Perceptions of the State of D&I and D&I Initiative in the ASF

Mariam Guizani
Oregon State University
Corvallis, Oregon, USA
guizanim@oregonstate.edu

Bianca Trinkenreich
Northern Arizona University
Flagstaff, AZ, USA
bt473@nau.edu

Aileen Abril Castro-Guzman
Oregon State University
Corvallis, OR, USA
castroga@oregonstate.edu

Igor Steinmacher
Northern Arizona University
Flagstaff, AZ, USA
igor.steinmacher@nau.edu

Marco Gerosa
Northern Arizona University
Flagstaff, AZ, USA
marco.gerosa@nau.edu

Anita Sarma
Oregon State University
Corvallis, OR, USA
anita.sarma@oregonstate.edu

ABSTRACT

Open Source Software (OSS) Foundations and projects are investing in creating Diversity and Inclusion (D&I) initiatives. However, little is known about contributors’ perceptions about the usefulness and success of such initiatives. We aim to close this gap by investigating how contributors perceive the state of D&I in their community. In collaboration with the Apache Software Foundation (ASF), we surveyed 600+ OSS contributors and conducted 11 follow-up interviews. We used mixed methods to analyze our data—quantitative analysis of Likert-scale questions and qualitative analysis of open-ended survey questions and the interviews to understand contributors’ perceptions and critiques of the D&I initiative and how to improve it. Our results indicate that the ASF contributors felt that the state of D&I was still lacking, especially regarding gender, seniority, and English proficiency. Regarding the D&I initiative, some participants felt that the effort was unnecessary, while others agreed with the effort but critiqued its implementation. These findings show that D&I initiatives in OSS communities are a good start, but there is room for improvements. Our results can inspire the creation of new and the refinement of current initiatives.

Lay Abstract: Open Source Software (OSS) is widely used in society (e.g., Linux, Chrome, and Firefox), and contributing to these projects helps individuals learn and showcase their skills, so much so that the history of contributions are increasingly being analyzed by hirers. However, the people who contribute to OSS are predominately men (about 90%). This means that women and other minorities lose out on job opportunities and OSS projects lose out on diversity of thought. OSS organizations such as the Apache Software Foundation (ASF) promote a variety of initiatives to increase diversity and inclusion (D&I) in their projects, but they are piece-meal and little is known about contributors’ perceptions about the usefulness and success of these initiatives. Here, we surveyed and interviewed ASF contributors to understand their perceptions about the state of D&I in the ASF and the effectiveness of existing D&I initiatives. Our findings show that individuals who are in the minority face challenges (e.g., stereotyping, lack of peer-network, and representation in decision making) and contributors’ perceptions of the D&I initiative are a mixed bag, ranging from commending the current efforts to considering them to be “lip service”. These findings suggest that current D&I initiatives in OSS communities are a good start, but much needs be done in terms of creating new successful initiatives and refining current ones.

KEYWORDS

Diversity, D&I initiative, Open Source Software

1 INTRODUCTION

Open Source Software (OSS) now plays a key role in software development as well as in workforce development, where contributors join projects to learn new skills [26], showcase their technical skills [45], and improve their career paths [67, 79]. However, despite its ubiquity, it is well documented that OSS has low diversity [39, 75].

Low diversity in OSS has unfortunate effects: (1) OSS projects miss out on the benefits of a diverse set of individuals and qualifications, or from the diversity of thought that these potential contributors could bring; (2) developers in the minority miss out on the learning and experience opportunities that OSS projects provide; and (3) job opportunities evade these individuals when OSS contributions are used to make hiring decisions [45, 67].

Gender diversity in OSS has been widely studied showing women are severely underrepresented. From an analysis of ten OSS projects, Bosu and Sultana [6] found less than 10% of contributors to be women and that women were rarely included in leadership positions. Other studies have reported similar low numbers (e.g., [6, 12, 27, 65, 81]). Researchers have also investigated how gender can differentiate team dynamics [80] and team perceptions [38, 80, 83], career progression [8], motivations to join [26] and type of contributions [79]. Others have reported on barriers that women face because of the technologies in use [47], in getting their contributions accepted [74] and participating in a generally hostile environment [51, 57].
More recently, research has started to look at other diversity attributes. For example, studies have investigated the effects of location and language on getting contributions accepted [24, 62, 63]. Others have analyzed age diversity by analyzing its impact on code reviews [50], or motivations to continue to contribute [13, 49]. The majority of research has focused on only one attribute of diversity, with a few emerging works extending the gender diversity analysis by incorporating another diversity attribute (e.g., location [53, 57], cognitive style [54]). Diversity, however, is a multi-dimensional construct that arises from attributes that differentiate people, demographic (e.g., age, gender, ethnicity) or otherwise (e.g., role, expertise, personality, cognitive styles) and focusing on a single attribute as the analysis lens, gives only a partial view of this complex phenomenon. Therefore, in this work we investigate:

**(RQ1.) How do contributors with different backgrounds perceive the state of D&I in their OSS community?**

The goal of RQ1 is to build and extend the current literature by analyzing the following six contributors background attributes: gender, education, English proficiency, seniority at ASF, compensation type and country of residence. We build on work by Lee and Carver [38] who investigated contributor perception of the state of D&I using the gender lens. Our study explores participant perceptions of role stereotyping and their ability to contribute using questions from [38]. While gender, country of residence and English proficiency attributes have been investigated by others, either in isolation or along with another lens, we are unaware of literature investigating education and compensation type. Given the changing landscape of OSS [71] where companies are increasingly employing developers to contribute to OSS and people are seeing OSS as a career stepping stone [26, 79], it's important to understand how these attributes play a role in creating inclusive, diverse OSS communities.

OSS projects are well aware of problems of toxic interactions that create a non-inclusive environment and low gender diversity [38]. To overcome these issues OSS projects often include a code of conduct [69] to manage communication expectations. Several OSS foundations have also started broader D&I initiatives such as, the Linux foundations’ Software Developer Diversity and Inclusion (SDDI) project [23], the D&I working group of the Community Health Analytics Open Source Software (CHA OSS) project [58], and the D&I initiative at the ASF [22]. These initiatives are tasked to improve diversity in different aspects of OSS such as, diversity in events, inclusive naming, and code of conduct.

While these OSS initiatives are important in their goals of fixing specific issues (e.g., toxic interactions, non-inclusive naming), they are siloed and we lack an understanding of how the community perceives them. Such an understanding is important to realise what is working and what is not to guide the initiatives.

**(RQ2.) How do contributors perceive D&I initiative(s) in their OSS communities?**

We answer these research questions by partnering with the ASF Diversity and Inclusion (D&I) committee [22] because of their interest in understanding their contributors’ perspective of the D&I state and the D&I initiative. We conducted an online survey with 600+ ASF respondents. The survey was designed in collaboration with the D&I committee and the ASF community at large. We followed up the survey results with 11 interviews to get a deeper understanding of contributors’ perspectives. We used mixed methods to analyze the survey and interview responses. Our results shed light on contributors’ background attributes that influence their perception of the state of D&I (see section 4.1) and their perception of the ASF D&I initiative, from its necessity to its efficacy and where it can be improved (see section 4.2).

## 2 BACKGROUND AND RELATED WORK

Diversity in OSS has gained considerable attention in recent years with OSS projects and foundations investing in efforts to create diverse and more inclusive communities. Research has also investigated the topic of low diversity and barriers to contributing to OSS. A majority of which has focused on one diversity aspect–gender, investigating gender distribution of women in OSS [6, 28, 35, 43, 53, 64, 65] and in leadership positions [8], perceptions of women contributors in OSS [38, 80], the impact of gender on productivity [81] and the barriers that women face [47, 51, 57, 74, 83].

For instance, Vasilescu et al. [80] used the gender lens to understand GitHub contributors’ perception of their team and awareness of their teammates’ backgrounds, with gender being the second-most noticed attribute. Other research focused on women’s experience in OSS and support systems in place to increase women’s participation [68, 69]. Singh and Brandon [69] found that only 12 out of 355 OSS websites have ‘women only’ sections and Lee and Carver [38] found that, while some contributors expressed a positive feeling about women’s participation in OSS, some were strongly opposed to their inclusion. Finally, researchers have investigated barriers that women face in tools and technology [47, 54], in getting pull requests accepted [74], and participating in discussions [51, 57].

Research has investigated the experience of ‘older’ contributors in OSS [13, 49, 50]. For instance, Murakami et al. [50] looked at how age can impact code reviews and found that age has no significant effect on code review correctness and efficiency. Morrison et al. [49] investigated the low participation of veteran software developers in OSS and how their contributions differ from those of their younger peers. Morrison et al. [49] results reflected that veteran OSS contributors are less socially motivated than their younger counterparts, which aligns with Davidson et al. [13] findings that older contributors face more social than technical challenges.

The impact of location on pull request acceptance has also been investigated. Furtado et al. [24] found that contributors from countries with low human development indexes face the most pull request rejections. Similarly, Rastogi et al. [63] investigated the top countries with highest and lowest pull request acceptance rates and Rastogi et al. [62] found pull request acceptance rate increases by 19% when the submitter and integrator are from the same country.

Recent studies have started to investigate diversity through the lens of multiple demographic attributes. For example, Prana et al. [57] investigated the difference in gender diversity between geographic regions and found that there has been a small improvement of gender diversity amongst contributors in Northern America and South-Eastern Asia. Catolino et al. [10] investigated community smells through the gender and experience lenses. Ortu et al. [53] also used a dual-lens approach and found that gender diversity increased productivity, while intra-team nationality diversity decreased the level of politeness.

In summary, existing research has largely investigated the topic of diversity using a single lens (e.g., gender, location, or age). Recently,
emerging studies have started investigating diversity by combining
gender with a second lens. Our work complements these works by
taking a multi-dimensional approach to investigate contributors’ per-
ception of the state of diversity using 6 demographic lenses—gender,
education, English proficiency, seniority at ASF, compensation type
and country of residence—that can impact contributors’ experiences
in their OSS community. We also investigate contributors’ perception
of the current D&I initiative, their critiques and suggestions to
improve it.

3 RESEARCH METHOD

We answer the research questions by focusing on a single large OSS
foundation (ASF) to get the perspective of contributors in a single,
mature OSS community instead of the broad OSS world. The ASF
is the world’s largest OSS Foundation with more than 460k people
and more than 350 mature projects and initiatives [20]. The ASF
serves as a good case study because of its relevancy and maturity, its
interest in improving D&I, and our collaboration with the ASF D&I
committee, who worked closely with us in the design and execution
of the study, allowing us to get legitimacy in the eyes of participants
and validate the interpretation of the study results.

We conducted an online survey and follow-up interviews with the
ASF contributors and used mixed methods to analyze the data.

3.1 Survey

Survey design: The goals of the survey were to understand ASF
contributors’ perception of: (1) the state of D&I in their community
and (2) the current D&I initiative, their critiques, and ideas for
improvements. The survey started with six demographic questions
followed by 12 Likert-scale items and one open-ended question (see
survey questions in supplemental material [29]). Table 1 presents
the Likert-scale questions, from contributors’ perception of role
stereotyping (Q1-Q3), their ability to contribute (Q4-Q8), being
represented within the community (Q9, Q10), and their perception
of the code of conduct (Q11, Q12).

We leveraged existing surveys when possible: Questions Q1, Q2,
Q4-Q7 were from Lee and Carver [38] who investigated contributor
perceptions of gender in OSS. 3 of the 6 demographics questions
(seniority at ASF, compensation, and gender identification) were
adapted from the 2016 ASF Committer Diversity survey [21] by
addressing the best practices recommended by the CHAOSS D&I
Working Group [58]. The other three background questions (country
of residence, English confidence, and education level) were adopted
from the Open Demographics Survey [14]. We used Lime Survey
licensed under GPLv2, which is the world’s leading open source
survey software, to conduct the survey.

We then engaged with the ASF D&I committee—composed of
18 experienced contributors with different roles including committ-
ters, Project Management Committee (PMC) members, and board
members—and the community at large to refine and pilot the survey
questions. An open-ended question on the state of the D&I initiative
was included based on the recommendation of the committee.

Survey Data Collection: In collaboration with the ASF D&I
committee, we invited the ASF committers to participate by sending
emails to every ‘apache.org’ email address and shared a link through
the ASF developer mailing lists. Participants were first presented
with a consent page that described the goal of the survey, the data col-
lection and handling policy, and who to contact (see supplemental
material [29]). The survey followed an opt-in strategy where par-
cipants started the survey if they agreed to voluntarily participate
after reading the consent form. The survey was open for 45 days. We
maintained the data confidentiality as per Apache Privacy Policies.
Identifiable information or IP addresses were not collected. When
participants gave email addresses for follow-up interviews, they were
stored separately from the responses.

We received 624 responses, resulting in a response rate of 8.5%
based on a considered total community size of 7500 contributors.
We received 130 responses to our optional open-ended question
about the D&I effort at the ASF. A majority of the 624 respondents
identified as men (88.47%) and 4.55% as women. We grouped to-
gether (see Table 2) the respondents who identified as non-binary,
trans-men/women, prefer to self-describe/not-to-say to preserve their
identity. A majority of respondents were volunteers (61.53%) and
senior contributors with three or more years of experience at the
ASF (73.35%). They resided in 53 different countries located in six
continents with the majority based in North America. Most of the
respondents reported some level of higher education (89.56%).

3.2 Interviews

We conducted follow-up interviews to get a deeper understanding
of contributors’ experience at the ASF and their perspective on the
D&I initiative. From the set of 69 respondents who agreed to being
contacted post-survey, we kept randomly selecting participants to
interview until reaching saturation of information. We ended up
interviewing 11 participants. Table 3 presents the demographics of
our interview participants.

Two researchers conducted the semi-structured interviews: one
led the interviews while the other observed and took notes. Before
each interview, we obtained the participant’s consent to audio record
for transcription purposes. The interview covered the participant’s
experience at the ASF and the mechanisms that help support the
contribution process at the ASF. The interview lasted between 30
minutes to one hour, after which, we thanked our participants and, as
a token of appreciation, we sent them a $50 gift card or its equivalent
in donation to the OSS project or organization of their choice.

3.3 Data Analysis

We used a mixed-method approach to answer our research questions.
We used ordinal logistic regression to analyze the answers to the
Likert-scale questions. For the open-ended question and interview
transcripts, we used open coding.

3.3.1 Quantitative: Ordinal Logistic Regression. To analyze the
Likert-scale responses, we used participants’ reported demographic
attributes (gender, seniority at ASF, English proficiency, compensa-
tion type, place of residence, and education level) as explanatory
variables and ran ordinal logistic regression for each one of the
twelve Likert-scale questions (Q1-Q12). The ordinal logistic regres-
sion model [31] is an extension of the logistic regression model
where the logits of a categorical response are linearly related to the
explanatory variables.

Let $Y$ be our ordinal outcome with $j$ categories (Likert-scale
values), then the ordinal logistic regression model is:
Table 1: Twelve-Likert scale questions about participants’ perception of the state of D&I and the code of conduct. The citation indicates that a question is a replication from Lee and Carver [38].

<table>
<thead>
<tr>
<th>Perception of Questions</th>
<th>Question ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other members of the project see me as a parental figure [38]</td>
<td>Q1. Parental figure</td>
</tr>
<tr>
<td>I am expected to take care of other members of the project more so than is usual [38]</td>
<td>Q2. Care taker</td>
</tr>
<tr>
<td>I feel some members of the community are patronizing to me</td>
<td>Q3. Patronized</td>
</tr>
<tr>
<td>I have an equal chance to get contributions accepted [38]</td>
<td>Q4. Equal chance</td>
</tr>
<tr>
<td>Nothing keeps me from contributing to the project [38]</td>
<td>Q5. No barriers</td>
</tr>
<tr>
<td>I have a solid network of open-source peers [38]</td>
<td>Q6. Network</td>
</tr>
<tr>
<td>It was easy to find a mentor with whom I felt comfortable [38]</td>
<td>Q7. Mentored</td>
</tr>
<tr>
<td>I have a hard time following discussions because of technical jargon</td>
<td>Q8. Tech jargon</td>
</tr>
<tr>
<td>The PMC represents a diverse set of people</td>
<td>Q9. Diverse PMC</td>
</tr>
<tr>
<td>I feel represented in the decision-making group</td>
<td>Q10. Represented</td>
</tr>
<tr>
<td>I was made aware of the code of conduct and how to report violations</td>
<td>Q11. Aware</td>
</tr>
<tr>
<td>I felt safer and more empowered to fully participate in this project because it followed the code of conduct</td>
<td>Q12. Empowered</td>
</tr>
</tbody>
</table>

Table 2: Demographics of the survey respondents (n=624).

<table>
<thead>
<tr>
<th>Demographics</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender: Male</td>
<td>547</td>
<td>86.47%</td>
</tr>
<tr>
<td>Gender: Woman</td>
<td>86</td>
<td>13.53%</td>
</tr>
<tr>
<td>Gender: Non-binary, trans-men/women, prefer to self-describe/not-to-say</td>
<td>41</td>
<td>6.66%</td>
</tr>
<tr>
<td>Seniority at the ASF: Less than 1 year</td>
<td>66</td>
<td>10.66%</td>
</tr>
<tr>
<td>Seniority at the ASF: 1 to 2 years</td>
<td>99</td>
<td>15.99%</td>
</tr>
<tr>
<td>Seniority at the ASF: 3 to 5 years</td>
<td>165</td>
<td>26.66%</td>
</tr>
<tr>
<td>Seniority at the ASF: 6 to 10 years</td>
<td>130</td>
<td>21.08%</td>
</tr>
<tr>
<td>Seniority at the ASF: Over 10 years</td>
<td>159</td>
<td>25.69%</td>
</tr>
<tr>
<td>Education: Ph.D degree</td>
<td>65</td>
<td>10.60%</td>
</tr>
<tr>
<td>Education: Master’s degree</td>
<td>264</td>
<td>42.63%</td>
</tr>
<tr>
<td>Education: Undergraduate degree</td>
<td>200</td>
<td>32.63%</td>
</tr>
<tr>
<td>Education: Technical training</td>
<td>33</td>
<td>5.38%</td>
</tr>
<tr>
<td>Education: High school</td>
<td>30</td>
<td>4.89%</td>
</tr>
<tr>
<td>Education: No formal education</td>
<td>1</td>
<td>0.16%</td>
</tr>
<tr>
<td>English Proficiency: Very confident</td>
<td>348</td>
<td>56.96%</td>
</tr>
<tr>
<td>English Proficiency: Confident</td>
<td>138</td>
<td>22.59%</td>
</tr>
<tr>
<td>English Proficiency: Average</td>
<td>81</td>
<td>13.26%</td>
</tr>
<tr>
<td>English Proficiency: Uncomfortable</td>
<td>13</td>
<td>2.13%</td>
</tr>
<tr>
<td>English Proficiency: Not confident</td>
<td>31</td>
<td>5.07%</td>
</tr>
<tr>
<td>Compensation: Paid work only</td>
<td>84</td>
<td>13.64%</td>
</tr>
<tr>
<td>Compensation: A mix, but mostly paid</td>
<td>153</td>
<td>24.84%</td>
</tr>
<tr>
<td>Compensation: Unpaid only</td>
<td>247</td>
<td>40.10%</td>
</tr>
<tr>
<td>Compensation: A mix, but mostly unpaid</td>
<td>80</td>
<td>12.99%</td>
</tr>
<tr>
<td>Compensation: An equal mix of paid and unpaid</td>
<td>52</td>
<td>8.44%</td>
</tr>
<tr>
<td>Continent: North America</td>
<td>250</td>
<td>40.60%</td>
</tr>
<tr>
<td>Continent: South America</td>
<td>5</td>
<td>0.83%</td>
</tr>
<tr>
<td>Continent: Europe</td>
<td>237</td>
<td>38.45%</td>
</tr>
<tr>
<td>Continent: Africa</td>
<td>7</td>
<td>1.16%</td>
</tr>
<tr>
<td>Continent: Asia</td>
<td>92</td>
<td>14.91%</td>
</tr>
<tr>
<td>Continent: Australia</td>
<td>10</td>
<td>1.61%</td>
</tr>
</tbody>
</table>

\[
\log \frac{P(Y \leq j)}{P(Y > j)} = \beta_0 + \eta_1 x_1 + \ldots + \eta_p x_p; \ j = 1, \ldots, j - 1
\]

where \( P(Y \leq j) \) is the cumulative probability of \( Y \) less than or equal to a specific category \( j \) of the response variable. The model has \( j - 1 \) intercepts denoted by \( \beta_0 \) and one parameter for each explanatory variable (the demographic attributes).

By definition, our dependent variables (Likert-scale responses) fit the first two assumptions of the ordinal logistic regression model: The dependent variables are ordered and one or more of the explanatory variables are either continuous, categorical or ordinal. We checked for the absence of multi-collinearity by generating a covariance matrix of our variables. The last assumption is the proportional odds which ensures that the relationship between each pair of outcome groups is the same, meaning that there is only one set of coefficients, which means that there is only one model. We used the Brant test to check for the proportional odds assumption. We conclude that the parallel assumption holds when the probability (p-values) for all variables are greater than \( \alpha = 0.05 \) and the Omnibus variable, which stands for the whole model, is also greater than \( \alpha \).

In the case where an explanatory variable failed the Brant test, we omitted that variable from the model (see grayed out cell in Tables 4, 5 and 6) to ensure that the model fits all the assumptions. We used R version 4.0.4 and \texttt{polr} in the “MASS” package for the analyses.

Note, for the purpose of this analysis, we divided each demographic attribute (e.g., seniority at ASF, see Table 2) into two segments. For instance, we grouped ASF contributors with low experience (<1 year, 1-2 years) as junior contributors and those with more experience (3-5, 6-10 and >10 years) as senior contributors. We also grouped the participants according to their gender: “gender-majority” for the ASF contributors who identified as men, and “gender-minority” for the ASF contributors who identified as women, non-binary, trans-men/women, prefer to self-describe/not-to-say.

3.3.2 Qualitative: Open Coding. We used open coding to analyze our open-ended question and interview data. First, we analyzed the survey open-ended question on the state of D&I (see survey questions in supplemental [29]). We inductively coded the answers, built post-formed codes as the analysis progressed and associated them to respective parts of the text. At this stage, our aim was to code contributors’ perception of D&I according to their discourse, and not according to any preconceived data. For each interviewee, we identified and coded each excerpt that presented a perception of the D&I. Once a week, the research team met to discuss the emerging categories and to refine their nomenclature. The coding process was conducted by one researcher and discussed with the other researchers until consensus about the resulting codes and quotes.

4 RESULTS

In the following, we present the analysis of how contributors perceive (1) the state of D&I in their community (RQ1, Section 4.1) and (2) the current D&I efforts (RQ2, Section 4.2).

4.1 Perceptions on the state of D&I

To answer RQ1, we analyzed the survey and interviews using mixed methods. Sections 4.1.1 to 4.1.3 detail contributors’ perceptions of different aspects of the state of D&I (see Table 1), and Section 4.1.4
Table 3: Interview participants’ demographics.

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>Seniority at the ASF</th>
<th>Education</th>
<th>English Proficiency</th>
<th>Compensation type</th>
<th>Residence</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Man</td>
<td>1 to 2 years</td>
<td>Undergraduate degree</td>
<td>Average</td>
<td>A mix, but mostly paid</td>
<td>Russia</td>
</tr>
<tr>
<td>12</td>
<td>Man</td>
<td>Over 10 years</td>
<td>Ph.D degree</td>
<td>Very confident</td>
<td>A mix, but mostly unpaid</td>
<td>US</td>
</tr>
<tr>
<td>13</td>
<td>Woman</td>
<td>3 to 5 years</td>
<td>Undergraduate degree</td>
<td>Confident</td>
<td>Unpaid only</td>
<td>Germany</td>
</tr>
<tr>
<td>14</td>
<td>Woman</td>
<td>Over 10 years</td>
<td>Undergraduate degree</td>
<td>Very confident</td>
<td>An equal mix of paid and unpaid</td>
<td>Ireland</td>
</tr>
<tr>
<td>15</td>
<td>*Other</td>
<td>Over 10 years</td>
<td>Undergraduate degree</td>
<td>Very confident</td>
<td>A mix but mostly paid</td>
<td>US</td>
</tr>
<tr>
<td>16</td>
<td>*Other</td>
<td>Over 10 years</td>
<td>Undergraduate degree</td>
<td>Very confident</td>
<td>Unpaid only</td>
<td>US</td>
</tr>
<tr>
<td>17</td>
<td>Man</td>
<td>3 to 5 years</td>
<td>Master’s degree</td>
<td>Not confident</td>
<td>A mix, but mostly unpaid</td>
<td>Italy</td>
</tr>
<tr>
<td>18</td>
<td>Man</td>
<td>Less than 1 year</td>
<td>Master’s degree</td>
<td>Very confident</td>
<td>Unpaid only</td>
<td>Japan</td>
</tr>
<tr>
<td>19</td>
<td>Woman</td>
<td>Less than 1 year</td>
<td>Master’s degree</td>
<td>Very confident</td>
<td>Unpaid only</td>
<td>Germany</td>
</tr>
<tr>
<td>100</td>
<td>Woman</td>
<td>3 to 5 years</td>
<td>Master’s degree</td>
<td>Very confident</td>
<td>An equal mix of paid and unpaid</td>
<td>US</td>
</tr>
<tr>
<td>110</td>
<td>Man</td>
<td>3 to 5 years</td>
<td>Master’s degree</td>
<td>Very confident</td>
<td>An equal mix of paid and unpaid</td>
<td>UK</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Elided to protect identity of these respondents.

illustrates contributors’ perceptions of the extent of the D&I issues. We segment the analysis based on the demographic attributes.

Tables 4, 5, and 6 show the ordinal logistic regression results for the survey questions. For instance, the model “Q2. Care taker” in Figure 1 shows how the demographic attributes explain the level of agreement on that question (see Section 3.3.1). The tables highlight the variables that show statistical significance at \( p < .05 \) or \( p < .01 \) post fixed by * or **, respectively. We also show the response distributions according to the Likert-scale for those demographic attributes that were statistically significant in our models (Figures 1, 2, 3). In the rest of this section, we discuss three different aspects of the contributors’ perceptions, namely role stereotyping, contributors’ ability to contribute, and their feeling of being represented.

4.1.1 Role Stereotyping. Past work has shown that gender can be a major source of bias in how people perceive others, which is linked to prescribing certain roles and traits to women [15, 84], and feminine attributes or qualities displayed by women tending to be devalued [4, 18]. Gender is also closely linked to two basic dimensions that individuals rely on to judge other people—when an individual meets someone, they intuitively make judgments of their warmth and competence [11, 19], subjecting women to role stereotyping. Women are stereotyped to be motherly, warm, and nurturing [9, 32, 46, 48]. Such stereotypes perpetuate behavioral expectations for women to assume the role of the parental figure and community care taker [36, 46].

Questions Q1-Q3 explore three types of role stereotyping, namely parental figure (Q1), care taker (Q2), and being patronized (Q3), and the extent to which individuals from different backgrounds perceive role stereotyping to take place at the ASF.

While women have been associated with the role stereotype of parental figure, we found that contributors who identified as gender-majority and those who associate with the gender-minority reported at similar rates being seen as a parental figure (Q1 in Table 4). This is inline with Lee and Carver [38]’s results. On the other hand, looking at different attributes, we found that senior ASF contributors, were 2.48 times more likely than juniors to feel pigeonholed into this role (Table 4, \( p < .01 \)). This difference can be seen in the distribution of responses in Figure 1—senior contributors agreed at a higher rate with the statement as compared to junior contributors, who disagreed at a higher rate. The results indicate that those more experienced with the ASF are being sought out for advice and mentoring.

Women have also been associated with care taker roles. Our results for this question (Q2 in Table 1) were similar to the aforementioned results. While the odds of contributors who associate with gender-minority feeling that they were expected to take care of others were 1.75 times higher than their counterparts, this difference was not statistically significant. The seniority lens, however, painted a different picture in which the odds of senior contributors feeling expected to take care of others was 1.86 times (\( p < .01 \)) higher than that of juniors (see Table 4, Figure 1). Non-confident
Table 4: Ordinal logistic regression for role stereotyping Likert-scale questions. The highlight corresponds to a statistically significant difference. OR refers to the odds ratio calculated as the exponential of the ordinal logistic regression value.

<table>
<thead>
<tr>
<th>Background attributes</th>
<th>Q1. Parental figure OR Std. err</th>
<th>Q2. Care taker OR Std. err</th>
<th>Q3. Patronized OR Std. err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender-minority vs. Gender-majority</td>
<td>1.09 0.29</td>
<td>1.75 0.29</td>
<td>3.36** 0.29</td>
</tr>
<tr>
<td>Senior vs. Junior</td>
<td>2.48** 0.21</td>
<td>1.86** 0.20</td>
<td>0.86 0.22</td>
</tr>
<tr>
<td>English non-confident vs. confident</td>
<td>0.79 0.23</td>
<td>1.67* 0.21</td>
<td>2.38** 0.23</td>
</tr>
<tr>
<td>College vs. No college</td>
<td>0.81 0.29</td>
<td>0.70 0.27</td>
<td>0.91 0.32</td>
</tr>
<tr>
<td>Paid vs. Non-paid</td>
<td>1.02 0.18</td>
<td>1.11 0.18</td>
<td>0.81 0.21</td>
</tr>
<tr>
<td>North America vs. Outside</td>
<td>0.81 0.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Odds ratio (OR) greater than 1 means that the first segment has greater chances of agreeing with the question than the second. Ratio less than 1 means the opposite. We tested the proportional odds assumption for each question using the Brant test. The grayed out cells correspond to the explainable variables that are omitted from the model to satisfy the proportional odds assumption (see section 3.3.1). A single * means that the variable is significant at $p < .05$ and a ** means that a variable is significant at $p < .01$.

English speakers were also more likely to report being expected to take on Care taker roles (1.67 times more than confident English speakers with $p < .05$, see Table 4). A possible explanation is that non-confident English speakers might have overcome language barriers when contributing to OSS and are sought out by others who face similar challenges.

Finally, we asked whether participants felt Patronized by other community members (Q3 in Table 1). Contributors who identified as gender-minority and who reported being non-confident English speakers felt to be so with odds ratios 3.36 ($p < .01$) and 2.38 ($p < .01$), respectively. This suggests that these minorities do not feel equally respected as their counterparts. The literature reports that those in the minority have to prove themselves and may not be taken seriously [38], perpetuating the social barriers they face [70].

4.1.2 Perceptions about ability to contribute. Past studies have reported that the “so called” meritocracy [51] in OSS is biased and include hostile environments [17] that create barriers to contribution [3, 70, 72, 80]. Lee and Carver [38] recently reported the perceptions of women contributors of their ability to contribute to OSS and the barriers they faced. Therefore, we aimed to understand how contributors with different backgrounds perceive their ability to contribute (questions Q4-Q8 in Table 5).

We found that none of the demographic attributes make a significant difference in participants’ perceptions of having an equal chance (Q4) to contribute and in finding a mentor (Mentored (Q7)). However, a more nuanced story emerges when considering the responses to questions regarding No barriers to contributing (Q5), having a solid network (Q6), and challenges in followingings depending on the demographic attribute. As Figure 2 shows, 74% of senior ASF contributors reported having a solid network of OSS peers with the odds 2.07 times ($p < .01$) higher than their junior counterparts (Table 5). This can probably be explained by senior contributors’ tenure and the connections built over the years, along with the fact that OSS communities can be hard for junior contributors to join [70].

While those who identified as gender-majority were more likely to report (1.92 times) higher than their junior counterparts (Table 5). This can probably be explained by senior contributors’ tenure and the connections built over the years, along with the fact that OSS communities can be hard for junior contributors to join [70].

With regards to Network, the respondents showed mixed feelings depending on the demographic attribute. As Figure 2 shows, 74% of senior ASF contributors reported having a solid network of OSS peers with the odds 2.07 times ($p < .01$) higher than their junior counterparts (Table 5). This can probably be explained by senior contributors’ tenure and the connections built over the years, along with the fact that OSS communities can be hard for junior contributors to join [70].

While those who identified as gender-majority were more likely to report (1.92 times) having a solid network of peers ($p < .01$), only 58% of gender-minority contributors agreed to the statement (Figure 2). This is worrisome since prior literature emphasizes the importance of same gender role models, peers in building a network, and sense of belonging [7, 56, 79].
Table 5: Ordinal logistic regression for ability to contribute.

<table>
<thead>
<tr>
<th>Background attributes</th>
<th>Q4. Equal chance OR</th>
<th>Std. err</th>
<th>Q5. No barriers OR</th>
<th>Std. err</th>
<th>Q6. Network OR</th>
<th>Std. err</th>
<th>Q7. Mentored OR</th>
<th>Std. err</th>
<th>Q8. Tech jargon OR</th>
<th>Std. err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender-minority vs. Gender-majority</td>
<td>0.59</td>
<td>0.34</td>
<td>0.41**</td>
<td>0.29</td>
<td>0.52**</td>
<td>0.29</td>
<td>1.54</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior vs. Junior</td>
<td>1.31</td>
<td>0.24</td>
<td>1.40</td>
<td>0.22</td>
<td>2.07**</td>
<td>0.20</td>
<td>0.96</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English non-confident vs. Confident</td>
<td>1.12</td>
<td>0.30</td>
<td>1.01</td>
<td>0.25</td>
<td>0.48**</td>
<td>0.24</td>
<td>2.58**</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>College vs. No college</td>
<td>0.44</td>
<td>0.49</td>
<td>0.81</td>
<td>0.36</td>
<td>1.15</td>
<td>0.31</td>
<td>1.28</td>
<td>0.32</td>
<td>0.91</td>
<td>0.37</td>
</tr>
<tr>
<td>Paid vs. Non-paid</td>
<td>0.78</td>
<td>0.23</td>
<td>1.01</td>
<td>0.21</td>
<td>0.85</td>
<td>0.20</td>
<td>0.87</td>
<td>0.19</td>
<td>0.70</td>
<td>0.25</td>
</tr>
<tr>
<td>North America vs. Outside</td>
<td>1.12</td>
<td>0.24</td>
<td>1.13</td>
<td>0.21</td>
<td></td>
<td></td>
<td>1.25</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While building a solid network can be challenging for a variety of reasons [70, 73], contributors who are not confident with their English speaking skills bear an additional communication burden. In fact, our results show that contributors who are not confident in English were 2.58 times more likely (p < .01) to report struggling to follow discussions because of technical jargon (see Table 5). The language barrier may impact contributors’ participation in communication channels, thus limiting their chances to build meaningful connections with their peers. Contributors who were confident English speakers were 2.08 times (Table 5, p < .01) more likely to report having a solid network of OSS peers.

**Ability to contribute.** Gender was relevant in the perception of barriers that keeps those in the gender-minority group from contributing or building a solid network of peers. Lack of proficiency in English is also associated with barriers in creating a network or following technical discussions that include jargon.

4.1.3 Perceptions of being represented. The feeling of being represented is important to being productive and satisfied [42]. Hagerty et al. [30] emphasize the importance of personal involvement in a system or environment, so that people feel as an integral part of it. Past work found that certain factors can impede the feeling of being represented, such as the lack of interpersonal relationships in the community [33, 76] and the perception that one’s voice is lost in an environment where the loudest voice prevails [51].

Table 6: Ordinal logistic regression for being represented.

<table>
<thead>
<tr>
<th>Background attributes</th>
<th>Q9. Diverse PMC OR</th>
<th>Std. err</th>
<th>Q10. Represented OR</th>
<th>Std. err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender-minority vs. Gender-majority</td>
<td>0.68</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior vs. Junior</td>
<td>0.86</td>
<td>0.20</td>
<td>1.67**</td>
<td>0.22</td>
</tr>
<tr>
<td>English non-confident vs. Confident</td>
<td>0.97</td>
<td>0.22</td>
<td>0.60**</td>
<td>0.25</td>
</tr>
<tr>
<td>College vs. No college</td>
<td>0.99</td>
<td>0.28</td>
<td>0.76</td>
<td>0.36</td>
</tr>
<tr>
<td>Paid vs. Non-paid</td>
<td>0.52**</td>
<td>0.18</td>
<td>0.93</td>
<td>0.22</td>
</tr>
<tr>
<td>North America vs. Outside</td>
<td>0.87</td>
<td>0.19</td>
<td>1.35</td>
<td>0.24</td>
</tr>
</tbody>
</table>

We asked contributors to rank their agreement on whether they felt that ASF has a DIVERSE PMC (Q9) and whether they felt being REPRESENTED (Q10) (see Table 1).

For DIVERSE PMC, we only found statistically significant difference when analyzing paid vs. non-paid contributors. More than half of volunteer contributors reported that the PMC was diverse as compared to 40% of the paid contributors (see Figure 3). Indeed, as Table 6 shows, the odds of non-paid contributors agreeing to DIVERSE PMC statement were 1.92 times higher than that of their paid peers (p < .01).

This highlights a possible tension in the relatively new hybrid-model of OSS, where paid employees making contributions do not perceive that the PMC is diverse enough. Since none of the other factors played a significant role, we assume here that the paid contributors did not see themselves being represented in the PMC.

Regarding the question about feeling REPRESENTED in the decision-making process, the seniority and the English confidence lens revealed differences. Senior ASF contributors were 1.67 times more likely (p < .05) to agree that they feel represented in decision-making processes. Similarly, contributors who are confident in their English speaking skills were 1.67 times more likely (p < .01) to agree than their counterparts (Table 6). In fact, 76% of participants who were confident in speaking English agreed to REPRESENTED as compared to 56% of those who were not confident in English. These findings are in line with past work that has found language and nationality to be a barrier in OSS [53, 57, 65].

**Being represented.** English speaking skills, compensation, and seniority attributes were relevant in participants’ perceptions of feeling represented in the PMC or decision making.
4.1.4 Perceptions of extent of D&I issues. To investigate contributors’ perceptions of D&I issues at the ASF, we qualitatively analyzed the open-ended survey question and interviews. Figure 4 presents a summary of the results, showing the contrasting perceptions of the extent of the D&I issues.

At one end of the spectrum, some contributors reported D&I issues are non-existent and that the community was diverse. For instance, S795 reported that “the ASF consists of maximum diversity” and that “diversity of physical attributes are both invisible and largely ignored [in] mailing lists.” Similarly, S36 said that “diversity has rarely been an actual issue” and that it is “used for political pressure that compromises free collaboration.” These responses reflect the belief that some feel D&I is an unsubstantiated, political issue.

![Figure 4: ASF contributors’ perspective on the extent of D&I issues.](image)

Some preferred to stay out of discussing D&I, such as S517, who shared “I have therefore withdrawn my active participation in all of these types of usually quite ‘loud’ discussions.”

On the other end of the spectrum, contributors reported D&I issues are prevalent, because of their first-hand experiences or that of others. Multiple participants who were in gender minority groups reported being affected by a biased committer selection process, as I5 shared: “it took a very, very long time for me to become a committer relative to other people who became committers...it’s also really, really hard to add other non-men as committers.” I3 (woman) had a similar experience explaining how it was hard for non-code contributors to become committers.

Some contributors described the community as an “old boys club” (S618) and reported noticing D&I issues. For instance, S259 reported: “I never experienced real negative emotions towards me, but I have seen people saying stupid things”. Some of them felt that the ASF was not a diverse community, when considering gender: “Of the 54 committers and 33 PMC members of [project name], just one (that I am aware of) identifies as a woman”(S753); or race: “the number of African American PMC members is even more unrepresentative of the US population”(S821).

![Figure 5: Responses to the 5-points Likert-scale item for ‘code of conduct’ (Q12 was significant).](image)

4.2 Perceptions on D&I initiative at the ASF

In this section, we answer RQ2 by discussing participants’ perceptions of the ASF D&I initiative. We start by looking at how they perceive the code of conduct specifically; then, we analyze their general perceptions of the ASF D&I initiative.

4.2.1 Code of conduct. A Code of Conduct (CoC) defines the expected behavior for the project’s community, which, when adopted, can help foster a positive social atmosphere [1]. Multiple communities, such as Django, Python, Ubuntu, Contributor Covenant, and Geek Feminism, adopted the code of conduct early on and contributed to its adoption in other communities [78]. It has now become one of the most adopted D&I effort in OSS [41]. The goal of questions Q11. Aware and Q12. Empowered (see Table 1) was to evaluate the perceptions of contributors about the ASF CoC.

Therefore, we asked the contributors whether they were aware of the ASF CoC and ways to apply it. There were no significant differences between segments across the different demographic attributes. The majority of contributors indicated that they were aware of the CoC (see Likert-scale figures in supplemental [29]). This is probably because CoCs are now widely adopted in most OSS communities.

![Table 7: Ordinal logistic regression for the code of conduct.](image)

We then followed up by asking contributors whether the presence of the CoC helped them feel safer and more empowered within their community. The compensation lens (paid vs. non-paid) was the only attribute that showed a statistically significant difference. 47% of non-paid contributors felt safer and more empowered because of CoC, as compared to 35% of paid contributors (see Figure 5). In fact, in our survey, non-paid contributors were 1.67 times more likely (p < .01) to agree that CoC made them feel empowered (Table 7).

**Code of conduct.** Participants are aware of the CoC and non-paid contributors reported being empowered by it.

4.2.2 D&I initiative at the ASF. As OSS communities have become aware of inclusivity problems, they have created different D&I initiatives and committees [22]. While this is a great start, it can backfire if the efforts are fragmented or incorrectly implemented. It is thus crucial to understand, early on, how contributors perceive these initiatives to ensure their success.
The open-ended survey question was focused on collecting contributors’ thoughts about the D&I initiative at the ASF. This question was answered by 130 (out of 624) participants—105 identified as men (gender-majority) and 25 were categorized as gender-minority. We triangulated the survey results with the interview results. Our analysis revealed 9 categories about their feedback on the D&I initiative and 4 categories of ways to improve this initiative. We built a perception model backed by the data, composed of these two perspectives (see Figure 6).

**Feedback on current D&I initiative.** Participants reported eight negative (or criticisms) and one positive (or supportive) perceptions.

The participants who provided negative feedback considered that the ASF and contributors are not aligned with the same goal. On one side, participants reported that the community is using D&I to defend their own interest, which can be a distortion manifested by either political or personal announcements—“using the cloak of diversity to fight their own fights” (S517). On the other side, participants voiced frustrations, raising that contributors are pushing back the initiative and “fighting to block D&I efforts at the ASF” (S38).

One possible reason for the lack of contributors’ support can be that the D&I initiative is not addressing all D&I issues, as “gender is only a tiny part of the picture” (S340). Participants felt that the ASF is applying a limited view of the many diversity aspects by considering only gender and dismissing “other factors that one would assume goes into the make-up of diversity” (S181). Another reason keeping contributors from being supportive is when they consider that the D&I initiative is not well-managed and organized. One participant reported that “nobody attends events more than once a week” and that the event frequency options were unrealistic, even for a person who “attends events more than almost anyone” (S340). Other participants mentioned that discussions occur in general mailing lists without adequate moderation, which makes these discussions “noisy, less focused, non-constructive, emotionally driven, and empathy-lacking” (S52, S353, S640). Thus, disorganization can be a problem if the ASF has D&I goals, but fails to implement it well.

A consequence of a lack of support from stakeholders is that the initiative is creating polarization and may “cause divisions” (S239) between contributors in minority groups and those who are not. S853 reported feeling “oppressed by sheer volume and intolerance in D&I movements in ASF...where anyone not in minority currently under protection is ganged upon and perceived as inherently inferior...feel unwelcome despite being caucasian heterosexual male”.

We found contradiction about the value of this initiative. Some contributors stated that the initiative is unnecessary or “irrelevant to ASF” (S745). This was echoed by other respondents saying that D&I is “getting way too much attention” (S492), and the ASF should not “spend over much time on D&I” (S392). Meanwhile, other participants considered that the community is not taking D&I seriously (S32, S703, S843, I10) and is not dedicating “a lot of momentum around the initiatives” (I10). This points that more effort is needed in building awareness in the broader ASF community about D&I issues and their impact on minority groups, as well as mechanisms to create allies from the majority groups. However, some participants provided positive feedback stating that the community is making efforts to improve D&I (S38, S39, S66, S94, S135, S157, S259, S349, I5, I11), commending the survey and the mentorship programs such as, Outreachy and Google Summer of Code (I5, I11).

**Ways to improve the D&I initiative.** Participants provided four ideas for the ASF to address two (out of the eight) negative feedback categories and improve the D&I initiative.

To address broader D&I issues, participants suggested the community to also consider other diversity aspects besides gender, like the contributors’ sexual orientation (S142, S622), skin color (S142), political affiliation (S142), religion (S142), access to technology (e.g., Internet, phones, computers) (S142), ethnicity (S181, S340, S398, S622, I2), economic factors (S142, S340), and languages spoken (S142, S181, S340). In order to put D&I into practice and have the community take it more seriously, participants claimed the community needs to be more proactive by “increas[ing] rates of contributor[s] from minority groups”, for example “getting more contributors who are not just males” (I11). Another idea was to have minority groups as speakers in events by having specific “calls for speaker proposals for people who are female or from lower ethnic groups” (I2, I11). Finally, participants reported that the ASF should be more inclusive to contributors who are already in, so “before trying to get more people, start with the people that are already in the community and help them to feel more included” (S334).

**D&I Initiative.** There are mixed feelings, while some appreciate the efforts that the ASF has undertaken to improve D&I, others are skeptical, feeling that current efforts are focused mainly on gender, is only ‘lip service’, and cause polarization.
5 DISCUSSION

Diversity is not only about gender. Participants in our study perceived that the ASF initiative to improve diversity is mostly focused on gender, missing other struggles of individuals with other diversity aspects such as, language skills, race, ethnicity, age, and other attributes that differentiate people (see Section 4.2.2). For example, contributors who were not confident English speakers were more likely to struggle in following technical discussions. This is in line with Steinmacher et al. [70]’s study, which showed that non-native English speakers face communication barriers. Considering that non-native English speakers can be discriminated against [13], they can lose their self-confidence, feel disengaged, and quit. This raises the importance of taking into account this minority (20% of our survey respondents are non-confident English speakers) and the language barrier they face which can hinder their participation and sense of belonging [70]. But proficiency in English is only one attribute that creates barriers—a recent work compared the acceptance of contributors from countries with different levels of human development [24]—research needs to also focus on other aspects of diversity. For instance, given the new landscape of OSS [71], projects are now counting on a mix of paid and volunteer contributors working either in standalone OSS projects or open source arms of commercial companies [44, 66], resulting in a variety of motivations and pathways for contributing [26, 40, 79, 82]. Therefore, compensation type is also a diversity attribute that needs to be studied. The hybrid OSS landscape is different from traditional OSS structure and philosophy as well as different from corporate settings studied by management literature [25, 52, 55]. Here, we take a first step in this direction.

Let’s talk intersectionality. By using more than one lens, our results shed light on potential intersectionality effect among different background attributes. Intersectionality can heighten the perception of being patronized, the expectation of being represented and the difficulty of building a network. Our results show that those contributors in the gender-minority group who are not confident in English report feeling patronized by their community (see Table 4). Other analyses could be done to understand how these factors interact. For example, how about gender minority contributors who are also not confident in English? Is their feeling of being patronized exacerbated by the interplay of stereotypes associated to both gender and English proficiency? Since diversity is a multidimensional construct, a unidimensional view does not paint a complete picture. We need to start looking at diversity from a multidimensional perspective, one that investigates the interplay between the different background attributes [2] and the consequences of such interplay on the D&I initiative.

More recently, research work in computing and collaborative work started investigating the intersectional experience of Black women [16, 61]. Similarly, recent work in HCI demonstrates how intersectionality can be applied to designing technology [60] and used as a model for exposing oppression systems and understanding the inner-works of power within the HCI field [16, 59, 77]. These works can guide future research on intersectionality in OSS.

Forming the full picture. Gender bias has been studied in different context in OSS showing that contributors who identify as women are subject to stereotypes, biases and social barriers [34, 37, 74]. While prior research has found women to be seen as motherly, warm, and nurturing [9, 32, 46, 48], our data showed that women (in our dataset) did not feel pigeonholed into these stereotypes (Section 4.1.1). These findings align with prior empirical work by Lee and Carver [38]. A possible explanation, as pointed by previous research on women acceptance rate [74], is survival bias. Given that our respondents are active contributors, most of them with 3+ years in the ASF (73.35%), our results may reflect the view of those who survive the barriers to contribution and biases. This points to a call for studies that focus on understanding the perception of those individuals from minority groups who left the community and those who do not think that contributing to OSS is feasible for them.

Hand-in-hand with contributors to define the D&I initiative. Our results showed that some contributors reported D&I to be a real problem, while others indicated that D&I should not receive much attention. Besides, while some contributors support and appreciate the D&I initiative implemented at the ASF, others have several critiques about the focus, management, and consequences of the initiative. This highlights the importance of making the community aware of the state of D&I and the impact that the lack of D&I can cause to people from the many different minority groups. This also highlights the importance of investing in designing and creating an appropriate D&I initiative that is based on the contributors’ needs and refined by getting feedback from different stakeholders. Further, why this initiative is created and how it would be implemented should be communicated back to the community with a process to collect constant feedback. OSS community managers can use the GQM+Strategies approach [5], and start by understanding the community needs and defining the D&I goals to be achieved, then cascading those goals into metrics and strategies. The use of metrics can help keep track and control the success or failure of the initiative and associated D&I goals through a measurement system.

6 THREATS TO VALIDITY

As any empirical study, this work also has limitations and threats to validity, which we present in this section.

Construct validity relates to the constructs used in our study. Incorrect questions in our survey or interviews can lead to incorrect measurements. To mitigate this threat, we reused questionnaires where we could [21, 38] and collaborated with the ASF D&I committee to design the instruments. Another threat can arise in the qualitative analysis process. To avoid misinterpretation in the qualitative coding of the data, we used the constant comparison method. As new codes emerged, we compared it with the existing code set and met frequently with the research team to discuss and clarify the codes. The code set generated from the survey results were then compared with the interview data.

Internal validity is related to our ability to capture the reality as close as possible, which in our case is accurately capturing contributors perspective of the state of D&I and the D&I initiative. The first threat might arise if we have a biased sampling of the ASF contributors. We believe this limitation was low as we deployed the survey widely receiving 600+ respondents who represented a wide set of demographics. We also leveraged mixed method and survey and interview data to better understand contributors’ perspectives.

Conclusion validity We focused our study on current ASF contributors. This might result in survival bias, potentially painting a
more “optimistic” picture. The perceptions of contributors who tried contributing and quit may differ. Another potential threat can be the imbalance in the number of participants from each background attribute, which might have resulted in fewer results were we found statistically significant differences between groups. Additionally, our results can suffer from self-selection biases and no-response bias as participation was voluntary we might have missed out on the point of view of contributors who did not choose to respond to the survey. It is also possible that those who were not proficient in English or those in the minority group knowing that the survey came from the ASF might not have volunteered.

External Validity Even though, the ASF is one of the largest OSS foundations, this study case might not be generalizable to all OSS projects and foundations. We believe however, that our results are representative since the ASF includes 300+ projects across multiple domains and thousands of contributors. The survey was answered by 624 ASF contributors resulting in a response rate of 8.5% based on a considered total community size of 7500 contributors.

7 CONCLUSION
In this work, we investigated contributor perceptions of the state of D&I and the D&I initiative at the ASF through 6 different lenses namely gender, education, English proficiency, seniority at ASF, compensation type and country of residence. Our findings show that in addition to gender, others lenses matter in contributors’ perception of the state of D&I (e.g., seniority at ASF, English proficiency, compensation type). Diversity is a multidimensional construct and research needs to use a broader lens to understand the state of D&I in OSS. Future work needs to also start to investigate how intersectionality plays a role for those who are in the minority.

Making substantive changes to a large, decentralized, heterogeneous community like the ASF is not easy. Some felt that D&I issues are a made-up construct and diversity attributes are invisible in OSS. While others recognized that “the Foundation operates largely as an ‘old boys club’”(S618). While some contributors appreciate the ASF’s D&I effort, some were more critical about the scope of the initiative, the polarization it creates, and the way it was put into action. “It [the D&I initiative] is in its infancy at the ASF, but I am glad it is happening”(S349).

ACKNOWLEDGMENTS
We thank all our survey respondents and our interview participants for their time and insight. This work is partially supported by the National Science Foundation (grants 1901031, 1815486, 1815503 and 1900903) and CNPq (grant #313067/2020-1).

REFERENCES
Andrew Kofink. 2015. Contributions of the under-appreciated: gender bias in social sciences communications.


Hypatia, 9, 2 (1994), 123–133.


Chaos/开支 diversity-inclusion/tree/master/focus-areas. [Online; accessed 2020-12-30].


Igor Steinmacher, Marco Aurelio Graciolito Silva, Marco Aurelio Gerosa, and David F Redmiles. 2015. A systematic literature review on the barriers faced by newcomers to open source software projects. Information and Software Technology.


