Towards More Gender-Inclusive Q&As: Investigating Perceptions of Additional Community Presence Information

PATRICK MARCEL JOSEPH DUBOIS, University of Manitoba, Canada
MAHYA MAFTOUNI, University of Manitoba, Canada
ANDREA BUNT, University of Manitoba, Canada

Online Question and Answer communities (Q&As) are popular spaces for learning and sharing knowledge. However, prior research suggests that Q&As may not be appealing to and inclusive of men and women, with absent social considerations listed as a potential contributing factor. We investigate how additional community presence information can affect users' perceptions of and engagement with a Q&A for graphic design software. Through a 10-day task-based field study with 30 participants (14 women, 14 men, 2 non-binary), we uncover how community presence information can humanize the Q&A and play a role in promoting an inclusive environment. On the other hand, some participants question if community presence information belongs in a Q&A and describe some privacy implications. The women in our sample also talked about the importance of diverse community demographics, while we did not observe this sentiment expressed by the men. Our findings contribute an understanding of how users perceive the role of community presence information within a Q&A. We also discuss how this information might impact women's future participation and engagement.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; *Empirical studies in HCI*; • **Human-centered computing → Collaborative and social computing**; *Empirical studies in collaborative and social computing*

Additional Key Words and Phrases: Gender; women; Q&As; social presence and awareness; homophily; community presence

1 INTRODUCTION

Online Question and Answer communities (Q&As) serve as important knowledge repositories for learning complex feature-rich software. Software users are encouraged to contribute their knowledge and/or to ask questions so that content is adapted to different contexts, such as knowledge levels, software versions and use cases. Contributors gain certain advantages, such as being able to showcase skills [66,93] and grow their professional networks [23,77,96,97]. There is

therefore a large body of research on ways to assist contributors in authoring content in various formats (e.g., [11,18,27,31,50,51]).

Unfortunately, recent research suggests that Q&As are not fully inclusive. Many women feel they do not receive the same level of support as men [26] and find the platforms and their communities unappealing [7]. Women do not participate as often, follow different contribution patterns, submit different types of content, and do not receive the same level of recognition as men do in Q&As [19,26]. On other types of platforms, consequences of gender imbalance include content being biased to a certain audience's perspectives due to missing views [14,69,88,90] and fewer quality contributions being made [9,20,83].

It is therefore necessary to work towards improving Q&As to be more gender-inclusive [9,20,83]. Previous work suggests that one deterrent for women in Q&As is a lack of attention paid to social considerations [19]. Research shows that women in particular benefit from social interactions when sharing knowledge [10,84], or when deciding to participate in a community [72]. Working towards improving social considerations in Q&As, we study how two components of socialization [71] can impact perceptions and use of Q&As. In particular, we introduce what we refer to as *community presence information* as a way to foster 1) homophily and 2) social presence and awareness (SPA).

To investigate how community presence information can impact Q&As, we created two Q&A interfaces, which we call Community Presence interfaces, with additional information showing 1) how community members are similar to a user (homophily) and 2) which questions community members have viewed (SPA). We deployed a mock graphic design software Q&A using these interfaces in a 10-day task-based field study where graphic design professionals and hobbyists answered questions using one of the Community Presence interfaces and a Baseline interface. Our findings suggest that community presence information can humanize a Q&A, potentially leading to increased empathy for other users and trust in the content, although humanization was perceived by some participants as going against the purpose of a Q&A. Our findings also suggest that homophily and heterophily can play complementary roles in promoting an inclusive environment particularly among women, and that community presence information can lead to both increased privacy concerns and awareness.

Through our research, we contribute two interface designs that display community presence information within a Q&A and findings from a 10-day task-based field study on how these designs affect user perceptions. We also contribute insight on how community presence information can potentially be used to work towards gender-inclusive Q&As.

1 RELATED WORK

1.1 Gender Differences in Q&As

A large body of work has studied the effects of gender imbalances (in terms of participation, inclusivity, support, etc.) in online communities in many spaces such as Wikipedia (e.g., [3,40,55,56,69,90]), OpenStreetMap (e.g., [14,82]), blogs (e.g., [58,65]) and discussion platforms (e.g., [30,91]). We overview findings related to Q&As, which are the focus of this work.

Gender differences in Q&As have been studied in-depth from a variety of different perspectives, including Q&A usage, content contributed, and received validation. Looking at which Q&A features are used, more men than women post content [19,89], and men post [19,52,89] and use upvoting and downvoting mechanisms [95] more frequently than women. Analyses of archival

data suggest that content-wise, men write longer answers, receiving more feedback than women as a result [73]. Men are also more willing than women to "play the game," authoring content likeliest to increase their reputation scores [89], while based on words used in authored content, women have a more cooperative than competitive attitude, as well as a supportive and collective outlook of the community with the aim of improving everyone's knowledge (using collective language like "we" instead of "I," and positive and encouraging words) [95]. Finally, in terms of validation, men earn more reputation points than women do [19,39,52,89]. Men are also likelier to receive upvotes based on the perceived competence of their posts and their popularity, while women tend to receive upvotes only based on their existing popularity [73]. In short, existing Q&As appear to have stronger appeal among men and men appear to have more advantages from participating compared to women.

Based on prior work, increasing the emphasis of a community's presence might be a promising path forward, as it appears that women prefer contributing community-oriented content. Men tend to ask and answer more informational questions (those that are factual in nature, e.g., "How can I unlock a layer?") while women prefer asking and answering more subjective, conversational ones (those that are opinion-oriented, e.g., "What are your favourite techniques to retouch a portrait?") [19,29]. It also appears that women are likelier to engage in discussions compared to men [95]. Linguistically, women use positive and encouraging language, while men tend to be neutral [39], women use more inclusive and team-oriented language [95], and women show more emotion through the use of emoticons [24].

As Q&A mechanisms appeal to men's competitiveness and women appear to have community-oriented attitudes, we look for ways interface design can be used to foster inclusive community dynamics. Specifically, we investigate how presenting SPA and homophily as community presence information can affect perceptions of and interaction with a Q&A.

1.2 SPA and Homophily in HCI

SPA and homophily are two concepts important to socializing [71]. Here, we define the two terms and the effects they have in different contexts, especially as they relate to improving interactions in and perceptions of online communities.

SPA is the idea of community members being mutually aware of each other and their activities: it is the "reciprocal awareness of others of an individual, and the individual's awareness of others within an interactive social space" [13]. In HCI, the idea of enhancing SPA has been used in various ways. It includes work such as feeds [22,35], alerts [46] and parallel interfaces [87,94] to keep people aware of community activities and project progress. Enhanced SPA tools also encourage user engagement and other positive social behaviours, which can potentially increase content quality. For example, teacher activity dashboards increase social interactions between peers [57], online shoppers' engagement with a task increases when aware of collaborators' activities, [92] and social network visualizations potentially increase learners' awareness of and engagement with knowledge sources and communities [75]. In contrast, poorly-used or presented SPA information can reduce work quality [21,76], emphasizing the need to develop and study implementations specific to community characteristics [57], such as Q&As.

The other social factor we are investigating in the context of Q&As is homophily [43]. It is the phenomenon describing people's preference to interact with other, similar individuals. Homophily has long been known to exist in various environments, including affecting elections [42], farmers [6], students [60], and more recently, MySpace users [85], Wikipedia editors [37],

and forum users [36]. Homophily can have several impacts. Wikipedia, Stack Overflow and Epinions users give higher evaluations to users or content authors similar to themselves [2]. In a commercial setting, consumers have higher levels of trust for product descriptions written by similar users as opposed to other user- or brand-generated content [44]. Specifically with Q&As, women tend to answer questions asked or already answered by other women [25,59], are likelier to view questions or upvote answers to questions asked by other women [24] and tend to interact with more women than men in general [8]. Communication between two people who share characteristics is more effective: with similar experiences and knowledge, less time is spent trying to understand one another [70]. It is no surprise that community leaders, event organizers, developers and others therefore create opportunities for encouraging homophilic behaviour to take advantage of these effects. For example, icebreakers [17] and robot characteristics [74] have been used to build more productive relationships by highlighting similarities between activity participants. Similarly, we investigate if we can prompt a sense of homophily in a Q&A by showing that a community has members similar to a user.

To summarize our survey of related work, findings from previous studies suggest that adding community presence information could potentially be a way to improve community dynamics in Q&As, something that is currently a deterrent to women's participation. SPA and homophily are two factors that facilitate social interaction in different contexts, including in online environments. However, studies of their effects in Q&As are limited in number, with some work identifying women showcasing naturally-occurring homophilic behaviour [8,24,25,59], but not explicitly prompted homophilic responses. We contribute to this space by using interface design to intentionally apply these concepts in a Q&A. We learn how a sense of SPA and of homophily prompted by community presence information can affect perceptions.

2 INCORPORATING ADDITIONAL COMMUNITY PRESENCE ON A Q&A

Prior work suggests that existing Q&As are unappealing to many women, resulting in women not fully benefiting from participating in Q&As (e.g., showcasing skills [66,93], growing their network [23,77,96,97]) and in content being of lower quality [9,20,83]. A factor deterring women from participating is the insufficiently social considerations in Q&A design [19]. Representing SPA and homophily as community presence information, we investigate how this information can impact users' perceptions of Q&As. In addition, we further consider how the degree of detail of community presence information affects responses from users: if the information needs to be detailed to impact one's sense of SPA and homophily or if it suffices to have a high-level summary.

To answer these questions, we created Community Presence interfaces by adding information about question viewers (a "view" being recorded when clicking on a question) to the question feed of a Q&A, focusing on promoting a sense of SPA and homophily. Aiming to increase the sense of SPA, the interfaces list information about question viewers, indicating their "presence" on a question. Meanwhile, to promote a sense of homophily, the interfaces display information about viewers' characteristics, prioritizing viewers with similarities to the user. We decided to display viewer information as opposed to active contributors mainly due to the 1-9-90 principle [61]; we felt that only showing information from contributors would not sufficiently populate the interface, while viewers present a much larger user subset to provide necessary information. To discern how the level of detail of community presence information can impact reactions, we implemented two Community Presence interfaces. We describe the interfaces, the Aggregated Views interface, and the Detailed Individuals interface in more detail.

The Aggregated Views interface (Figure 1) provides a high-level summary of the community presence information. In addition to information commonly found in a Q&A (question's title, question asker's name, date and time the question was asked, number of answers and views the question has received), the interface lists each of the characteristics for which the user provided information in their profile. For each of the characteristics, the percentage of question viewers sharing that characteristic is shown. With this interface, the user cannot see any detailed information (including characteristics) about individual viewers, the user can only see the aggregated community percentages for characteristics identical to their own. The intent is for users to see where viewers most like themselves are present in the Q&A.
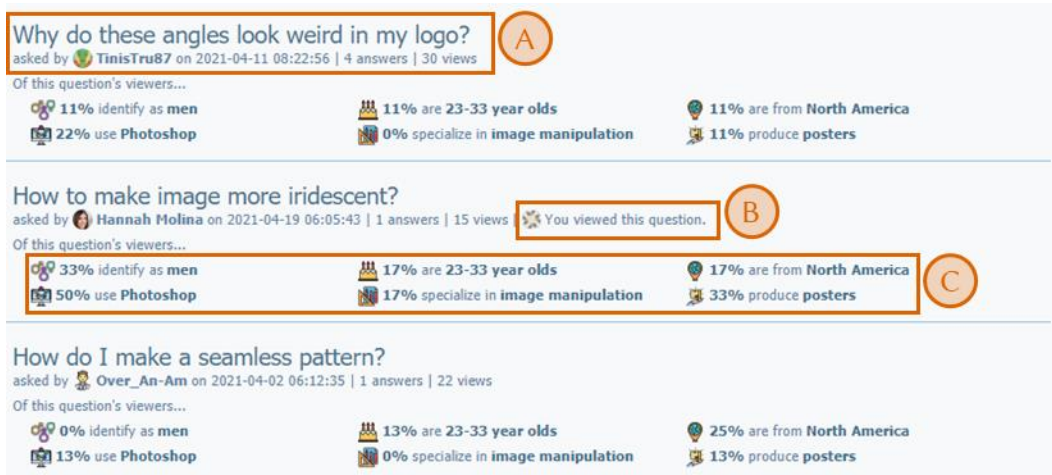

Figure 1: The Aggregated Views interface. A) Standard question information. B) An indicator showing the user has viewed this question. C) For each characteristic, the percentage of question viewers who have the user's characteristic. Here, the user is a North American man, so he can see how many North Americans and how many men have viewed each question (17% and 33% respectively in the highlighted example).

In comparison, the Detailed interface (Figure 2) displays information about individual question viewers. Each question displays two lists of question viewers: viewers that the user has followed and viewers that share some similarity with the user. In our implementation, these "similar viewers" must share at least one characteristic with the user, are sorted from most to least similar, and a maximum of six are displayed (due to space limitations). When the user hovers over a viewer's icon, the viewer's profile is displayed, showing all their characteristics (of those specified in the viewer's profile). Any characteristics that are the same as the viewer's are highlighted. The user can choose to follow or unfollow users from their profiles. This way, the user can see individual question viewers specifically, with a detailed view of their profile. With this interface, users do not see community-wide statistics.

Both Community Presence interfaces make use of and display user characteristics. Characteristics can be divided into two groups: external and internal characteristics [28]. External characteristics can generally be obtained without much interaction with a person and often represent a "surface-level" view of one's identity (e.g., name, place of birth, age, gender). On the other hand, deeper interactions with a person are needed to understand their internal characteristics (e.g., values, philosophies, preferences, attitudes). As the type of a characteristic that people have in common can impact the strength of each other's impressions [28], we selected six characteristics as a

starting point for our investigation: we treat gender, age and region the user identifies with as external characteristics, and preferred software, specialization and product type produced as internal characteristics.

Figure 2: The Detailed Individuals interface. A) Standard question information. B) Viewers that the user follows. C) Viewers who share characteristics with the user, ordered from most to least similar, fading out as they become less similar. D) The user can hover over profile icons (PosterCreator's in this screenshot), to see a viewer's characteristics. Any characteristics shared by the user are bolded (both the user and PosterCreator produce posters in this example), and the user can choose to follow/unfollow the viewer.

3 INVESTIGATING THE EFFECTS OF ADDITIONAL COMMUNITY PRESENCE

To see how community presence information can impact a user's perceptions of and engagement with a Q&A, and to see if there are any gender differences, we conducted a 10-day task-based field study. We used the interactions with the Community Presence interfaces to ground interviews on how community presence information can be used to promote inclusive environments.

For the purposes of this study, we chose graphic design as our Q&A topic. Graphic design has many different complex software, allowing for a variety of content, including content related to workflows, troubleshooting and opinions [19]. It is also a field that has near-parity in the number of men and women practitioners [81], something important to consider since imbalanced gender composition can impact level of engagement [38].

3.1 Participants

We recruited people who have graphic design knowledge and aimed to balance the number of men and women participants. We also sought to include the perspective of non-binary participants. We advertised through word-of-mouth and on social media sites like Reddit and Facebook, specifically targeting graphic design, women in graphic design and tech, and GSRM (gender, sexual and romantic minority) groups. We asked potential participants to share what kind of graphic design work they do, and at what frequency to ensure they had knowledge about graphic design software. We offered to conduct the study in English or in French, although all participants chose to participate in English. Participants received $25 CAD after signing the consent form, and $125 CAD upon completion of the study.

We recruited 30 participants: 14 men, 14 women and two non-binary participants. Participant ages ranged from 20-45, with the median age at 24.5. Participants reported how long they did graphic design work professionally or as a hobby: three indicated less than a year, twelve 1-5 years, nine 6-10 years and six over 10 years. Participants also identified their primary job title and field of work/study: 21 identified as graphic, UI or web designers, three participants had jobs indirectly related to design (i.e., marketing), six participants did not identify anything related to design as their profession (three in business management positions, two in healthcare, one in hospitality). Our call for participation did not require participants to use Q&A sites as we wanted to include perspectives of those who do not use Q&As. In all, 21 participants visit Q&A sites, with eight men and seven women having asked or answered questions.

3.2 Conditions

As a comparison point to our two Community Presence interfaces, we developed a Baseline interface that only has information that is commonly found in Q&As (see Figure 3). Every participant used the Baseline interface and one of the Community Presence interfaces so that we could see if gender differences in usage patterns or perceptions would emerge. Therefore, our study had a mixed design with one within-subject factor (*Interface Type*: Baseline and Community Presence interface) and two between-subject factors (*Gender*: man, woman and non-binary; *Enhancement Type*: Aggregated Views and Detailed Individuals interfaces). The order of interface and enhancement types was counterbalanced between genders.

is there a technique when using the paint bucket to get the color to blend into the black outlines well?
asked by 🔵 **Jak** at 2021-04-05 20:21:22 | 2 answers | 22 views | ❇️ You answered this question.

List of RGB quaternary (and beyond) colors?
asked by 😺 **Rataxana54** on 2021-04-14 12:11:16 | 4 answers | 14 views

What is the best way to enhance images
asked by ♟️ **Cy** at 2021-04-24 11:10:48 | 3 answers | 19 views | ❇️ You viewed this question.

Critique my 10y/o logo and my redesign attempt
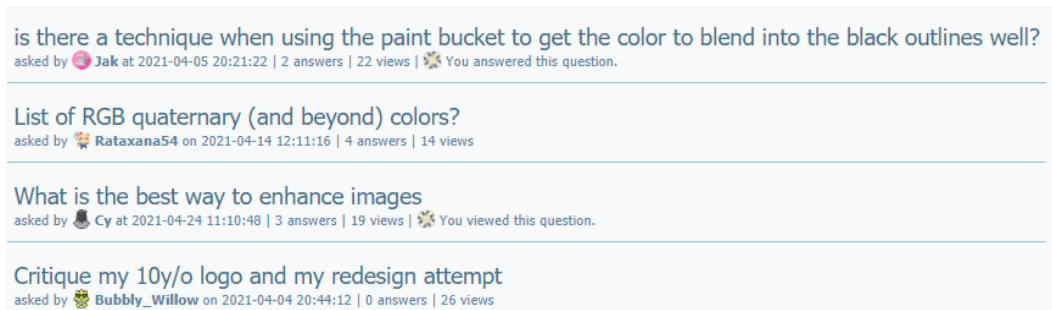asked by 🐻 **Bubbly_Willow** on 2021-04-04 20:44:12 | 0 answers | 26 views

Figure 3: The Baseline interface. Each question only has the information commonly found in Q&As and indicators showing which questions the user has viewed.

3.3 Q&A Content

As explained previously, graphic designers use a variety of software with the potential for a variety of questions. To increase ecological validity, we considered many factors when populating the Q&A. For practical purposes, we had to ensure that the Q&A contained questions about something every participant could answer. Furthermore, previous work suggests that content type and source can impact a person's willingness to answer, with different question types appealing to men and women differently [19], so we varied and balanced the types and sources of content.

In all, we collected two groups of 40 questions about graphic design from Graphic Design Stack Exchange[1] and Quora[2] each by using various graphic design software as filters/tags/search terms.

---

[1] https://graphicdesign.stackexchange.com/

[2] https://www.quora.com/

We had ten questions (five from Graphic Design Stack Exchange and five from Quora) from each of the following types of questions: seeking factual/troubleshooting information, workflows, opinions, and examples. Although the original questions usually requested answers for specific software, we made all the questions software agnostic to allow participants to answer any question regardless of known software.

Previous work also suggests that the number of existing answers can affect contributors' willingness to answer questions [15,25]. We randomized the number of answers per question from 0-5, ensuring that each question group described in the previous paragraph had the same total number of answers. Our answers were sourced from the original questions' answers.

We created 50 users for the Q&A, 25 per group of questions. Based on suggestions that they may impact participation and engagement [34], we invented usernames and selected profile icons using various styles found online (realistic names, fictional character names, fantasy names, with and without numbers and symbols; photographs, cartoons, images of popular media, people, animals, symbols). As we wanted to see if homophily can play a role in participant behaviour and perspectives when using the Community Presence interfaces, we generated characteristics for each of the users, ensuring a variety of combinations of characteristics. The users' characteristics were randomized with the following constraints: there were 40 men, 40 women, 10 non-binary users for gender; a distribution in age from 18-65 (the curve leaning towards the younger end of the range), with randomized regions, software, product and specialization (where each one had a minimum of one and a maximum of nine). We randomly assigned each question and answer to a user as its asker/answerer and randomly assigned users as question viewers.

The question groups and their associated answers and viewers were assigned to participants in a round-robin fashion.

3.4 Procedure

After confirming the participant's graphic design abilities, the participant met the first author for an initial meeting. The author explained the study's purpose and the procedure. The participant filled in a consent form, a demographics questionnaire (which included identifying the characteristics needed for the Community Presence interfaces) and received $25 CAD.

Following the initial meeting, the participant used the first assigned interface for a period of five days to ensure that they got familiar with the interface, was able to fully explore it, and had time to do research if needed to answer Q&A questions without researcher pressure. The author instructed the participant to try to use the Q&A every day and to answer at least one question per day (but to answer as many as they felt comfortable answering). To reduce participant posting anxiety and workload, and as reassurance that their content would not be judged or be misleading to potential readers, the author told the participant that the community was constructed and static, but that the participant should otherwise pretend that it was real and to try to behave as they do with live online communities. After the five days, the participant filled in a post-interface questionnaire. Once the questionnaire was submitted, the participant could access the second interface, where they followed the same procedure as with the first interface.

After filling the second post-interface questionnaire, the participant met the first author once more for a semi-structured interview. These interviews ranged from 30-60 minutes, and focused on the participant's thoughts on the Community Presence interfaces and the participant's feelings and perceptions of the community as they were using the Q&A. As part of the interviews, the

author introduced the interface that the participant did not use to elicit initial comparative reactions. The participant received $125 CAD at the end of this session.

3.5 Data Collection and Analysis

Our primary data-collection method for participant interface preferences and perspectives was the semi-structured interviews. In addition, the prototype logged which questions the participant viewed/clicked on and answered, and we administered two post-interface questionnaires per participant. The questionnaire had 35 Likert-like questions, taken from questionnaires used to measure SPA [5,33,47], homophily [53], engagement [62] and sense of community [1,18].

All interviews were fully transcribed, and participants could request to review their transcripts. The transcripts were then analyzed thematically: the first author grouped participant quotes by similar topic and feeling, then assigned thematic labels. A second author reviewed the groupings and themes, rearranging the quotes according to her interpretations and modifying the labels. Together, the two authors then went over the quotes and themes again, discussing their interpretations and revising the labels until they were in agreement [54].

Participants rated their agreement to the 35 questions on the post-interface questionnaires using scales of 1 (strongly disagree) to 5 (strongly agree). The answers were then summarized into three summative values: sense of SPA, sense of homophily and sense of engagement (the scales for sense of community having been combined into the others). We compared quantitative means (number of questions answered and clicked on, the three summative values) using a two-way ANOVA and report results as significant if $p < 0.05$.

4 RESULTS

Our task-based field study revealed mixed participant reception and perceptions of having community presence information available in a Q&A. All participants had a clear preference between the Baseline interface and their assigned Community Presence interface. As Figure 4 shows, overall, participant preferences are generally evenly spread out with no clearly preferred interface. Our findings shed light on how and why participant perceptions diverged.
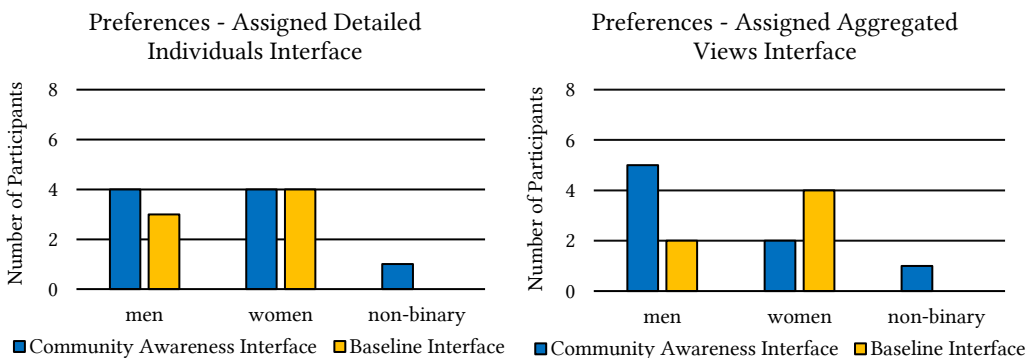

Figure 4: The distribution of participant interface preferences.

4.1 Perceptions of Potential Uses and Usefulness of Community Presence Information

Although we saw consistency in participant responses on the potential uses of the community presence information, their opinions differed on their perceived usefulness in a Q&A. Our

interview themes shed light on what participants saw as important factors when determining their position in terms of humanizing the Q&A, of using homophily and heterophily to promote inclusive environments and of privacy implications. We also found some differences between men and women, particularly about the potential role of heterophily in inclusivity. When presenting interview quotes, participant identifiers that start with M are men, W are women, and N are non-binary participants.

5.1.1 Community Presence Information Humanizes the Q&A

Almost every participant explained how the community presence information shifted how they viewed the Q&A; greater focus was placed on the users as opposed to the content. Participants had mixed feelings towards this change of focus. Some participants suggested that this was a good thing and felt like it was easier to engage and interact with the Q&A.

> I think because of the percentages below each question, it made me think more about the community [with the Aggregated Views interface]. [With the Baseline interface], I just skimmed through the previous answers. – M01

> The [Baseline interface] was more like a Q&A with people you don't know. The [Detailed Individuals interface] for me looked more like a forum, that you have people that are similar, always there, hanging out, helping each other. The [Baseline interface], I didn't feel like that, even though the [users] were the same. – M08

Many of the participants who viewed this shift positively described this effect as humanizing the Q&A. They felt like the displayed characteristics gave life to the other users and made them more "real," to the point that some participants wondered if the other users were in fact, other participants in the study (despite understanding that they did not see content from other participants). In contrast, participants often used the word "bot" to describe the users that appeared on the Baseline interface.

> It was nice because as I said, it felt like someone real. It felt real. It felt like a real community. [...] I could get some information from this person, not just this person has this question, or just someone with another username and nothing attached to it. – W02

> For the [Detailed Individuals interface], [the users] didn't feel like bots. They felt like real people. So, because they had names that I knew, they had interests that I could have read. – M13

Participants receptive of the humanization of the Q&A saw positive changes to their perception of the Q&A. M08 described that he felt greater empathy for users asking questions. He felt a greater desire to be helpful as opposed to his usual behaviour.

> [I] started to care about those icons that said [I was] related to a real person. [...] I would [do research] for [people]. In the [Baseline interface] maybe I would just put a "Google it" you know? "Google it, man, it's an easy question, it shouldn't be there. It's so easy that it's not good for the community to be that naive." But in the [Detailed Individuals interface], maybe I would be more empathetic about that. I would do the extra mile for them. – M08

M04 and W12 additionally described that humanization affected their trust in the content: M04 described the users with community presence information (i.e., characteristics) as more trustworthy and reliable, while W12 described the content on the Baseline interface as being less trustworthy.

> I think more interest is developed, you have that authenticity, I don't know the exact word. But, it feels more reliable, it doesn't feel like a bot is answering or something like that. – M04

> I wouldn't trust the answers [on the Baseline interface] nearly as much [...] versus on the [Detailed Individuals interface]. – W12

Other participants felt negatively about shifting the focus from the content to the users. A few participants expressed they would feel undesirable pressure to participate or be more self-conscious about their participation if their presence were made known to other users, especially if they did not answer questions.

> If I open the question [and] did not answer, I would not like for people to see that it was me. Maybe [they would] feel ignored or something like that. – W06

> [I]t also kind of makes me feel like [I'm being rude] if I'm on the site and then I view [a question], but then I don't answer it, so there it creates this weird interaction I feel. Where you know it makes me self-conscious when I shouldn't be self-conscious and there's no reason to be. – M02

Finally, some participants (men and women) expressed that community presence information simply does not belong on a Q&A despite any advantages it brings and understanding that other people might find it useful. They explained that Q&As are for knowledge-sharing and that socializing is not important for this task.

> I kinda don't care about people. Ok, I want to help and be helped but when I [am in a] graphic design community […], I'm just in for the content, not [to] make friendships. – M05

> For me I'm just [using this kind of platform] to help teach, whereas others might actually use it for this kind of [community presence] information. – W13

Based on our interviews, participants consistently explained that their perception of the Q&A changed because of the community presence information: a greater focus was placed on the users and the community as opposed to the content. Participant reception of such a shift in focus was quite mixed, however, with participants diverging on the value it can provide to a knowledge-sharing platform.

5.1.2 Homophily and Heterophily Can Both Promote Inclusive Environments

In general, participants, regardless of gender, felt like they had an easier time relating to and feeling included in the communities when using the Community Presence interfaces than when using the Baseline interface because they could see active users similar to themselves. For example, W11 explains how knowing that someone is from the same place as her makes it easier to connect with them, W03 describes how it is easier to fit in with others like her and N01 describes how they chose to see content posted by people similar to themselves:

> [W]hen someone's from the same place as me for some reason it makes me feel like we connect in some way. – W11

> The thing that caught my eye […] the first time [is the list of similar users]. I will look for the similarities. […] I would look for that because it gives you a sense of fitting in. – W03

> [I was] more interested in seeing what people like me would say. Yes, so if there were two like similar questions [I] picked the one that had more people in my demographic – N01

The above responses assume that there are similar users viewing questions to begin with. One woman mentioned that she might consider viewing questions with no women viewers to ensure that women are represented, suggesting that there needs to be a "starting point" that could welcome future users.

> [I]f it was like 0% of the people who answered this question or viewed this question are women and I would be like, "hey we need some representation," I might view the question. – W07

Although our goal was to use homophily to affect perceptions of the Q&A, some women talked about the effects of heterophily, while men and non-binary participants did not. As opposed to homophily, where people are likely to engage with others with similarities, heterophily is the

phenomenon where people are likely to engage with others with differences [70]. With the Detailed Individuals interface, women welcomed seeing users and their characteristics, whether the characteristics were the same as theirs or not. On the other hand, women could only see how many users matched their characteristics with the Aggregated Views interface and wanted these numbers to also include viewers that were unlike them. For example, W07 and W05 appreciated seeing geographically and gender diverse users.

> If you ask a question, you assume everyone is this white dude from North America. And so it was really nice to see like there is like a diversity of people answering. – W07

> [I]t's important to feel like I'm in a global melting pot. [People] have different qualities of life around the world as well. So, you may be thinking that like everyone is using Photoshop, but for someone in Zambia, where I am, or in South Africa, where I'm from, Photoshop might be expensive. – W05

To summarize, most participants saw value in seeing viewers with similarities to themselves. However, women in particular desired to see (more) information about diverse viewers. Homophily and heterophily can potentially both be used to promote inclusive Q&As, as participants said seeing similar users to themselves helped them feel like they fit in, and women described seeing diverse users as a signal that everyone is welcome.

5.1.3 Integrating Community Presence Information has Privacy Implications

Participants' heightened awareness of other people viewing the Q&A came with some costs, especially as information is needed to populate the user profiles. Some participants using the Detailed Individuals interface had privacy concerns with sharing information about themselves, even if they recognized the information as not being personally identifying; they did not trust how other users would use or interpret such information. No participants using the Aggregated Views interface had such concerns, perhaps because the data was aggregated.

> [It's] common on social media that [people] send, I don't know, some not good things to your profile when you can chat with them when they know you're a boy or a girl. – M10

> I'd like to see something like that, how others are seeing me right now. Maybe it's difficult because of the data is personalized so they're going to see different from me. I don't know how they're going to see me. I was worried about that. Because if I wanted to make an impression, I didn't know how they're going to see me. – M08

There was even one participant concerned that some users might put in false information in their profile for deceptive purposes.

> Like on [social media] they'll say they're a woman when they're a man. People feel more inclined to like get close to them. There's a lot of deception on the Internet. – W05

Some participants phrased their heightened privacy concerns using the Community Presence interfaces as an advantage. Since they could see other viewers' information, they had a clear idea of how their own information was being shared. With the Baseline interface, they knew information was being collected and shared, but nothing precise.

> Because of the [...] percentages that were shown [on the Aggregated Views interface], I was more conscious about [the characteristics] that [I was publicly sharing]. – M01

> I think with the [Detailed Individuals interface] I just kind of had this sense of like [...] I'm in a group, and these are the only other peers. I feel this is what they're sharing, like it was kind of like I had less privacy concerns because it felt like other people were also sharing stuff – W12

In short, although participants did have privacy concerns about having to share more information, there was an upside, that community presence information showed how that information was being used.

4.2 Impacts of Community Presence Information on Behaviour and Questionnaire Responses

In our interviews, the majority of participants reported intentionally trying to behave identically using both interfaces (e.g., answer a precise number of questions per day), and tried to give the same answers both times they filled the questionnaires. We believe this behaviour can partially be explained by the study description we gave participants: an exploration that could help us better understand the role of interfaces in working towards gender inclusivity, rather than a study comparing different interfaces. Some participants explained that they behaved similarly across their assigned interfaces because they believed that otherwise, they would be inconsistent and misrepresent themselves or even their gender. In addition, we were cognizant of anchoring effects when participants fill the same questionnaires multiple times [12]. Therefore, when comparing quantitative means, we only consider data from participants' first-assigned interface and so have *Gender* (man vs. woman) and *Interface Type* (Community Presence interface vs. Baseline interface) as between-subject factors. We did not compare *Enhancement Type* (Aggregated Views interface vs. Detailed Individuals interface) nor include non-binary participants due to data sparsity.

We logged the number of answers posted and questions clicked. We also calculated three summative scores from the questionnaire administered after each condition (but as described above, only analyzed data from the first condition). In all, we analyze the means of seven men and six women using the Baseline interface and seven men and eight women using a Community Presence interface. All quantitative data were analyzed using a two-way ANOVA. In this section, we describe the results in detail, see Table 1 for a summary.

Table 1: Results from interface usage and questionnaire data divided by *Interface Type* and *Gender*. Statistically significant main and interaction effects are bolded.

Variable	Baseline interface		Community Presence interface		Effects		
	Men's Mean (*SE*)	Women's Mean (*SE*)	Men's Mean (*SE*)	Women's Mean (*SE*)	*Gender*	*Interface Type*	*Gender* ✕ *Interface Type*
Answers posted	16.143 (3.247)	10.500 (1.875)	17.714 (2.714)	8.000 (1.195)	$F_{1,24}$ **= 10.454** p **= 0.004**	$F_{1,24}$ = 0.038 p = 0.847	$F_{1,24}$ = 0.735 p = 0.400
Questions clicked on	22.143 (3.225)	22.333 (2.848)	26.286 (3.006)	13.375 (2.203)	$F_{1,24}$ **= 5.051** p **= 0.034**	$F_{1,24}$ = 0.724 p = 0.403	$F_{1,24}$ **= 5.358** p **= 0.029**
SPA min: 13, max: 65	43.429 (1.938)	42.500 (3.233)	44.429 (2.644)	41.875 (2.601)	$F_{1,24}$ = 0.441 p = 0.513	$F_{1,24}$ = 0.005 p = 0.944	$F_{1,24}$ = 0.096 p = 0.759
Homophily min: 11, max: 55	35.143 (2.064)	40.833 (3.506)	38.857 (3.188)	31.125 (1.807)	$F_{1,24}$ = 0.150 p = 0.702	$F_{1,24}$ = 1.290 p = 0.267	$F_{1,24}$ **= 6.470** p **= 0.018**
Engagement min: 11, max: 55	42.571 (2.359)	44.333 (3.127)	44.571 (2.716)	38.125 (1.726)	$F_{1,24}$ = 0.909 p = 0.350	$F_{1,24}$ = 0.733 p = 0.400	$F_{1,24}$ = 2.791 p = 0.108

5.2.1 Participant Usage and Contributions

To get a sense if participant usage would change between interfaces, we compared the number of answers they posted and the number of questions they clicked on between interfaces.

We saw a main effect on *Gender* for the number of answers posted ($F_{1,24}$ = 10.454, p = 0.004), where men (Baseline: 16.143, Community Presence: 17.714) posted more answers than women (Baseline: 10.500, Community Presence: 8.000), replicating previous findings (e.g., [19,52,89]) but we saw no statistically significant effect on *Interface Type* ($F_{1,24}$ = 0.038, p = 0.847) and no interaction effect of *Gender × Interface Type* ($F_{1,24}$ = 0.735, p = 0.400).

As for question clicks, we also saw a statistically significant main effect of *Gender* ($F_{1,24}$ = 5.051, p = 0.034) but not of *Interface Type* ($F_{1,24}$ = 0.724, p = 0.403). We did see a statistically significant interaction effect of *Gender × Interface Type* ($F_{1,24}$ = 5.358, p = 0.029). As illustrated by Figure 5 (left), the men clicked on more questions using the Community Presence interface (26.286) compared to the Baseline interface (22.143), but the women clicked on almost half the number of questions using the Community Presence interface (13.375) compared to the Baseline interface (22.333). This could mean that women did not need to click on as many questions to obtain the information they needed to pick which questions to answer, with the information on the question feed (i.e., the community presence information) sufficing. It is also possible that the women did not use the interface as much, but the other findings do not lend much support for this interpretation.
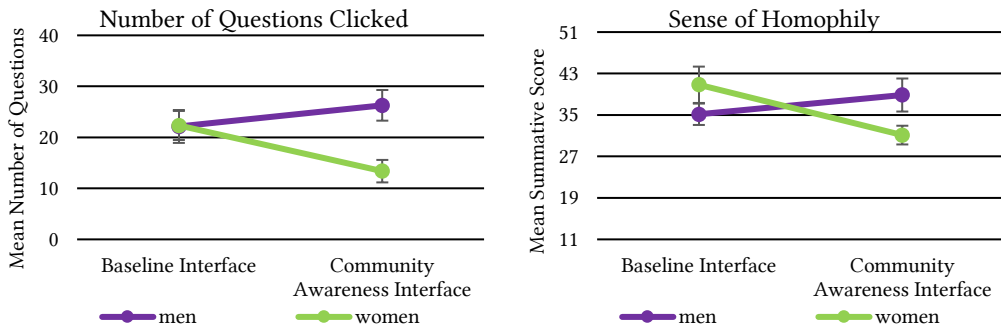

Figure 5: Mean number of questions participants clicked on (left) and participants' sense of homophily (right). Error bars represent *SE*.

In summary, we did not see many changes in behaviour. However, this potentially suggests that the community presence information did not prove to be too distracting from the task.

5.2.2 Participant Sense of SPA, Homophily and Engagement

To compare participants' sense of SPA, homophily and engagement, we used three summative scores from the questionnaires. The summative scores could range from 13-65 for the sense of SPA and 11-55 for both sense of homophily and engagement, where a low score denotes a low sense of SPA/homophily/engagement. In calculating the summative scores, we inverted the scores for any negative statements (e.g., a 1 for "I felt as if I was alone in the community" was converted to a 5 for the SPA score).

Looking at sense of SPA, men (Baseline: 43.429, Community Presence: 44.429) and women (Baseline: 42.500, Community Presence: 41.875) had similar scores: we did not see a statistically

significant main effect of *Gender* ($F_{1,24} = 0.441$, $p = 0.513$) or of *Interface Type* ($F_{1,24} = 0.005$, $p = 0.944$), nor an interaction effect of *Gender* × *Interface Type* ($F_{1,24} = 0.096$, $p = 0.759$).

As for scores for sense of homophily, we did not see statistically significant main effects of *Gender* ($F_{1,24} = 0.150$, $p = 0.702$) or *Interface Type* ($F_{1,24} = 1.290$, $p = 0.267$). However, we did see a statistically significant interaction effect of *Gender* × *Interface Type* ($F_{1,24} = 6.470$, $p = 0.018$). Looking at Figure 5 (right), there is a slight increase for men comparing the Baseline interface (35.143) to the Community Presence interface (38.857) and a decrease for women between the Baseline interface (40.833) and the Community Presence interface (31.125). We can find some support for this difference from our interviews: women seemed to be more sensitive to heterophily and diversity; with the Community Presence interfaces making this information more easily available, heterophilic responses could have been stronger than homophilic ones.

Finally, for sense of engagement, men (Baseline: 42.571, Community Presence: 44.571) and women (Baseline: 44.333, Community Presence: 38.125) again had similar scores, with no statistically significant effect of *Gender* ($F_{1,24} = 0.909$, $p = 0.350$), of *Interface Type* ($F_{1,24} = 0.733$, $p = 0.400$), or of *Gender* × *Interface Type* ($F_{1,24} = 2.791$, $p = 0.108$).

6 DISCUSSION

Our study has shown that community presence information has potential to change user perceptions of a Q&A by humanizing it, by promoting a more inclusive environment and by increasing user privacy awareness. On the other hand, some participants felt that such information does not belong in a Q&A because it detracts from the content, and some were concerned about privacy implications. Below we elaborate on our findings and promising directions for future research. We also discuss limitations of our study.

5.1 Community Presence Information: Choosing When and How to Reveal Personal Characteristics

Participants mentioned that community presence information humanized the Q&As. In particular, the focus shifted from the content to the Q&A users, potentially increasing participant trust of content and empathy towards other users. Participants also mentioned feeling a heightened sense of homophily seeing that they share characteristics in common with viewers. We saw these perceptions expressed with both Community Presence interfaces, despite their differing levels of detail.

The women in our study expressed some feelings of heterophily, which is generally not well supported by the Aggregated Views interface, as it only shows viewers with the user's characteristics. We considered showing summaries of all characteristics on the Aggregated Views interface but were concerned about presenting an overwhelming amount of information that would de-emphasize the similar viewers. To address this concern, some participants suggested that the information could be available on demand. However, these same participants then explained that this might start transforming the interface from showing community presence to an analytical tool to better understand community composition and interests. In short, it appears that organizing user characteristics into user profiles is the preferred method for showcasing community diversity.

Some participants expressed privacy concerns with community presence information, including that too much information might be shared, and that the information is shared upon question viewing (as opposed to when answering a question). For the study, we assured participants that

any information they provided (including their characteristics) would be anonymized and used for research purposes only. We did ask participants which of the characteristics they would be (un)comfortable sharing given a live deployment. While many participants did not have strong feelings on the matter, some participants were concerned that this information could potentially be used maliciously, for example, to harass women, which is a well-known phenomenon on the internet (e.g., [32,45,68]). Furthermore, some participants explained that showing their activity to others could cause undesirable pressure to participate. Sharing characteristics and activity might make some users feel even more vulnerable than they already do, potentially widening the participation gap we would like to close.

We chose to show viewers as opposed to active contributors on the interfaces to highlight a wider range of community activity. Showing only question askers and/or answerers might help users feel a higher degree of control over their presence: when asking or answering a question, a user has already consciously decided to signal their presence. On the other hand, showing only active contributors would skew community presence to those comfortable contributing content. One potential compromise would be to allow viewers to opt-in/opt-out of having their presence displayed. This would increase flexibility at the cost of additional interface complexity. It would also impact how the community is represented.

In our implementation, participants were required to provide data for all six of the characteristics used by the interfaces. This decision was primarily for study purposes: we wanted the interfaces to compare participants with our fake users with the intent of prompting homophilic responses. However, it is unclear if Q&A users would be willing to create and maintain profiles in the first place. Both participants that did and did not see potential in the approach said they would be willing to provide their characteristics, but they might have said so just to please the researcher [16]. Compounded with the privacy issues mentioned above, it would be reasonable to assume hesitance from Q&A users to provide personal information.

Future versions could give users the option to not provide data for all characteristics or hide their choices from others. While this might alleviate some of the privacy concerns, the benefits of our interfaces in promoting homophilic and heterophilic responses do rely on sufficient community data. Another possible solution would be to reconsider the list of characteristics to ensure that there is enough for users to feel comfortable supplying at least some data. For example, participants offered ideas ranging from graphic design skill level to specifying favourite books, movies, and food. Users could therefore tailor how the interface supports homophilic and heterophilic responses. Another approach would be to collect user characteristics, but not display them to others. However, users may become mistrustful of a system that does not explain how it determines similarity [86]. Future research could also consider different approaches for prompting homophilic responses, for example, by matching content authors using their writing styles.

We populated our interfaces with fake users and ensured all characteristics were represented with no strong imbalances (e.g., 40 men, 40 women, 10 non-binary users). Our goals in doing so were twofold: we wanted to ensure that a participant with any characteristic would be able to see users with the same characteristics; and large community imbalances in representation can negatively impact participation [38]. Real Q&As tend not to have such balanced representation. For example, two popular Q&As, Quora and Stack Overflow, have more men than women users [67,80]. It is possible that explicitly displaying characteristics could increase the risk of creating a vicious cycle: seeing that a community has low representation might in fact discourage someone from

participating. These issues are more evident with the Aggregated Views interface than with the Detailed Individuals interface. The Aggregated Views interface provides a global view of community demographics, including any imbalances, whereas the Detailed Individuals interface show only select viewers. The composition of this set of viewers has the potential to deemphasize imbalances (e.g., a woman user could potentially see only other women viewers in this list). In fact, it is possible that "similar, underrepresented viewers" could end up being overrepresented in the Detailed Individuals interface, but we do not consider this to be a problem, as the interfaces are meant to emphasize user similarities. Additionally, prior work about skewed community demographics tend to refer to posted content only; viewer demographics are more challenging to collect, but might not be as skewed as active contributors. Further research is be needed to understand these impacts.

5.2 Using Community Presence Information to Improve Women's Participation

Our motivations for including community presence information for promoting a sense of SPA and homophily include improving social dynamics in Q&As, and eventually working towards solutions that increase participation from women. Our study suggests that this approach is a promising step towards these goals, despite some resistance from participants.

Using community presence information, we aimed to prompt a homophilic response from participants. Our intent in showing that there are users similar to participants using the Q&A was to create stronger feelings of belongingness, especially since gender homophily is stronger among women than men [4,41]. Although our findings support this assertion, women clarified that seeing a diversity of users has the potential to encourage them to keep using the Q&A. How an interface might simultaneously support both phenomena is an interesting design problem. For example, one might consider showing similar community members to a new user, but gradually increasing the number of diverse members to encourage retention.

The women participants clicked on fewer questions using the Community Presence interfaces. They did not exhibit this behaviour on the Baseline interface, nor did men on any interface. This suggests that women obtained the information they needed to pick which questions to answer from the question feed, potentially showing stronger interest in the community presence information compared to the men. The other gender-related quantitative results did not show strong effects, apart from the number of questions answered (which is unsurprising given prior work [19,52,89]). Thus, our study did not demonstrate strong positive gender effects, but also did not show signs of the extra awareness information causing a significant distraction. Given that participants were required to answer at least one question per day, it is also possible that we have not captured participants' natural sense of engagement. As for SPA, most of the questions asked about participants' awareness of user activity, and it might have been difficult for participants to look past the mock, static community we created. A longer-term, more open-ended deployment is needed to better understand the impacts on women's participation and engagement.

In our interviews, one participant clearly expressed that he felt more empathy for people asking questions when there was community presence information and we saw similar inklings from other participants. These sentiments suggest that community presence information could potentially assist in creating a collaborative or collectivist environment. Currently, Q&As are quite individualistic and competitive [49], which is a deterrent to participation, especially among women [19]. Our findings suggest that if community presence information can lead to collaborative environments, Q&As might have greater appeal among women, who already display

supportive and community-oriented values [95]. In addition, we might see increased participation from others with collectivist outlooks who do not feel included in Q&As [64].

Although community presence information has potential to improve women's participation in Q&As, some men and women participants felt that this information does not belong in Q&As. They felt that the Q&A would shift from a knowledge-sharing community to a social networking one. A recent experiment where Stack Overflow added a "thank you" button [78] had similar feedback, with users describing this experiment as a slippery slope towards social networking [79]. Many users go to Q&As for knowledge-sharing, and some might think that exclusively social features distract from this objective. It is possible to hide social features from users who do not want them, but social dynamics could still change and cannot be "toggled off." We argue, however, that it is not certain if Q&As would "further become" social networks: Q&As are social in nature and additional social features enhance this aspect of the interaction, they do not replace the goal of asking and answering questions. This desire to maintain the status quo highlights that Q&As have structural issues that need to be studied.

5.3 Moving Beyond Community Presence Information and Q&As

In this paper, we investigated how community presence information can be used to work towards improving feelings of socialness in Q&As. Social presence and awareness and homophily are not the only two components of social interaction, nor is community presence information the only way for an interface to take advantage of social phenomena. For example, one can consider allowing users to use alternative feedback mechanisms rewarding good social behaviours [48] or to form sub-communities to support one another in the sometimes overwhelmingly large Q&As [26]. Combining and comparing solutions, so that they interact as a whole will be an interesting challenge, particularly as feature awareness in Q&As is a barrier reported by more women than men [26]. It is important that any new features integrate well with existing mechanics and are easy to discover, understand and use.

Furthermore, while our findings potentially extend beyond Q&As, different types of platforms have different community dynamics and norms that would have to be considered. Q&As are for asking and answering questions about a particular topic and so we could leverage information gathered through these tasks (e.g., question views). Other platforms may have different uses and/or contexts, which may not provide the same or any information that could be collected implicitly. For example, it may not make sense to use "views" as community presence information for community-contributed software tutorial comments (which are usually all displayed on the same page) or for in-software help (which do not have traditional browsing interfaces). Q&As are also composed of questions, which are discrete entities that we can use to identify user presence. In general, exploring how to share community presence information on other knowledge-sharing communities present interesting future work.

5.4 Limitations

In our 10-day task-based field deployment study, we populated our interfaces with ecologically valid content, however, they were not powered by live communities. Although this was done intentionally, to reduce participant concerns about community criticism, it might have increased their confidence beyond what they would naturally exhibit. Additionally, to guarantee that participants would have sufficient exposure to the interfaces, they were asked to answer at least one question a day. The most significant impact of these two decisions would be on the number of answers that participants posted: it is likely that it is higher than what participants would have

posted in a live Q&A. Long-term field studies would allow participants to engage with community presence information in a natural setting, raising ecological validity. Long-term studies are also needed to see permanent effects on user behaviour [63].

Furthermore, our interfaces did not contain all features that are commonly found in Q&As. Popular features include comments on questions and answers, and reputation systems. We did not include these features to focus the study on the community presence information. These other features, however, do have a significant impact on gender participation [73,89,95]. It is possible that the design of these features excludes certain people, among them women, that inclusive elements, potentially such as community presence information, may not be enough to overcome this exclusion. Future research should examine how integrating community presence information in fully-featured Q&As might help encourage more gender-diverse participation.

Finally, our study had a relatively low number of participants considering the between-subjects analysis we used for the quantitative data. A study with a larger number of participants would give a clearer picture of usage differences between genders and interfaces. We also had an insufficient number of non-binary participants to include this group in the quantitative analyses. For studies to be fully inclusive, the number of non-binary participants should ideally be balanced with the number of men and women.

6 SUMMARY

We presented an investigation into using community presence information to prompt Q&A users with a sense of social presence and awareness, and homophily, to work towards balancing gender participation. Based on interview and questionnaire data from a 10-day task-based field study, we found that community presence information can humanize a Q&A, promote an inclusive environment, and increase privacy awareness. These findings suggest many different potential avenues for further researching how Q&As can use community presence information to foster healthy social dynamics and encourage participation from women.

REFERENCES

[1]   Dagmar Abfalter, Melanie E. Zaglia, and Julia Mueller. 2012. Sense of Virtual Community: A Follow Up on its Measurement. *Computers in Human Behavior* 28, 2: 400–404. https://doi.org/10.1016/j.chb.2011.10.010

[2]   Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2012. Effects of User Similarity in Social Media. *Proceedings of the 5th ACM International Conference on Web Search and Data Mining - WSDM'12*: 703–712. https://doi.org/10.1145/2124295.2124378

[3]   Judd Antin, Raymond Yee, Coye Cheshire, and Oded Nov. 2011. Gender Differences in Wikipedia Editing. *Proceedings of the Symposium on Wikis and Open Collaboration - WikiSym'11*: 11–14. https://doi.org/10.1145/2038558.2038561

[4]   Chen Avin, Barbara Keller, Zvi Lotker, Claire Mathieu, David Peleg, and Yvonne Anne Pignolet. 2015. Homophily and the Glass Ceiling Effect in Social Networks. *Proceedings of the 6th Innovations in Theoretical Computer Science - ITCS'15*: 41–50. https://doi.org/10.1145/2688073.2688097

[5]   Frank Biocca, Chad Harms, and Jenn Gregg. 2001. The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity. *4th Annual International Workshop on Presence*, December: 1–9. Retrieved from http://astro.temple.edu/~lombard/ISPR/Proceedings/2001/Biocca2.pdf

[6]   Santi Priya Bose. 1967. Social Interaction in an Indian Village. *Sociologia Ruralis* 7, 2: 156–175.
[7]   Siân Brooke. 2019. "Condescending, Rude, Assholes": Framing Gender and Hostility on Stack Overflow. *Proceedings of the Third Workshop on Abusive Language Online*: 172–180. https://doi.org/10.18653/v1/w19-3519
[8]   Siân Brooke. 2021. Trouble in Programmer's Paradise: Gender-Biases in Interacting on Stack Overflow. *7th International Conference on Computational Social Science IC²S²*: 1–3.
[9]   Mason A. Carpenter. 2002. The Implications of Strategy and Social Context for the Relationship Between Top Management Team Heterogeneity and Firm Performance. *Strategic Management Journal* 23, 3: 275–284. https://doi.org/10.1002/smj.226
[10]   Sangmi Chai, Sanjukta Das, and H. Raghav Rao. 2011. Factors Affecting Bloggers' Knowledge Sharing: An Investigation Across Gender. *Journal of Management Information Systems* 28, 3: 309–341. https://doi.org/10.2753/MIS0742-1222280309
[11]   Parmit K. Chilana, Andrew J. Ko, and Jacob O. Wobbrock. 2012. LemonAid: Selection-Based Crowdsourced Contextual Help for Web Applications. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'12*: 1549–1558. https://doi.org/10.1145/2207676.2208620
[12]   Andy Cockburn and Carl Gutwin. 2019. Anchoring Effects and Troublesome Asymmetric Transfer in Subjective Ratings. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'19*: 12 pages. https://doi.org/10.1145/3290605.3300592
[13]   Richard H. Cutler. 1995. Distributed Presence and Community in Cyberspace. *Interpersonal Computing and Technology* 3, 2: 12–32.
[14]   Maitraye Das, Brent Hecht, and Darren Gergle. 2019. The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'19*: 1–14. https://doi.org/10.1145/3290605.3300793
[15]   David Dearman and Khai N. Truong. 2010. Why Users of Yahoo! Answers Do Not Answer Questions. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'10*: 329–332. https://doi.org/10.1145/1753326.1753376
[16]   Nicola Dell, Vidya Vaidyanathan, Indrani Medhi, Edward Cutrell, and William Thies. 2012. "Yours is Better!" Participant Response Bias in HCI. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'12*: 1321–1330. https://doi.org/10.1145/2207676.2208589
[17]   Julie S. Dixon, Heather Crooks, and Karen Henry. 2006. Breaking the Ice: Supporting Collaboration and the Development of Community Online. *Canadian Journal of Learning and Technology / La revue canadienne de l'apprentissage et de la technologie* 32, 2.
[18]   Patrick Dubois, Volodymyr Dziubak, and Andrea Bunt. 2017. Tell Me More! Soliciting Reader Contributions to Software Tutorials. *Proceedings of Graphics Interface - GI'17*: 16–23.
[19]   Patrick Marcel Joseph Dubois, Mahya Maftouni, Parmit K. Chilana, Joanna McGrenere, and Andrea Bunt. 2020. Gender Differences in Graphic Design Q&As: How Community and Site Characteristics Contribute to Gender Gaps in Answering Questions. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2. https://doi.org/10.1145/3415184
[20]   P. Christopher Earley and Elaine Mosakowski. 2000. Creating Hybrid Team Cultures: An Empirical Test of Transnational Team Functioning. *Academy of Management Journal* 43, 1: 26–49. https://doi.org/10.2307/1556384
[21]   Alberto Espinosa, Jonathan Cadiz, Luis Rico-Gutierrez, Robert Kraut, William Scherlis, and Glenn Lautenbacher. 2000. Coming to the Wrong Decision Quickly: Why Awareness Tools Must be Matched with Appropriate Tasks. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'00* 2, 1: 392–399. https://doi.org/10.1145/332040.332463
[22]   Umer Farooq, Craig H. Ganoe, John M. Carroll, Isaac G. Councill, and C. Lee Giles. 2008. Design and Evaluation of Awareness Mechanisms in CiteSeer. *Information Processing and Management* 44, 2: 596–612. https://doi.org/10.1016/j.ipm.2007.05.009
[23]   Holly Fawcett. 2012. 3 Unusual Q&A Sites to Source IT Talent From - Quora, Github, StackOverflow. *SocialTalent*. Retrieved April 2, 2020 from https://www.socialtalent.com/blog/technology/3-unusual-qa-sites-to-source-it-talent-from-quora-github-stackoverflow
[24]   Alejandro Figueroa. 2017. Male or Female: What Traits Characterize Questions Prompted by Each Gender in Community Question Answering? *Expert Systems with Applications* 90: 405–413. https://doi.org/10.1016/j.eswa.2017.08.037
[25]   Denae Ford, Alisse Harkins, and Chris Parnin. 2017. Someone Like Me: How Does Peer Parity Influence Participation of Women on Stack Overflow? *Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing - VLHCC'17*: 239–243. https://doi.org/10.1109/VLHCC.2017.8103473
[26]   Denae Ford, Justin Smith, Philip J. Guo, and Chris Parnin. 2016. Paradise Unplugged: Identifying Barriers for Female Participation on Stack Overflow. *Proceedings of the ACM Symposium on Foundations of Software Engineering - FSE'16*: 846–857. https://doi.org/10.1145/2950290.2950331
[27]   C. Ailie Fraser, Joy O. Kim, Alison Thornsberry, Scott Klemmer, and Mira Dontcheva. 2019. Sharing the Studio: How Creative Livestreaming can Inspire, Educate, and Engage. *Proceedings of the ACM Conference on Creativity and Cognition - C&C'19*: 144–155. https://doi.org/10.1145/3325480.3325485
[28]   Senhui Fu, Qing Yan, and Guangchao Charles Feng. 2018. Who Will Attract You? Similarity Effect Among Users on Online Purchase Intention of Movie Tickets in the Social Shopping Context. *International Journal of Information Management* 40: 88–102. https://doi.org/10.1016/j.ijinfomgt.2018.01.013

[29]   Giovanni Gardelli and Ingmar Weber. 2012. Why Do You Ask This? Using Toolbar Data to Identify Common Patterns of Q&A Users. *Proceedings of the 21st Annual Conference on World Wide Web Companion - WWW'12*: 815–822. https://doi.org/10.1145/2187980.2188205

[30]   Tamar Ginossar. 2008. Online Participation: A Content Analysis of Differences in Utilization of Two Online Cancer Communities by Men and Women, Patients and Family Members. *Health Communication* 23: 1–12. https://doi.org/10.1080/10410230701697100

[31]   Floraine Grabler, Maneesh Agrawala, Wilmot Li, Mira Dontcheva, and Takeo Igarashi. 2009. Generating Photo Manipulation Tutorials by Demonstration. *ACM Transactions on Graphics - TOG'09* 28, 3: 66:1-66:9. https://doi.org/10.1145/1531326.1531372

[32]   Kishonna L. Gray. 2012. Intersecting Oppressions and Online Communities: Examining the Experiences of Women of Color in Xbox Live. *Information Communication and Society* 15, 3: 411–428. https://doi.org/10.1080/1369118X.2011.642401

[33]   Charlotte N. Gunawardena and Frank J. Zittle. 1997. Social Presence as a Predictor of Satisfaction within a Computer-mediated Conferencing Environment. *International Journal of Phytoremediation* 21, 1: 8–26. https://doi.org/10.1080/08923649709526970

[34]   Cheng Guo and Kelly Caine. 2021. Anonymity, User Engagement, Quality, and Trolling on Q&A Sites. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1. https://doi.org/10.1145/3449215

[35]   William R. Hazlewood, Kevin Makice, and William Ryan. 2008. Twitterspace: A Co-developed Display Using Twitter to Enhance Community Awareness. *Proceedings of the 10th Anniversary Conference on Participatory Design: Experiences and Challenges - PDC'08*: 230–233. https://doi.org/10.1145/1795234.1795284

[36]   Elina H. Hwang, Param Vir Singh, and Linda Argote. 2015. Knowledge Sharing in Online Communities: Learning to Cross Geographic and Hierarchical Boundaries. *Organization Science* 26, 6: 1593–1611. https://doi.org/10.1287/orsc.2015.1009

[37]   Daniela Iosub, David Laniado, Carlos Castillo, Mayo Fuster Morell, and Andreas Kaltenbrunner. 2014. Emotions under Discussion: Gender, Status and Communication in Online Collaboration. *PLoS ONE* 9, 8. https://doi.org/10.1371/journal.pone.0104880

[38]   Christopher F. Karpowitz, Tali Mendelberg, and Lee Shaker. 2012. Gender Inequality in Deliberative Participation. *American Political Science Review* 106, 3: 533–547. https://doi.org/10.1017/S0003055412000329

[39]   Onur Kucuktunc, B. Barla Cambazoglu, Ingmar Weber, and Hakan Ferhatosmanoglu. 2012. A Large-Scale Sentiment Analysis for Yahoo! Answers. *Proceedings of the ACM Conference on Web Search and Data Mining - WSDM'12*: 633–642. https://doi.org/10.1145/2124295.2124371

[40]   Shyong (Tony) K. Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R. Musicant, Loren Terveen, and John Riedl. 2011. WP:Clubhouse? An Exploration of Wikipedia's Gender Imbalance. *Proceedings of the Symposium on Wikis and Open Collaboration - WikiSym'11*: 1–10. https://doi.org/10.1145/2038558.2038560

[41]   David Laniado, Yana Volkovich, Karolin Kappler, and Andreas Kaltenbrunner. 2016. Gender Homophily in Online Dyadic and Triadic Relationships. *EPJ Data Science* 5, 1. https://doi.org/10.1140/epjds/s13688-016-0080-6

[42]   Paul F. Lazarsfeld, Bernard Berelson, and Hazel Gaudet. 1968. *The People's Choice.* Columbia University Press, New York Chichester, West Sussex.

[43]   Paul F. Lazarsfeld and Robert K. Merton. 1954. Friendship as a Social Process: A Substantive and Methodological Analysis. *Freedom and Control in Modern Society* 18, 1: 18–66.

[44]   James M. Leonhardt, Todd Pezzuti, and Jae Eun Namkoong. 2020. We're Not So Different: Collectivism Increases Perceived Homophily, Trust, and Seeking User-Generated Product Information. *Journal of Business Research* 112, March: 160–169. https://doi.org/10.1016/j.jbusres.2020.03.017

[45]   Karen Levy and Solon Barocas. 2017. Designing Against Discrimination in Online Markets. *Berkeley Technology Law Journal* 32, 3: 1183–1238.

[46]   Olivier Liechti, Mark Sifer, and Tadao Ichikawa. 1999. A Non-Obtrusive User Interface for Increasing Social Awareness on the World Wide Web. *Personal and Ubiquitous Computing* 3, 1–2: 22–32. https://doi.org/10.1007/bf01305317

[47]   Guan-Yu Lin. 2004. Social Presence Questionnaire of Online Collaborative Learning: Development and Validity. *2004 Annual Proceedings - Chicago* 1: 588–591. Retrieved from http://files.eric.ed.gov/fulltext/ED499961.pdf

[48]   Mahya Maftouni, Patrick Marcel Joseph Dubois, and Andrea Bunt. 2022. "Thank You for Being Nice": Investigating Perspectives Towards Social Feedback on Stack Overflow. *Proceedings of Graphics Interface - GI'22.*

[49]   Lena Mamykina, Bella Manoim, Manas Mittal, George Hripcsak, and Björn Hartmann. 2011. Design Lessons from the Fastest Q&A Site in the West. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'11*: 2857–2866. https://doi.org/10.1145/1978942.1979366

[50]   Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2011. IP-QAT: In-Product Questions, Answers & Tips. *Proceedings of the ACM Symposium on User Interface Software and Technology - UIST'11*: 175–184.

[51]   Justin Matejka, Wei Li, Tovi Grossman, and George Fitzmaurice. 2009. CommunityCommands: Command Recommendations for Software Applications. *Proceedings of the ACM Symposium on User Interface Software and Technology - UIST'09*: 193–202. https://doi.org/10.1145/1622176.1622214

[52]   Anna May, Johannes Wachs, and Anikó Hannák. 2019. Gender Differences in Participation and Reward on Stack Overflow. *Empirical Software Engineering*: 1997–2019. https://doi.org/10.1007/s10664-019-09685-x

[53]   Linda L. McCroskey, James C. McCroskey, and Virginia P. Richmond. 2006. Analysis and Improvement of the Measurement of Interpersonal Attraction and Homophily. *Communication Quarterly* 54, 1: 1–31. https://doi.org/10.1080/01463370500270322

[54]   Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction - CSCW'19* 3, 39. https://doi.org/10.1145/3359174

[55]   Amanda Menking and Ingrid Erickson. 2015. The Heart Work of Wikipedia: Gendered, Emotional Labor in the World's Largest Online Encyclopedia. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'15*: 207–210. https://doi.org/10.1145/2702123.2702514

[56]   Amanda Menking, Ingrid Erickson, and Wanda Pratt. 2019. People Who Can Take It: How Women Wikipedians Negotiate and Navigate Safety. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI'19*: 1–15. https://doi.org/10.1145/3290605.3300702

[57]   Konstantinos Michos and Davinia Hernández-Leo. 2018. Supporting Awareness in Communities of Learning Design Practice. *Computers in Human Behavior* 85: 255–270. https://doi.org/10.1016/j.chb.2018.04.008

[58]   Heather Molyneaux, Susan O'Donnell, Kerri Gibson, and Janice Singer. 2008. Exploring the Gender Divide on YouTube: An Analysis of the Creation and Reception of Vlogs. *American Communication Journal* 10, 2: 1–14. https://doi.org/10.3168/jds.S0022-0302(91)95784-X

[59]   Savannah Morgan. 2017. How Are Programming Questions from Women Received on Stack Overflow? A Case Study of Peer Parity. *Proceedings Companion of the 2017 ACM SIGPLAN International Conference on Systems, Programming, Languages, and Applications: Software for Humanity - SPLASH'17*: 39–41. https://doi.org/10.1145/3135932.3135952

[60]   Quan Nguyen, Oleksandra Poquet, Christopher Brooks, and Warren Li. 2020. Exploring Homophily in Demographics and Academic Performance Using Spatial-Temporal Student Networks. *Proceedings of The 13th International Conference on Educational Data Mining - EDM'20*, Edm: 194–201. Retrieved from https://github.com/quan3010/EDM20_Nguyen

[61]   Jakob Nielsen. 2006. The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities. Retrieved January 11, 2022 from https://www.nngroup.com/articles/participation-inequality/

[62]   Heather L. O'Brien, Paul Cairns, and Mark Hall. 2018. A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form. *International Journal of Human Computer Studies* 112, July 2017: 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

[63]   Heather L. O'Brien, Emma Morton, Andrea Kampen, Steven J. Barnes, and Erin E. Michalak. 2020. Beyond Clicks and Downloads: A Call for a More Comprehensive Approach to Measuring Mobile-Health App Engagement. *BJPsych Open* 6, 5: 1–3. https://doi.org/10.1192/bjo.2020.72

[64]   Nigini Oliveira, Michael Muller, Nazareno Andrade, and Katharina Reinecke. 2018. The Exchange in StackExchange: Divergences between StackOverflow and its Culturally Diverse Participants. *Proceedings of the ACM Conference on Human-Computer Interaction - CSCW'18* 2, 30.

[65]   Sarah Pedersen and Caroline MacAfee. 2007. Gender Differences in British Blogging. *Journal of Computer-Mediated Communication* 12, 4: 1472–1492. https://doi.org/10.1111/j.1083-6101.2007.00382.x

[66]   Dan Perkel and Becky Herr-Stephenson. 2008. Peer Pedagogy in an Interest-Driven Community: The Practices and Problems of Online Tutorials. *Proceedings of the Media@lse Conference on Media, Communication and Humanity*: 1–30.

[67]   Quora. What is the Ratio of Male to Female Users on Quora? Retrieved September 13, 2019 from https://www.quora.com/What-is-the-ratio-of-male-to-female-users-on-Quora

[68]   Joseph Reagle. 2013. "Free as in Sexist?" Free Culture and the Gender Gap. *First Monday* 18, 1.

[69]   Joseph Reagle and Lauren Rhue. 2011. Gender Bias in Wikipedia and Britannica. *International Journal of Communication* 5: 21.

[70]   Everett M. Rogers and Dilip K. Bhowmik. 1970. Homophily-Heterophily: Relational Concepts for Communication Research. *Public Opinion Quarterly* 34, 4: 523–538. https://doi.org/10.1086/267838

[71]   Alfred P. Rovai. 2002. Building Sense of Community at a Distance. *The International Review of Research in Open and Distributed Learning* 3, 1. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/79/152

[72]   Alfred P. Rovai and Jason D. Baker. 2005. Gender Differences in Online Learning: Sense of Community, Perceived Learning, and Interpersonal Interactions. *The Quarterly Review of Distance Education* 6, 1: 31–44.

[73]   Răzvan Rughiniş, Alina Petra Marinescu-Nenciu, Ştefania Matei, and Cosima Rughiş. 2014. Computer-Supported Collaborative Questioning. Regimes of Online Sociality on Quora. *Iberian Conference on Information Systems and Technologies - CISTI'14*. https://doi.org/10.1109/CISTI.2014.6876946

[74]   Elaheh Sanoubari, Stela H. Seo, Diljot Garcha, James E. Young, and Verónica Loureiro-Rodríguez. 2019. Good Robot Design or Machiavellian? An In-the-Wild Robot Leveraging Minimal Knowledge of Passerby's Culture. *ACM/IEEE International Conference on Human-Robot Interaction - HRI'19*: 382–391. https://doi.org/10.1109/HRI.2019.8673326

[75]   Amy Soller, Renata Guizzardi, Alessandra Molani, and Anna Perini. 2004. SCALE: Supporting Community Awareness, Learning, and Evolvement in an Organizational Learning Environment. *Proceedings of the 6th International Conference on Learning Sciences - ICLS'04*: 489–496.

[76] Michael B. Spring and Vichita Vathanophas. 2003. Peripheral Social Awareness Information in Collaborative Work. *Journal of the American Society for Information Science and Technology* 54, 11: 1006–1013. https://doi.org/10.1002/asi.10305

[77] Stack Overflow. Stack Overflow Jobs. Retrieved April 2, 2020 from https://stackoverflow.com/jobs

[78] Stack Overflow. 2020. Saying Thanks: Testing a New Reactions Feature. Retrieved January 10, 2022 from https://stackoverflow.blog/2020/06/17/saying-thanks-testing-a-new-reactions-feature/

[79] Stack Overflow. 2020. Feature Test: Thank You Reaction. Retrieved January 10, 2022 from https://meta.stackoverflow.com/questions/398367/feature-test-thank-you-reaction

[80] Stack Overflow. 2021. Developer Survey Results 2021. Retrieved January 14, 2022 from https://insights.stackoverflow.com/survey/2021

[81] Statistics Canada. 2018. Major Field of Study - Classification of Instructional Programs (CIP) 2016 (432). *2016 Canadian Census.* Retrieved July 22, 2019 from https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/dt-td/Rp-eng.cfm?APATH=3&DETAIL=0&DIM=0&FL=A&FREE=0&GC=0&GID=0&GK=0&GRP=1&LANG=E&PID=111838&PRID=10&PTYPE=109445&S=0&SHOWALL=0&SUB=0&THEME=123&Temporal=2016&VID=0&VNAME E=&VNAMEF=

[82] Monica Stephens. 2013. Gender and the GeoWeb: Divisions in the Production of User-Generated Cartographic Information. *GeoJournal* 78, 6: 981–996. https://doi.org/10.1007/s10708-013-9492-z

[83] Katrin Talke, Søren Salomo, and Alexander Kock. 2012. Top Management Team Diversity and Strategic Innovation Orientation: The Relationship and Consequences for Innovativeness and Performance. *Strategic Direction* 28, 4: 1–11. https://doi.org/10.1108/sd.2012.05628daa.008

[84] W. Andrew Taylor. 2004. Computer-Mediated Knowledge Sharing and Individual User Differences: An Exploratory Study. *European Journal of Information Systems* 13, 1: 52–64.

[85] Mike Thelwall. 2009. Homophily in MySpace. *Journal of the American Society for Information Science and Technology* 60, 2: 219–231. https://doi.org/10.1002/asi

[86] Nava Tintarev and Judith Masthoff. 2007. A Survey of Explanations in Recommender Systems. *Proceedings of the International Conference on Data Engineering*: 801–810. https://doi.org/10.1109/ICDEW.2007.4401070

[87] Christoph Treude and Margaret-Anne Storey. 2010. Awareness 2.0: Staying Aware of Projects, Developers and Tasks using Dashboards and Feeds. *Proceedings of the International Conference on Software Engineering* 1: 365–374. https://doi.org/10.1145/1806799.1806854

[88] Francesca Tripodi. 2021. Ms. Categorized: Gender, Notability, and Inequality on Wikipedia. *New Media & Society*: 146144482110237. https://doi.org/10.1177/14614448211023772

[89] Bogdan Vasilescu, Andrea Capiluppi, and Alexander Serebrenik. 2012. Gender, Representation and Online Participation: A Quantitative Study of StackOverflow. *Proceedings of the Conference on Social Informatics*: 332–338. https://doi.org/10.1093/iwc/iwt047

[90] Claudia Wagner, David Garcia, Mohsen Jadidi, and Markus Strohmaier. 2015. It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *Proceedings of the AAAI Conference on Web and Social Media - ICWSM'15*: 454–463.

[91] Youcheng Wang and Daniel R. Fesenmaier. 2004. Modeling Participation in an Online Travel Community. *Journal of Travel Research* 42, 3: 261–270. https://doi.org/10.1177/0047287503258824

[92] Jie Wei, Stefan Seedorf, Paul Benjamin Lowry, Christian Thum, and Thimo Schulze. 2017. How Increased Social Presence Through Co-Browsing Influences User Engagement in Collaborative Online Shopping. *Electronic Commerce Research and Applications* 24: 84–99. https://doi.org/10.1016/j.elerap.2017.07.002

[93] Lei Xu, Tingting Nian, and Luís Cabral. 2020. What Makes Geeks Tick? A Study of Stack Overflow Careers. *Management Science* 66, 2: 587–604. https://doi.org/10.1287/mnsc.2018.3264

[94] Qiang Alex Zhao and John T. Stasko. 2002. What's Happening?: Promoting Community Awareness through Opportunistic, Peripheral Interfaces. *Proceedings of the Workshop on Advanced Visual Interfaces - AVI'02*: 69–74. https://doi.org/10.1145/1556262.1556271

[95] Elijah Zolduoarrati and Sherlock A. Licorish. 2021. On the Value of Encouraging Gender Tolerance and Inclusiveness in Software Engineering Communities. *Information and Software Technology* 139, June: 106667. https://doi.org/10.1016/j.infsof.2021.106667

[96] Do StackOverflow Users get Job Offers? *Meta Stack Exchange.* Retrieved April 2, 2020 from https://meta.stackexchange.com/questions/122725/do-stackoverflow-users-get-job-offers/134855

[97] Will High Reputation in Stack Overflow Help to Get a Good Job? *Software Engineering Stack Exchange.* Retrieved April 2, 2020 from https://softwareengineering.stackexchange.com/questions/20407/will-high-reputation-in-stack-overflow-help-to-get-a-good-job