UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Trouble in programmer's paradise

gender-biases in sharing and recognising technical knowledge on Stack Overflow

Brooke, S.J.

DOI
10.1080/1369118X.2021.1962943

Publication date
2021

Document Version
Final published version

Published in
Information Communication and Society

License
CC BY

Link to publication

Citation for published version (APA):
Brooke, S. J. (2021). Trouble in programmer's paradise: gender-biases in sharing and recognising technical knowledge on Stack Overflow. *Information Communication and Society*, *24*(14), 2091-2112. https://doi.org/10.1080/1369118X.2021.1962943

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:12 Sept 2026

INFORMATION, COMMUNICATION & SOCIETY
2021, VOL. 24, NO. 14, 2091–2112
https://doi.org/10.1080/1369118X.2021.1962943

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Trouble in programmer's paradise: gender-biases in sharing and recognising technical knowledge on Stack Overflow

S. J. Brooke [a,b]

[a]Department of Methodology, London School of Economics and Political Science, London, UK; [b]Oxford Internet Institute, University of Oxford, Oxford, UK

ABSTRACT

This paper examines gender-biases on Stack Overflow, the world's largest question-and-answer forum of programming knowledge. Employing a non-binary gender identification built on usernames, I investigate the role of gender in shaping users' experience in technical forums. The analysis encompasses 11-years of activity, across levels of expertise, language, and specialism, to assess if Stack Overflow is really a paradise for programmers. I first examine individual users, asking if there are gender differences in key user metrics of success, focusing on reputation points, user tenure, and level of activity. Second, I test if there are gender-biases in how technical knowledge is recognised in the question-answer format of the platform. Third, using social network analysis I investigate if interaction on Stack Overflow is organised by gender. Results show that sharing and recognising technical knowledge is dictated by users' gender, even when it is operationalised beyond a binary. I find that feminine users receive lower scores for their answers, despite exhibiting higher effort in their contributions. I also show that interaction on Stack Overflow is organised by gender. Specifically, that feminine users preferentially interact with other feminine users. The findings emphasise the central role of gender in shaping interaction in technical spaces, a necessity for participation in the masculine-dominated forum. I conclude the study with recommendations for inclusivity in online forums.

ARTICLE HISTORY
Received 8 January 2021
Accepted 24 July 2021

KEYWORDS
Gender; programming; social networks; anonymity; sexism; technology

Introduction

Often called a 'programmer's paradise', Stack Overflow (SO) is a question and answer-based website focused on writing computer code. It is the largest online community of programming knowledge, boasting over 14 million registered users and more than 50 million monthly visitors (SO, 2020). Its convenience and accessibility are typical of the open nature of coding culture, so often proclaimed to be a meritocracy (Tanczer, 2016). SO covers a range of topics and levels of technical familiarity, from complete

CONTACT S. J. Brooke s.j.brooke@lse.ac.uk Department of Methodology, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

beginners to seasoned programmers dealing with questions of algorithmic complexity. With low material barriers to entry, in theory anyone can join the forum, allowing enthusiasts and professionals to meet with a shared passion for 'open' technology. SO (2020) is a nexus of computer-mediated communication and programming knowledge where participation can have real consequences for employment, drawing people from all walks of life to hone their coding skills.

In turning to existing scholarship on SO, the role of gender in metrics of success and activity has been analysed on an individual level (Ford et al., 2016; May et al., 2019; Vasilescu et al., 2012), however the effect of gender on interaction dynamics and evaluation of contributions has yet to be studied. Scholars have employed social network analysis to examine geographies of participation on SO (Stephany et al., 2020), but not gender. I therefore aim to shed new light on how gender-biases dictate the recognition and sharing of technical knowledge. I further contribute to the literature on gender and technology with a methodological strategy that incorporates a non-binary operationalisation of gender. I focus on 'gender salience', or intensity of gendered self-expression, spanning five-categories: *masculine, mostly masculine, anonymous, mostly feminine, feminine*. This framing permits the analysis to capture if gender-biases on technical forums are intensified with extreme poles (masculine/feminine) of gender representation.

This study problematises the notion that SO is a paradise for programmers. Rather, I propose that gender-biases dictate patterns of sharing and recognising technical knowledge on the platform. This paper first aims to characterise gender gaps on the forum, asking what differences are present in key platform metrics, second if the recognition of technical knowledge in answers is gendered, and third if interaction on SO is organised by gender. The research questions are:

(RQ1) Are there gender differences in user metrics of reputation, tenure, and activity?

(RQ2) Are there gender-biases in how technical answers are scored?

(RQ3) Is interaction on Stack Overflow organised by gender?

To contextualise these research questions, I first introduce SO as a platform. Next, I turn to related literature outlining Risman's (2004) framing of gender as social structure, and relevant perspectives from studies on gender and technology. I then delineate the methodological approach, including the non-binary gender identification procedure and how I will operationalise effort and test gender-biases in the scoring of technical answers. Following this, I will turn to the findings where the results subsections corresponding to the three research questions. Finally, I discuss the implications of the results, concluding with recommendations for gender inclusivity on SO. The findings of this study have consequences for how we understand sexism in the wider technology community, particularly in the sharing and recognition of technical knowledge.

Stack Overflow

The image of an open, anonymous forum is fitting with early, utopian visions of the internet (Wajcman, 2010). Users on SO are required to create an account to engage with the platform's basic features and accumulate 'reputation' – a metric reflecting how much a user has contributed to the platform. As reputation increases, users gain

features ('privileges') which alters how much administrative power they have. Each contribution to the platform receives a 'score', the total number of 'upvotes' (approval) minus 'downvotes' (disapproval). If a user's question or answer is upvoted they receive +10 reputation. If an answer is marked as 'accepted' (+15) or if a 'bounty' of points (< + 1000) is placed on a question, a user gains additional reputation for their contribution. Through usernames and scoring systems the platform enacts norms of interaction, encouraging particular forms of self-presentation (van der Nagel, 2017). As a research site, the 'open' community of SO forms a unique opportunity to understand women as both subjects (users) and creators (programmers) of the digital world.

Despite its popularity, SO has a reputation for hostility and elitism that reflects the sexism of wider technology culture (Ford et al., 2017). Given the widespread use of SO as educational tool, the platform exposes even beginner programmers to the gender-biases of technical communities. Evidence of such gender-biases can be found in the annual SO Developer's Survey (2020), the largest and most comprehensive effort to understand the perspectives of programmers. As gender is not available on user profiles, the Developers Survey (2020) is the only point where SO collects information on demographics, revealing that 91.5% of the users identify as male. In 2019, SO (2020) asked ~90,000 users what they would most like to change about the site. Men wanted new features, using terms such as 'official', 'complex', 'algorithm', but in contrast, developers who were women wanted to see a change in the culture, calling attention to 'condescending', 'rude', 'assholes' (Brooke, 2019b). In 2020, this data was collected for a second time. Men described their participation on SO with the terms 'guys', 'big', 'valuable', whereas women emphasised their experiences of a 'toxic', 'rude', 'culture' (SO, 2020). The Developers Survey thus indicates that user's experience of interaction differs based on gender, with evidence that women are subject to greater levels of hostility. The immense popularity of SO entails that gender-biases on the platform are not only representative of technical forums, but of programming practise more generally.

Related scholarship

Structuring and signalling gender

Despite groundings in critical, feminist literature, researchers in the social sciences often operationalise gender as a binary variable (Risman, 2004). In computational work, the distinction between biological and social conceptions of gender can be all but completely lost. Labels of one/zero, male/female are exchanged for masculine/feminine, without challenging the essentialism of binary distinctions between people. Risman (2004) instead conceives of gender as a social structure, on the same analytic plane as politics and economics. Such structures exist outside of individual desires or motivations, and norms are developed when actors occupy similar positions in the social structure and evaluate their own options vis-à-vis the position of analogous others (Risman, 2004). From this standpoint, a person's actions, desires, and ability to make certain choices are patterned, but not determined entirely, by social structure (Risman, 2004). For Risman (2004), the power of gender is in defining who can be compared with whom; as human beings are binarised into types, differentiation diffuses expectations for gender equality. Much inquiry into gender differences in technical forums has focused on

'gender gaps' in individual user activity masking larger cultural biases. For instance, whilst May et al. (2019) observe that women are less successful on SO in terms of reputation score, they attribute the difference to level and type of activity. Although they do acknowledge the sexism of technology culture, May et al. (2019, p. 2010) propose that men thrive on programming forums as they are 'generally more competitive'. Such statements invoke notions of personal responsibility through gendered traits, accrediting women's exclusion to their own behaviour. Rather, Nguyen et al. (2016) argue that in bridging disciplines computational researchers must focus on the dominant role of social structures such as gender. This is not to ignore agency, but to account for how it is restrained and restricted by gender. In challenging the assumption of binary sex-difference, I will examine gendered differences in user's reputation, platform tenure, and activity, not just between men and women, but across levels of gender salience.

Gender, belonging, and participation

As men are seen to be 'naturally' more gifted with technology (Wajcman, 2010), gender can shape assumptions of ability and whether a person can imagine themselves participating in technology culture. As the physical identity markers of who is contributing are rendered invisible on anonymous platforms, 'male-centricity' can be amplified to 'male-by-default' (Tanczer, 2016). Risman's (2004, p. 432) gender as social structure highlights the consequences of assuming who is occupying a space; women cannot compare themselves to 'similarly situated others' to evaluate their capacity for action. Usernames can be a valuable resource for gender visibility, allowing users to distinguish themselves from one another and even highlight aspects of their identity (van der Nagel, 2017). The norms and cultural affinities of a platform are expressed through usernames, encouraging particular forms of self-presentation and tokens of belonging (van der Nagel, 2017). For example,'SpockEnterprise1993' indicates a fondness for Star Trek, 'Pink_Pixel_Princess' displays femininity, or prefixes such as 'Mr' or 'Miss' explicitly signal gender. In anonymous forums feminine usernames can be met with hostility, perceived as a demand for preferential treatment and chivalry from the masculine-majority userbase (Brooke, 2019a). Though they can be used to distinguish, the absence of identity cues in usernames leads to the assumption that the individual is part of the culturally dominant group (van der Nagel, 2017). As a venue for self-expression usernames thus have the potential to challenge or confirm perceptions of who participates on anonymous forums. As detailed in the Methods section, the gender identification procedure of this study is based on usernames, reflecting how users see each other and interact on SO.

Beyond a masculine-majority population, the structure of programming forums can disadvantage women and preclude a sense of belonging. Marwick (2013) points to how the norms and values of programmers are built into the sites they produce. She (2013) argues that such platforms are orientated around a strategic application of business logics. Here, user's reputation or the points a post receives ('scores') are prioritised as unbiased metrics, as they are rewarded for competitive displays of knowledge and often at the expense of others (Ford et al., 2017), traits commonly associated with masculinity (May et al., 2019). On technical platforms, behaivour identified as feminine – such as encouragement or friendliness – are not rewarded by the scoring system and thus deemed to be valueless (Marwick, 2013). Previous scholarship finds that women

are only successful on programming forums when their gender is obscured, but are penalised and their contributions devalued when their gender is known (Terrell et al., 2017). The second research question tests such devaluing on SO, investigating if gender-biases are present in how technical answers are scored on the platform. Thus, the organisation of such technology sites favour anonymous self-expression and scoring systems that are most tolerable to young, white, men (Marwick, 2013), as the architects of technical forums.

The masculine-majority population and design structure of SO mutually contribute to women's lack of participation. On SO women have been shown to be less aware of the platform's features than men, deterred from participating by the site's intimidating (male-dominated) community size (Ford et al., 2016, 2017). Ford et al. (2016) cite three features of SO that deter women from contributing: (1) anonymity was seen to encourage blunt (direct) and argumentative responses on posts, (2) invisibility of women leads to the site feeling like a 'boys club' (Ford et al., 2016, p. 6), (3) large communities are intimidating, and not possible in the same way offline. Ford et al. (2016) determined that the main barrier to women's participation is feeling that they lack the adequate technical qualifications to contribute. In a follow-up study, Ford et al. (2017) discovered that women are more likely to participate in a conversation on SO if they see other women already taking part, or even if women are just a visible presence. They referred to the positive influence of similar others as *peer parity* (Ford et al., 2017). Building on Risman (2004) and Ford et al. (2017) the final element of this study asks if SO is organised by gender, in particular, the effect of analogous others. Using social network analysis to test for peer parity and reciprocity I will investigate if gender shapes how users communicate on the platform. Drawing together the work of Risman (2004), Marwick (2013), and Ford (2016, 2017), this study investigates gender-biases in sharing and recognising technical knowledge on SO.

Method

Data and sampling

The SO data was retrieved from the Stack Exchange Data Dump, a quarterly upload of all the site's content hosted on the Internet Archive (https://archive.org/details/stackexchange). The dataset spans from the founding of the site in 2008 to the most recent quarterly 'dump' when the data was accessed in November 2019. The sampling frame included America, Canada, and the UK. The common cultural heritage of these countries means that they broadly share conceptions of gender, allowing for a measure of generalisation. However, this is not to say the project is a complete picture of gender in technology culture. A focus on the Anglosphere means that I limit the study to a Western and white conception of gender. Previous work in this field has shown that user location can be a significant mediator of how contributions on SO are evaluated (Stephany et al., 2020). I am not able to foreground intersectionality in this work, with ethnicity largely obscured in the data. I acknowledge this absence not to justify it, but rather to propose that the methods I use and the findings I present be taken forward.

In addition to location, users were required to have a tenure on the platform of longer than 7-days to filter out temporary accounts, without penalising new users. Tenure was

Table 1. Overall dataset sizes (matched).

Dataset	*n*
Users (Total)	560,106
Networked users	171,653
Questions	2,115,887
Accepted answers	1,091,629
Answers (total)	6,485,035
Comments	5,018,703

defined as the period between when the account was created, and when they last interacted on the platform. The selection procedure resulted in a sample of 560,106 users prior to building the network. Table 1 shows the complete breakdown of users, questions and answers, and comments that will be used to answer the first and second research questions. 'Accepted answers' are those that were deemed to be 'best' by the original poster of a question.

Networked users are those that were connected in the SO network to address the third research question. These connections consisted of questions, answers, and comments between users which fit the location and platform tenure criteria.

Gender identification procedure

In computational research design processes, gender is often produced from a masculine perspective, where women are characterised by an absence of male traits (Brooke, 2019b). This has led to much discussion on what is considered 'fair' in computational approaches to the study of inequality (Corbett-Davies & Goel, 2018). These debates, however, rarely engage with how social identity is embedded in experience and context (Corbett-Davies & Goel, 2018). If a researcher's method only allows them to collect data from a binary model of gender, they will ultimately produce research that presumes gender is a binary. In using inclusive methods, researchers can rely on self-disclosed information. As SO profiles do not disclose gender, this project focuses on usernames, a user-centred alternative to explicit identification. Gender is inferred as masculine or feminine only if a SO user's display name is a 'real name' (given name, such as Jane Smith), where that name is 'male' or 'female'. Work into gender and SO most commonly uses *genderComputer,* a tool developed by Vasilescu et al. (2012).[1] genderComputer is written in Python, built on a database of first names from 73 countries, as well as *gender.c*, an open-source C program for name-based gender inference. This library is used by a wide range of studies into gender and SO (Ford et al., 2016; Lin & Serebrenik, 2016; May et al., 2019). Inferring a user's gender from their name and location, genderComputer is specifically built from Drupal, WordPress, and SO.

I expanded the library of names that genderComputer uses to infer gender, from ~5000 (per gender) to 639,824 masculine and 959,835 feminine names using USA birth records from 1939 to 2019. For coherence, I followed the procedure of Vasilescu et al. (2012) and required a name to be used twice as frequently by men than women to infer masculinity, and the same procedure reversed for femininity. By default, genderComputer also parses for gender-specific words, such as 'Mr', 'sir', 'girl'. I extended this list to include more gender-specific and slang terms, such as 'bro'. I also expanded the feminine-associated terms, as this list was a third of the size of the default masculine

words. There was some room for manoeuvre here, with the algorithm's categorisations also including *mostly male*, *mostly female*, and *anonymous*. Even with the difficulties presented by SO data (such as Samuel being written as S4mu31), *genderComputer* is shown to automatically infer a person's gender with high precision (~95%). Though recall suffers due to ambiguity (~60%) the inclusion of *unknown* (anonymous) identification category improved this metric (~90%). About 5000 of the gender inferred names were manually inspected and the results confirmed satisfactory inter-coder reliability.

This study uses a five-level labelling system for gender inference to expand gender beyond a binary understanding, whilst still leaving room for identity to be purposefully obscured. Testing the classification on the manually labelled data, the classifier's accuracy was above 80% for each category, which was adequate. The total name-based gender identification is shown in Table 2. Whilst 'anonymous' is positioned in the centre of the table, this does not mean that they are considered to be a midpoint on the masculine-feminine scale. Rather, anonymous is positioned centrally to reflect how it encompasses multiple gender identities.

It may initially seem simplistic to focus on names, but this mirrors the experience of the average user with the platform, where display names are the fundamental identity marker. Given the masculine nature of technology culture (Wajcman, 2010), inferring that male names correspond to male identities is also contextually relevant. As we seek to understand gender differences, biases, and peer parity on SO, usernames are a simple but effective site of gendered identity.

Testing bias and operationalising effort

The second element of the analysis will examine if there is gender-bias in how answers are scored. To test if SO is equitable, I will explore gender-biases in the readability and effort of answers, and how this aligns to average scores for each gender category. Answer effort was operationalised from the limited guidance on effectively answering questions. I focused on answers as there is more reputation to be gained here and because asking questions on SO is highly structured. A user is walked through the stages of asking a good question, and even shown 'Similar Questions' which have already been answered as they type out the title to their submission. They are supplied with a detailed step-by-step guide including encouragement to: (1) summarise the problem, (2) describe what they've tried, and (3) show some code. Question-askers are also encouraged to provide a minimum reproducible example, which is a workable snippet of code to demonstrate the bug or problem. In comparison, negligible guidance is offered for answers as they are seen to be judged on merit. The quality of an answer is assessed purely on the

Table 2. Breakdown of users' visible gender.

	Total users		Networked users	
	n	%	*n*	%
Masculine	310,785	55	99,933	58
Mostly Masculine	142,302	25	45,169	26
Anonymous	13,800	2	3864	2
Mostly Feminine	12,035	2	3412	2
Feminine	80,184	14	19,275	11
Total	560,106		171,653	

score it receives, for instance a user is awarded the 'Good Answer' badge (platform award) if their post receives a score of over 25.

Established SO users have worked to fill the gap left by the platform and offer guidance for providing effective answers on SO. The clearest example of this is Jon Skeet, a senior software engineer at Google who is a cult figure on the platform due to his contribution of > 35,000 answers and having the highest reputation score of all time. In his blog, Skeet (2009) provided guidelines on how to 'answer technical questions helpfully', which is often shared across SO. The principles of a good answer are (1) reading the question, (2) code is king, (3) highlighting side issues, (4) providing links to related resources with context, and (5) style matters (including text formatting). I first investigate if simple text formatting, referred to as readability, is a contributing factor to the score that answers receive. Then, building on the guidance provided by Skeet (2009), I examine if contextually relevant features of effort explain the distribution of scores. Table 3 operationalises technical answer effort, the component features, example, and count of features in the dataset.

Answer effort is therefore a standardised score comprised of the frequency counts of code, resources, and format in Table 3. In total, I analysed 733,434 posts, sampling answers where both the question asker and responder where in the dataset. In operationalising answer effort, I test if SO is a meritocracy in which users are evaluated equally, regardless of gender.

Results and analysis

User's reputation, tenure, and activity

I first aim to explore if there are gender differences in users' reputation, tenure on the platform, and level of activity, referencing RQ1. I extend previous work by incorporating non-binary gender identification into the analysis of these key metrics. Beginning with reputation, I focus on (1) users with at least 15 points to account for basic privileges, (2) 100 points and above to extend May et al.'s (2019) analysis. Any user who gains 15 reputation points gains the ability to upvote questions and answers, facilitating a gender analysis of new or less engaged users.

Table 3. Answering technical questions: operationalisation of 'Effort'.

Element	Feature	Example	*n*
Code	Code	*Fragment of Computer Code*	366,183
	Formatted code	*Snippet of code with line breaks and spacing*	452,520
	Code with comments	*Programmer-readable explanation or annotation in the code snippet*	7167
Resources	Internal URL	*URL to somewhere else on SO, with context*	27,737
	External URL, popular	*URL to a site external to SO (i.e., GitHub, Wikipedia, MSDN). URL domain must be used more than 50 times*	22,541
	External URL, with context	*URL to a site external to SO (i.e., GitHub, Wikipedia, MSDN) with context.*	19,646
Format	Emphasis	*Bold or strong, italic or emphasis. Headings used*	137,160
	Structured text	*Use of lists (ordered, unordered) and/or bullet points*	57,112
	Image	*answers that contain an image*	20,298
	Sectioning	*Update or edit section added to text*	13,523

Table 4. Reputation of users (> 15).

Gender	% of group	Median	Mean	Std.	Max.	n
Masculine	38	135	1531	10,489	888,089	118,524
Mostly masculine	37	137	1739	11,262	1,128,126	53,760
Unknown/anonymous	32	120	1209	9630	532,305	4546
Mostly feminine	33	115	1179	4852	118,999	4042
Feminine	27	101	1018	6518	376,439	21,975

Reputation

The users here are the complete sample ($n = 560{,}106$) showing the percentage of users of each gender category in Tables 4 and 5.

In first looking at users with > 15 reputation, the analysis (OLS with Bonferroni correction) found $f = 21.49$, $p = > 0.001$. This indicates that there is a significant difference in group (gender) means across reputation. The results show that a significant amount of variance is explained by gender $F(4{,}116{,}078) = 21.49$, $p < 0.05$. Moreover, the model demonstrates that while there is a statistically significant difference between the mean reputation of feminine users and masculine ($\beta = 512^{***}$, std = 69) and mostly masculine ($\beta = 720^{***}$, std = 82) users, there is not a significant difference between feminine users and mostly feminine ($\beta = 512$, std = 75) and unknown users ($\beta = 191$, std = 167). Initial analysis shows that users identified as mostly feminine and feminine have lower reputation than more masculine users.

In moving to users with > 100 reputation and comparison with May et al. (2019), analysis found that there is still a significant difference across non-binary genders for > 100 reputation ($f = 12.12$, $p = >0.001$). Again, feminine users are significantly different from masculine ($\beta = 617^{***}$, std = 137) and mostly masculine ($\beta = 953^{***}$, std = 148) users, but are not significantly different from mostly feminine ($\beta = 117$, std = 312) and unknown ($\beta = 189$, std = 298) users. There is a statistically significant difference in reputation points across gender levels, with feminine categories having lower reputation as a metric of success on SO.

User tenure

Tenure on SO (Table 6) is an important indicator of participation, implying an investment in interacting on the forum.

In modelling (Welch's ANOVA), there is a significant difference in group (gender) means across the duration of activity for user accounts. I also considered the possibility of an interaction effect between gender and user tenure on reputation. I am testing H_1: *There is an interaction between user tenure and gender*. The tenure and reputation were $\log_{10}$ transformed to account for skewness. The overall MANOVA model was found to be significant ($f = 6.832$, $p = 0.0001$).

Table 5. Reputation of users (> 100).

Gender	% of Group	Median	Mean	Std.	Max.	n
Masculine	22	506	2604	13,660	888,098	68,870
Mostly Masculine	22	532	2941	14,602	1,128,126	31,465
Unknown/Anonymous	18	465	2176	12,944	532,305	2486
Mostly Feminine	18	484	2104	6389	118,999	2228
Feminine	13	420	1987	9096	376,439	11,034

Table 6. User tenure by gender, > 15 reputation.

Gender	*Median*	*Mean*	*Std.*	*Max.*	*n*
Masculine	1907	1882	1060	4047	118,524
Mostly masculine	1927	1908	1107	4047	53,760
Unknown/anonymous	1818	1792	1044	4041	4546
Mostly feminine	1860	1829	1012	4046	4042
Feminine	1625	1660	1050	4044	21,975

Table 7 shows that both gender and user tenure are associated with reputation. Moreover, the interaction between gender and user tenure is also significant, indicating that the relationship between tenure and reputation is mediated by gender. Post-hoc testing (Welch's *t*-test with Bonferroni correction) revealed that user tenure positively affects reputation for each level of gender. I thus find that there is a gender difference in the relationship between user tenure and reputation, with more masculine users seeing the highest return from the time they invest.

Level of activity

In their study of SO, May et al. (2019) find that the gender differences in reputation between users with > 100 points are due to differences in level and form of activity. They argue that men answer more questions, and this is more highly rewarded by the SO scoring system (May et al., 2019). I test if these findings hold with non-binary gender identification, examining the average frequency and score of contributions by user gender. Table 8 shows the average question count by gender. Users who did not post a question, answer, or comment are not included in these tables.

The most prolific user in terms of number of questions asked was labelled as feminine and posted 1069 questions over an 11-year period. They are in the top 0.08% of all users of the site and have answered 92 questions themselves, predominantly focusing on JavaScript and C#. However, we can see this user is clearly an outlier. The presence of highly successful individuals in technology does not negate the negative experiences of the majority of women. In general, counter to May et al. (2019), Table 8 shows that there are not substantial gender differences in the number of questions asked by users.

Table 9 illustrates the average number of answers by users in the dataset. It indicates that there is limited gender difference in the average (median) number of answers offered by users. The average number of comments on a question and answer by gender category are shown in Tables 10 and 11, respectively.

Next, Table 12 highlights the average question score of users. The average (mean) score falls with increased femininity. The highest score is mostly masculine, which is likely due to outliers (as shown in the max column). These results indicate that more masculine users receive higher scores on their questions compared to anonymous and feminine users on average.

Table 7. MANOVA table, reputation user tenure * gender.

	SS_M	*df*	*F*	*PR(> F)*
Gender	147	4	77.84	< 0.001
User tenure	28,357	1	60,194.80	< 0.001
User tenure * gender	117	4	62.06	< 0.001
Residual	95,554	202,841	–	–

Table 8. Average question count by gender.

	Mean	Median	Max	Std.	Count
Masculine	5.89	2	900	14.94	607,323
Mostly masculine	6.05	2	822	15.83	283,940
Anonymous	5.55	2	395	14.05	23,039
Mostly feminine	5.63	2	255	13.66	20,893
Feminine	5.17	2	1069	16.19	110,863
All users	5.84	2	1069	15.18	1,046,058

Table 9. Average answer count by gender.

	Mean	Median	Max	Std.	Count
Masculine	8.77	2	4106	42.84	808,231
Mostly masculine	9.53	2	5914	46.42	398,155
Anonymous	7.16	2	746	25.09	24,832
Mostly feminine	8.22	2	813	26.55	25,014
Feminine	6.80	2	950	25.08	109,156
All users	8.73	2	5914	41.82	1,365,388

Table 10. Average comment (question) count posted by gender.

	Mean	Median	Max	Std.	Count
Masculine	11.19	4	3696	46.59	397,748
Mostly masculine	12.12	4	3557	45.81	195,704
Anonymous	10.19	3	788	32.43	13,408
Mostly feminine	10.19	3	736	28.17	12,592
Feminine	9.47	3	1076	29.33	61,399
All users	11.22	4	3696	44.31	680,851

Table 11. Average comment (answers) count posted by gender.

	Mean	Median	Max	Std.	Count
Masculine	9.36	3	2958	38.33	510,820
Mostly masculine	10.16	3	5521	44.97	256,844
Anonymous	8.01	3	534	23.35	15,871
Mostly feminine	8.65	3	594	22.43	15,692
Feminine	7.89	3	993	24.76	75,171
All users	9.39	3	5521	38.63	874,398

Table 12. Average question score by gender.

	Mean	Median	Max	Std.	Count
Masculine	2.92	1	6391	23.03	1,189,227
Mostly masculine	3.00	1	7243	22.98	545,289
Anonymous	2.42	1	827	12.73	46,843
Mostly feminine	2.65	1	3892	23.544	41,380
Feminine	2.46	1	3277	20.03	224,395
All users	2.88	1	7243	22.53	2,047,134

Overall, these results demonstrate that feminine users have a significantly lower reputation and tenure on SO than more masculine users. There is evidence for a gendered difference in the relationship between user tenure and reputation, where more masculine

users see the highest return. However, there is no significant gender difference in average frequency of questions, answers, or comments. As there is no disparity in frequency, these results indicate that users' contributions may be valued and scored differently. As answers are a lucrative source of reputation points compared to other forms of contributions, they will be the focus of analysis in the next section to explain the presence of the gendered difference.

Gender and scoring technical knowledge

In the introduction I highlighted how platforms of technology culture are often framed as open meritocracies. Whilst I have shown that there is a gendered difference in reputation as a metric of users' success, the causal mechanism of the disparity requires further investigation. The subsequent analysis explores if the content of answers accounts for lower reputation for more feminine users. First, I will examine if there is a gender difference in the answer scores. Next, I test if there are gendered differences in the organisation of text to account for basic readability. Following this, I build a context-specific measure of 'answer effort' from resources posted and celebrated on SO. By incorporating the novel operationalisation of effort, I can empirically test if there are gender-biases in how answers are scored addressing RQ2.

Scoring answers

In testing if SO is meritocratic, I first examine if there is a gendered difference in answer scores. In an equitable framework, equal effort leads to equal outcome, measured here in terms of the answer score (upvotes – downvotes). Users with the highest reputation on SO tend to accrue reputation through answering questions. May et al. (2019) find that men on SO answer more questions and are scored more on average than women. Whilst I did not find gendered disparities in average number of answers by users, there is a significant difference between gender groupings in the scores they receive ($f = 5.935^{***}$). Post-hoc testing (Welch's t-test, Bonferroni correction), confirmed that there was a significant difference in the scores for feminine ($\bar{x}$ = 1.91, std = 9.83) and masculine ($\bar{x}$ = 2.06, std = 12.29) users; $t = 3.24$, $p = 0.001$, as well as feminine and mostly masculine ($\bar{x}$ = 2.16, std = 15.52) users, $t = 4.66$, $p = 0.001$. Therefore, masculine and mostly masculine users receive the highest scores for their answers. As I previously found no significant difference in the frequency of interactions, biases in scoring may explain variations in reputation based on gender. To test if the difference in scoring technical knowledge is a result of gender-bias, the following section examines if gender differences in score can be attributed to readability or answer effort.

Gender and readability

A potential explanation for gender difference in reputation is the text formatting or readability of interactions and submissions to SO. Specifically the length of answers, number of paragraphs, and text to code ratio. Figure 1 illustrates the average answer length by gender category, counting the plain-text unigrams (words) only. As suggested in the boxplot, further testing indicated that there was no difference in answer length by gender group.

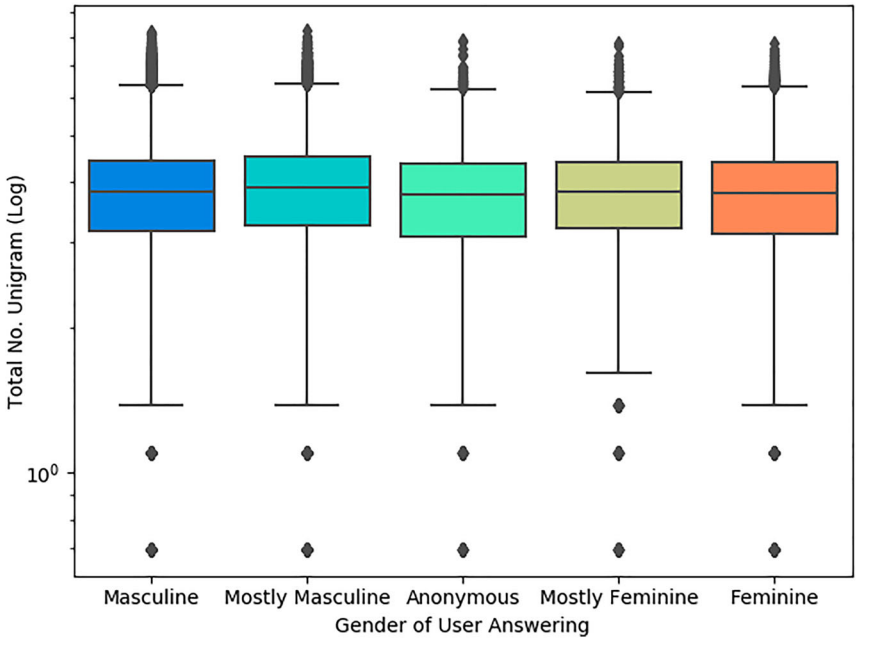

Figure 1. Readability: average answer length boxplot.

Moving beyond plain-text in examining readability, Skeet (2009) highlights the importance of including example code as formatted snippets. I examined simple ratio of code snippets to text-paragraphs, where an answer consisted of more than a single paragraph, as this is considered bad practice according to the established referenced guidelines and indictive of low effort (Skeet, 2009). Figure 2 shows the distribution of the ratio of text-paragraphs to code snippets for all question-answer pairs of users in the dataset. The average across all users was three text-paragraphs to one snippet on code. Figure 2 indicates there is not a clear difference between genders in the proportion of text to code, which was confirmed with ANOVA testing. Whilst text structure is a straightforward approach to testing gender-biases in scoring, it does show that the readability of answers is not a contributor to gendered differences in evaluating answers on SO.

Gender and answer effort

As there are no significant gender differences in readability, the next stage of analysis is to test the assumption of technical meritocracy and examine features of answer effort. The component features of answer effort were (1) code; (2) resources; and (3) formatting, as expanded on in the Methods section in Table 3. Overall, there was a significant difference (using Welch's ANOVA) between the gender category of the user providing the answer and the answer effort. Post-hoc testing (Welch's *t*-test, Bonferroni correction) revelled that the significant differences were between masculine and feminine users ($t = 12.08$, $p = >0.001$), masculine and mostly feminine users ($t = 5.16$, $p = 0.005$), and masculine and anonymous users ($t = 7.25$, $p = >0.001$). I therefore conclude that there are more markers of answer effort in non-masculine (anonymous, mostly feminine, and feminine) than more masculine users. Consequently, I find support for the presence of gender-

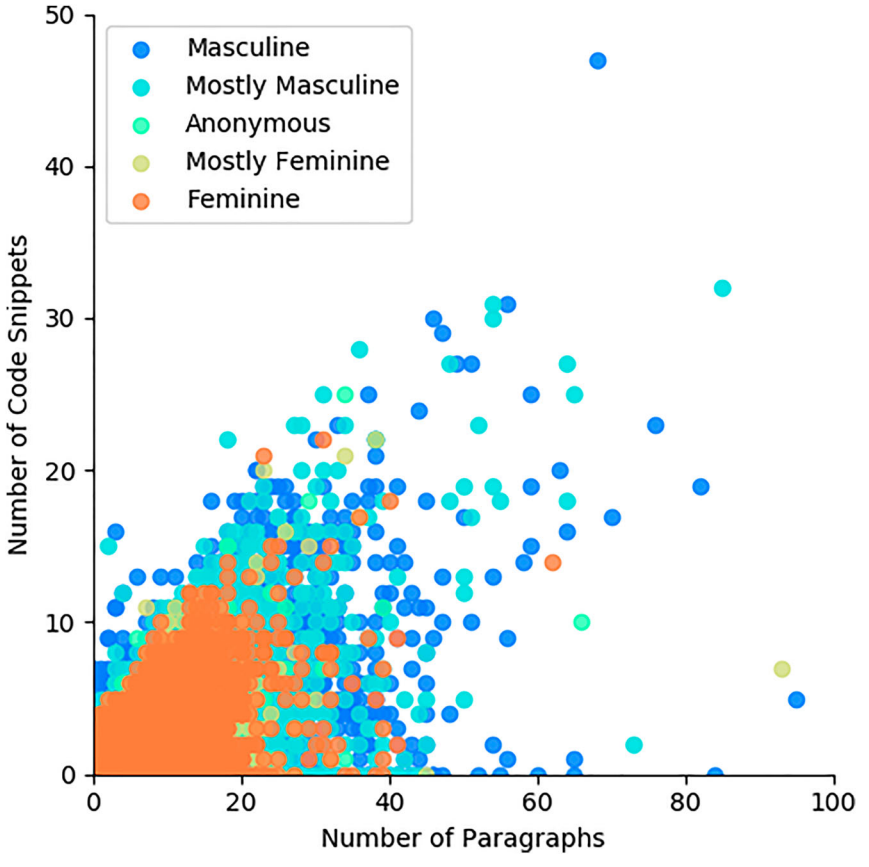

Figure 2. Readability: paragraph to code ratio by gender.

biases in how answers are scored on SO. Taken with the result that masculine users receive higher scores for their answers, I find empirical evidence that signalling masculinity in usernames is sufficient to positively benefit users. It is important here to recognise that the operationalisation of answer effort is rudimentary, but the result is nonetheless informative. I find that feminine users receive lower scores for their answers, which has consequences for their reputation as a metric of success on SO. This difference is not a matter of readability or effort, which leads to the conclusion that there is measurable gender-bias in the scoring of answers. Masculine users benefit from signalling their gender are not shown the analysis to produce higher effort answers.

Mapping interaction on SO

In the final portion of this study, I examine if interaction on SO is organised by gender. Social network analysis allows an examination of patterns communication beyond

Table 13. Basic network descriptors.

Nodes	173,053
Edges	1,616,800
Average degree	9.34

frequency counts. Building on Ford et al.'s (2017) study, I test for evidence of peer parity and reciprocity, investigate if gender organises how users communicate on the platform, in reference RQ3 (Table 13).

Describing the network

In examining the SO network, I first determine if gender organises users by level of connectivity. The focus here is on the degree as the measure of the total number of edges adjacent to any given node, or the number of interactions of a given user. In a Facebook social graph, this could be characterised as the number of friends a user would have. In the SO network, degree represents an interaction pair (e.g., question-answer, answer-comment). The degree k is the number of connections that an individual has, and p_k is that fraction of individuals in the network that have exactly k connections. Figure 3 is the degree distribution (p_k) for masculine and feminine users (nodes) and Figure 4 is across the whole network for all users. The plots represent a randomly selected subsample of users for ease of visualisation, with no information lost.

The graphs in Figures 3 and 4 indicate that the degree distribution is consistent across gender groupings. Degree distribution is represented as $p_{k \propto k}{}^{-a}$, where $a > 0$. Plotting on a logarithmic scale reveals the long tail of the degree distribution, where most users have a small degree and thus minimal interactions. The presence of hubs of users that are considerably larger in degree than most nodes are a characteristic of 'power law' networks. Users who are highly connected will become more connected more quickly than less connected users. This is also referred to as a preferential attachment process (or cumulative advantage). Indeed, most individuals have few connections to others within the network, whilst a much smaller population are a lot more active. These figures suggest that

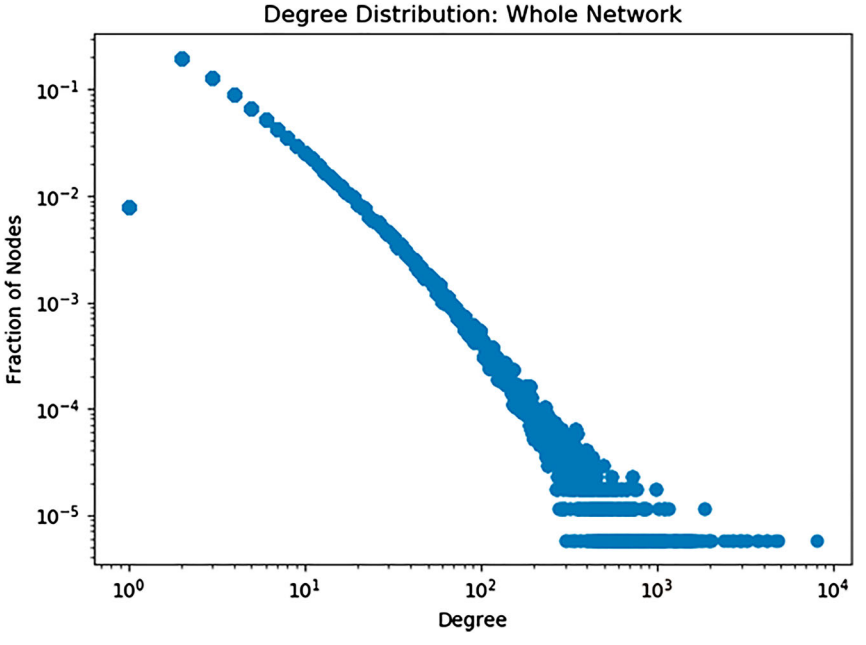

Figure 3. Degree distribution log-log scale: complete network (sample).

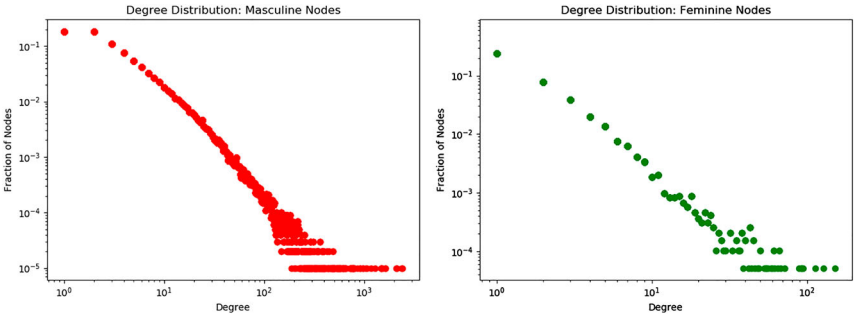

Figure 4 .#Degree distribution on log-log scale: by gender (sample).

patterns of connectivity hold across gender, but we can get a clearer picture by examining the average degree by gender in Table 14.

Table 14 shows that the masculine users have the highest average number of interactions, which decreases as the salience of masculinity decreases. Feminine users are the least connected, with the least average number of degrees by all measurements. Looking to the mean and median scores, some heavily connected nodes skew the average degree. Nonetheless, both the mean and median degree scores show that the more masculine a user is, the more communication they have on the platform. As I also found evidence of cumulative advantage the higher average connectivity of more masculine users implies that they will become more connected, more quickly. Below, Table 15 shows the average in degree and out degree of users by gender grouping. Masculine or mostly masculine usernames are more likely to receive a comment or answer on their post in comparison to other users.

The average in degree of masculine and mostly masculine users is marginally higher that their out degree, indicating productive participation on SO. In comparison, anonymous, mostly feminine, and feminine users answer and comment more on the contributions of others than they receive. Feminine users are particularly disadvantaged, receiving only 80% of what they contribute. These findings hint at the presence of gender preferences in interacting, which will be investigated in more depth later.

In addition to average connectivity, I examine if the most connected users are organised by gender through the k-core degeneracy. The k-core of a graph is that maximal subgraph that is a network of the most connected users. Starting from the complete network, I iteratively remove nodes with degree equal to or less than k. First removing users with one or less connections ($k = 1$), then two or less ($k = 2$), then three or less ($k = 3$), and so on. At each value of k, the portion of nodes for each gender category is recorded. This

Table 14. Average degree count by gender.

	Mean	*Median*	*Max*	*Std.*	*n*
Masculine	16.33	5	5001	60.45	99,933
Mostly Masculine	16.30	5	8066	65.55	45,169
Anonymous	12.94	4	1188	40.77	3864
Mostly Feminine	14.08	4	1527	44.05	3412
Feminine	11.67	3	1644	41.15	19,275
All Users	16.14	5	8066	59.31	171,653

Table 15. Average in/out degree by gender.

	Average in degree	*average out degree*	*% Inward*
Masculine	7.36	7.27	101
Mostly Masculine	8.10	7.65	106
Anonymous	5.84	6.65	88
Mostly Feminine	6.61	6.94	95
Feminine	4.99	6.24	80

process is stopped when we reach the *main core*, a homomorphic image where all nodes have the same degree of connectivity (number of connections). The largest degree in the network was 59. The fraction of core by gender for a given value of *k* is shown in Figure 5.

Figure 5 shows that the proportion of users of each gender identification generally holds across levels of connectivity. The overall sparsity of the SO graph implies that users are far from each other in the SO network. However, even with a sampling frame based in location and tenure, user connectivity is evidently organised by gender. The low proportion of women users holds across the least and most connected users.

Gendered peer parity

Previous research by Ford et al. (2017) proposes that peer parity is a significant feature of interaction on SO, where women are more likely to participate if they see that women are already taking part. I empirically tested this finding in the framework of non-binary gender categories. Building on the earlier indications of gendered preferences in interaction, assortativity is the preference of users to be attached to (interact with) other users of the same gender identification, a useful metric of peer parity. Assortativity provides a metric that accounts for differences in gendered homophily but is most useful as a descriptive guide in ascertaining if a network is assorted by gender – a quantification of the association between nodes. The overall assortativity coefficient based on gender for the SO network is 0.33, which suggests the presence of gendered peer parity. This indicates that SO is organised by gender as users tend to respond to users of the same or similar gender categorisation.

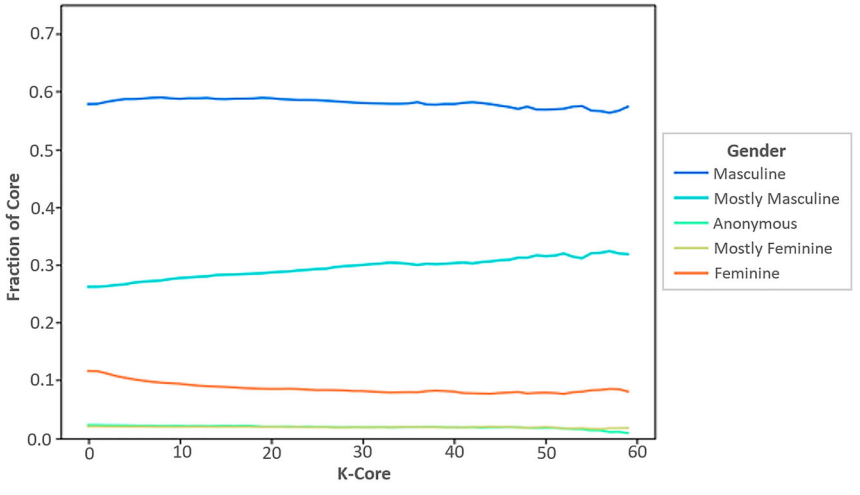

Figure 5. *k*-core degeneracy by gender.

To further understand how SO is organised by gender, we can compare the probability of following a randomly selected edge and arriving at a user with of a particular gender category. These probabilities are represented by $p(g)$. Here, I am computing the conditional probability $p(g'|g)$ that a random neighbour (connected user) of individuals with gender g has gender g', denoting (abbreviating) the gender category to M, F. For neighbours of masculine users, $p(F|M) = 0.08$ and $p(M|M) = 0.60$. For feminine users $p(M|F) = 0.55$ and $p(F|F) = 0.14$. In all cases, the random neighbour is more likely to be masculine. Yet, a random neighbour for a feminine user is almost twice as likely to be another feminine user than for masculine users. This is evidence of gendered peer parity in the organisation of the SO network. Supporting Ford et al. (2016, 2017), I conclude that feminine users will preferentially interact with other saliently feminine users. Overall, I find empirical evidence that the SO network is organised by gender in terms of assortativity and peer parity.

Gendered reciprocity

The final element of analysing the SO network is to examine if interaction is more likely to be reciprocated by users of the same gender category in comparison to others. In this evaluation I look at the reciprocity scores for each gender category, that is, the percentage of interactions that are returned by users. Overall, the percentage of ties that were reciprocated was 24%, across gender categories. In Table 16 this is broken down by gender, where again the rows are the categories of node from, and the columns are the categories of node too**.**

In Table 16, the diagonal axis shows clear evidence of gendered reciprocity. For each gender category, the highest level of returned interaction is from the same gender. The margins for this majority are large, up to three times the reciprocity for any other group. The highest levels of reciprocity are for users identified as mostly feminine and anonymous, whilst the comparatively lower reciprocity of feminine users is likely due to lower levels of reciprocated participation in general. The high levels of reciprocity for feminine and mostly feminine users, supporting the earlier findings that gendered peer parity is a significant feature of more feminine user's interaction on SO. Therefore, I find evidence that patterns of interaction on are SO network organised by gender. As users are inclined to communicate within gendered boundaries masculine users benefit from being the majority group, challenging the notion of 'open' technology culture on SO.

Conclusion

This study has characterised gender-biases in sharing and recognising technical knowledge on SO. The evidence provided highlights three important components of gender-

Table 16. Reciprocity by gender.

		To				
From		Masculine	Mostly Masculine	Anonymous	Mostly Feminine	Feminine
	Masculine	**34.35%**	26.21%	24.88%	25.03%	27.33%
	Mostly Masculine	27.44%	**42.65%**	26.91%	27.45%	29.62%
	Anonymous	21.42%	22.36%	**58.21%**	22.87%	23.19%
	Mostly Feminine	23.71%	24.53%	23.93%	**68.10%**	25.19%
	Feminine	21.65%	22.38%	20.44%	21.47%	**61.75%**

biases on the platform. First, encompassing a non-binary operationalisation, I extend the work of May et al. (2019) and show that there are significant gender differences in key user metrics. Feminine and mostly feminine users have the lowest reputation score on average, the central measure of success on the platform. I show that reputation is not a simple function of user' tenure on SO, but the relationship is mediated by users' gender. Given that I find no evidence of gender difference in the frequency of activity, the second element of the analysis aimed to explore the divergence in reputation points. I asked if there are gender biases in how technical knowledge is scored, focusing on answers as a particularly lucrative source of reputation points. Challenging the framing of technical forums as open meritocracies, I found that gender determines how users' answers are scored. I also find that this difference is not a matter of readability or effort, but gender-bias in the scoring of answers. Despite evidence of higher effort, the contributions of feminine users are undervalued, in comparison to their masculine and anonymous counterparts. This finding supports Marwick's (2013) thesis that the scoring systems of platforms permit the functioning of sexism in subtler forms, inherently biasing contextual success towards masculine identities and behaviour. Whilst gaps in reputation indicate the presence of sexism on SO, gender-biases in answer scoring indicate that difference is not merely the result of user tenure or level of activity. Based on these results I conclude that women's negative experiences in technical spaces are the result of structurally supported gender-biases, not a lack of effort or knowledge.

Third, I applied social network analysis to determine if the SO network is organised by gender. I showed that masculine or mostly masculine usernames are more likely to receive a comment or answer on their post in comparison to anonymous of more feminine users. In comparison, feminine users receive less answers and comments than they contribute and remain underrepresented across levels of connectivity. Building on Ford et al. (2017), I show that peer parity is a significant feature of interaction on SO, and that feminine users are more likely to participate if they see that women are already taking part. Quantifying this preference, feminine users are twice as likely to interact with feminine-labelled individuals than other gender signalled categories. Additionally, the SO network shows clear evidence of gendered reciprocity, with users replying to individuals of the same gender identification than others. The consequences of the gendered organisation of the SO network are that feminine users are required to work harder to legitimately participate and be recognised on the platform.

Coupling Risman's (2004) conception of gender with Marwick's (2013) critique on the functioning of platforms, I find that the structure of SO limits the agency of women. Feminine users receive the least responses, are scored lower on average,

Table 17. Recommendations for gender inclusivity on SO.

Feature	Recommendation
Code of conduct	(1) Banner of the Code of Conduct across the platform (2) Show users specifics of the Code of Conduct when posting
User gender	(3) Collect user's gender (4) Give users the option to show gender on their profile
Voting	(5) Do not show score on posts. Prevents herding behaivour
Posts (Reputation)	(6) Equal reputation awarded for questions and answers (7) New users gain higher reputation
Sexism flag	(8) Allow users to specifically report a comment as sexist. (9) This can be expanded to explicitly flag racism.

and interact within largely within gendered boundaries. However, gendered peer parity offers some hope for greater inclusivity. If the number of women visibly participating in technical forums can be increased, the results of this study indicate that more women will be encouraged to contribute. To do so, SO needs to become a more welcoming environment, addressing its gender-biases and how its reputation systems and structure perpetuates masculine dominance. Based on the findings of this paper I can make several recommendations for gender inclusivity, outlined in the table below (Table 17).

Supplementing the empirical conclusions, this study also demonstrates how social data science can operationalise gender beyond a binary, the framing of gender in five categories is a step towards computational research which can more faithfully represent social conceptions of gender. Future research should add depth to these findings, characterising nuances of gender-biases within a specific specialties or programming languages. Beyond further categorising the specific dimensions of sexism on programming forums, we need to consider how prejudice can be challenged; what informal interventions already show evidence of success? What can technical forums do to visibly challenge sexism? In addressing these questions, we can pursue a technology culture which is inclusive and diverse, not merely in specialisms but in people.

Supporting previous work on technology culture and gender (Terrell et al., 2017), this study finds that signalling femininity in usernames disadvantages users on SO. Through user moderation and quantified scores SO obscures the sexism that runs rampant on the platform. Women are not discounted because they are less competitive or lack technical expertise, they are excluded because technology culture sees technical knowledge as fundamentally incompatible with femininity. In this context, the numerical basis of platform scoring systems is not evidence of a proven meritocracy, but the functioning of established norms. This paper concludes that women are not disadvantaged by their own actions; they are penalised by a scoring structure which conceals sexism and disregarded by a masculine-majority userbase. Far from programmers' paradise, gender-biases dictate the sharing and recognition of technical knowledge on SO.

Note

1. Code available at https://github.com/tue-mdse/genderComputer

Acknowledgments

My heartfelt gratitude to the referees and editors who reviewed this study. Their insights have considerably improved the final paper. I would also like to express my sincere thanks to Gina Neff, Dong Nguyen, Theodora Sutton for their generosity, guidance, and patience. For providing comments on earlier drafts, thanks to Chico Camargo, Julia Slupska, and Felix Brooke.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Dr. S. J. Brooke's research interests lie in the intersection of critical research and computational methods. She employs computational linguistics, machine learning, and ethnography to study inequality in online communities. Dr Brooke's current work investigates gender inequality in the cultures associated with technology. [email: s.j.brooke@lse.ac.uk].

ORCID

S. J. Brooke http://orcid.org/0000-0001-7558-7924

References

Brooke, S. (2019a). There are no girls on the internet: Gender performances in advice animal memes. *First Monday*, *24*(10), 1–1. https://doi.org/10.5210/fm.v24i10.9593

Brooke, S. (2019b). *Condescending, rude, assholes. Framing Gender and Hostility on Stack Overflow*. *ACL Anthology*, Proceedings of the Third Workshop on Abusive Language Online (1), 172–180. https://doi.org/10.18653/v1/W19-3519

Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of Fair machine learning. arXiv: cs.CY. *Computers and Society*, http://arxiv.org/abs/1808.00023

Ford, D., Harkins, A., & Parnin, C. (2017). Someone like me: How does peer parity influence participation of women on Stack Overflow?. Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC. 239–243. https://doi.org/10.1109/VLHCC.2017.8103473.

Ford, D., Smith, J., Guo, P. J., & Parnin, C. (2016). Paradise unplugged: Identifying barriers for female participation on Stack Overflow. *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering - FSE* 2016, 846–857. https://doi.org/10.1145/2950290.2950331

Lin, B., & Serebrenik, A. (2016). Recognizing gender of Stack Overflow users. *MSR Conf*, to appear. https://doi.org/10.1145/2901739.2901777

Marwick, A. (2013). *Status update: Celebrity, publicity, and branding in the age of social media*. Yale.

May, A., Wachs, J., & Hannák, A. (2019). Gender differences in participation and reward on Stack overflow. *Empirical Software Engineering*, *24*(4), 1997–2019. https://doi.org/10.1007/s10664-019-09685-x

Nguyen, D., Doğruöz, A. S., Rosé, C. P., & de Jong, F. (2016). Computational sociolinguistics: A survey. *Computational Linguistics*, *42*(3), 537–593. https://doi.org/10.1162/COLI_a_00258

Risman, B. J. (2004). Gender as a social structure: Theory wrestling with activism. *Gender and Society*, *18*(4), 429–450. https://doi.org/10.1177/0891243204265349

Skeet, J. (2009). Answering technical questions helpfully. *Jon Skeet's Coding Blog*. https://codeblog.jonskeet.uk/2009/02/17/answering-technical-questions-helpfully/

Stack Overflow. (2020). Stack *Overflow developer survey* 2020. https://insights.stackoverflow.com/survey/2020#community-developers-perspectives-by-gender

Stephany, F., Braesemann, F., & Graham, M. (2020). Coding together–coding alone: The role of trust in collaborative programming. *Information Communication and Society*, https://doi.org/10.1080/1369118X.2020.1749699

Tanczer, L. M. (2016). Hacktivism and the male-only stereotype. *New Media and Society*, *18*(8), 1599–1615. https://doi.org/10.1177/1461444814567983

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. *PeerJ Computer Science*, *3*, e111. https://doi.org/10.7717/peerj-cs.111

van der Nagel, E. (2017). From usernames to profiles: The development of pseudonymity in Internet communication. *Internet Histories*, *1*(4), 312–331. https://doi.org/10.1080/24701475.2017.1389548

Vasilescu, B., Capiluppi, A., & Serebrenik, A. (2012). Gender, representation and online participation: A quantitative study. *The British Computer Society*, https://doi.org/10.1093/iwcomp/xxxxxx

Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, *34*(1), 143–152. https://doi.org/10.1093/cje/ben057