the Zxing project: *"The biggest 'bottleneck' would have probably been the slow pace in communication… if you have a deadline a few days every now and then it can be quite bothersome"* [N8].

Table 3. Evidences of Reception Issues per Source

Barriers	Literature	Students Feedback	Open Question			Interviews		
			Less than 6 months	6 months to 3 years	More than 3 years	Dropouts	Newcomers	Experienced
Not receiving an answer	• (5)	• (1)						• (1)
Delayed answers	• (2)	• (1)					• (1)	• (2)
Impolite answers	• (2)	• (1)					• (1)	• (2)
Receiving answers with too advanced/ complex contents			• (1)				• (1)	• (1)

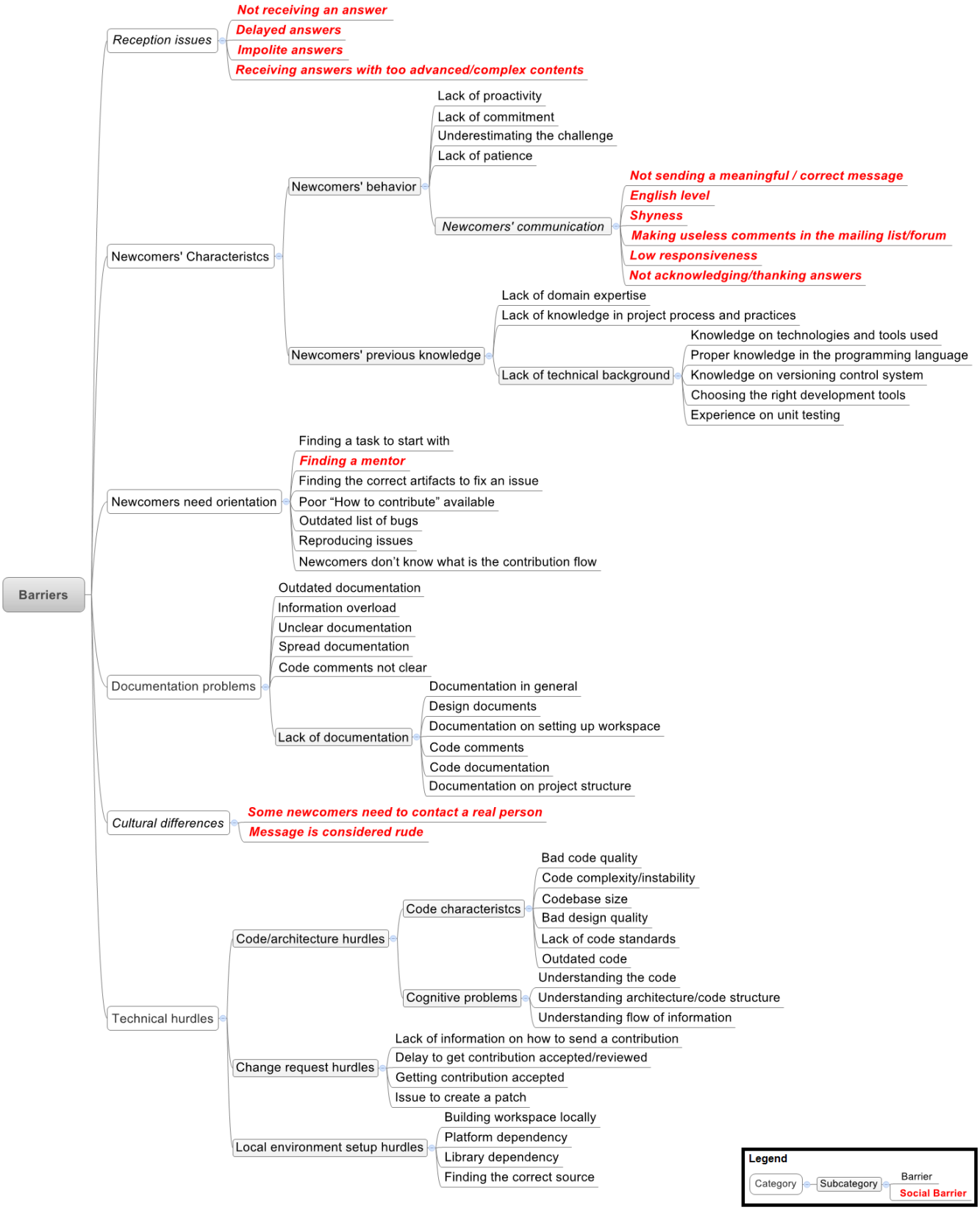

Figure 2. Model of barriers for newcomers to OSS

In the SLR, we identified evidence of this barrier in two studies [36, 88]. Jensen et al. [36] analyzed the mailing lists of four OSS projects and found that "*None of the newbies who failed to receive a reply within 24 hours of posting their first question were still posting to the mailing lists beyond the study period.*" However, we could not find any specific evidence of delayed or late reply in the complementary CSCW literature analyzed for this study. Additional study may be necessary to verify whether this is an issue in other domains.

Impolite answers (2-1-0-3, 4 projects) also appeared in the students' feedback and in the interviews. For example, an experienced member reported: "*…and of course one more thing is the developers' attitude. Some developers may not be suitable for receiving newcomers, they may get angry pretty quickly and kill the interest of the newcomers. Very few of the newcomers know how to behave against this kind of tough developers*" [E9].

CSCW literature, including the studies related to OSS, confirms this barrier. One of the OSS studies [73] reported that non-returning newcomers can be attributed to receiving a condescending response. In the context of Wikipedia, Farzan and Kraut [20] reported demotivated newcomers due to experienced editors' hostile behavior. Some students complained of reversions and deletions that occurred without proper/polite explanation. Zhu et al. [89], Suh et al. [80], and Halfaker et al. [28] found that negative feedback reduced motivation in Wikipedia. And Suh et al. [80] reported that the excessive number of newcomer reverts evidence the growing resistance from the Wikipedia community to new content, especially from occasional and new editors.

Receiving answers with too advanced/complex contents (0-0-1-2, 3 projects) was another barrier evidenced in our data obtained from practitioners. In some cases, newcomers receive answers that required in-depth knowledge about the project and technologies. For example, a newcomer reported: "*The reason I didn't find the reply helpful is due to they talk a little bit out of my understanding of the project*" [O2]. Another newcomer reported a similar problem: "*I found it awesome to get a quick and nice reaction, but the suggestions I could do seemed a bit far fetched to give to a beginner*" [N8]. We could not find any literature reporting or supporting this specific finding.

From the analysis of these 4 barriers, we noticed that reception could result in a smaller number of returning newcomers. Open collaboration communities therefore ought to be attentive to newcomer reception. Indeed, welcome messages, assistance, and constructive criticism increase the retention of newcomers [13]. The literature presented automated strategies that can be used to soften reception issues [21, 27], which were mainly focused on automated answer/feedback to newcomers' contributions. We are not aware of any approach like this in OSS communities, but this strategy's applicability and effectiveness in these settings should be explored and assessed.

Newcomers' Communication Behavior

Newcomers' characteristics may be barriers to project contribution. Newcomers are expected to possess a minimum requirement of previous technical background to perform a development task. In addition, the community also expects newcomers to have certain social skills. We could evidence 6 barriers under this category as presented in Table 4. Notably, only experienced members reported these barriers, whereas the newcomers did not mention their own communication behavior as possible problems.

Three experienced members reported not sending a correct / meaningful message (2-0-0-3, 2 projects) as a barrier. Community members may not answer a message if they do not understand it: "*in general I answer the questions that are well written … some people post things that… I don't know how to answer. So, I wait until someone else makes an attempt and see if the original poster will make a better effort the second time to post something that I can understand*" [E4].

Singh [73] studied this problem in OSS forums. She demonstrated that the community responds better to informative subject lines, comprehensible posts, and correct messages, which is also studied in other CSCW domains. Burke et al. [7], for example, analyzed Usenet communities, finding that self-introductions can double the odds of receiving answers. Arguello et al. [1] also analyzed Q&A history and found that on-topic messages and use of vernacular language increased reply likelihood. Joyce and Kraut [41] also analyzed Usenet communities and reported that newcomers were more likely to receive a response if they asked a question or wrote a longer post.

A more specific and related barrier regards English level (0-0-0-4, 3 projects). English is adopted in most OSS projects. After reporting problems with incorrect messages, E4 amended: *"for some people it is due to English as a second language, I understand that but still..."* Other participant reinforced this issue: "*Having a decent English is needed*" [E1]. It is difficult to address this issue, but providing guidelines or asking for clarification may help. For larger communities, matching a newcomer to a member that speaks the same language might be helpful.

Table 4. Evidences of Newcomers' Communication Behavior per Source

Barriers	Literature	Students Feedback	Open Question			Interviews		
			Less than 6 months	6 months to 3 years	More than 3 years	Dropouts	Newcomers	Experienced
Not sending a meaningful/correct message	• (2)							• (3)
English level								• (4)
Shyness								• (2)
Making useless comments in the mailing list/forum								• (1)
Low responsiveness								• (1)
Not acknowledging/thanking answers received								• (1)

Shyness (0-0-0-2, 2 projects) was also reported as a problem. An experienced member reported that once he gave up contributing because he was shy to ask the community: "*I was trying to solve a bug… by myself. I was kind of shy to ask for help"* [E11]. Preece et al. [62] analyzed the MSN bulletin board and found that 28.3% of the lurkers gave shyness as a reason for not posting. A possible approach to deal with this issue would be "breaking the ice." As soon as newcomers subscribe, a member could approach them; automatic greetings could also be used.

For this category of barriers, we found that CSCW literature mainly studies and provides evidence on how, in order to be well received, newcomers should behave when sending their first messages. The evidence found relies on studies conducted with historical data of Q&A communities. For the other barriers, we could not find any evidence in the literature.

Finding a mentor

In accordance with Dagenais et al.'s [15] work on software projects, we found that newcomers often face unfamiliar and rugged landscapes when starting to contribute to an OSS project. Consequently, they need proper orientation to find their way into the project and to contribute correctly.

Finding a mentor (3-3-2-5, 7 projects) was identified as a social barrier for newcomers to OSS. This was evidenced in all the sources we examined, and was reported both by newcomers and experienced members, as presented in Figure 2. One newcomer gave up because of the difficulty of finding a mentor: *"… basically, see, I was not active contributor at that time ... if some meaningful direction could be provided then I would have started. This direction I didn't get … someone with my profile ... who want to do some stuff with open source project probably some basic handholding would have helped*." [D2]

Ease in locating an expert or a mentor was also evidenced in the systematic review. Cubranic et al. [14] reported, "*It can be difficult for newcomers to join such groups [OSS projects] because it is hard to obtain effective mentoring.*" To alleviate this problem, Canfora et al. [9] proposed a tool that recommends mentors to newcomers. They evaluated the tool by surveying project members, finding that mentoring is important to newcomers. There are also some mentorship-related studies in open collaboration communities. For instance, Musicant et al. [56] qualitatively analyzed data from Wikipedia's program Adopt-a-user and found that several key mentor functions are missing or not consistently fulfilled. Most adopters focus on establishing their legitimacy rather than proactively guiding, protecting, and supporting the long-term growth of adoptees. Choi et al. [13] analyzed Wikipedia socialization tactics and found that they rarely assign new members a mentor or provide clear guidance about how to behave in a project.

We observed that newcomers look for mentorship when placing their first contributions. However, as reported by Musicant et al. [56], the goal of experienced members seems to be self-promotion rather than coaching and supporting. Discovering what motivates experienced members to properly mentor newcomers is therefore likely a fruitful future research area.

Cultural Differences

Differences related to individuals' cultural backgrounds are a known problem in distributed software development [75], including OSS development [32]. Cultural differences exist in open collaboration; volunteers have diverse national, organizational, and professional backgrounds, resulting in different values, perceptions, and work behaviors. According to Herbsleb and Moitra [31], this can lead to serious misunderstandings and conflicts. In our study, barriers from this category appeared only during the interviews, as reported by two experienced members and one newcomer.

Some newcomers may consider a message rude (0-0-0-2, 2 projects) due to cultural interpretation: *"All the community is very nice. Of course, there are some 'German' guys. One time, a guy was rude with me, but, you know, we Brazilians are not used to the 'German way to talk directly'"* [N13]. In the OSS scenario, projects rely only on textual communication and often involve people who do not want to spend time writing careful messages. In addition, there are strong egos involved.

Another kind of barrier evidenced concerns for the need of a personal contact (0-0-0-2, 2 projects) to create a bridge, or a stronger connection, to the project. One experienced member reported the specific case of his compatriots: "*…people behave more emotionally in our country, I mean, newcomers need some special attention. Maybe since we are Mediterranean people, I don't know, but I think this is not the case in many projects in the world."* [E9]. An experienced member from another project also reports the same issue in a more general context: "*Although it may be a cultural aspect of open source that people prefer to make initial contact with a real person, so I don't have a problem with that*" [E7].

Although the literature underscores cultural differences among participants as an important aspect [38, 46, 58, 61], the systematic review conducted by Steinmacher et al. showed that few studies investigated or dealt with cultural issues in distributed software development [75]. This category of barriers could be another fruitful research topic.

DISCUSSION

For purposes of clarity, we present most of our results discussion alongside each of the barriers in the previous section. In this section, we highlight some of the findings and present some higher-level discussion.

In general, this study discovered empirical evidence of the barriers faced by newcomers to OSS projects when placing their first contribution. This empirical evidence is important as many studies are motivated by or deal with anecdotal evidence. This paper brings evidence from reality, which is rarely precisely documented.

The model presented in Figure 2 groups social and non-social barriers into six categories, illustrating the high number of technical barriers compared to social barriers. This imbalance can be explained by the characteristics of OSS communities, which demand contributors and tasks with specific technical skills and knowledge requirements.

In the previous section, we related our findings to those reported in the literature on other open collaboration communities, mainly Wikipedia. We observed that the social barriers identified are fairly similar to those evidenced in other domains. Thus, some of the solutions and mitigation strategies used in those contexts could be tried on OSS communities and vice versa.

Some barriers are well reported and analyzed by the literature; the most explored barriers are those related to reception issues. For example, not receiving an answer is well evidenced quantitatively in both OSS and Q&A literature, and these results were in line with our results. Receiving impolite answers was also largely studied in the CSCW literature, mainly on the analysis of reverts in Wikipedia. The proposed strategies of automated answers and feedback used in Wikipedia can be adapted and then evaluated in OSS context.

On the other hand, some barriers identified by our study are neglected in the literature. One of them is receiving answers with too advanced/complex contents. We could not identify any study dealing with this barrier. There are opportunities to investigate this kind of barrier in different domains, and provide a deeper understanding of this issue.

The barriers related to newcomers' communication behavior are also under-studied. These barriers were evidenced by the community members' perspective, which is rarely investigated. Sometimes, the newcomers themselves are creating the barriers when they post useless comments or do not acknowledge an answer, which heightens the need to better understand what the community expects from the newcomers, and how these expectations affect or impact the newcomers' reception issues.

Cultural differences also deserve to be highlighted here. There are some possibilities related to this barrier that can be followed up in different CSCW domains. For example, how do collaborators observe cultural differences? Do they consider such differences when dealing with their peers? Do these barriers cause issues related to trust, misunderstandings, etc.?

In general, having substantiated and characterized the barriers faced by newcomers in a particular domain, we believe that the CSCW community can benefit from our research. Future research ought to focus on searching for commonalities and differences among barriers faced in different domains in order to build models and theories about joining processes in open collaboration communities.

OSS researchers can benefit from these results by using them to conceive strategies for newcomer support. By including the context of the CSCW and related research literature, we provide a starting point to conceive such support. To achieve this, it would be necessary to put more effort on specific research topics, such as understanding (and creating ways to measure) the influence of the barriers in newcomers' experience, identifying and creating different strategies to lower each barrier, and proposing metrics to grade the support offered for each barrier.

OSS practitioners can take advantage of adapting the existing strategies used in other domains to design their own tool-based support. Assuming that, as stated by Dagenais et al. [15], "newcomers are explorers who must orient themselves within an unfamiliar landscape," the model of barriers can be used by open collaboration communities to place the proper signs and maps to help newcomers orient themselves, and alert them about the obstacles that they might face.

Are the barriers always a problem?

Although we considered the barriers as a negative aspect of onboarding, some barriers can be used as filters. Findings from some studies [18, 27] revealed some of these entry barriers led to improved contributions in the future. Moreover, research conducted in the OSS domain demonstrated that socialization barriers are useful for maintaining community integration and the quality of the community's product.

Since our focus was on newcomers placing their first contribution, regardless of their intention to become a community member, we were not concerned about overall socialization issues or future contributions, thus distinguishing our study from the existing literature. However, a clear direction for future work is to explore how the communities perceive these barriers and how they impact the quality of contributions from newcomers.

Are newcomers good?

Even knowing that a continual influx of newcomers is often related to the OSS project success [3, 11, 63, 74], there are some downsides. These negative aspects revolve around the two main premises that underlie the Brooks' Law: productivity (the *ramp up* problem) and communication/ coordination overhead [5].

In the context of this research, we do not consider the *ramp up* to reach high productivity as a threat. We understand that most newcomers do not want to become core developers. Regarding coordination overhead, we believe that by providing the right support to newcomers this overhead can be reduced. Besides, Eric Raymond, in his well-known paper "The Cathedral and Bazaar" [64], claims that the effects of Brooks' Law regarding communication/coordination overhead does not apply as the more peripheral developers work on *separable parallel subtasks and interact with each other very little.* Therefore, coordination overhead has its effects mainly within the core group [10, 39, 45, 69].

LIMITATIONS

Although we analyzed data from a variety of sources and from different projects, we likely did not discover all possible barriers or provide full explanations of the barriers. We are aware that each project has its singularities and that the OSS universe is huge, meaning the level of support and the barriers can differ according to the project or the ecosystem. Our strategy to consider different projects and different developer profiles aimed to alleviate this issue, identifying recurrent mentions of barriers from multiple perspectives.

Another threat to the results' validity is the data classifications' subjectivity. We used Grounded Theory procedures to mitigate this threat, given that Grounded

Theory requires the entire analysis to be grounded in the data collected. Additionally, we discussed the analysis process along with two other researchers to encourage a better validation of the interpretations through mutual agreement.

During interviews, experienced members were asked to answer questions regarding barriers faced by newcomers when they were seeking to place their first contribution, but due to memory effects they may have referred to the whole joining process. To avoid this kind of situation, we reinforced the focus of the research and verified some answers.

Since we sent open invitations to a mailing list, sampling bias affects our interviewees and open question respondents, namely self-selection bias and social desirability bias. However, we tried to counteract this effect by seeking out different sources and analyzing the answers in context to identify specifics.

CONCLUSION

Numerous communities are maintained by volunteers, who can easily drop out [8]. Explicit effort is necessary to mitigate obstacles and problems in these communities [83]. In this study, we identified, organized, and discussed social barriers that hinder newcomers' first contribution to OSS projects. The barriers emerged from a literature review on OSS newcomers, as well as from interviews conducted with, and questionnaires answered by, developers in different stages of the OSS community joining process. This study differs from the existing literature by focusing on the barriers newcomers face when trying to place their first contribution to a project, rather than focusing on the entire joining process. Additionally, our study followed a qualitative approach based on practitioner feedback. The qualitative view complements the existing dominant quantitative-based research on newcomers' joining process [27, 28, 36, 84, 88, 89].

This study's results provides insights for communities that want to smooth newcomer onboarding and lays a foundation for building better onboarding support tools. In addition, researchers can use the model to plan further qualitative and quantitative studies to investigate specific barriers, their interplay, and their in-practice implications. In particular, some barriers, such as cultural differences, receiving answers with too advanced/complex contents, and newcomers' shyness, were evidenced by the practitioners, but not found in the systematic literature review. These barriers certainly warrant further investigations.

Identifying and understanding the initial barriers was a first step towards better orienting newcomers' first steps. OSS projects can benefit from additional contributions if they offer the right support to newcomers who are trying to contribute to the project for the first time. A smooth first contribution may increase the total number of successful contributions made by single contributors and, hopefully, the number of long-term contributors.

ACKNOWLEDGEMENTS

The authors would like to thank Fundação Araucária, CNPq (process 477831/2013-3), FAPEAM (processes 062.00146/2012; 062.00600/2014; 062.00578/2014; and 01135/2011), NAPSoL-PRP-USP, NAWEB, and FAPESP for their financial support. Igor Steinmacher received grants from CAPES (BEX 2038-13-7) during the development of this work.

REFERENCES

[1] Arguello, J., Butler, B.S., Joyce, E., Kraut, R., Ling, K.S., Rosé, C. and Wang, X. 2006. Talk to Me: Foundations for Successful Individual-group Interactions in Online Communities. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, 2006), 959–968.

[2] Beecham, S., Baddoo, N., Hall, T., Robinson, H. and Sharp, H. 2008. Motivation in Software Engineering: A systematic literature review. *Information & Software Technology*. 50, 9-10 (2008), 860–878.

[3] Bird, C., Gourley, A., Devanbu, P., Swaminathan, A. and Hsu, G. 2007. Open Borders? Immigration in Open Source Projects. *ICSE Workshops MSR '07. Fourth International Workshop on Mining Software Repositories, 2007* (2007), 6–6.

[4] Bonaccorsi, A. and Rossi, C. 2004. Altruistic individuals, selfish firms? The structure of motivation in Open Source software. *First Monday*. 9, 1 (2004).

[5] Brooks, F.P. 1995. *The Mythical Man-Month: Essays on Software Engineering*. Addison-Wesley Professional.

[6] Bryant, S.L., Forte, A. and Bruckman, A. 2005. Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia. *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work* (Sanibel Island, Florida, USA, 2005), 1–10.

[7] Burke, M., Joyce, E., Kim, T., Anand, V. and Kraut, R. 2007. Introductions and Requests: Rhetorical Strategies That Elicit Response in Online Communities. *Communities and Technologies 2007*. C. Steinfield, B. Pentland, M. Ackerman, and N. Contractor, eds. Springer London. 21–39.

[8] Burke, M., Marlow, C. and Lento, T. 2009. Feed Me: Motivating Newcomer Contribution in Social Network Sites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA, 2009), 945–954.

[9] Canfora, G., Di Penta, M., Oliveto, R. and Panichella, S. 2012. Who is Going to Mentor Newcomers in Open Source Projects? *Proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering* (Cary, North Carolina, 2012), 44:1–44:11.

[10] Capiluppi, A. and Adams, P. 2009. Reassessing Brooks' Law for the Free Software Community. *Open Source Ecosystems: Diverse Communities Interacting*. Springer Berlin Heidelberg. 274–283.

[11] Capiluppi, A. and Michlmayr, M. 2007. From the Cathedral to the Bazaar: An Empirical Study of the Lifecycle of Volunteer Community Projects. *Open Source Development, Adoption and Innovation*. J. Feller, B. Fitzgerald, W. Scacchi, and A. Sillitti, eds. Springer Boston. 31–44.

[12] Chengalur-Smith, I.N., Sidorova, A. and Daniel, S.L. 2010. Sustainability of Free/Libre Open Source Projects: A Longitudinal Study. *Journal of the Association for Information Systems*. 11, 11 (2010).

[13] Choi, B., Alexander, K., Kraut, R.E. and Levine, J.M. 2010. Socialization Tactics in Wikipedia and Their Effects. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA, 2010), 107–116.

[14] Cubranic, D., Murphy, G.C., Singer, J. and Booth, K.S. 2005. Hipikat: a project memory for software development. *IEEE Transactions on Software Engineering*. 31, 6 (2005), 446–465.

[15] Dagenais, B., Ossher, H., Bellamy, R.K.E., Robillard, M.P. and Vries, J.P. de 2010. Moving into a new software project landscape. *32nd International Conference on Software Engineering* (2010), 275–284.

[16] David, P.A. and Shapiro, J.S. 2008. Community-based production of open-source software: What do we know about the developers who participate? *Information Economics and Policy*. 20, 4 (2008), 364–398.

[17] Dittrich, Y., John, M., Singer, J. and Tessem, B. 2007. Editorial: For the Special Issue on Qualitative Software Engineering Research. *Inf. Softw. Technol.* 49, 6 (Jun. 2007), 531–539.

[18] Ducheneaut, N. 2005. Socialization in an Open Source Software Community: A Socio-Technical Analysis. *Computer Supported Cooperative Work (CSCW)*. 14, 4 (2005), 323–368.

[19] Fang, Y. and Neufeld, D. 2009. Understanding Sustained Participation in Open Source Software Projects. *J. Manage. Inf. Syst.* 25, 4 (Apr. 2009), 9–50.

[20] Farzan, R. and Kraut, R.E. 2013. Wikipedia Classroom Experiment: Bidirectional Benefits of Students' Engagement in Online Production Communities. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France, 2013), 783–792.

[21] Faulkner, R., Walling, S. and Pinchuk, M. 2012. Etiquette in Wikipedia: Weening New Editors into Productive Ones. *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration* (Linz, Austria, 2012), 5:1–5:4.

[22] Fogel, K. 2013. *Producing Open Source Software: How to Run a Successful Free Software Project*. O'Reilly Media.

[23] Forte, A. and Lampe, C. 2013. Defining, Understanding, and Supporting Open Collaboration: Lessons From the Literature. *American Behavioral Scientist*. 57, 5 (2013), 535–547.

[24] Franca, A.C.C., Gouveia, T.B., Santos, P.C.F., Santana, C.A. and Silva, F.Q.B. da 2011. Motivation in software engineering: A systematic review update. *Evaluation Assessment in Software Engineering (EASE 2011), 15th Annual Conference on* (Apr. 2011), 154–163.

[25] França, A.C.C., Silva, F.Q.B. da, L. C. Felix, A. de and Carneiro, D.E.S. 2014. Motivation in software engineering industrial practice: A cross-case analysis of two software organisations. *Information & Software Technology*. 56, 1 (2014), 79–101.

[26] Fugelstad, P., Dwyer, P., Filson Moses, J., Kim, J., Mannino, C.A., Terveen, L. and Snyder, M. 2012. What Makes Users Rate (Share, Tag, Edit...)?: Predicting Patterns of Participation in Online Communities. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA, 2012), 969–978.

[27] Halfaker, A., Geiger, R.S., Morgan, J. and Riedl, J. 2013. The Rise and Decline of an Open Collaboration System: How Wikipedia's reaction to sudden popularity is causing its decline. *American Behavioral Scientist*. 57, (2013).

[28] Halfaker, A., Kittur, A. and Riedl, J. 2011. Don't Bite the Newbies: How Reverts Affect the Quantity and Quality of Wikipedia Work. *7th Intl. Symposium on Wikis and Open Collaboration* (2011), 163–172.

[29] Hannebauer, C., Book, M. and Gruhn, V. 2014. An Exploratory Study of Contribution Barriers Experienced by Newcomers to Open Source Software Projects. *Proceedings of the 1st International Workshop on CrowdSourcing in Software Engineering* (Hyderabad, India, 2014), 11–14.

[30] Hars, A. and Ou, S. 2001. Working for free? Motivations of participating in open source projects. *34th Annual Hawaii Intl. Conference on System Sciences* (2001), 9 pp.

[31] Herbsleb, J.D. and Moitra, D. 2001. Global software development. *Software, IEEE*. 18, 2 (Mar. 2001), 16–20.

[32] Herraiz, I., Robles, G., Amor, J.J., Romera, T., Barahona, J.M.G. and Carlos, J. 2006. The processes of joining in global distributed software projects. *2006 International Workshop on Global Software Development for the Practitioners* (2006), 27–33.

[33] Hertel, G., Niedner, S. and Herrmann, S. 2003. Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel. *Research Policy*. 32, 7 (2003), 1159–1177.

[34] Hsieh, G., Hou, Y., Chen, I. and Truong, K.N. 2013. Welcome!": Social and Psychological Predictors of Volunteer Socializers in Online Communities. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA, 2013), 827–838.

[35] Jalali, S. and Wohlin, C. 2012. Systematic Literature Studies: Database Searches vs. Backward Snowballing. *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement* (Lund, Sweden, 2012), 29–38.

[36] Jensen, C., King, S. and Kuechler, V. 2011. Joining Free/Open Source Software Communities: An Analysis of Newbies' First Interactions on Project Mailing Lists. *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (2011), 1–10.

[37] Jergensen, C., Sarma, A. and Wagstrom, P. 2011. The Onion Patch: Migration in Open Source Ecosystems. *19th ACM SIGSOFT Symposium and the 13th European Conf. on Foundations of Software Engineering* (2011), 70–80.

[38] Ji, Y.G., Hwangbo, H., Yi, J.S., Rau, P.L.P., Fang, X. and Ling, C. 2010. The Influence of Cultural Differences on the Use of Social Network Services and the Formation of Social Capital. *International Journal of Human-Computer Interaction*. 26, 11-12 (2010), 1100–1121.

[39] Johnson, J.P. 2001. Economics of Open Source Software. *Unpublished paper*. (May. 2001).

[40] Jørgensen, N. 2007. Developer autonomy in the FreeBSD open source project. *Journal of Management & Governance*. 11, 2 (2007), 119–128.

[41] Joyce, E. and Kraut, R.E. 2006. Predicting Continued Participation In Newsgroups. *Journal of Computer-Mediated Communication*. 11, (2006), 2006.

[42] Ke, W. and Zhang, P. 2010. The Effects of Extrinsic Motivations and Satisfaction in Open Source Software Development. *J. AIS*. 11, 12 (2010).

[43] Kitchenham, B. and Brereton, P. 2013. A systematic review of systematic review process research in software engineering. *Information and Software Technology*. 55, 12 (Dec. 2013), 2049–2075.

[44] Kitchenham, B. and Charters, S. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. Technical Report #EBSE 2007-001. Keele University and Durham University.

[45] Koch, S. 2004. Profiling an Open Source Project Ecology and Its Programmers. *Electronic Markets*. 14, 2 (2004), 77–88.

[46] Koh, J., Kim, Y.-G., Butler, B. and Bock, G.-W. 2007. Encouraging Participation in Virtual Communities. *Commun. ACM*. 50, 2 (Feb. 2007), 68–73.

[47] Kraut, R.E., Burke, M., Riedl, J. and Resnick, P. 2012. The Challenges of Dealing with Newcomers. MIT Press. 179–230.

[48] Krogh, G. von, Haefliger, S., Spaeth, S. and Wallin, M.W. 2012. Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development. *MIS Quarterly*. 36, 2 (2012), 649–676.

[49] Krogh, G. von, Spaeth, S. and Lakhani, K.R. 2003. Community, joining, and specialization in open source software innovation: A case study. *Research Policy*. 32, 7 (2003), 1217–1241.

[50] Von Krogh, G., Spaeth, S. and Lakhani, K.R. 2003. Community, joining, and specialization in open source software innovation: A case study. *Research Policy*. 32, 7 (2003), 1217–1241.

[51] Lakhani, K.R. and Wolf, R.G. 2005. Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects. *Perspectives on Free and Open Source Software*. MIT Press.

[52] Lampe, C. and Johnston, E. 2005. Follow the (Slash) Dot: Effects of Feedback on New Members in an Online Community. *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work* (Sanibel Island, Florida, USA, 2005), 11–20.

[53] Maalej, W., Happel, H.-J. and Rashid, A. 2009. When users become collaborators: towards continuous and context-aware user input. *Proceeding of the 24th ACM SIGPLAN conference companion on Object oriented programming systems languages and applications* (Orlando, Florida, USA, 2009), 981–990.

[54] Marlow, J., Dabbish, L. and Herbsleb, J. 2013. Impression Formation in Online Peer Production: Activity Traces and Personal Profiles in Github. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA, 2013), 117–128.

[55] Meirelles, P., Santos, C., Miranda, J., Kon, F., Terceiro, A. and Chavez, C. 2010. A study of the relationships between source code

metrics and attractiveness in free software projects. *2010 Brazilian Symposium on Software Engineering (SBES)*, (2010), 11–20.

[56] Musicant, D.R., Ren, Y., Johnson, J.A. and Riedl, J. 2011. Mentoring in Wikipedia: A Clash of Cultures. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (Mountain View, California, 2011), 173–182.

[57] Nakakoji, K., Yamamoto, Y., Nishinaka, Y., Kishida, K. and Ye, Y. 2002. Evolution Patterns of Open-source Software Systems and Communities. *Proceedings of the International Workshop on Principles of Software Evolution* (Orlando, Florida, 2002), 76–85.

[58] Nguyen, D.T. and Fussell, S.R. 2013. Effect of Message Content on Communication Processes in Intercultural and Same-culture Instant Messaging Conversations. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA, 2013), 19–32.

[59] Oreg, S. and Nov, O. 2008. Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values. *Computers in Human Behavior*. 24, 5 (2008), 2055–2073.

[60] Pham, R., Singer, L., Liskin, O., Figueira Filho, F. and Schneider, K. 2013. Creating a Shared Understanding of Testing Culture on a Social Coding Site. *Proceedings of the 2013 International Conference on Software Engineering* (San Francisco, CA, USA, 2013), 112–121.

[61] Preece, J. 2001. Sociability and usability in online communities: Determining and measuring success. *Behaviour & Information Technology*. 20, 5 (2001), 347–356.

[62] Preece, J., Nonnecke, B. and Andrews, D. 2004. The top five reasons for lurking: improving community experiences for everyone. *Computers in Human Behavior*. 20, 2 (2004), 201–223.

[63] Qureshi, I. and Fang, Y. 2011. Socialization in Open Source Software Projects: A Growth Mixture Modeling Approach. *Org. Res. Methods*. 14, 1 (2011), 208–238.

[64] Raymond, E.S. 1999. *The Cathedral and the Bazaar*. O'Reilly & Associates, Inc.

[65] Roberts, J.A., Hann, I.-H. and Slaughter, S.A. 2006. Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects. *Manage. Sci.* 52, 7 (Jul. 2006), 984–999.

[66] Santos, C., Kuk, G., Kon, F. and Pearson, J. 2013. The Attraction of Contributors in Free and Open Source Software Projects. *J. Strateg. Inf. Syst.* 22, 1 (Mar. 2013), 26–45.

[67] Scacchi, W. 2002. Understanding the requirements for developing open source software systems. *IEE Proceedings Software*. 149, 1 (2002), 24–39.

[68] Schilling, A., Laumer, S. and Weitzel, T. 2012. Who Will Remain? An Evaluation of Actual Person-Job and Person-Team Fit to Predict Developer Retention in FLOSS Projects. *Proceedings of the 2012 45th Hawaii International Conference on System Sciences* (Washington, DC, USA, 2012), 3446–3455.

[69] Schweik, C.M., English, R.C., Kitsing, M. and Haire, S. 2008. Brooks' Versus Linus' Law: An Empirical Test of Open Source Projects. *Proceedings of the 2008 International Conference on Digital Government Research* (Montreal, Canada, 2008), 423–424.

[70] Seaman, C.B. 1999. Qualitative methods in empirical studies of software engineering. *Software Engineering, IEEE Transactions on*. 25, 4 (Jul. 1999), 557–572.

[71] Shah, S.K. 2006. Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. *Manage. Sci.* 52, 7 (2006), 1000–1014.

[72] Sharp, H., Baddoo, N., Beecham, S., Hall, T. and Robinson, H. 2009. Models of motivation in software engineering. *Information & Software Technology*. 51, 1 (2009), 219–233.

[73] Singh, V. 2012. Newcomer integration and learning in technical support communities for open source software. *Proceedings of the 17th ACM international conference on Supporting group work* (New York, NY, USA, 2012), 65–74.

[74] Sinha, V.S., Mani, S. and Sinha, S. 2011. Entering the circle of trust: developer initiation as committers in open-source projects. *Proceedings of the 8th Working Conference on Mining Software Repositories* (Waikiki, Honolulu, HI, USA, 2011), 133–142.

[75] Steinmacher, I., Chaves, A. and Gerosa, M. 2013. Awareness Support in Distributed Software Development: A Systematic Review and Mapping of the Literature. *Computer Supported Cooperative Work (CSCW)*. 22, 2-3 (2013), 113–158.

[76] Steinmacher, I., Silva, M.A.G. and Gerosa, M.A. 2014. Systematic review on problems faced by newcomers to open source projects. *10th International Conference on Open Source Software* (2014), 10pp.

[77] Steinmacher, I., Wiese, I., Chaves, A.P. and Gerosa, M.A. 2013. Why do newcomers abandon open source software projects? *International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)* (2013), 25–32.

[78] Steinmacher, I., Wiese, I.S. and Gerosa, M.A. 2012. Recommending mentors to software project newcomers. *Third International Workshop on Recommendation Systems for Software Engineering (RSSE)* (Washington, DC, USA, Jun. 2012), 63–67.

[79] Strauss, A. and Corbin, J.M. 1998. *Basics of Qualitative Research : Techniques and Procedures for Developing Grounded Theory*. SAGE Publications.

[80] Suh, B., Convertino, G., Chi, E.H. and Pirolli, P. 2009. The Singularity is Not Near: Slowing Growth of Wikipedia. *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (Orlando, Florida, 2009), 8:1–8:10.

[81] Vora, P., Komura, N. and Team, S.U. 2010. The n00b Wikipedia Editing Experience. *6th Intl. Symposium on Wikis and Open Collaboration* (2010), 36:1–36:3.

[82] Wang, J. and Sarma, A. 2011. Which bug should I fix: helping new developers onboard a new project. *Proceedings of the 4th International Workshop on Cooperative and Human Aspects of Software Engineering* (Waikiki, Honolulu, HI, USA, 2011), 76–79.

[83] Wang, L.S., Chen, J., Ren, Y. and Riedl, J. 2012. Searching for the Goldilocks Zone: Trade-offs in Managing Online Volunteer Groups. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA, 2012), 989–998.

[84] Wang, Y.-C., Kraut, R. and Levine, J.M. 2012. To Stay or Leave?: The Relationship of Emotional and Informational Support to Commitment in Online Health Support Groups. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA, 2012), 833–842.

[85] Wolff-Marting, V., Hannebauer, C. and Gruhn, V. 2013. Patterns for tearing down contribution barriers to FLOSS projects. *12th Intl. Conf. on Intelligent Software Methodologies, Tools and Techniques* (2013), 9–14.

[86] Ye, Y. and Kishida, K. 2003. Toward an Understanding of the Motivation Open Source Software Developers. *Proceedings of the 25th International Conference on Software Engineering* (Portland, Oregon, 2003), 419–429.

[87] Yu, S. and Ming, W. 2009. Research on individual motivation model of software engineering. *Journal of Communication and Computer*. 6, 11 (2009), 12.

[88] Zhou, M. and Mockus, A. 2012. What make long term contributors: Willingness and opportunity in OSS community. *Software Engineering (ICSE), 2012 34th International Conference on* (Jun. 2012), 518–528.

[89] Zhu, H., Kraut, R. and Kittur, A. 2012. Effectiveness of Shared Leadership in Online Communities. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA, 2012), 407–416.

[90] Zhu, H., Zhang, A., He, J., Kraut, R.E. and Kittur, A. 2013. Effects of Peer Feedback on Contribution: A Field Experiment in Wikipedia. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France, 2013), 2253–2262.