Perceived diversity in software engineering: a systematic literature review

Gema Rodríguez-Pérez[1] (ORCID) **· Reza Nadri[2] · Meiyappan Nagappan[2]**

Abstract

We define perceived diversity as the diversity factors that individuals are born with. Perceived diversity in Software Engineering has been recognized as a high-value team property and companies are willing to increase their efforts to create more diverse work teams. The current diversity state-of-the-art shows that gender diversity studies have been growing during the past decade, and they have shown the benefits of including women in software teams. However, less is known about how other perceived diversity factors such as race, nationality, disability, and age of developers are related to Software Engineering. Through a systematic literature review, we aim to clarify the research area concerned with perceived diversity in Software Engineering. Our goal is to identify (1) what issues have been studied and what results have been reported; (2) what methods, tools, models, and processes have been proposed to help perceived diversity issues; and (3) what limitations have been reported when studying perceived diversity in Software Engineering. Furthermore, our ultimate goal is to identify gaps in the current literature and create a call for future action in perceived diversity in Software Engineering. Our results indicate that the individual studies have typically had a gender diversity perspective focusing on showing gender bias or gender differences instead of developing methods and tools to mitigate the gender diversity issues faced in SE. Moreover, perceived diversity aspects related to SE participants' race, age, and disability need to be further analyzed in Software Engineering research. From our systematic literature review, we conclude that researchers need to consider a wider set of perceived diversity aspects for future research.

Keywords Perceived diversity · Systematic literature review · Software engineering

✉ Meiyappan Nagappan
  mei.nagappan@uwaterloo.ca

  Gema Rodríguez-Pérez
  gerope@mail.ubc.ca

  Reza Nadri
  rnadri@uwaterloo.ca

[1] University of British Columbia, Kelowna, Canada

[2] University of Waterloo, Waterloo, Canada

1 Introduction

Team diversity refers to the individual differences between members of a team. It can be present on various dimensions such as value diversity (e.g., beliefs, goals, values), information diversity (e.g., experience, knowledge, background), and social diversity (e.g., gender, age, race) (Jehn et al. 1999). Inspired by the social diversity dimension (Jehn et al. 1999), in this paper we define "perceived diversity" as the perceived internal diversity aspects that individuals are born with (e.g., gender, age, race, and nationality). The perception of these diverse aspects from a person can activate the prejudices, stereotypes, or biases that other individuals might have against that person (Evans 2003; Heiniger and Mercie 2018; Bertrand and Mullainathan 2004). For example, when the person A believes that they have perceived these diverse aspects from person B, it might activate whatever biases person A might have about people with that diversity aspects.

Some of the problems associated with diversity in working teams can be explained by the *Similarity-Attraction theory* and the *social categorization perspective*. The *Similarity-Attraction theory* postulates that individuals working in groups prefer working with others similar to them (Byrne 1971). The *social categorization perspective* predicts that group members are more prone to like, trust, and co-operate with similar others (Homan et al. 2007). Many companies are aware of the lack of diversity in their organizations and this has resulted in a wave of efforts to increase the diversity of employees in worldwide tech organizations.[1] [2] [3]

The diversity of a team is essential beyond ethical reasons. In the past decades, diversity has increasingly been recognized as an essential feature of a team (Page Scott 2007). For example, without gender diversity, teams may focus more on doing things faster and less on doing new things (Østergaard et al. 2011); and without race or nationality diversity, teams might not benefit from multiple points of view, availability of knowledge and skills, and constructive conflict (Shachaf 2008). Since Software Engineering (SE) activities involve teams of developers, it is interesting to understand the way how members of SE teams interact with their similar and dissimilar peers when developing software products.

Recent studies have shown that SE teams have problems associated with perceived diversity aspects in both industrial and Open Source Software (OSS) environments. Blincoe et al. found that 12% of men working in IT teams have admitted to having a conscious bias against women [SLR(Blincoe et al. 2019)].[4] Davidson et al. found that OSS contributors have witnessed discrimination towards others, especially against non-native English speakers and women [SLR(Davidson et al. 2014)].

However, while researchers have made progress in showing that gender diversity increases innovation and productivity (Østergaard et al. 2011; Tourani et al. 2017) [SLR(Vasilescu et al. 2015)], reduces turnover and conflict within teams [SLR(Vasilescu et al. 2015)], and produces more user-friendly software [SLR(Burnett et al. 2016)], comparatively less research has examined other perceived diversity aspects. Hence, through a Systematic Literature Review (SLR) we analyze the previous studies that have been published until May 2020 on the topic of perceived diversity in SE. **Our aim** is to identify what has been studied and discovered related to perceived diversity aspects in SE, what means

---

[1]https://diversity.google/

[2]https://www.microsoft.com/en-us/diversity

[3]https://diversity.fb.com/read-report/

[4]Following the guidelines described in Hall et al. (2011), publications included in our Systematic Literature Review (SLR) are cited using the following format: [SLR#ref]. Other citations use the normal format [#ref].

have been proposed to mitigate perceived diversity issues, and what needs to be done in SE to increase the knowledge about perceived diversity. Our results help researchers be aware of all the work that has been done and all the work that needs to be done so that they take further actions against these perceived diversity issues. Also, our results may help practitioners to identify which tools and methods they can use to make effective decisions on assessing the inclusivity and diversity of both their teams and their products.

There are five literature reviews for some perceived diversity aspects in SE previous to our SLR. Three reviews focused on gender diversity (Canedo et al. 2019; Spichkova et al. 2017; Silveira and Prikladnicki 2019), one on cultural diversity (Fazli and Bittner 2017), and one about the characteristics of diversity in SE (Menezes and Prikladnicki 2018). This last literature review is similar to ours but it was a work in progress that included only 29 studies. Our SLR differs from the previous SLR because (i) we analyze studies that address not only gender and cultural – in our SLR, culture is a sub-dimension inside nationality – diversity but other software perceived diversity aspects (e.g., age, and race), and (ii) we summarize the outcomes studied and identify gaps in the literature. We make four significant contributions by presenting:

– A summary of 131 studies addressing perceived diversity in SE. Researchers in SE can use these studies as the basis of future investigations into perceived diversity.
– A subset of 41 studies on inclusivity that report tools, models, and practices proposed to help assess perceived diversity in SE. Researchers and practitioners can use these results to assess perceived diversity in their teams.
– An identification of important limitations and threats to validity when analyzing perceived diversity. Researchers studying perceived diversity can use this outcome to mitigate or avoid known threats in their studies.
– An identification of gaps in the current perceived diversity research that suggests areas such as race for further investigation. Researchers might use the gaps identified to carry out further research.

This paper is organized as follows. Section 2 presents the perceived diversity background and previous related SLR studies published in the SE field. In Section 3, we present our SLR methodology, research selection criteria, and process. Section 4 shows the results of applying our assessment criteria to 131 studies and synthesizes our results. Section 5 discusses the results associated with the different perceived diversity aspects. Section 6 identifies the threats to the validity of this SLR. Finally, in Section 7 we summarize and present our conclusions.

2 Background and Related Work

2.1 Perceived Diversity Aspects

Software Engineering (SE) concerns not only technical aspects of how to build and develop software but also concerns human aspects (Lenberg et al. 2015) as software products are created by developers (Hongyun et al. 2009).

When software developers are working in face-to-face teams, they can perceive perceived diversity aspects (e.g., gender, race, and nationality) from other members of the team. Similarly, when software developers are working in online teams, they can infer the perceived diversity aspects from others based on the developers' names, photos, pronouns,

English fluency in their comments, and their online profiles. For example, developers in online communities are aware of the gender, ethnicity, and age of most of their team members [SLR(Vasilescu and Filkov 2015)].

For this systematic literature review, we considered four perceived diversity aspects that we believe to be the most relevant for the Software Engineering community. **Gender** is related to an individual's own gender identity typically as man, woman, or non-binary (Usher 2006). In our SLR, gender-related studies encompass topics such as gender identity and gender perception. **Age** is the biological age of people. In our SLR, age-related papers study the age of developers. **Race** is a social construct used to categorize diverse populations, it is linked with physical characteristics such as skin color. In our SLR, race-related papers encompass previous studies that analyzed different racial and ethnic groups such as Black, Hispanic, Asian Pacific Islander, or White, among others. **Nationality** is the country of origin, language, or culture that characterizes social groups. The nationality papers in this SLR study the communication between teams, stereotypes that can trigger bias against developers, and national beliefs that can influence SE practices and practitioners.

2.2 Related Work

Previous to our SLR, five literature reviews have analyzed gender diversity in SE.

Canedo et al. (2019) conducted a systematic literature review in the context of Open Source Software (OSS) communities. Their study analyzed 24 papers intending to find factors that could help increase the engagement of women to contribute to OSS. This SLR identifies some factors among women that cause a lack of interest, and some possible solutions. The main findings of Dias Canedo et al.'s SLR (2019) indicate that women are underrepresented in the OSS community with less than 10% of the total developers. Furthermore, they stated that the reason for this under-representation may be associated with women's workplace conditions, which may support men's gender bias. Our SLR expands Canedo et al.'s SLR (2019) because it focuses not only on women and OSS but on five perceived diversity dimensions in OSS and industry.

Spichkova et al. (2017) conducted a literature review of gender diversity aspects within the field of Software Architecture. The authors only found one paper published in this field. Therefore, their results indicated a big gap in Software Architecture literature as the majority of the publications on gender diversity aspects within SE were not focused on this field.

Silveira and Prikladnicki (2019) conducted a systematic mapping study in the context of SE and Agile Methodologies to identify how diversity is discussed in SE. They identified 221 qualified papers in their systematic mapping. These qualified papers studied Gender (129), LGBTQI (2), Age (10), Race (7), Cultural (67), and Disabilities (1). Our SLR differs from the systematic mapping because our study largely explains the outcome of the papers. While Silveira and Prikladnicki (2019) report the frequency of papers per year, per diversity, and per conference, we study the outcome, the methods and tools proposed, whether these studies show any type of bias or inclusivity-efforts and the limitations and threats found in previous studies. Furthermore, we have investigated why Silveira and Prikladnicki (2019) systematic mapping identifies 221 qualified papers whereas our SLR identifies 131 qualified papers. After comparing both outcomes we have noticed that some of the studies in Silveira and Prikladnicki (2019) did not qualify for the inclusion criteria in our SLR as the publications were not specifically from SE journals or conferences. Another difference is that the systematic mapping includes papers related to software management and agile methodologies but our SLR does not.

Fazli and Bittner (2017) conducted a systematic literature review to identify the impact of national cultural factors on collaborative software development approaches. Fazli et al.'s SLR (2017) analyze 20 papers and its results indicate that there are differences in communication, interaction, and decision-making during collaborations. Such differences may cause problems in distributed projects because of some cultural ignorance issues.

The closest study to ours is the SLR by Menezes and Prikladnicki (2018) which aims to analyze the characteristics of diversity in SE through a systematic literature review. Their SLR included 29 papers and their findings relate the types of perceived diversity with a SE domain. Our SLR includes the 29 papers from Menezes and Prikladnicki's SLR (2018) and goes further to identify not only the SE domain, but also the type of study, the tools, methods, and processes proposed. Furthermore, our SLR also summarizes the outcomes studied so far.

3 Systematic Literature Review Methodology

We conducted a Systematic Literature Review to review the literature on the perceived diversity in Software Engineering. We followed the SLR approach identified by Kitchenham and Charters (2007). This approach presents appropriate guidelines that have been derived from guidelines in medical research and adjusted to suit software engineering. Besides these guidelines, we also followed the structure of Hall et al.'s work (2011) for conducting the review and presenting the results.

All the steps of our SLR (i.e., research questions, inclusion and exclusion criteria, and research selection process) are documented in this section. The research questions help us formulate the aim of this systematic review. The inclusion and exclusion criteria help us assess each potential study. The research selection process helps us find as many studies relating to the research question as possible. The related data are available online[5] for further validation and replication.

3.1 Research Questions

This SLR aims to summarize the existing evidence concerning the perceived diversity in Software Engineering. Our purpose is to help researchers and practitioners to identify what has been studied so far, what has been proposed to help foster perceived diversity in SE, and what are the limitations and threats faced in previous SE perceived diversity studies.

– **RQ1: What are the types of perceived diversity research studies in SE?**

   Motivation With this question our goal is to depict an overview of the perceived diversity state-of-the-art in SE and help researchers and practitioners to better understand what has been studied and what are the results of these studies. Based on the results from RQ1, we can identify gaps in the current literature and create a call for future action in perceived diversity in SE. For that, we identified the frequency of the papers of each perceived diversity aspect per year, the most frequent venues which publish perceived diversity studies, and the number of papers published for each perceived

---

[5]https://vault.cs.uwaterloo.ca/s/Zc5qDwoDmKzA7gG

diversity aspect. Our primary contribution in this RQ is the summary of the outcomes for each perceived diversity aspect in SE.

– **RQ2: What the perceived diversity research has proposed to foster diversity-inclusiveness in SE?**

Motivation With this question we want to identify and describe the tools, methods, and practices proposed by previous studies. For that, we identified the papers that describe and prove inclusivity efforts and summarize their outcomes. Researchers can use these results to further study perceived diversity issues. Furthermore, practitioners may use these results to assess or foster diversity and inclusiveness in SE.

– **RQ3: What are the challenges faced by SE researchers when studying perceived diversity?**

Motivation With this question we aim to help researchers identify the threats to validity and limitations that previous studies have faced when studying perceived diversity in SE. That way, researchers examining perceived diversity can easily identify whether their study presents known threats to validity. For that, we summarized the threats and limitation sections of the papers included in the SLR.

3.2 Inclusion and Exclusion Criteria

The inclusion criteria for studies to be included in our SLR are:

– We included studies that analyze any perceived diversity aspect related to SE.
– We included studies that analyze any perceived diversity aspect related to SE in student teams that simulate SE activities.
– We included studies that are written in English.
– We included studies published as either Journal papers, conference proceedings, or Workshop proceedings.
– We included studies published in computer science venues.
– We included studies published only in mature venues where SE papers are typically published.
– We included peer-reviewed studies.

The exclusion criteria for studies to be included in our SLR are:

– We removed book chapters, work in progress papers, poster, and Master or Ph.D. theses because they may be published as papers as well and we did not want to double count.
– We removed short versions of long version papers to mitigate duplicate results.
– We removed discussion papers, i.e., those papers studying the perceived diversity in a psychological, social, or philosophical context because they were not related to SE activities.
– We removed software engineering education papers that describe studies to increase the gender diversity in SE courses. For example, papers that explained the results of implementing certain criteria, frameworks, models, or practices at Universities. We are aware that this is an important discussion that we are not including in this SLR and that should be analyzed in future research, but this SLR focuses on perceived diversity aspects when engineering software instead of when studying software engineering.
– We removed papers that analyze perceived diversity from a software management perspective (i.e., planning, scheduling, resource allocation, execution, delivery) and a team

management perspective (i.e., communication) because these papers were not related to perceived diversity in the software development.

– We removed papers that describe machine learning techniques for identifying some perceived diversity aspect in people because they are not related to SE per se.

– We also removed literature review papers because they are synthesizing the results from previous studies.

3.3 Research Selection Process

We looked in the IEEE database and Google Scholar. We use Google Scholar because it has high accuracy in locating citations and it provides more results than other search engine databases. We also use the IEEE database because it is a solid academic source. To identify the tentative articles we used the following search string. We have limited the search of the strings only to the title:

Google Scholar:
1. allintitle: diversity ("software engineering" OR "software development" OR "open source" OR floss OR foss)
2. allintitle: (gender OR female OR woman OR women OR sexism) AND ("software engineering" OR "software development" OR "open source" OR floss OR foss)
3. allintitle: (ethnicity OR ethnic OR race) AND ("software engineering" OR "software development" OR "open source" OR floss OR foss)
4. allintitle: (age OR adults OR young) AND ("software engineering" OR "software development" OR "open source" OR floss OR foss)
5. allintitle: (diversity OR gender OR LGBTQ OR race OR ethnicity OR culture OR age OR Location) AND ("software engineering" OR "software development" OR "open source" OR floss OR foss)

IEEE database:
("Software engineer" OR "Software Engineering" OR "software development" OR "software developer") AND (diversity OR gender OR LGBTQ OR race OR ethnicity OR culture OR age OR Location)"

The search in Google Scholar and IEEE database resulted in a total of 5671 papers. During the cleaning process, we removed 2477 entries because they were duplicated entries, master and Ph.D. thesis entries, unrelated entries (e.g., keynotes, panels, welcome message), and entries from unrelated venues to software or engineering computing. During the first interaction, we read the title and the venue of publication and removed 2822 entries based on the exclusion criteria. When the title was insufficient to decide, we looked over the abstract. During the second iteration, we read the abstract of the tentative studies and removed 138 papers based on the exclusion criteria. When the abstract was insufficient to decide, we looked over the paper. On the third iteration, we examined the tentative papers thoroughly and removed 168 papers based on the exclusion criteria. On the fourth iteration, we investigated the references to the resulting set of publications – a process known as snowballing. This iteration was repeated until no newer relevant studies were found. We added 51 new

papers to the set of publications. In the fifth iteration, we identified who are the prolific authors in perceived diversity aspects from the set of publications, and we contacted them to assess the 117 publications. Based on the prolific authors' recommendations we included another 14 papers to our SLR list. Note that 8 out of the 14 recommended papers were published after May 2020, therefore they were not found in our first search. Table 1 presents the number of tentative papers after each iteration.

Until the fourth iteration, the first and second authors of this paper divided the resulting set of publications in half and excluded the papers based on the exclusion and inclusion criteria. Furthermore, when the first or second author of this paper had doubts about the inclusion or exclusion of a paper, they annotated it as an undecided paper. Then, both authors discussed the undecided papers with each other until they reached an agreement to include or exclude these papers. In the fifth iteration, the first and third authors discussed the prolific authors' recommendations until reach an agreement. These procedures help improve the reliability of the results.

4 Results

We first describe the data that we extracted from the 131 qualified papers. We then descriptively synthesized the data with a quantitative summary.

Demographic Data The first paper studying perceived diversity in our dataset of 131 papers was published in 2003. The 131 papers were published by more than 490 unique authors from different countries within Africa, Australia, Asia, Europe, North America, and South America. Figure 1 shows a yearly trend of the paper counts grouped by the perceived diversity aspects. We can observe an increasing trend of studies in perceived diversity aspects from 2013 with peaks in 2018 and 2019.

The type of publication with more papers published was conferences (65%), followed by journals (28%), and workshops (7%). The top three conferences with more publications were: "IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)" (15%), "International Conference on Software Engineering (ICSE)" (13%), and "ACM Conference on Human Factors in Computing Systems (CHI)" (12%). The Journal publications were more spread between different Journals. The top three journals with more publications were: "IEEE Software" (30%), "IEEE Transactions on Software Engineering (TSE)" (13%), and "Empirical Software Engineering (EMSE)" (8%). Finally, the workshop with

Table 1 Resulting set of publications after each iteration

| Selection process	# of papers remaining
Search: Google Scholar + IEEE database	5671
Cleaning Data: removing duplicates,	3194
First Iteration: exclusion based on the title	372
Second Iteration: exclusion base on the abstract	234
Third Iteration: exclusion based on a thoroughly read	66
Fourth Iteration: adding snowballing papers	117
Fifth Iteration: adding experts' recommendations	131
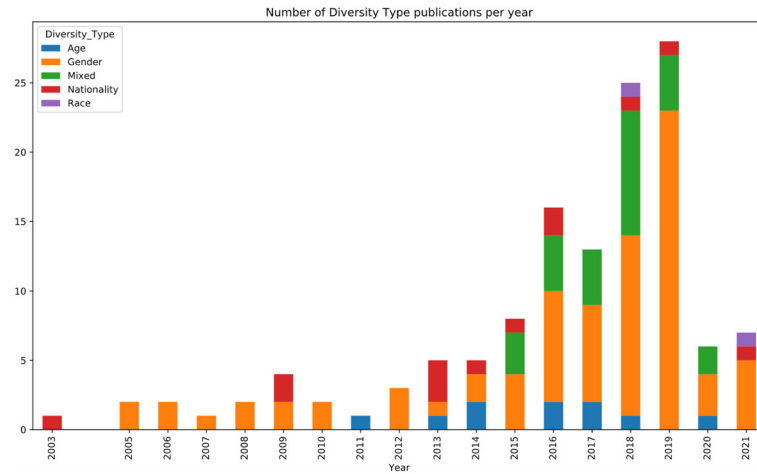
Fig. 1 Yearly trend of the papers count grouped by the perceived diversity aspects

a higher number of publications was: "ICSE-Workshop on Gender Equality in Software Engineering" (67%).

SE knowledge area We classified each paper based on the knowledge area that the authors were addressing in the papers. For identifying the knowledge area, we used the "SoftWare Engineering Body Of Knowledge (SWEBOK)" framework (Bourque et al. 2014). Table 2 shows the descriptions of the SWEBOK that we found in the papers and the percentages of papers within each knowledge area.

Table 2 Description and percentages of the knowledge area studied in the 131 papers included in our SLR according to the SWEBOK framework

SWEBOK	Description	%
Software Design	The paper studies perceived diversity aspects in the process of defining the architecture, components, interfaces, and other characteristics of a system or component.	13%
Software Construction	The paper studies perceived diversity aspects in a detailed creation of working software through a combination of detailed design, coding, unit testing, integration testing, debugging, and verification.	14%
Software Testing	The paper studies perceived diversity aspects in the dynamic verification of the behavior of a program against expected behavior on a finite set of test cases.	4%
Software Quality	The paper studies perceived diversity aspects in software quality characteristics, the value and cost of software quality, or software quality improvement.	12%
Software Requirements	The paper studies an perceived diversity issue related to elicitation, negotiation, analysis, specification, and validation of software requirements.	1%
Professional Practice	The paper studies perceived diversity aspects in work dynamics activities.	19%
Not Applicable	If SWEBOK categories does not define the area of knowledge in the paper	37%

Purpose data We identified two main purposes between the 131 papers: (1) papers that show perceived diversity biases or differences in SE, and (2) papers that describe or present inclusivity-efforts to tackle perceived diversity in SE. While 69% of the papers show perceived diversity bias or differences in SE activities, only 31% of the papers describe inclusivity-efforts to assess the perceived diversity in SE.

Research method of studies We classified the 131 papers based on the research method used in the paper following the five most relevant research empirical methods described by Easterbrook et al. (2008). In addition to these five methods, we have added two more categories ("Mixed study" and "Not applicable") as some papers cannot be easily classified into the five methods. The description of each research method and the percentage of papers that fall into each category is presented in Table 3.

Perceived diversity aspects We classified the papers based on the five perceived diversity aspects: Gender, Race, Age, and Nationality.

Gender diversity was the dimension most studied with 61% of the papers. These papers encompass gender identity studies, gender perception studies, and studies on transgender developers. Nationality diversity was the second dimension most studied with 10% of the papers. Nationality papers analyze the communication between teams, stereotypes that can trigger bias against developers, and cultural/national beliefs that can influence SE practices and practitioners. The third dimension most studied was age diversity with 8% of the papers. Finally, race diversity studies accounted for 2%. The remaining 19% of the papers studied a mixture of perceived diversity aspects, the most common combinations were Gender-Age (6%), Gender-Race (3%), and Gender-Nationality (3%).

Outcome Data We extracted the main results of the papers and the threats to validity related to perceived diversity and, when it was possible, the dependent and independent variables used in the studies. Figure 2 shows the paper counts grouped by the SWEBOK activity and the study methodology used within the different perceived diversity aspects. Most of

Table 3 Description of the research method used on the studies included in our SLR

| Category	Description	%
Action Research	It is commonly used for improving conditions and practices in a specific context.	3%
Survey Research	It commonly for collecting information from a pool of respondents by asking multiple survey questions.	17%
Case Study	It us commonly used for in-depth understanding on how and why a certain phenomena occur. It can be a exploratory or a confirmatory case studies.	39%
Controlled Experiment	It us commonly used for (1) collecting data that aims to observe, interact and understand a phenomenon; and/or (2) for determining in precise terms how the variables are related and whether a cause-effect relationship exists between them.	16%
Ethnographic Study	It is commonly used for understanding how technical communities build a culture that enables them to perform technical work collaboratively.	2%
Mixed Study	If the study defines, or it can be derived, more than one research method.	11%
Not Applicable	If the paper is a position paper or if the study does not define the applied research method and it can not be derived or interpreted from reading the paper.	12%
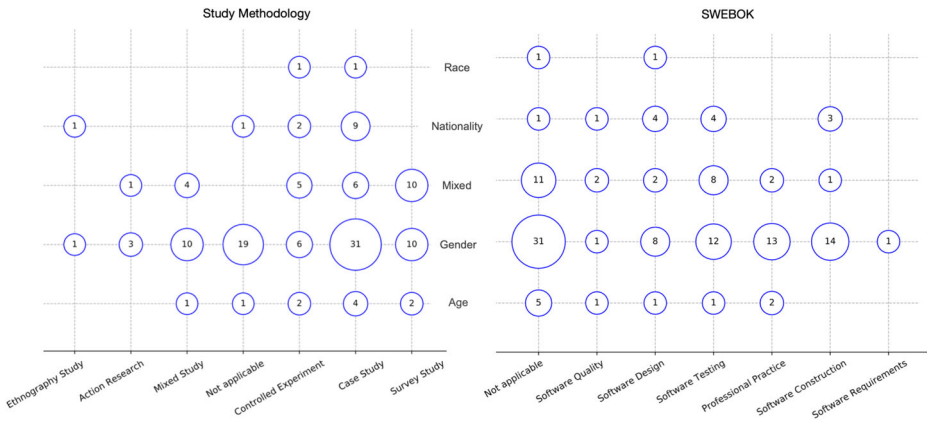
Fig. 2 Perceived diversity aspects grouped by the SWEBOK activity (right) and the study methodology (left) used in the 131 papers from our SLR

the papers in our SLR have a gender diversity dimension, are case studies, survey studies, or controlled experiments, and the SWEBOK is not clearly defined (not applicable) or professional practice.

Table 4 presents the synthesis of the 131 papers based on the data extracted previously. Each paper is classified based on the diversity type, the research methodology, and the purpose. Notice that we did not use a statistical meta-analysis method to combine numerical results because the different nature of the 131 papers increases the complexity to carry out a meta-analysis.

4.1 RQ1: What are the Types of Perceived Diversity Research Studies in SE?

Since the findings from the 131 papers are very broad, we present what has been studied and discovered related to the perceived diversity in SE in different sections. These sections summarized the main findings and report (1) differences within the perceived diversity aspects; (2) relationships between the perceived diversity aspects and SE metrics; (3) SE practitioners' perceptions about diversity aspects; and (4) challenges, barriers, and motivations faced by SE participants within the perceived diversity aspects.

4.1.1 Differences Within the Perceived Diversity Aspects

From the 131 studies in our SLR, we found results related to differences in gender, nationality, age, and race.

Gender differences We report gender differences based on women and men as the studies included in this SLR only reported differences based on these two genders. Findings from previous studies show both statistically and no statistically significant gender differences between men and women in different fields. We first describe the findings related to statistically significant gender differences and then the findings related to no statistically significant gender differences.

Table 4 Synthesis of the social data form the 131 papers

Study methodology	Gender	Nationality	Age	Race	Mixed
Bias/Difference Studies Ethnography Study	[SLR (Nafus 2012)] 1	[SLR (Shah and Harrold 2013)] 1	–	–	–
Action Research	–	–	–	–	[SLR (Irrgang 2018)] 1
Mixed-Study	[SLR (Beckwith et al. 2006; Subrahmaniyan et al. 2008; Burnett et al. 2010; Vasilescu et al. 2015; Ford et al. 2016; Qiu et al. 2019)] 6	–	–	–	[SLR (Thomas et al. 2018)] 1
Controlled-Experiment	[SLR (Beckwith et al. 2005; Fisher et al. 2006; Beckwith et al. 2007; Chintakovid and Wiedenbeck 2009; Cao et al. 2010; Zohreh Sharafi et al. 2012; Beckwith et al. 2006; Gómez et al. 2017; Metaxa-Kakavouli et al. 2018; Kaur Kuttal et al. 2019; Gralha et al. 2019; Ford et al. 2019; Pretorius et al. 2020)] 13	[SLR (Anderson et al. 2019)] 1	[SLR (Murakami et al. 2017)] 1	–	–
Survey Study	[SLR (Adikaram and Wijayawardena 2015; Morgan 2017; James et al. 2017; Singh 2019b; Blincoe et al. 2019; Catolino et al. 2019a; Murphy-Hill et al. 2019; Lee and Carver 2019; Machado et al. 2020)] 9	–	[SLR (Kanij et al. 2011; Davidson et al. 2014)] 2	–	[SLR (Choi and Pruett 2015; Vasilescu and Filkov 2015; Robert 2016; Robles et al. 2016; Bollmann et al. 2017; Schloegel et al. 2018; Hyrynsalmi 2019; Wurzelová et al. 2019; Hyrynsalmi and Hyrynsalmi 2019)] 9

Table 4 (continued)

	Study methodology	Gender	Nationality	Age	Race	Mixed
	Case Study	[SLR (Kuechler et al. 2012; Gila et al. 2014; Marques 2015; Terrell et al. 2017; Lyon et al. 2018; Kirton and Robertson 2018; Wang et al. 2018; Balali et al. 2018; Wang 2018; Catolino et al. 2019b; Izquierdo et al. 2018; Bosu and Sultana 2019; Imtiaz et al. 2019; Guzman and Rojas 2019; Singh 2019a; Bastarrica and Simmonds 2019; Vedres and Vasarhelyi 2019; Razavian and Lago 2015; Paul et al. 2019; Vasilescu et al. 2014; Zacchiroli 2020)] 21	[SLR (Daniel et al. 2013; Holtkamp et al. 2015; Alfaro and Chandrasekaran 2015; Dong et al. 2016; Rastogi et al. 2018; Al-Ani et al. 2013; Borchers 2003; Casey 2009; Furtado et al. 2020)] 9	[SLR (Morrison and Murphy-Hill 2013; Morrison et al. 2016; Kopeć et al. 2018)] 3	[SLR (Nadri et al. 2020)] 1	[SLR (Aué et al. 2016; Ortu et al. 2017; Pieterse and van Eekelen 2018; Mendez et al. 2018; Buffardi 2016)] 5
	Not Applicable	[SLR (Grigoreanu et al. 2009; Brun and Meliou 2018; Ford et al. 2019)] 3	[SLR (Cunha et al. 2009)] 1	[SLR (Davidson et al. 2014)] 1	[SLR (Thomas et al. 2018)] 1	–
Inclusivity- Efforts Studies	Ethnography Study	–	–	–	–	–
	Action Research	[SLR (Burnett et al. 2017; Ahmadi et al. 2018; Hilder-brand et al. 2020)] 3	–	–	–	–

Table 4 (continued)

Study methodology	Gender	Nationality	Age	Race	Mixed
Mixed-Study	[SLR (Chatterjee et al. 2021; Hui and Farnham 2016; Qiu et al. 2019; Padala et al. 2020)] 4	–	[SLR (Schloegel et al. 2016)] 1	–	[SLR (Mendez et al. 2019; German et al. 2018; Filippova et al. 2017)] 3
Controlled-Experiment	[SLR (Beckwith et al. 2005; Grigoreanu et al. 2008; Hill et al. 2016; Hill et al. 2017; Yi and Redmiles 2019; Yi and Zhang 2020)] 6	–	–	–	–
Survey Study	[SLR (Prado et al. 2020)] 1	–	–	–	[SLR (Storey et al. 2016)] 1
Case Study	[SLR (Moon 2013; Burnett et al. 2016; Draude and Maaß 2018; Mendez et al. 2018; Singh and Brandon 2019; Cunningham et al. 2016; Burnett et al. 2016; Ford et al. 2017; May et al. 2019; Vorvoreanu et al. 2019)] 10	–	[SLR (Baltes et al. 2020)] 1	–	[SLR (Altiner and Ayhan 2018)] 1
Not Applicable	[SLR (Judy 2012; Mendez et al. 2018; Burnett et al. 2018)] 3	[SLR (Sharp et al. 2016)] 1	[SLR (Kopeć et al. 2018)] 1	–	[SLR (Galhotra et al. 2017; Angell et al. 2018; Aydemir and Dalpiaz 2018; Ford 2020; Gren 2018)] 5

The count of papers in each column-row is highlighted in a gray box

Statistically significant gender differences The studies from this SLR report gender differences in the way how developers solve problems [SLR(Beckwith et al. 2006)], how developers tinker [SLR(Beckwith et al. 2006)], how developers use debugging strategies [SLR(Grigoreanu et al. 2009)], and when developers use those strategies successfully [SLR(Grigoreanu et al. 2009)]. For example, findings show (1) significant gender differences in the features elected to use by developers and the willingness to tinker and explore features [SLR(Burnett et al. 2010)], (2) men tend to switch more frequently between debugging strategies [SLR(Cao et al. 2010)], (3) women tend toward underconfidence and men tend to use more unfamiliar software features [SLR(Beckwith et al. 2005)], (4) software environments are often aligned with the needs of men rather than women [SLR(Beckwith et al. 2005)]; and (5) some end-user tools for debuggers may not fully support women debugging strategies [SLR(Subrahmaniyan et al. 2008)]. Similarly, other studies report gender differences during code review practices indicate that (1) accepted pull requests submitted by women and men provide similar descriptions in terms of length, and generate a similar number of discussions [SLR(Imtiaz et al. 2019)], (2) when both genders are known, women tend to have contributions accepted more often than men when they the contributors are from insiders to a project, but men's acceptance rates are higher when the contributors are from outsiders to a project, [SLR(Terrell et al. 2017)], (3) while women concentrate their work across fewer projects and organizations, men contribute to a higher number of projects and organizations [SLR(Imtiaz et al. 2019)], (4) there are gender differences in the use of positive opinion words, emoticons, and expletives during the code review as while men tend to express more positive/negative sentiments, women tend to express more neutral comments instead of expressing strong sentiments [SLR(Paul et al. 2019)], and (5) men and women follow different comprehension strategies when reading source code [SLR(Zohreh Sharafi et al. 2012)]. There are also gender differences in pair programming related to coordination, communication, and collaboration. Same-gender pairs tend to be democratic but mixed-gender pairs tend to have one authoritarian partner, and women preferred women partners [SLR(Kaur Kuttal et al. 2019)]. However, the findings indicate that productivity in pair programming is not affected by differences in gender pair types [SLR(Gómez et al. 2017)]. Finally, there are gender differences within the OSS community and the SO community. OSS studies found that (1) there is gender bias in the authorship of contributions to OSS projects from the past 50 years since while men have contributed with more than 92% of OSS code in the past 50 years, women reached 10% of the yearly contributions in 2019 [SLR(Zacchiroli 2020)], and (2) the perceptions and preferences of women and men developers in GitHub are statistically significant. Women developers significantly prefer in-person communication and are more frequently aware of the gender, ethnicity, and age of most of their team members [SLR(Vasilescu and Filkov 2015)]. Studies on SO reported that (1) men and women posted equally as often in that platform, but men post statistically significant a higher number of posts [SLR(Kuechler et al. 2012)], and (2) there are significant gender differences in the participation and success in SO. While women are asking more questions, men are obtaining more votes because they answer more questions [SLR(Wang 2018),(May et al. 2019)].

No statistically significant gender differences Findings indicate that there are no statistically significant gender differences between men and women in terms of (1) the ability to learn the new features when debugging [SLR(Beckwith et al. 2007)], (2) the difficulty of the strategies used for program comprehension [SLR(Fisher et al. 2006)], (3) the time using debugging strategies [SLR(Cao et al. 2010)], (4) the individual productivity in OSS projects [SLR(Bosu and Sultana 2019)], (5) accuracy, required time, and effort in

source code reading [SLR(Zohreh Sharafi et al. 2012)], and (6) timing of the review, content category, and character length in the mobile app's feedback [SLR(Guzman and Rojas 2019)].

Nationality differences The results from the SLR indicate nationality differences in an academic, industrial, and OSS context. Findings in the academic context reported that while student teams formed by homogeneous cultures perform better initially student teams formed by heterogeneous cultures perform better in the long term [SLR(Anderson et al. 2019)]. In an industrial context, the national differences of two vendor teams of the same organization were responsible for the differences in the software-testing approaches (concerning the testing-team structure, thought process, expectations, the primary focus, and trust levels) [SLR(Shah and Harrold 2013)]. Finally, in an OSS context, GitHub contributors from countries with low human development indexes (HDI) face more rejections than other contributors from or high HDI countries [SLR(Furtado et al. 2020)].

Age differences The results from the SLR indicate that some developers tend to use stereotypes and do not include older populations, and older adults prefer to participate as consultants because they may face some confidence barriers [SLR(Kopeć et al. 2018)]. In SO, the users' reputation scores increase relative to age well into the 50's, users in their 30's tend to focus on fewer topics than both younger and older users in age [SLR(Morrison and Murphy-Hill 2013)].

Racial differences The results from the SLR indicate that there are no traces of explicit racism in the written comments left by GitHub developers in code review practices. However, while the pull requests from perceptible non-white developers tend to be closed and non-merged without a reason, the pull requests from perceptible white developers tend to have a reason [SLR(Nadri et al. 2020)].

4.1.2 Relationships Between the Perceived Diversity Aspects and SE Metrics

We have identified the relationship between dependent and independent variables that have been studied in the papers and summarized the empirical evidence, i.e., the hypothesis tested, assumptions, or bias, that have been reported so far with relation to the perceived diversity aspects. The independent variables are the different perceived diversity aspects. The dependent variables are related to SE metrics.

Gender The studies have reported negative and positive relationships between gender diversity in SE and SE metrics in different online communities and industrial workplaces. We first describe the results related to online communities and then present the results related to industrial workplaces.

Online communities Studies analyzing the SO community indicate that (1) the scoring SO uses is designed to rewards practices more common among men (posting and answering questions) than among other people [SLR(Vasilescu et al. 2014)], (2) women who encountered other women in SO were more likely to engage sooner than those who did not [SLR(Ford et al. 2017)]; and (3) it is uncommon for a woman that posts a question to receive an answer or comment from another woman, but most of the women interaction exists when the post is initiated by a woman [SLR(Morgan 2017)]. Similarly, studies

analyzing the GitHub community indicate that (1) women are more likely to continue participating in a project when the diversity of the team members' expertise is higher [SLR(Qiu et al. 2019)], (2) the presence of women in development teams generally reduces community smells [SLR(Catolino et al. 2019b)], (3) gender diversity has a significant positive effect on productivity [SLR(Vasilescu et al. 2015)], and (4) code review developers spent the majority of their time reviewing the source code in a pull request, but they also spent a considerable amount of time looking at technical and social signals of the pull requester [SLR(Ford et al. 2019)]. Finally, studies analyzing gender diversity in online communities generally state that (1) the majority of developers in library OSS are men between 30 and 49 years old [SLR(Choi and Pruett 2015)], (2) women participation and women leadership in OSS is not representative of the true population as less than 10% of the developers contributing in OSS are perceived as women [SLR(Bosu and Sultana 2019)], (3) there are more women participation in library OSS than general OSS projects [SLR(Choi and Pruett 2015)], (4) women start to participate in FLOSS projects at a later age than men, mainly with tasks different from coding, and have different reasons to start and remain contributing in FLOSS projects [SLR(Robles et al. 2016)]; and (5) women are often hesitant to make contributions to a new project in GitHub [SLR(Wang et al. 2018)].

Industrial workplace Studies analyzing industrial workplaces indicate that (1) the gender of software engineers was a statistically significant predictor of self-rated productivity as software engineers who identified themselves as women and who selected to write their gender identity reported significantly higher self-rated productivity than men [SLR(Murphy-Hill et al. 2019)]; (2) gender diversity has a significant positive effect on performance [SLR(Gila et al. 2014)], (3) while the self-perception in men's end-user debuggers is positively correlated with the debugging performance, the self-perception in women's end-users is not correlated [SLR(Chintakovid and Wiedenbeck 2009)], (4) women do not seem to be willing to acknowledge their own performance in software development teams, but their performance is highly regarded by their teammates [SLR(Bastarrica and Simmonds 2019)], (5) all professional software developers, both men and women, have implicit gender biases which may impact their decision-making on the evaluation of other developers' contributions [SLR(Yi and Redmiles 2019)], and (6) there is a gender pay gap of 18% in a UK company in which only 11% of senior positions were taken by women [SLR(Kirton and Robertson 2018)]. Finally, studies performed by students simulating industrial software teams indicate that (1) gender diverse teams are more effective and coordinated than non-gender diverse teams [SLR(Marques 2015)], and (2) gender alone, and cognitive style and gender together in teams were positively associated with feature novelty. Women designed significantly more novel software features [SLR(Pretorius et al. 2020)].

Nationality Studies from this SLR reported that diversity in the language of team members (1) has a negative impact on the quality of the work [SLR(Pieterse and van Eekelen 2018)], (2) a negative impact on community engagement [SLR(Daniel et al. 2013)], (3) it might lead to difficulties in the communication between GitHub users [SLR(Vasilescu et al. 2015)], and (4) together with the country of residence has a positive effect on market success [SLR(Daniel et al. 2013)]. Findings related to diversity in the country of residence indicates that developers' country (1) has a statistically significant relationship to the project's success, but a small effect size [SLR(Aué et al. 2016)], (2) is linked with lower team politeness [SLR(Ortu et al. 2017)], and (3) have a statistically significant relationship with the acceptance of pull requests [SLR(Rastogi et al. 2018)]. Findings related to diversity in the culture of developers indicate that (1) people working together in a team will form ties

with each other despite their cultural differences [SLR(Dong et al. 2016)], (2) cultures with high levels of inequality in the hierarchical structure of software development organizations may have negative effects to form and retain software development teams [SLR(Borchers 2003)], and (3) the communication in distributed software development teams is negatively affected when there is a high cultural diversity in the team [SLR(Casey 2009)].

Age Studies from this SLR reported that (1) the age of code reviewers does not affect the correctness and efficiency of their reviews [SLR(Murakami et al. 2017)], (2) age diversity has a low correlation with team efficiency [SLR(Altiner and Ayhan 2018)], (3) fully involving older adult participation improves the process of developing apps for the older population, and (4) developers older than 40 years and with more than 15 years of experience can contribute to increasing the knowledge of software development in SO [SLR(Morrison et al. 2016)].

Race Studies from this SLR reported that (1) the R community has an under-representation of non-white attendees in their UserR! conference [SLR(Bollmann et al. 2017)], (2) there are gender and racial stereotypes behavior embedded in modern software [SLR(Brun and Meliou 2018)], and (3) the racial diversity in student software development teams does not have an impact on how well the team members cooperated to develop software products [SLR(Pieterse and van Eekelen 2018)].

4.1.3 Participants' Perceptions About Perceived Diversity in SE

We have summarized the perceptions, experiences, and reflections of participants who work and contribute to software engineering activities. These perceptions, experiences, and reflections are related to the perceived diversity aspects and have been reported in the 131 paper.

Gender Studies related to FLOSS communities have found that many participants express a positive sentiment towards including women in the participation of FLOSS projects. However, women contributors did face sexism or had encountered a sexism incident [SLR(Lee and Carver 2019)]. In fact, harsh and sexist treatment faced by women is said to be "as constant as it is extreme" and harsh discussions about which source code piece should get accepted or merged in the software leads to the project behave as a "pushyocracy" instead of a meritocracy, a prime reason why women leave these communities [SLR (Nafus 2012)]. Previous findings also report that (1) the majority of older adults claimed to have witnessed discrimination towards others in FLOSS, especially against non-native English speakers and women [SLR(Davidson et al. 2014)], (2) successful OSS women believe that having codes of conduct in the projects would help to increase and retain women and allies [SLR(Singh 2019b)], and (3) transgender developers have expressed discomfort when participating in Hackathons and they are concerned about LGBQTPhobia in these events [SLR(Prado et al. 2020)]. Similarly, studies related to the industrial work environment have found that (1) women are less satisfied with the spirit of their team at their company [SLR(James et al. 2017)], (2) 12.3% of the men admitted to having a conscious bias against women and 40.2% of women have felt gender discrimination during their career [SLR(Blincoe et al. 2019)], (3) developers perceived gender diversity of a team as being less important than developer experience or team size to mitigate community smells [SLR(Catolino et al. 2019a)], (4) gender minorities in teams formed predominantly by men feel less included when the team uses audio channels or visual communication channels to discuss work [SLR(Hui and Farnham 2016)], (5) transgender developers think that working remotely can foster a more inclusive

environment as they can control their professional identity [SLR(Ford et al. 2019)], and (6) men or women software architects who exhibit "feminine expertise"[6] are perceived as successful when in dealing with software architecting's human aspects [SLR(Razavian and Lago 2015)].

Nationality Studies from this SLR reported that (1) teams with perceptions of higher levels of diversity are more willing to cooperate when they are geographically dispersed [SLR(Robert 2016)], (2) intercultural competencies are not specifically important for team participants in the coding phase of the software product. But, intercultural competencies are seen as very important in tasks with high levels of communications [SLR(Holtkamp et al. 2015)], and (3) some cultures can lead to strong stereotypical beliefs that challenge women in teams dominated by men. These women have perceived difficulties having their work recognized, proving themselves, and feeling members of their team [SLR(Adikaram and Wijayawardena 2015)].

Age Studies from this SLR reported that (1) software testing teams think that the age diversity of the team is not an important factor for improving the performance when compared with personality diversity and experience diversity [SLR(Kanij et al. 2011)], and (2) developers' performance expectations in China, Germany, Poland, and Bulgaria are biased towards middle-aged employees over younger and older employees. These developers have negative age stereotypes toward older or younger employees depending on the country [SLR(Schloegel et al. 2018)].

Race Studies from this SLR reported that Black women have experienced complete isolation in the field of computer science and they do not know if this negative experience is because of their gender or their race [SLR(Thomas et al. 2018)].

4.1.4 Challenges, Motivations, and Barriers Faced by SE Participants Within the Perceived Diversity Aspects

We have summarized the challenges, motivations, and barriers of participants who work and contribute to software engineering activities. These challenges, motivations, and barriers are related to the perceived diversity aspects and have been reported in the 131 paper.

Gender Studies focused on community-based environments have identified that (1) while the most common barriers among women in FLOSS are related to social factors such as not being taken seriously, needed to prove themselves, or found it hard to find a mentor or attract attention, the most common barriers among men in FLOSS are related to entry factors such as the built environment, tooling, license terms, and the attitude of other participants [SLR(Lee and Carver 2019)], (2) there are five hidden barriers identified in SO that forestall women from contributing [SLR(Ford et al. 2016)], (3) newcomer women developers face gender bias related to multiple problem-solving facets in tool/infrastructure [SLR(Padala et al. 2020)]. In fact, 73% of the barriers in tools and infrastructure faced by OSS newcomers had some form of gender bias [SLR(Mendez et al. 2018)], (4) the competence-confidence gap is a threat to women's contribution in GitHub [SLR(Wang et al. 2018)], and (5) while stereotypical web interfaces can trigger

---

[6]the paper defines feminine expertise as "a combination of traits (characteristics and behavior attributed to femininity) and skills (learned abilities in playing the gender roles)"

a less sense of belonging in women exposed to these interfaces, a neutral web interface does not negatively affect the sense of belonging of men or women [SLR(Metaxa-Kakavouli et al. 2018)]. Similarly, studies focused on industrial contexts have identified that (1) adult women in Finland were motivated to make a change in career into the software industry because of the variety of new jobs offered, the equal wage, and a flexible work climate [SLR(Hyrynsalmi and Hyrynsalmi 2019)], (2) women face challenges related to their self-esteem when talking about their skills [SLR(Hyrynsalmi and Hyrynsalmi 2019; Hyrynsalmi 2019)] and find the male-dominated industry sometimes challenging [SLR(Hyrynsalmi 2019)], and (3) women and men are affected by different types of interruptions during social isolation. While most of the challenges faced by women are related to the support with housework and child care responsibilities, men most of the challenges faced by men are related to the work space [SLR(Machado et al. 2020)].

Nationality GitHub developers have reported challenges related to social barriers, i.e., language, economic reasons, or political reasons, in the communications channels they used for work in GitHub [SLR(Storey et al. 2016)]. These challenges have been also reported by mentors in OSS as the main barriers faced by their newcomer mentees in OSS [SLR(Balali et al. 2018)].

Age Studies from this SLR reported that (1) developers older than 40 years and with more than 15 years of experience are less motivated to interact in SO than their younger peers [SLR(Morrison et al. 2016)], (2) the top three motivations of older adults to contribute to OSS are intrinsic motivation, community identification, and altruism [SLR(Davidson et al. 2014)], (3) the top 3 barriers perceived by older adults in their first time in OSS are lack of communication, installation issues, and documentation issues [SLR(Davidson et al. 2014)], (4) older adults face challenges on their first contribution and more social challenges in more recent contributions [SLR(Davidson et al. 2014)], and (5) the media in United Stated has pictured ageism as a major barrier in the process of hiring developers in large software companies [SLR(Baltes et al. 2020)].

4.2 RQ2: What the Perceived Diversity Research has Proposed to Foster Diversity-Inclusiveness in SE?

To answer RQ2 we have identified the tools, processes, models, and practices reported in the papers whose purpose was Inclusivity-Efforts. These papers help to mitigate bias in SE by proposing different elements to assess the perceived diversity of teams, or by identifying perceived diversity-inclusiveness issues. The results from RQ2 can be used by both researchers in perceived diversity and practitioners willing to increase diversity.

4.2.1 Tools, Models, and Theories to Foster Perceived Diversity-Inclusiveness in SE

GenderMag It is a software inspection method to solve gender biases in software developers' problem-solving experiences [SLR(Burnett et al. 2016; Burnett et al. 2018)]. Previous research shows that GenderMag is very effective in identifying gender-inclusiveness issues when evaluating technologies [SLR(Burnett et al. 2016; Hill et al. 2016; Vorvoreanu et al. 2019; Burnett et al. 2017)] and was found a powerful tool to uncover potential gender biases in system functionality and interface design [SLR(Cunningham et al. 2016)]. Although GenderMag is partially based on gendered personas, it did not promote gender stereotyping [SLR(Hill et al. 2017)]. Recently, Hilderbrand et al. have reported real-world

experiences on how software teams can embed GenderMag into their development processes [SLR(Hilderbrand et al. 2020)]. To help practitioners, GerderMag has a **GenderMag Recorder's Assistant**, which is a semi-automated visual tool that reduces the cognitive load for practitioners that work with GenderMag [SLR(Mendez et al. 2018)].

InclusiveMag It is a generalization of the GenderMag method that assessed whether the software supports a particular dimension of diversity. InclusiveMag helps researchers to develop new inclusive design methods and evaluate the software product in terms of inclusiveness [SLR(Mendez et al. 2019)].

AID Tool It is an automated detector for gender-inclusivity bugs in OSS project pages. This tool automates the GenderMag method for detecting gender-inclusivity bugs in OSS projects [SLR(Chatterjee et al. 2021)].

Themis It is an open-source tool that helps discover and debug discrimination in different perceived diversity dimensions (race, age, and gender) in software products. This tool automatically generates tests to detect and measure discrimination [SLR(Angell et al. 2018; Galhotra et al. 2017)].

Gender-Extended Research and Development (GERD) It is a model that combines gender studies approaches and computer science thinking. The model identifies seven core processes or phases of research and development and aims to contribute to the more inclusive development of technology. It also encourages reflection on how the construction and use of software systems are socially embedded [SLR(Draude and Maaß 2018)].

The intergroup contact theory This theory was proposed by Gordon W. Allport (Allport et al. 1954) and poses that increasing interactions between different groups of people would possibly be helpful to reduce biases under some conditions. A recent study has proved that this theory helps reduce implicit gender biases in both general and SE-specific contexts [SLR(Yi and Zhang 2020)].

Ethics-aware SE It is a model that captures, analyzes, and reflects the ethical values of different stakeholders in the SE processes and software specifications. This model assists in the creation of ethical software development by ensuring (1) that developers create software for organizations that are aligned with their ethical values and (2) that organizations follow their ethical principles. [SLR(Aydemir and Dalpiaz 2018)].

A framework to recognize fair and unfair modern code reviews This framework is based on fairness theory and can be used to study and manage social behavior in modern code reviews [SLR(German et al. 2018)].

Support for Participant Involvement in Rapid and Agile software development Labs (SPIRAL) SPIRAL is a method that provides strategies for the direct involvement of older adults in the development process [SLR(Kopeć et al. 2018)].

Redesign of SO reward system The scoring SO uses is designed to reward practices more common among men than among other people. users. Thus, May *et al.* have proposed an alternative scoring system that equalizes the rewards obtained in SO and does not penalize

any group of users in absolute terms. With this redesign, the median woman is marginally more successful than the median man [SLR(May et al. 2019)].

4.2.2 SE practices that Help to Foster Diversity

1. A formal definition of software fairness testing and causality-based measure of discrimination [SLR(Galhotra et al. 2017)];
2. Codes of conduct and women-only spaces facilitate women engagement and retention in online communities [SLR(Singh and Brandon 2019; Singh 2019a)];
3. Politeness helps to involve women in FLOSS projects [SLR(Moon 2013)];
4. Brainstorming techniques help participants to feel satisfied with the team's process. The use of these strategies supports the satisfaction and outcomes of minority participants working in a team [SLR(Filippova et al. 2017)];
5. Interventions to reduce age stereotypes in software development: An awareness-based intervention and a cooperation-based workshop [SLR(Schloegel et al. 2016)];
6. Ethnography studies as a useful and usable approach to empirical software engineering research [SLR(Sharp et al. 2016)];
7. Agile values of collaboration as a framework for addressing gender diversity issues in a team [SLR(Judy 2012)];
8. The inclusion of two features in an end-user software development environment. The first feature provides a way to express less confident judgments and the second feature provides better explanations in the learning process of the end-user system [SLR(Beckwith et al. 2005)]. Furthermore, these features are effective in removing unintended barriers that affect the performance of women without harming men performance [SLR(Grigoreanu et al. 2008)];
9. Living Labs: This approach can help women professionals in their work as it offers opportunities to share experiences, provide insights, and create social changes in a collaborative environment [SLR(Ahmadi et al. 2018)];
10. Guidelines for encouraging older adults (< 50 years) in FOSS. These guidelines focus on providing feedback by projects and improving the communication by first-time older contributors and by projects [SLR(Davidson et al. 2014)];
11. Two proposed approaches that can address subjects' implicit gender biases at individual and organizational levels [SLR(Yi and Redmiles 2019)]. The first approach is to design continuous training courses to learn about people's own implicit biases. The second approach states that software development organizations should encourage and hire women to take counter-stereotypical roles such as technical leadership;
12. Research approaches such as perspective alignment, community trust-building, narrative research to understand community building [SLR(Ford 2020)];
13. Five recommendations for gender-inclusive hackathon events: (1) have a gender-inclusive organizing team; (2) foster inclusive communication; (3) make safety visible; (4) provide good working conditions to participants; and (5) showcase transgender people in the event [SLR(Prado et al. 2020)].
14. Gender discrimination in software development can be avoided using participatory and interdisciplinary approaches [SLR(Irrgang 2018)]; and
15. Empirical SE research on intersectionality to better understand software engineering in general and obtain more realistic and diverse studies regarding culture, gender, and ethnicity [SLR(Gren 2018)].

4.3 RQ3: What are the Challenges Faced by SE Researchers when Studying Perceived Diversity?

To answer RQ3 we have identified the threats to validity and limitations reported by the authors in the papers. The results are presented based on the limitation and threats found per the type of study methodology of the study and can be used by researchers on perceived diversity in SE to mitigate the possible threats of validity that their studies can have. Notice that some papers did not present a threat to the validity section and some threats were not directly related to perceived diversity.

Students as subjects The low number of women students in class when compared to men students can influence the results [SLR(Bastarrica and Simmonds 2019)].

Measuring cultural team diversity One of the limitations when measuring the cultural diversity of a team with this Hofstede's framework (Hofstede 2005) is the underlying assumption that few dimensions can explain cultural beliefs and values (Shachaf 2008). Therefore, the use of ethnographic methods to discover the cultural beliefs and values of a team may be a better approach.

Gender identification Men and women can misrepresent their genders in their online profiles [SLR(Ford et al. 2016)]. Indeed, around 30% of the surveyed participants who masked their gender by choosing gender-neutral pseudonyms were women participants [SLR(Lee and Carver 2019)]. When this occurs, the results of studies based on labels extracted from online profiles might not be accurate [SLR(Terrell et al. 2017)]. Furthermore, the tools and heuristics used to identify the gender of developers might not give reliable information when users do not use their real names [SLR(Qiu et al. 2019)]. Another limitation is the assumption that users/developers can be classified into only three gender groups (men, women, and non-binary), and then focus the study on only two genders (men and women) [SLR(Guzman and Rojas 2019)]. Risks implied by reducing gender studies to men and women are related to the inherent marginalization of non-binary individuals [SLR(Izquierdo et al. 2018)]. Researchers should acknowledge the possible limitations in their sampling and/or gaps in the generalizability of their study if they have only included people of binary genders (Scheuerman et al. 2021).

Unbalanced representation of real demographics When doing a qualitative study in FLOSS, the number of women's responses is normally very low (1-5% (Ghosh et al. 2002),[SLR(Storey et al. 2016)], 10.9% [SLR(Lee and Carver 2019)]). This is something that researchers should be aware of and try to balance the women's responses in these studies [SLR(Lee and Carver 2019)].

Generalization When survey respondents and interview participants come from the same geographical location or culture, the results from the study may not be generalized [SLR(Blincoe et al. 2019)]. Furthermore, when analyzing OSS projects, authors cannot generalize their results because OSS projects can be extremely different in terms of product, participant type, community structure, and governance [SLR(Bosu and Sultana 2019)]. The culture of participants in a case study can threat the generalization of results too [SLR(Gilal et al. 2016)]. Also, a recent study states that considering intersectionality in research studies is important to understand the needs of different individuals as the experiences of Black

women are often different from the experiences of both Black men and non-Black women in the United States (Ross et al. 2020).

5 Discussion

This SLR shows that perceived diversity studies in SE have critically reflected upon practices, methods, and even tools concerning inequality issues in both the development of software and the social aspects of the software. Specifically, while most of the studies analyzed in this SLR have shown social issues related to gender, less has been done to research how racial, age, and nationality diversity relates to SE.

Thus, we believe that more research initiatives have to be done to help the SE research community to understand the challenges and needs that other groups are facing in SE when participating in software engineering. In the following subsections, we discuss the findings of this SLR for each one of the perceived diversity aspects studied in SE.

5.1 Gender Diversity in SE

The study of gender diversity in SE has been essential to demonstrate the benefits that gender diversity brings to the workplace. Previous studies have proven that gender diversity is a factor associated with an increase in productivity [SLR(Vasilescu et al. 2015)] and innovation (Østergaard et al. 2011) in projects. However, as can be seen in Table 4, much of this gender research has focused on identifying bias through case studies, rather than proposed models, tools, or processes to assess gender diversity in software development teams. From the 80 studies on gender diversity, 53 (66%) have analyzed bias in SE and 27 (34%) have described Inclusivity-Efforts.

Furthermore, although 60% of the papers analyzed in this SLR have focused on gender, only two have analyzed transgender developers [SLR(Ford et al. 2019),(Prado et al. 2020)]. The transgender population has experienced and is still experiencing inexcusably high levels of harassment and violence (Lombardi et al. 2002). In the United States, 46% of transgenders had experienced verbal harassment (James et al. 2016). Therefore, we believe that gender identity, i.e., transgender and non-binary people, should be further studied in SE to understand their needs and to foster safer spaces for everyone.

Recently, LinkedIn reported that women in the technology area represent 24.4% of their users, while men represent 77.6% (Hall and Durruthy 2020). This gap is also present in the results from our SLR since women developers are still underrepresented in OSS and industrial companies after more than a decade of research [SlR(Robles et al. 2016; Izquierdo et al. 2018)]. Women still face discrimination between their peers [SlR(Blincoe et al. 2019)]. This result may indicate that more research into inclusivity-efforts in gender diversity should be done. OSS and industry should invest in reducing the gender gap. Likely, these efforts may have to be coordinated with other efforts to continue removing existing barriers and walking through an equitable integration in developer teams.

Currently, many companies are aware of the gender gap and are actively trying to reduce it. Specifically, big companies such as Google, Microsoft, and Mozilla are encouraging women to work in the software industry through some initiatives, i.e., Google's Diversity[7],

---

[7]https://diversity.google

Microsoft's Global Diversity and Inclusion program,[8] and Mozilla's Diversity and Inclusion Strategy.[9] That way, these companies aim to create a better workplace free from prejudice and discrimination.

> The gender diversity in SE has been the dimension most studied in our SLR. Currently, the papers show the dearth of women in SE, bias against women in some communities, and sometimes negative perceptions about women working in teams. Furthermore, we identify that gender identity should be further considered and studied in SE.

5.2 Racial Diversity in SE

Unfortunately, people might hold unconscious beliefs about various social groups that can be triggered by the perceived race derived from one's physical appearance or one's name. This unconscious bias can impact some races on behalf of others. Some studies on social psychology have demonstrated that to get one callback when applying for a job, if African-American people are identified from the names, they need to send almost double the resumes than people associated with White names (Bertrand and Mullainathan 2004). Furthermore, companies are less willing to hire a Mexican-Spanish-accented applicant than the standard American English-accented applicant for software engineering jobs (Hosoda et al. 2012).

Hence, the study of racial diversity in SE can help to identify the benefits of tackling problems from different perspectives (Shachaf 2008). We strongly believe that understanding the needs and barriers of different developers may help to avoid and mitigate stereotyping in the workplace. The results of this SLR indicate a gap in the literature when studying racial diversity in SE as we found only a few papers addressing racial diversity in SE. There is a lack of empirical evidence about race issues in SE. Some initial work reports on the negative experiences of Black women in the field of computer science [SLR(Thomas et al. 2018)]. Additionally, previous research has specifically not considered the perceptions of other ethnic subgroups such as Hispanics or indigenous people. Quesenberry and Trauth argue that it is not enough to simply increase the numbers of underrepresented minorities, organizations must adapt their practices to support the needs of these minorities (Quesenberry and Trauth 2012). Thus, we believe that to foster a healthier SE community by understanding and addressing minorities' needs, researchers should further study the racial diversity in SE.

Furthermore, Thomas et al. report that Black women do not know if their negative experience was because of their gender or their race (Thomas et al. 2018). Therefore, we also believe that a better understanding of the intersectionality issues faced by SE developers would be essential to avoid negative experiences.

> This SLR show that SE research in racial diversity is very limited. Although some inclusivity-efforts (i.e., brainstorming, themis tool, and ethics aware SE) have been proposed, a few studies have focused on racial diversity in SE. Therefore, we would like to create a "call for future action" to research quantitatively and qualitatively racial diversity in SE.

---

[8]https://www.microsoft.com/en-us/diversity/default.aspx
[9]https://wiki.mozilla.org/Diversity_and_Inclusion_Strategy

5.3 Nationality Diversity in SE

The study of nationality diversity can help to identify the benefits of developing software in a cross-cultural environment and how to solve cultural issues. Specifically, in online collaborative environments is where the teams' geographical diversity can be observed as developers are often from geographically separated localities [SLR(Vasilescu et al. 2015)]. Previous studies have demonstrated the cross-cultural issues faced by developers in distributed software development (Krishna et al. 2004). For example, Mishra and Mishra identified the communication, collaboration, and coordination issues as the most negative issues in distributed software development (Mishra and Mishra 2014). Shachaf found that cultural and language differences negatively impact cost, trust, cohesion, and team identity (Shachaf 2008). Paris et al. demonstrated that despite papers' quality, papers with authors from some regions receive fewer citations than papers from authors from other regions (Paris et al. 1998).

In this SLR, most of the studies about nationality diversity have shown that collaboration between industrial organizations from different cultures may raise misunderstandings, discomfort, or cultural issues that may end with people leaving the organization. For example, Casey [SLR(Casey 2009)] studied the implications of misunderstanding and not addressing cultural differences in an Irish multinational organization that off-shored part of their software development process to Malaysia. Casey [SLR(Casey 2009)] found that some cultural differences force key Malaysian personnel to leave the organization which resulted in serious implications for the success of the project. Managers from both Irish and Malaysian organizations recognized the requirement for cultural training. Therefore, this SLR indicates the need for cultural training to avoid misunderstandings and discomfort between organizations.

However, less is known about what are the challenges that different nationalities are facing in OSS development. In online collaborative environments, there are a huge diversity of nationalities as developers often are from different geographically separated localities [SLR(Vasilescu et al. 2015)]. We found only one paper that studied the relationship between the country of residence of developers and the acceptance or rejection of their contributions to OSS projects [SLR(Rastogi et al. 2018)]. Other papers have studied geographical location but as a combination factor with culture. Therefore, we believe that the collaboration in OSS may face challenges since multiple different nationalities may collaborate in one project. Researchers should further investigate nationality diversity in online collaborative environments for OSS. For example, researchers can use qualitative ethnographic methods to discover how the significant components of nationality may affect software development.

> This SLR has shown that the nationality diversity in SE has been studied from an industrial context. These studies have reported how different cultures may raise misunderstandings, discomfort, or cultural issues. Therefore, we believe the SE community needs more research on nationality diversity from an OSS point of view and more research on training developers about different nationalities.

5.4 Age Diversity in SE

The study of age diversity in SE development teams can help to understand the myriad of skills and talents that each generation can bring to the team. For example, younger developers

might have greater familiarity with the cutting edge of technology tools compared with their more mature counterparts. But mature SE developers might have exceptional interpersonal skills and experience leading software teams. However, previous social psychological studies have reported negative stereotypes toward older and younger employees across industries. While older employees are seen as poorer, less adaptive, and less innovative performers (Posthuma and Campion 2009), younger employees are seen as being disloyal to their employer, less supportive to their colleagues, and not willing to work with older employees (Hertel et al. 2013; O'higgins 2001). These stereotypes might lead to issues in attracting or keeping highly qualified developers, and ultimately, harm the project outcome.

This SLR has found a few papers addressing age diversity in SE. These papers report that age diversity does not affect the team efficiency (Altiner and Ayhan 2018) or the performance of developers [SLR(Murakami et al. 2017)]. However, some cultures may have stereotypes that damage both young and old developers in an industrial context [SLR(Schloegel et al. 2018)]. Thus, it was very interesting to find a paper that demonstrates how to reduce age stereotype bias through two interventions [SLR(Schloegel et al. 2016)]. The first intervention was an awareness-based workshop with 56 employees and the second intervention was a cooperation-based workshop conducted with 76 employees.

This SLR also indicates that the number of old developers participating in FOSS or SO is worrisome [SLR(Robles et al. 2016; Morrison et al. 2016)]. Thus, researchers should further study how to engage and retain older developers in these communities. That way, the SE community can avoid the loss of knowledge.

> This SLR has shown that developers may have stereotypes toward older and younger employees. However, these studies have been carried out in an industrial context. This SLR has also indicated a low number of old developers participating in OSS. Therefore, we believe that the SE community would benefit from further research on age diversity from an OSS point of view.

5.5 Recommended Papers From Prolific Authors in Perceived Diversity Aspects in SE

In this section, we describe briefly studies related to perceived diversity aspects in SE that were not included in this SLR because they do not meet the inclusion criteria but were suggested by the prolific authors in perceived diversity in SE. The studies are discussed according to the following categories:

Disability Although there are many dimensions of disability in SE, we briefly discuss two dimensions of disability: blind or visually impaired software developers and neurodiverse software developers. We believe that future work should analyze further what has been published about disability in SE.

The perceptions and experiences of blind and visually impaired software developers have reported that these developers perceived some barriers using screen readers when writing code (Mealin and Murphy-Hill 2012). The biggest challenges are related to workplace dynamics, project management dynamics, and tool accessibility (Huff et al. 2020). Although neurodiverse software engineers may desire some accommodations, most of them do not disclose their diagnosis to their company (Morris et al. 2015). Finally, a remote video game

coding camp can improve the communication skills of autistic college students (Begel et al. 2021)

Intersectional studies There is a need for intersectional studies as they can help us obtain a more realistic picture of software engineering with regards to culture, gender, and ethnicity [SLR(Gren 2018)]. Recently, three Black women academics shared their lived intersectional experiences during a global pandemic. They also created a call for action to " stand in solidarity with Blacks in computing; and acknowledge, disavow, and dismantle Whiteness and oppressive power structures in the field of computing" (Erete et al. 2021). A study showed how the experiences of Black women are often different from the experiences of both Black men and non-Black women in the United States. In fact, Black women, are less likely to be introduced to computer science by their family during their school than other women, and they have fewer friends in computer science than both Black men and non-Black women (Ross et al. 2020). Finally, a study has shown how a virtual summer camp that focuses on informal computer science learning opportunities can increase the confidence of both Black and Latina girls (Braswel et al. 2021).

OSS projects for social good A recent study by Huang et al. identified what are the motivations of software developers to contribute to OSS projects with a societal goal. They found that developers that participate in OSS projects for social good do not focus on their professional benefit but on leaving their mark on society. Developers in OSS projects for social good also evaluate the owners of these projects significantly more than other OSS developers (Huang et al. 2021).

6 Threats to Validity

We present the threats to validity according to Wohlin et al. (2012), who discuss four main threats in software engineering: internal, construct, external, and conclusion.

Internal validity threats Our SLR may present some internal validity threats due to our choice of Google Scholar and IEEE database as the search engine for all publications. This is because other publications may exist that are not indexed by these engines. Another threat is that our search was done based on keywords appearing in the title. Therefore, relevant papers not using our keywords might be missing. To minimize this threat we followed the procedures for performing SLRs described by Kitchenham and Charters (2007) and we offered a replication package so that third parties can inspect our sources. Another factor affecting internal validity may be the maturity of the perceived diversity issues in SE, in the sense that our study may have been done at too early a stage to conclude. However, the first study we found was in 2003, thereby we believe that more than 15 years is enough time to allow us to extract valid lessons from the literature. Although we believe we have analyzed the most important papers in perceived diversity in SE so far, we may have missed some relevant studies due to language constraints, and thus underestimate the extent of perceived diversity research in SE. Also, we decided to exclude papers published in non-CS venues. Hence, our results must be qualified as applying only perceived diversity research published in mature international software engineering conferences or journals. Although we might have missed some important papers due to the search strings we used in Google Scholar and IEEE database, we have contacted field experts in the perceived diversity area to assess our SLR and minimize the number of important studies missed.

Construct validity threats Our SLR may have construct validity threats as we have manually analyzed hundreds of papers for specific and sometimes very detailed aspects, so human errors may have occurred. For example, two of the authors selected the candidate studies independently and extracted the data. To mitigate human errors, before include or exclude a problematic paper for which the authors were unsure, they discussed the paper. Also, the first author has checked the data extraction from the second author and they have discussed the paper until they reach an agreement. The second author checked 25% of the data extracted by the first author and they agreed on the extraction. Our replication package does not remove the human errors that we may have incurred, but it offers the possibility for others to check and improve our work.

External validity threats Although we have analyzed more than one hundred papers that addressed some perceived diversity issues in SE, we cannot claim that our results can be generalized outside the SE community. Our results present a general overview of the perceived diversity state-of-the-art in SE. We identify gaps that should be addressed, but we cannot claim that our results can be generalized in individual or specific SE communities. Thus, we believe that further investigation on perceived diversity issues is necessary for specific SE communities.

Conclusion validity threats For the analyses of our SLR, we have extracted as much data as possible from the 131 papers: publication year, authors, venue, type of publication, SE activity done in the paper, and purpose. However, there might be other undiscovered valuable data in the papers that we did not extract. Not having used this undiscovered data can be a threat to the conclusion validity of the study since the results may be different. This SLR summarizes the findings of previous SE papers in perceived diversity. Therefore, it might be possible that some results about gender diversity are not generalized as previous studies might have considered all developers to be cisgender developers

7 Conclusion

Software Engineering is a collaborative field where a variety of diverse software developers work together to develop software. The higher diversity in a software development team, the greater possibility to understand and represent the final user's needs (Muller and Kuhn 1993). However, as developers are human beings, they can hold conscious or unconscious bias when working with their peers. Such bias can be triggered in person A when person B believes that they have perceived diversity aspects (i.e., gender, age, and race) from person A.

Hence, to better understand how these perceived diversity characteristics are related to SE we have carried out this SLR. We first defined the four perceived diversity aspects (gender, age, race, and nationality) considered in this SLR and then analyzed what 131 previous studies have discovered and proposed to increase the awareness and reduce the bias of these four perceived diversity aspects in SE.

With this SLR we aim 1) to assist the SE community with the understanding of the different factors that may influence the engagement and permanence of diverse developers working in software engineering; 2) to identify different methods used to improve perceived diversity in teams, and 3) to be aware of the threats to validity and limitations that previous studies have.

In conclusion, this study highlights the gaps in the current literature and creates a call for future action in perceived diversity in Software Engineering. In addition, we found that gender diversity has been widely studying. Previous studies demonstrate that women increase productivity, performance, and efficiency. But, unfortunately, some developers still have a strong bias against women in OSS and industry. Nationality and age diversity has been the second and third most studied aspects. The papers indicate the differences that can occur when different cultures work together and that some old developers are seldom found in OSS and SO communities. Finally, our SLR indicates that race has been the perceived diversity aspect least studied in SE and they should be further studied.

7.1 Replication Package

The replication package can be found online.[10]

References

Morgan S. (2017) How are programming questions from women received on stack overflow? a case study of peer parity. In: Proceedings Companion of the 2017 ACM SIGPLAN International conference on systems, programming, languages, and applications: Software for Humanity. pp 39–41

Adikaram AS, Wijayawardena K (2015) What happens to female employees in skewed it project teams in sri lanka? revisiting kanter. South Asian J Human Resour Manag 2(1):37–57

Ahmadi M, Weibert A, Ogonowski C, Aal K, Gäckle K, Marsden N, Wulf V (2018) Challenges and lessons learned by applying living labs in gender and it contexts. In: Proceedings of the 4th Conference on gender & IT, pp 239–249

Al-Ani B, Marczak S, Prikladnicki R, Redmiles D (2013) Revisiting the factors that engender trust of global systems engineers. In: 2013 IEEE 8th international conference on global software engineering. IEEE, pp 31–40

Alfaro Iván, Chandrasekaran R (2015) Software quality and development speed in global software development teams. Bus Inf Syst Eng 57(2):91–102

Allport GW, Clark K, Pettigrew T (1954) The nature of prejudice

Altiner S, Ayhan MB (2018) An approach for the determination and correlation of diversity and efficiency of software development teams. S Afr J Sci 114(3-4):1–9

Anderson G, Keith M, Albrecht C, Spruill A, Pettit C (2019) Optimizing software team performance with cultural differences. In: Proceedings of the 52nd Hawaii International conference on system sciences

Angell R, Johnson B, Brun Y, Meliou A (2018) Themis: Automatically testing software for discrimination. In: Proceedings of the 2018 26th ACM Joint meeting on european software engineering conference and symposium on the foundations of software engineering. pp 871–875

Aué J, Haisma M, Tómasdóttir KF, Bacchelli A (2016) Social diversity and growth levels of open source software projects on github. In: Proceedings of the 10th ACM/IEEE International symposium on empirical software engineering and measurement. pp 1–6

Aydemir FB, Dalpiaz F (2018) A roadmap for ethics-aware software engineering. In: 2018 IEEE/ACM International workshop on software fairness (FairWare). IEEE, pp 15–21

Balali S, Steinmacher I, Annamalai U, Sarma A, Gerosa MA (2018) Newcomers' barriers... is that all? an analysis of mentors' and newcomers' barriers in oss projects. Computer Supported Cooperative Work (CSCW) 27(3-6):679–714

Baltes S, Park G, Serebrenik A (2020) Is 40 the new 60? how popular media portrays the employability of older software developers. IEEE Softw 37(6):26–31

---

[10] https://vault.cs.uwaterloo.ca/s/Zc5qDwoDmKzA7gG

Bastarrica MC, Simmonds J (2019) Gender differences in self and peer assessment in a software engineering capstone course. In: 2019 IEEE/ACM 2nd international workshop on gender equality in software engineering (GE). IEEE, pp 29–32

Beckwith L, Burnett M, Grigoreanu V, Wiedenbeck S (2006) Gender hci: What about the software? Computer 39(11):97–101

Beckwith L, Burnett M, Wiedenbeck S, Cook C, Sorte S, Hastings M (2005) Effectiveness of end-user debugging software features: Are there gender issues? In: Proceedings of the SIGCHI Conference on human factors in computing systems, pp 869–878

Beckwith L, Inman D, Rector K, Burnett M (2007) On to the real world: Gender and self-efficacy in excel. In: IEEE Symposium on visual languages and human-centric computing (VL/HCC 2007). IEEE, pp 119–126

Beckwith L, Kissinger C, Burnett M, Wiedenbeck S, Lawrance J, Blackwell A, Cook C (2006) Tinkering and gender in end-user programmers' debugging. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, pages 231–240

Beckwith L, Sorte S, Burnett M, Wiedenbeck S, Chintakovid T, Cook C (2005) Designing features for both genders in end-user programming environments. In: 2005 IEEE Symposium on visual languages and human-centric computing (VL/HCC'05). IEEE, pp 153–160

Begel A, Dominic J, Phillis C, Beeson T, Rodeghero P (2021) How a remote video game coding camp improved autistic college students' self-efficacy in communication

Bertrand M, Mullainathan S (2004) Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. Am Econ Rev 94(4):991–1013

Blincoe K, Springer O, Wrobel MR (2019) Perceptions of gender diversity's impact on mood in software development teams. IEEE Softw 36(5):51–56

Bollmann S, Cook D, Dumas J, Fox J, Josse J, Keyes O, Strobl C, Turner H, Debelak R (2017) A first survey on the diversity of the r community

Borchers G (2003) The software engineering impacts of cultural factors on multi-cultural software development teams. In: 25th International conference on software engineering, 2003. Proceedings. IEEE, pp 540–545

Bosu A, Sultana KZ (2019) Diversity and inclusion in open source software (oss) projects: Where do we stand? In: 2019 ACM/IEEE international symposium on empirical software engineering and measurement (ESEM), pp 1–11. IEEE

Bourque P, Fairley RE et al (2014) Guide to the software engineering body of knowledge (SWEBOK (R)) :Version 3.0. IEEE Computer Society Press, Washington

Braswel KM, Johnson J, Brie'anna B, Jamie Payton. (2021) Pivoting during a pandemic: Designing a virtual summer camp to increase confidence of black and latina girls. In: the 52 ACM Technical symposium on computer science education

Brun Y, Meliou A (2018) Software fairness. In: Proceedings of the 2018 26th ACM joint meeting on european software engineering conference and symposium on the foundations of software engineering, pp 754–759

Buffardi K (2016) Localized open source software projects : Exploring realism and motivation. In: 2016 11th International conference on computer science & education (ICCSE). IEEE, pp 382–387

Burnett M, Counts R, Lawrence R, Hanson H (2017) Gender hcl and microsoft: Highlights from a longitudinal study. In: 2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). IEEE, pp 139–143

Burnett M, Fleming SD, Iqbal S, Venolia G, Rajaram V, Farooq U, Grigoreanu V, Czerwinski M (2010) Gender differences and programming environments: across programming populations. In: Proceedings of the 2010 ACM-IEEE international symposium on empirical software engineering and measurement, pp 1–10

Burnett M, Peters A, Hill C, Elarief N (2016) Finding gender-inclusiveness software issues with gendermag: a field investigation. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 2586–2598

Burnett M, Sarma A, Mendez C, Oleson A, Hilderbrand C, Steine-Hanson Z, Ko AJ (2018) Gender biases in software for problem-solving. In: Designing technologies to support human problem solving: workshop at VL/HCC. IEEE

Burnett M, Stumpf S, Macbeth J, Makri S, Beckwith L, Kwan I, Peters A, Jernigan W (2016) Gendermag: A method for evaluating software's gender inclusiveness. Interact Comput 28(6):760–787

Byrne DE (1971) The attraction paradigm, vol 11. Academic Press, Cambridge

Canedo ED, Tives HA, Marioti MB, Fagundes F, de Cerqueira JAS (2019) Barriers faced by women in software development projects. Information 10(10):309

Cao J, Rector K, Park TH, Fleming SD, Burnett M, Wiedenbeck S (2010) A debugging perspective on end-user mashup programming. In: 2010 IEEE Symposium on visual languages and human-centric computing. IEEE, pp 149–156

Casey V (2009) Leveraging or exploiting cultural difference? In: 2009 Fourth IEEE international conference on global software engineering, pp 8–17. IEEE

Catolino G, Palomba F, Tamburri DA, Serebrenik A, Ferrucci F (2019) Gender diversity and community smells: insights from the trenches. IEEE Softw 37(1):10–16

Catolino G, Palomba F, Tamburri DA, Serebrenik A, Ferrucci F (2019) Gender diversity and women in software teams: How do they affect community smells? In: 2019 IEEE/ACM 41st International conference on software engineering: software engineering in society (ICSE-SEIS), pp 11–20. IEEE

Chatterjee A, Guizani M, Stevens C, Emard J, May ME, Burnett M, Ahmed I, Aid AnitaSarma. (2021) An automated detector for gender-inclusivity bugs in oss project pages. In: 2021 IEEE/ACM 43th International conference on software engineering (ICSE). IEEE

Chintakovid T, Wiedenbeck S (2009) User perceptions and gender in end-user debugging: How do they affect outcomes? In: 2009 IEEE symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 217–224

Choi N, Pruett JA (2015) The characteristics and motivations of library open source software developers: An empirical study. Libr Inf Sci Res 37(2):109–117

Cunha AB, Canen AG, Capretz MAM (2009) Personalities, cultures and software modeling: Questions, scenarios and research directions. In: 2009 ICSE Workshop on cooperative and human aspects on software engineering. IEEE, pp 23–31

Cunningham SJ, Hinze A, Nichols DM (2016) Supporting gender-neutral digital library creation; A case study using the gendermag toolkit. In: International conference on asian digital libraries. Springer, pp 45–50

Daniel S, Agarwal R, Stewart KJ (2013) The effects of diversity in global, distributed collectives: A study of open source project success. Inf Syst Res 24(2):312–333

Davidson JL, Mannan UA, Naik R, Dua I, Jensen C (2014) Older adults and free/open source software: A diary study of first-time contributors. In: Proceedings of the international symposium on open collaboration. pp 1–10

Davidson JL, Naik R, Mannan UA, Azarbakht A, Jensen C (2014) On older adults in free/open source software: reflections of contributors and community leaders. In: 2014 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 93–100

Dong W, Ehrlich K, Macy MM, Muller M (2016) Embracing cultural diversity: Online social ties in distributed workgroups. In: Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing, pp 274–287

Draude C, Maaß S (2018) Making it work: integrating gender research in computing through a process model. In: Proceedings of the 4th conference on gender & IT. pp 43–50

Easterbrook S, Singer J, Storey Margaret-Anne, Damian D (2008) Selecting empirical methods for software engineering research. In: Guide to advanced empirical software engineering. Springer, pp 285–311

Erete S, Rankin YA, Thomas JO (2021) I can't breathe: Reflections from black women in cscw and hci. Proc ACM Human-Comput Interact 4(CSCW3):1–23

Evans JSTBT (2003) In two minds: dual-process accounts of reasoning. Trends Cognit Sci 7(10):454–459

Fazli F, Bittner EAC (2017) Cultural influences on collaborative work in software engineering teams. In: Proceedings of the 50th Hawaii international conference on system sciences

Filippova A, Trainer E, Herbsleb JD (2017) From diversity by numbers to diversity as process: supporting inclusiveness in software development teams with brainstorming. In: 2017 IEEE/ACM 39th International conference on software engineering (ICSE). IEEE, pp 152–163

Fisher M, Cox A, Zhao L (2006) Using sex differences to link spatial cognition and program comprehension. In: 2006 22nd IEEE international conference on software maintenance. IEEE, pp 289–298

Ford D (2020) Approaches to understanding a socio-cultural experience of global knowledge workers. In: Proceedings of the 2020 CHI Conference on human factors in computing systems

Ford D, Behroozi M, Serebrenik A, Parnin C (2019) Beyond the code itself: how programmers really look at pull requests. In: 2019 IEEE/ACM 41st International conference on software engineering :software engineering in society (ICSE-SEIS). IEEE, pp 51–60

Ford D, Harkins A, Parnin C (2017) Someone like me: How does peer parity influence participation of women on stack overflow? In: 2017 IEEE symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 239–243

Ford D, Milewicz R, Serebrenik A (2019) How remote work can foster a more inclusive environment for transgender developers. In: 2019 IEEE/ACM 2nd International workshop on gender equality in software engineering (GE). IEEE, pp 9–12

Ford D, Smith J, Guo PJ, Parnin C (2016) Paradise unplugged: Identifying barriers for female participation on stack overflow. In: Proceedings of the 2016 24th ACM SIGSOFT International symposium on foundations of software engineering, pp 846–857

Furtado L, Cartaxo B, Treude C, Pinto G (2020) How successful are open source contributions from countries with different levels of human development? IEEE Softw

Galhotra S, Brun Y, Meliou A (2017) Fairness testing: testing software for discrimination. In: Proceedings of the 2017 11th Joint meeting on foundations of software engineering. pp 498–510

German D, Robles G, Poo-Caamaño G, Yang X, Iida H, Inoue K (2018) "was my contribution fairly reviewed?" a framework to study the perception of fairness in modern code reviews. In: 2018 IEEE/ACM 40th International conference on software engineering (ICSE). IEEE, pp 523–534

Ghosh RA, Glott R, Krieger B, Robles G (2002) Free/libre and open source software. Survey and study

Gila AR, Jaafa J, Omar M, Tunio MZ (2014) Impact of personality and gender diversity on software development teams' performance. In: 2014 International conference on computer, communications, and control technology (I4CT). IEEE, pp 261–265

Gilal AR, Jaafar J, Omar M, Basri S, Waqas A (2016) A rule-based model for software development team composition Team leader role with personality types and gender classification. Inf Softw Technol 74:105–113

Gómez O. S, Solari M, Pardo CJ, Ledezma AC (2017) A controlled experiment on productivity of pair programming gender combinations: Preliminary results. In: CIbSE, pp 679–692

Gralha C, Goulao M, Araujo J (2019) Analysing gender differences in building social goal models: a quasi-experiment. In: 2019 IEEE 27th International requirements engineering conference (RE). IEEE, pp 165–176

Gren L (2018) On gender, ethnicity, and culture in empirical software engineering research. In: 2018 IEEE/ACM 11th International workshop on cooperative and human aspects of software engineering (CHASE). IEEE, pp 77–78

Grigoreanu V, Brundage J, Bahna E, Burnett M, ElRif P, Snover J (2009) Males' and females' script debugging strategies. In: International symposium on end user development. Springer, pp 205–224

Grigoreanu V, Cao J, Kulesza T, Bogart C, Rector K, Burnett M, Wiedenbeck S (2008) Can feature design reduce the gender gap in end-user software development environments?

Guzman E, Rojas AP (2019) Gender and user feedback: An exploratory study. In: 2019 IEEE 27th International requirements engineering conference (RE). IEEE, pp 381–385

Hall T, Beecham S, Bowes D, Gray D, Counsell S (2011) A systematic literature review on fault prediction performance in software engineering. IEEE Trans Softw Eng 38(6):1276–1304

Hall C, Durruthy R (2020) Workforce-diversity-report. https://careers.linkedin.com/diversity-and-inclusion/workforce-diversity-report. Accessed: 2020-05-05

Heiniger S, Mercie H (2018) National bias of international gymnastics judges during the 2013-2016 olympic cycle

Hertel G, van der Heijden BIJM, de Lange AH, Deller J, Rauschenbach C, Krumm S, Thielgen M (2013) Age and work-related stress: a review and meta-analysis. J Manag Psychol

Hilderbrand C, Perdriau C, Letaw L, Emard J, Steine-Hanson Z, Burnett M, Sarma A (2020) Engineering gender-inclusivity into software: Ten teams' tales from the trenches. In: International conference on software engineering, 2020. Proceedings. IEEE

Hill C, Ernst S, Oleson A, Horvath A, Burnett M (2016) Gendermag experiences in the field The whole, the parts, and the workload. In: 2016 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 199–207

Hill C, Haag M, Oleson A, Mendez C, Marsden N, Sarma A (2017) Margaret Burnett. Gender-inclusiveness personas vs. stereotyping: Can we have it both ways? In: Proceedings of the 2017 chi conference on human factors in computing systems, pp 6658–6671

Hofstede G (2005) Cultures and organizations: software of the mind/geert hofstede and gert jan hofstede

Holtkamp P, Jokinen JussiPP, Pawlowski JM (2015) Soft competency requirements in requirements engineering, software design, implementation, and testing. J Syst Softw 101:136–146

Homan AC, Van Knippenberg D, Van Kleef GA, De Dreu CarstenKW (2007) Bridging faultlines by valuing diversity: diversity beliefs, information elaboration, and performance in diverse work groups. J Appl Psychol 92(5):1189

Hongyun Yu, Bao X, Yang S (2009) Research and improvement of team software process. In: 2009 WRI World congress on computer science and information engineering, vol 7. IEEE, pp 654–658

Hosoda M, Nguyen LT, Stone-Romero EF (2012) The effect of hispanic accents on employment decisions. J Manag Psychol 27(4):347–364

Huang Y, Ford D, Zimmermann T (2021) Leaving my fingerprints: Motivations and challenges of contributing to oss for social good. In: 2021 IEEE/ACM 43th international conference on software engineering (ICSE). IEEE

Huff EW, Boateng K, Moster M, Rodeghero P, Brinkley J (2020) Examining the work experience of programmers with visual impairments. In: 2020 IEEE International conference on software maintenance and evolution (ICSME). IEEE, pp 707–711

Hui JS, Farnham SD (2016) Designing for inclusion Supporting gender diversity in independent innovation teams. In: Proceedings of the 19th lork. pp 71–85

Hyrynsalmi S (2019) The underrepresentation of women in the software industry: thoughts from career-changing women. In: 2019 IEEE/ACM 2nd International workshop on gender equality in software engineering (GE). IEEE, pp 1–4

Hyrynsalmi S, Hyrynsalmi S (2019) What motivates adult age women to make a career change to the software industry? In: 2019 IEEE International conference on engineering, technology and innovation (ICE/ITMC). IEEE, pp 1–8

Imtiaz N, Middleton J, Chakraborty J, Robson N, Bai G, Murphy-Hill E (2019) Investigating the effects of gender bias on github. In: 2019 IEEE/ACM 41st international conference on software engineering (ICSE). IEEE, pp 700–711

Irrgang M (2018) Hands-on participatory and interdisciplinary design in computer science: an example. In: Proceedings of the 4th Conference on gender & IT. pp 29–33

Izquierdo D, Huesman N, Serebrenik A, Robles G (2018) Openstack gender diversity report. IEEE Softw 36(1):28–33

James T, Galster M, Blincoe K, Miller G (2017) What is the perception of female and male software professionals on performance, team dynamics and job satisfaction? insights from the trenches. In: 2017 IEEE/ACM 39th International conference on software engineering :software engineering in practice track (ICSE-SEIP). IEEE, pp 13–22

James S, Herman J, Rankin S, Keisling M, Mottet L, Anafi M (2016) The report of the 2015 us transgender survey

Jehn KA, Northcraft GB, Neale MA (1999) Why differences make a difference: A field study of diversity, conflict and performance in workgroups. Adm Sci Q 44(4):741–763

Judy KH (2012) Agile values, innovation and the shortage of women software developers. In: 2012 45th Hawaii International conference on system sciences. IEEE, pp 5279–5288

Kanij T, Merkel R, Grundy J (2011) A preliminary study on factors affecting software testing team performance. In: 2011 International symposium on empirical software engineering and measurement. IEEE, pp 359–362

Kaur Kuttal S, Gerstner K, Bejarano A (2019) Remote pair programming in online cs education: Investigating through a gender lens. In: 2019 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 75–85

Kirton G, Robertson M (2018) Sustaining and advancing IT careers: Women's experiences in a uk-based IT company. J Strateg Inf Syst 27(2):157–169

Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering

Kopeć W, Balcerzak B, Nielek R, Kowalik G, Wierzbicki A, Casati F (2018) Older adults and hackathons: a qualitative study. Empir Softw Eng 23(4):1895–1930

Kopeć W, Nielek R, Wierzbicki A (2018) Guidelines towards better participation of older adults in software development processes using a new spiral method and participatory approach. In: Proceedings of the 11th International workshop on cooperative and human aspects of software engineering. pp 49–56

Krishna S, Sahay S, Walsham G (2004) Managing cross-cultural issues in global software outsourcing. Commun ACM 47(4):62–66

Kuechler V, Gilbertson C, Jensen C (2012) Gender differences in early free and open source software joining process. In: IFIP International conference on open source systems. Springer, pp 78–93

Lee A, Carver JC (2019) Floss participants' perceptions about gender and inclusiveness: a survey. In: 2019 IEEE/ACM 41st International conference on software engineering (ICSE). IEEE, pp 677–687

Lenberg P, Feldt R, Wallgren LG (2015) Behavioral software engineering: A definition and systematic literature review. Journal of Systems and software 107:15–37

Lombardi EL, Wilchins RA, Priesing D, Malouf D (2002) Gender violence: Transgender experiences with violence and discrimination. J Homosex 42(1):89–101

Lyon LA, Clayton C, Green E (2018) Tinkering in the wild: What leads to success for female end-user programmers? In: 2018 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 125–129

Machado LS, Caldeira C, Perin M, Souza CleidsonRBde (2020) Gendered experiences of software engineers during the covid-19 crisis. IEEE Softw

Marques M (2015) Software engineering education—does gender matter in project results?—a chilean case study. In: 2015 IEEE frontiers in education conference (FIE). IEEE, pp 1–8

May A, Wachs J, Hannák A (2019) Gender differences in participation and reward on stack overflow. Empir Softw Eng 24(4):1997–2019

Mealin S, Murphy-Hill E (2012) An exploratory study of blind software developers. In: 2012 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 71–74

Mendez C, Hanson ZS, Oleson A, Horvath A, Hill C, Hilderbrand C, Sarma A, Burnett M (2018) Semi-automating (or not) a socio-technical method for socio-technical systems. In: 2018 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 23–32

Mendez C, Letaw L, Burnett M, Stumpf S, Sarma A, Hilderbrand C (2019) From gendermag to inclusivemag; An inclusive design meta-method. In: 2019 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 97–106

Mendez C, Padala HS, Steine-Hanson Z, Hilderbrand C, Horvath A, Hill C, Simpson L, Patil N, Sarma A, Burnett M (2018) Open source barriers to entry, revisited: A sociotechnical perspective. In: Proceedings of the 40th International conference on software engineering, pp 1004–1015

Mendez C, Sarma A, Burnett M (2018) Gender in open source software: what the tools tell. In: Proceedings of the 1st International workshop on gender equality in software engineering. pp 21–24

Menezes Álvaro, Prikladnicki R (2018) Diversity in software engineering. In: Proceedings of the 11th International workshop on cooperative and human aspects of software engineering. pp 45–48

Metaxa-Kakavouli Danaë, Wang K, Landay JA, Hancock J (2018) Gender-inclusive design: sense of belonging and bias in web interfaces. In: Proceedings of the 2018 CHI Conference on human factors in computing systems, pages 1–6

Mishra A, Mishra D (2014) Cultural issues in distributed software development: A review. In: OTM Confederated international conferences" On the Move to Meaningful Internet Systems". Springer, pp 448–456

Moon E (2013) Gendered patterns of politeness in free/libre open source software development. In: 2013 46th Hawaii international conference on system sciences. IEEE, pp 3168–3177

Morris MR, Begel A, Wiedermann B. (2015) Understanding the challenges faced by neurodiverse software engineering employees: Towards a more inclusive and productive technical workforce. In: Proceedings of the 17th International ACM SIGACCESS Conference on computers & accessibility. pp 173–184

Morrison P, Murphy-Hill E (2013) Is programming knowledge related to age? In: In companion to the working conference on mining software repositories. Citeseer, pp 1–4

Morrison P, Pandita R, Murphy-Hill E, McLaughlin A (2016) Veteran developers' contributions and motivations: An open source perspective. In: 2016 IEEE Symposium on visual languages and human-centric computing (VL/HCC). IEEE, pp 171–179

Muller MJ, Kuhn S (1993) Participatory design. Commun ACM 36(6):24–28

Murakami Y, Tsunoda M, Uwano H (2017) Wap: Does reviewer age affect code review performance? In: In 2017 IEEE 28th International symposium on software reliability engineering (ISSRE). IEEE, pp 164–169

Murphy-Hill E, Jaspan C, Sadowski C, Shepherd D, Phillips M, Winter C, Knight A, Smith E, Jorde M (2019) What predicts software developers' productivity? IEEE Trans Softw Eng

Nadri R, Perez GR, Nagappan M (2020) Insights into nonmerged pull requests in github: Is there evidence of bias based on perceptible race? IEEE Softw

Nafus D (2012) 'patches don't have gender': What is not open in open source software. New Media Soc 14(4):669–683

O'higgins N (2001) Youth unemployment and employment policy: A global perspective. O'Higgins

Ortu M, Destefanis G, Counsell S, Swift S, Tonelli R, Marchesi M (2017) How diverse is your team? investigating gender and nationality diversity in github teams. J Softw Eng Res Dev 5(1):1–18

Østergaard CR, Timmermans B, Kristinsson K (2011) Does a different view create something new? the effect of employee diversity on innovation. Res Pol 40(3):500–509

Padala S, Mendez C, Dias LF, Steinmacher I, Hanson ZS, Hilderbrand C, Horvath A, Hill C, Simpson L, Burnett M, et al. (2020) How gender-biased tools shape newcomer experiences in oss projects. IEEE Trans Softw Eng

Page Scott E (2007) How the power of diversity creates better groups, firms schools and societies

Paris G, Leo GD, Menozzi P, Gatto M (1998) Region-based citation bias in science. Nature 396(6708):210

Paul R, Bosu A, Sultana KZ (2019) Expressions of sentiments during code reviews Male vs. female. In: 2019 IEEE 26th international conference on software analysis, evolution and reengineering (SANER). IEEE, pp 26–37

Pieterse V, van Eekelen M (2018) Cultural diversity and the performance of student software engineering teams. In: Southern African Computer Lecturers Association (SACLA), vol 2018, p 120

Posthuma RA, Campion MA (2009) Age stereotypes in the workplace: Common stereotypes, moderators, and future research directions. J Manag 35(1):158–188

Prado R, Galeno WM, Gama KS, Pinto G (2020) How trans-inclusive are hackathons? IEEE Softw

Pretorius C, Razavian M, Eling K, Langerak F (2020) Combined intuition and rationality increases software feature novelty for female software designers. IEEE Softw

Qiu HS, Li YL, Padala S, Sarma A, Vasilescu A (2019) The signals that potential contributors look for when choosing open-source projects. Proc ACM Human-Comput Interact 3(CSCW):1–29

Qiu HS, Nolte A, Brown A, Serebrenik A, Vasilescu B (2019) Going farther together: The impact of social capital on sustained participation in open source. In: 2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE). IEEE, pp 688–699

Quesenberry JL, Trauth EM (2012) The (dis) placement of women in the it workforce: an investigation of individual career values and organisational interventions. Inf Syst J 22(6):457–473

Rastogi A, Nagappan N, Gousios G, van der Hoek André (2018) Relationship between geographical location and evaluation of developer contributions in github. In: Proceedings of the 12th ACM/IEEE International symposium on empirical software engineering and measurement, pp 1–8

Razavian M, Lago P (2015) Feminine expertise in architecting teams. IEEE Softw 33(4):64–71

Robert LP (2016) Far but near or near but far? the effects of perceived distance on the relationship between geographic dispersion and perceived diversity. In: Proceedings of the 2016 CHI Conference on human factors in computing systems. pp 2461–2473

Robles G, Reina LA, González-Barahona JM, Domínguez SD (2016) Women in free/libre/open source software: The situation in the 2010s. In: IFIP International conference on open source systems. Springer, pp 163–173

Ross M, Hazari Z, Sonnert G, Sadler P (2020) The intersection of being black and being a woman: Examining the effect of social computing relationships on computer science career choice. ACM Trans Comput Educ (TOCE) 20(2):1–15

Scheuerman MK, Spiel K, Haimson OL, Foad H, Branham SM (2021) HCI Guidelines for Gender Equity and Inclusivity. https://www.morgan-klaus.com/gender-guidelines.html. Accessed: 2021-02-15

Schloegel U, Stegmann S, Maedche A, Van Dick R (2016) Reducing age stereotypes in software development: The effects of awareness-and cooperation-based diversity interventions. J Syst Softw 121:1–15

Schloegel U, Stegmann S, Van Dick R, Maedche A (2018) Age stereotypes in distributed software development The impact of culture on age-related performance expectations. Inf Softw Technol 97:146–162

Shachaf P (2008) Cultural diversity and information and communication technology impacts on global virtual teams: An exploratory study. Inf Manag 45(2):131–142

Shah H, Harrold MJ (2013) Culture and testing: What is the relationship? In: 2013 IEEE 8th International conference on global software engineering. IEEE, pp 51–60

Sharp H, Dittrich Y, Souza CleidsonRBD. (2016) The role of ethnographic studies in empirical software engineering. IEEE Trans Softw Eng 42(8):786–804

Silveira KK, Prikladnicki R (2019) A systematic mapping study of diversity in software engineering: a perspective from the agile methodologies. In: 2019 IEEE/ACM 12th International workshop on cooperative and human aspects of software engineering (CHASE). IEEE, pp 7–10

Singh V (2019) Women-only spaces of open source. In: 2019 IEEE/ACM 2nd International workshop on gender equality in software engineering (GE). IEEE, pp 17–20

Singh V (2019) Women participation in open source software communities. In: Proceedings of the 13th European conference on software architecture-Volume 2, pp 94–99

Singh V, Brandon W (2019) Open source software community inclusion initiatives to support women participation. In: IFIP International conference on open source systems. Springer, pp 68–79

Spichkova M, Schmidt H, Trubiani C (2017) Role of women in software architecture: an attempt at a systematic literature review. In: Proceedings of the 11th European conference on software architecture : companion proceedings. pp 31–34

Storey Margaret-Anne, Zagalsky A, Filho FF, Singer L, German DM (2016) How social and communication channels shape and challenge a participatory culture in software development. IEEE Trans Softw Eng 43(2):185–204

Subrahmaniyan N, Beckwith L, Grigoreanu V, Burnett M, Wiedenbeck S, Narayanan V, Bucht K, Drummond R, Fern X (2008) Testing vs. code inspection vs. what else? male and female end users' debugging strategies. In: Proceedings of the SIGCHI Conference on human factors in computing systems. pp 617–626

Terrell J, Kofink A, Middleton J, Rainear C, Murphy-Hill E, Parnin C, Stallings J (2017) Gender differences and bias in open source: Pull request acceptance of women versus men. Peer J Comput Sci 3:e111

Thomas JO, Joseph N, Williams A, Burge J et al (2018) Speaking truth to power: Exploring the intersectional experiences of black women in computing. In: 2018 Research on equity and sustained participation in engineering, computing, and technology (RESPECT). IEEE, pp 1–8

Tourani P, Adams B, Serebrenik A (2017) Code of conduct in open source projects. In: 2017 IEEE 24th International conference on software analysis, evolution and reengineering (SANER). IEEE, pp 24–33

Usher R (2006) North American lexicon of transgender terms, GLB Publishers, San Francisco

Vasilescu B, Capiluppi A, Serebrenik A (2014) Gender, representation and online participation: A quantitative study. Interact Comput 26(5):488–511

Vasilescu B, Filkov V (2015) Alexander Serebrenik. Perceptions of diversity on git hub: A user survey. In: 2015 IEEE/ACM 8th International workshop on cooperative and human aspects of software engineering. IEEE, pp 50–56

Vasilescu B, Posnett D, Ray B, Brand MarkGJvanden, Serebrenik A, Devanbu P, Filkov V (2015) Gender and tenure diversity in github teams. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 3789–3798

Vedres B, Vasarhelyi O (2019) Gendered behavior as a disadvantage in open source software development. EPJ Data Sci 8(1):25

Vorvoreanu M, Zhang L, Huang Yun-Han, Hilderbrand C, Steine-Hanson Z, Burnett M (2019) From gender biases to gender-inclusive design: An empirical investigation. In: Proceedings of the 2019 CHI Conference on human factors in computing systems. pp 1–14

Wang Y (2018) Understanding the reputation differences between women and men on stack overflow. In: 2018 25th Asia-pacific software engineering conference (APSEC). IEEE, pp 436–444

Wang Z, Yi W, Redmiles D (2018) Competence-confidence gap: A threat to female developers' contribution on github. In: 2018 IEEE/ACM 40th International conference on software engineering : software engineering in society (ICSE-SEIS). IEEE, pp 81–90

Wohlin C, Runeson P, Höst M, Ohlsson MC, Regnell B, Wesslén A (2012) Experimentation in software engineering. Springer Science & Business Media, Berlin

Wurzelová P., Palomba F, Bacchelli A (2019) Characterizing women (not) contributing to open-source. In: 2019 IEEE/ACM 2nd International workshop on gender equality in software engineering (GE). IEEE, pp 5–8

Yi W, Redmiles D (2019) Implicit gender biases in professional software development: An empirical study. In: 2019 IEEE/ACM 41st International conference on software engineering: software engineering in society (ICSE-SEIS). IEEE, pp 1–10

Yi W, Zhang M (2020) Reducing implicit gender biases in software development: does intergroup contact theory work? In: Proceedings of the 28th ACM Joint meeting on european software engineering conference and symposium on the foundations of software engineering, pp 580–592

Zacchiroli S (2020) Gender differences in public code contributions: a 50-year perspective. IEEE Softw

Zohreh Sharafi, Soh Z, Guéhéneuc Y-G, Antoniol G (2012) Women and men—different but equal: On the impact of identifier style on source code reading. In: 2012 20th IEEE International conference on program comprehension (ICPC). IEEE, pp 27–36

Gema Rodríguez-Pérez is an assistant professor in the Irving K. Barber Faculty of Science at the University of British Columbia. Her research interests include empirical software engineering, mining software repositories, bug prediction and detection, and diversity in software engineering.. Dr. Rodriguez-Perez received her Ph.D. from the University Rey Juan Carlos, Madrid, Spain.

Reza Nadri is a Software and Data Architect at RideFlag, Missis-sauga, Ontario, Canada. Nadri was aresearch assistant at the University of Waterloo. His research interests include empirical software engineering, mining software repositories, and diversity in software engineering. Nadri received a master's degree from the Cheriton School of Computer Science, University of Waterloo. Contact him at rnadri@uwaterloo.ca.

Meiyappan Nagappan is an associate professor and the Current Ross & Muriel Cheriton Faculty Fellow in the David R. Cheriton School of Computer Science at the University of Waterloo, Water-loo, Ontario, N2L 3G1, Canada. His research is centered around the use of large-scale software engineering data to address the concerns of the various stakeholders (for example, developers, operators, and managers). Nagappan received a Ph.D. in computer science from North Carolina State University. Further information about him can be found at http://www.mei-nagappan.com. Contact him at mei.nagappan@uwaterloo.ca.