

Voice of the Users: A Demographic Study of Software Feedback Behaviour

James Tizard¹, Tim Rietz², Kelly Blincoe¹

¹University of Auckland, New Zealand

²Karlsruhe Institute of Technology, Germany

jtiz003@aucklanduni.ac.nz, tim.rietz@kit.edu, k.blincoe@auckland.ac.nz

Abstract—User feedback on mobile app stores, product forums, and on social media can contain product development insights. There has been a lot of recent research studying this feedback and developing methods to automatically extract requirement-related information. This feedback is generally considered to be the “voice of the users”; however, only a subset of software users provide online feedback. If the demographics of the online feedback givers are not representative of the user base, this introduces the possibility of developing software that does not meet the needs of all users. It is, therefore, important to understand who provides online feedback to ensure the needs of underrepresented groups are not being missed.

In this work, we directly survey 1040 software users about their feedback habits, software use, and demographic information. Their responses indicate that there are statistically significant differences in who gives feedback on each online channel, with respect to traditional demographics (gender, age, etc). We also identify key differences in what motivates software users to engage with each of the three channels. Our findings provide valuable context for requirements elicited from online feedback and show that considering information from all channels will provide a more comprehensive view of user needs.

I. INTRODUCTION

Software users give large volumes of feedback online about the products they are using [1], [2]. Previous research has found that software feedback in app stores, social media, and on product forums contains valuable product development insights that can be used to guide the evolution of the software being discussed [2], [3], [4]. This feedback is generally considered to be the “voice of the users” [5]. There has been a recent surge of requirements engineering research developing ways to make this feedback more accessible to product development teams (e.g. [3], [6], [7], [8]).

However, only a subset of software users provide online feedback. If online feedback is being used to drive product development decisions, the concerns and desires of only the vocal users are being considered. If the demographics of the vocal users are not representative of the overall set of users, this introduces the possibility of developing biased software that does not meet the needs of all users. Therefore, it’s important to understand which software users do give online feedback and in doing so identify groups whose views may be underrepresented.

Yet, very little research has investigated who is giving online feedback for software products with respect to the demographics of the users. This may be due to the fact that demographic information of feedback givers is not readily available. On some feedback channels, even the full name of the person providing the feedback is unavailable.

Some preliminary studies have investigated the gender and geographic location of users providing feedback on app stores [5], [9]. These studies found that men were more likely than women to provide feedback on the Apple app store. However, these results are obtained by approximating gender based on usernames, since actual gender identity of the feedback givers is not available on app stores.

In this work, we overcome the online data sparsity problem by directly surveying 1040 software users about their software use, feedback habits, and their demographic information. We collect demographic information of feedback givers on three popular channels: app stores, product forums and social media. By directly collecting demographic information, we were able to examine feedback habits across a wide range of demographics, including gender, age, education, and ethnicity. We also investigated what motivates feedback givers and if their software usage habits relates to their feedback giving habits.

Our study was guided by the following research questions:

RQ1: *What are the demographics of software users who report to give online written feedback?*

RQ2: *What motivates software users to give online feedback and are there differences across demographics?*

RQ3: *Does the likelihood of giving online written feedback vary based on the type of software used and the duration of software usage?*

The contributions of this paper are insights about who gives online feedback and what motivates them. Specifically: (1) We show that there are differences in the feedback habits of software users based on traditional demographics. For gender, men reported giving more written feedback than women. With age, distinct patterns emerged with respondents between 35 and 45 reporting to give the most written feedback on all channels. (2) We show that user groups have different motivations to give feedback and these motivations vary across each of the three feedback channels. Respondents also reported differences in the success of in-app prompts between eliciting app ratings and written feedback, and differences in the frequency individual feedback givers write on app stores, product forums and social media. (3) We present evidence that software users feedback habits also vary with respect to the way they use software. Respondents who spend more hours each day on their phone or computer report giving more written feedback about the software they’re using. The software platform being used also has a relationship to written feedback rates, with Linux (computer) and Android

(phone) users reporting to give more feedback than those using other platforms.

Our findings give valuable insights into which software users give online feedback and what motivates them to do so. Additionally, the findings emphasise the need to mine all three feedback channels in order to get the most representative requirements from the user base when leveraging online user feedback.

The paper is structured as follows: Section II reviews the related work that informed our research. In section III, we present our research methodology. The results are presented in section IV and discussed in section V, including a discussion on the threats to validity. Finally, section VI concludes the paper.

II. RELATED WORK

A. User feedback in requirements engineering

Researchers have found requirements relevant information in feedback on several prominent online channels, including: app stores, social media, and product forums [2], [3], [4]. These channels can contain large volumes of valuable information. Pagano and Maalej [2] found that approximately a third of user reviews on app stores contain information related to software requirements. User feedback that contains bug reports or feature requests (and more) can be used by developers to address the needs and desires of their users which is critical to the ongoing success of their software.

Manually eliciting software requirements from online feedback can be extremely time intensive due to the large volumes and varying quality of text language that comes from highly distributed user bases [10]. Much recent research has investigated methods to automatically mine requirements in user feedback on app stores, Twitter and product support forums [1], [3], [6], [7], [8], [11], [12].

B. Demographics of software user who give feedback

There has been limited research to understand which software users give online feedback and what motivates them. Guzman et al. [5] looked at the difference between men and women who give feedback on the Apple app store. They manually approximated the gender of each person leaving a written review based on their username. They found a slight majority (57%) of the reviews were written by men. However, there were differences in this ratio when geographic region was considered. For example, in India 83% of feedback givers were men. In Australia, women wrote the majority (67%) of the reviews. They did not find any statistically significant differences in review sentiment, content, and rating between genders.

Another study investigated differences in feedback from the Apple app stores of eight countries [9]. This study found that feedback characteristics such as sentiment, content, rating, and length significantly varied between the countries.

These studies were both limited to the Apple app store. In addition, since demographic information like gender is not available for app store users, gender was only approximated and other demographics like age could not be studied. This study provides a more thorough investigation into demographics of feedback givers across three prominent

types of online feedback channels (app stores, product forums, and social media).

III. METHODOLOGY

To better understand which software users give online feedback and what motivates them, we ran an online questionnaire. This enabled us to survey a large number of software users (1040) to determine if reported feedback habits and motivations differed across demographics.

A. Survey Design

The survey consisted of 24 multiple-choice questions, shown in Table I. The survey consisted of five main sets of questions. The first three sets of questions asked about the feedback the participant provides in the three feedback channels under investigation: app stores (Q1-5), social media (Q6-9), and product forums (Q10-13). The remaining two sets of questions collect software usage information (Q15-18) and demographic information (Q19-24). Descriptions of what was meant by app store and product forum feedback were given within the survey to help participants understand the question context. Questions eliciting details on feedback habits were asked before software usage and demographic questions to highlight the propose of the study and to maintain participant interest.

The sets of questions on the three feedback channels each follow the same general format. First, the participant is asked if they have given feedback on that channel. Next, if applicable, they are asked how frequently they give feedback, the type of feedback given (e.g. reporting a bug), and their motivation for providing feedback on this channel. These questions were all multiple choice. The answer options for the type of feedback provided and the motivation for providing feedback were based on findings from recent research studies on each of these feedback channels. The participants were also asked about their perceptions on the impact of their feedback on influencing changes in the software products (Q14). For some questions, participants could select more than one answer choice (e.g. motivation for giving feedback). The full list of questions and answer choices is shown in Table I. Abbreviated answers for each question are given in the table, an unabbreviated copy of the survey can be found on Zenodo¹.

The software usage questions asked participants how they interact with software products including the types of devices they use, the types of software they use, and the hours spent on devices each day. The answer choices for the types of software were obtained from the categories of apps on popular app stores.

The demographic questions collected information on the participants age, gender, ethnicity, education, and employment. These questions and their associated answer choices were informed by traditional marketing demographic categories [13] as well as the New Zealand census (2018) [14].

B. Ethics Approval

The survey had ethics approval from the University of Auckland's Human Participants Ethics Committee.

¹<https://zenodo.org/record/3674076#.XkxNFygzZPY>

TABLE I
SURVEY QUESTIONS

| Question | RQ | Topic | Question | Answer Source |
|------------|-----|--|---|----------------------------|
| Q1. | All | App store | What review types have you given to mobile apps in the past? (choose all that apply) (None / Prompted rating / Prompted written review / Direct rating / Direct written review) | - |
| Q2. | RQ2 | App store | How many times have you given mobile apps you use a star rating in the last year? (None / 1-4 times / 5-12 times / 13-26 times / 27-52 times / 53 or more times) | - |
| Q3, 7, 11. | All | App store (Q3), Product forum (Q7), Social media (Q11) | How many times have you written (or given a review) on this channel in the last year? (None / 1-4 times / 5-12 times / 13-26 times / 27-52 times / 53 or more times) | - |
| Q4, 8, 12. | RQ1 | App store (Q4), Product forum (Q8), Social media (Q12) | What types of posts (or reviews) have you written about software (or apps)? (choose all that apply) (Praise (all channels) / Report bug (all channels) / Request feature (all channels)) / Ask a question (all channels) / Recommend to others (app stores, social media) / Dissuade others (app stores, social media) / Discuss shortcoming (app stores, social media) / Dispraise or criticise (app store, product forum) / Discuss a helpful situation (app stores) / Discuss specific feature (app stores) / Assist others (product forums) / Other, please specify (all channels)) | Q4[2], Q8[3], Q12[4] |
| Q5, 9, 13. | RQ2 | App store (Q5), Product forum (Q9), Social media (Q13) | What was your motivation(s) to write on this channel in the past? (choose all that apply) (Show appreciation / Show dissatisfaction / Influence improvement / Recommend / Discourage others / Connect or socialise about software / No specific motivation / Other, please specify) | Q5[2], Q9[3], Q13[4] |
| Q6. | All | Product forums | How have you used software product forums in the past? (choose all that apply) (I haven't / Reading and viewing / Written posts) | - |
| Q10. | All | Social media | Have you used social media (E.g. Twitter, Facebook) to discuss software products you are using? (choose all that apply) (I haven't / Reading and viewing / Written posts) | - |
| Q14. | RQ2 | App store, Product forum, Social media | How likely do you think it is for an app/software product to change based on your online reviews? (Definitely will / Probably will / Might or might not / Probably won't / Definitely won't) | [15] |
| Q15. | RQ3 | Software usage | What type of mobile phone do you currently use? (choose all that apply) (iPhone / Android (E.g. Samsung, Pixel) / I don't use a mobile phone / Other, please specify) | - |
| Q16. | RQ3 | Software usage | What type of computer do you currently use? (choose all that apply) (Windows / Mac (Apple) / Linux / I don't use a computer / Other, please specify) | - |
| Q17. | RQ3 | Software usage | How many hours per day do you use your phone? (Less than 1 hour / 1-4 hours / 4-8 hours / More than 8 hours) | - |
| Q18. | RQ3 | Software usage | How many hours per day do you use your computer? (Less than 1 hour / 1-4 hours / 4-8 hours / More than 8 hours) | - |
| Q19. | RQ1 | Demographics | Do you work or have you previously worked in the software industry? (No / I work or have worked in software / Other, please specify) | - |
| Q20. | RQ1 | Demographics | How old are you? (Under 18 years old / 18-24 years old / 25-34 years old / 35-44 years old / 45-54 years old / Over 55 years old) | [14] |
| Q21. | RQ1 | Demographics | What is your gender? (Man / Woman / Prefer not to say / Prefer to self-specify (please specify)) | [14] |
| Q22. | RQ1 | Demographics | What is your ethnicity? (White (European) / Asian / Pacific people / African / Middle Eastern / Latin American / Other, please specify) | [14] |
| Q23. | RQ1 | Demographics | What is your highest level of education completed? (Secondary school / Post secondary, Vocational training / 1-2 year tertiary education / Bachelor degree (3-4 years) / Master degree (postgraduate), Doctoral (postgraduate) / Other, please specify) | [16] |
| Q24. | RQ1 | Demographics | What is your current employment status? (Employed full-time (> 40 hours) / Employed part-time (< 40 hours) / Currently unemployed / Student / Retired / Self-employed / Unable to work / Other, please specify) | [14] |

TABLE II
RESPONDENT DEMOGRAPHICS

| Demographic Type | Group | Number of Respondents | Percentage of Respondents |
|-------------------|------------------------|-----------------------|---------------------------|
| Gender | Men | 571 | 54.90 |
| | Women | 454 | 43.65 |
| | Other | 16 | 1.54 |
| Age | Under 18 years | 61 | 5.87 |
| | 18 - 24 years | 571 | 54.90 |
| | 25 - 34 years | 285 | 27.40 |
| | 35 - 44 years | 50 | 4.81 |
| | 45 - 54 years | 29 | 2.79 |
| | Over 55 | 44 | 4.23 |
| Ethnicity | White/European | 790 | 75.96 |
| | Asian | 149 | 14.33 |
| | Middle Eastern | 26 | 2.50 |
| | Latin American | 24 | 2.31 |
| | Pacific and Maori | 18 | 1.73 |
| | African | 7 | 0.67 |
| | Other | 27 | 2.60 |
| Education | Secondary school | 411 | 39.52 |
| | Vocational Training | 14 | 1.35 |
| | 1-2 year Tertiary | 62 | 5.92 |
| | Bachelor degree | 390 | 37.50 |
| | Master degree | 129 | 12.40 |
| | Doctoral degree | 25 | 2.40 |
| | Other | 9 | 0.87 |
| Employment | Full time (> 40 hours) | 215 | 20.67 |
| | Part time (< 40 hours) | 78 | 7.50 |
| | Student | 644 | 61.92 |
| | Self-employed | 28 | 2.69 |
| | Currently unemployed | 39 | 3.75 |
| | Retired | 15 | 1.44 |
| | Unable to work | 4 | 0.38 |
| | Other | 18 | 1.73 |

C. Recruiting Participants

To recruit participants, convenience sampling was used [17]. This was chosen as the best method to engage a good number of survey participants in a reasonable time period. The possible sources of bias from our sampling methodology are discussed in section V-A. Survey participation was advertised via several avenues (described below) and incentivised with the chance to win a \$200/€120 cash prize. The survey was primarily made available online through the Qualtrics survey platform [18].

A link to the Qualtrics survey was shared on Facebook and Twitter by the authors (and their colleagues). In addition, we recruited from a pool of university participants using the hroot software [19]. The pool includes nearly 3500 participants who registered online to be invited to and participate in scientific studies, either on-site or online. This pool is mainly advertised at the Karlsruhe Institute of Technology, so the pool contains primarily students between the ages of 18 and 30. Through hroot, 2570 participants were invited. Hardcopies of the survey were also distributed in public areas of Auckland city, during December 2019. The completed hardcopy survey responses were manually consolidated with the online survey responses. The survey was open to anyone 16 years or older.

TABLE III
USER FEEDBACK WITH AGE, SIGNIFICANCE TESTS

| | App Store | | Product Forums | | Social Media | |
|------------------------------|---------------|--------------|----------------|-------------------|---------------|-------------------|
| | Chi2 | p | Chi2 | p | Chi2 | p |
| <i>Under 18 < 18 - 24</i> | 4.96 | 0.026 | 5.79 | 0.016 | 0.017 | 0.896 |
| <i>Under 18 < 25 - 34</i> | 3.600 | 0.058 | 11.419 | 0.001 | 0.165 | 0.685 |
| <i>Under 18 < 35 - 44</i> | 11.760 | 0.001 | 17.087 | < 0.001 | 8.264 | 0.004 |
| <i>18 - 24 < 25 - 34</i> | 0.27 | 0.603 | 9.04 | 0.003 | 0.936 | 0.333 |
| <i>18 - 24 < 35 - 44</i> | 5.603 | 0.018 | 12.183 | < 0.001 | 26.509 | < 0.001 |
| <i>25 - 34 < 35 - 44</i> | 6.570 | 0.01 | 2.169 | 0.141 | 14.214 | < 0.001 |
| <i>35 - 44 > 45 - 54</i> | 0.997 | 0.318 | 0.049 | 0.825 | 1.90 | 0.168 |
| <i>45 - 54 > Over 55</i> | 0.571 | 0.450 | 0.437 | 0.508 | 0.010 | 0.919 |
| <i>35 - 44 > Over 55</i> | 5.487 | 0.019 | 0.770 | 0.380 | 4.815 | 0.028 |

Note: statistically significant results are bolded

D. Survey Participants

Across all collection channels, 1040 participants fully completed the survey. All respondents reported having used software on computer or mobile, therefore all respondents are software users. The make up of the survey respondents with respect to gender, age, ethnicity, education and employment are shown in Table II.

Regarding the highest level of education obtained, we noticed that many respondents reported secondary school (411) and bachelor degree (390). Given the hroot software recruited from a pool of university participants, we suspected education level could be associated with the age of the participants. We saw that 90.02% of secondary school educated reported to be under 25, compared to only 41.61% of those who have higher education. After controlling for age, we did not see any significant differences in feedback habits in regard to education level. Thus, we do not report results considering education level.

E. Survey Analysis

To answer our research questions, we analysed the ratio of respondents in each user group (based on demographics or software usage) that reported a particular behaviour, e.g. giving feedback on a particular feedback channel or having a certain motivation. Chi-squared tests, which tests for differences in proportion between two groups [20], were used to find if differences in reported behaviours between user groups are statistically significant.

IV. RESULTS

A. Demographics

RQ1: What are the demographics of software users who report to give online written feedback?

In this section we present the percentage of written feedback givers in each demographic group.

Feedback across online channels: Overall, 30.96% of survey respondents reported having written feedback, on any of the three online channels. The most survey respondents reported having written feedback on app stores (18.16%), then on product forums (13.45%) and least on social media (7.11%). The majority of feedback giving respondents gave feedback to only one channel (77.64%), 19.57% had written on two channels, with 2.80% writing on all three (Fig. 1). A Chi-squared test showed the higher rate of respondents using only one feedback channel over multiple channels is statistically significant ($p < 0.001$).

Age: Under 18's, reported to have given the least feedback of all ages, across all channels (*app store 6.6%*,

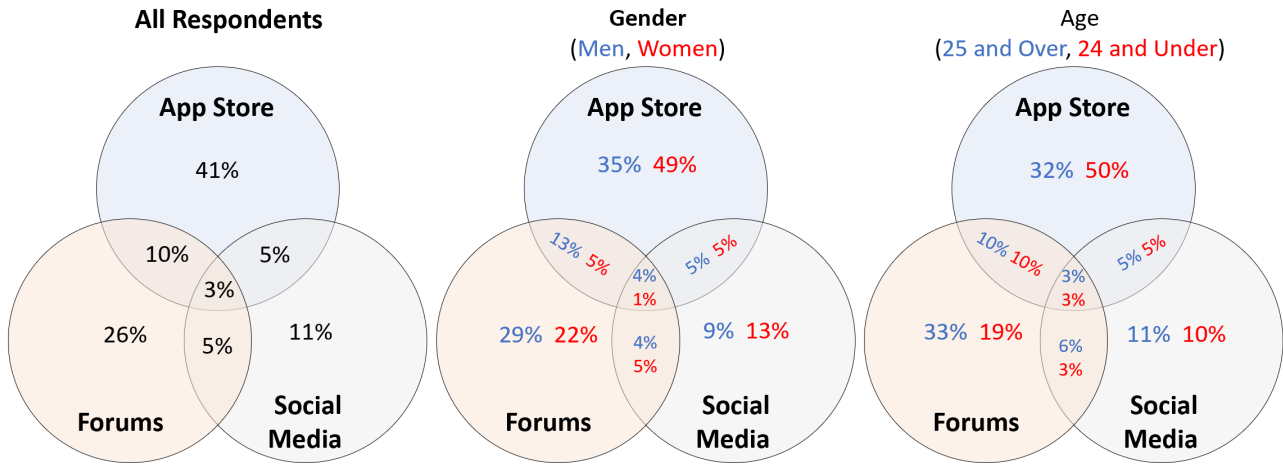


Fig. 1. Overview of the proportion of feedback givers who use each online channel

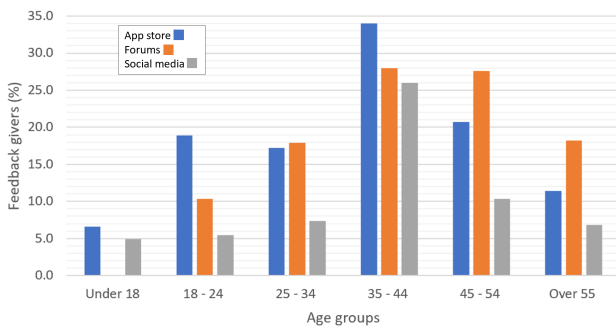


Fig. 2. User feedback with age

TABLE IV
COMPARING APP STORE AND FORUM FEEDBACK WITH AGE

| | App Store (%) | Product Forums (%) |
|-----------------------|---------------|--------------------|
| Under 25 years old | 17.72 | 9.34 |
| 25 years old and over | 18.87 | 19.85 |

TABLE V
USER FEEDBACK WITH GENDER

| | Number of Respondents | App Store (%) | Product Forums (%) | Social Media (&) |
|-------|-----------------------|---------------|--------------------|------------------|
| Men | 571 | 20.32 | 18.04 | 8.23 |
| Women | 454 | 14.54 | 8.15 | 5.73 |

forums 0.0%, social 4.9%) (Fig. 2). Conversely, 35-45 year old's (50 respondents), reported to give the most feedback across all channels (app store 34.0%, forums 28.0%, social 26.0%). Chi-squared tests show there are statistically significant differences between ages (shown in Table III).

Also of note, respondents under 25 preferred to give feedback to the app store over product forums, shown in bold in table IV. Under 25's preference for app stores was shown to be statistically significant using a chi-squared test ($p < .001$). Respondents 25 and over used app stores and forums more equally, with those over 44 reporting more forum use. However, the differences in channel use for those 25 and above was not found to be significant.

Gender: Men, reported to give more feedback than woman, across all channels, shown in table V. On apps

TABLE VI
USER FEEDBACK TYPE WITH GENDER

| Feedback Type | App Store (%) | | Forums (%) | | Social Media (%) | |
|------------------------|---------------|--------------|--------------|--------------|------------------|--------------|
| | Men | Women | Men | Women | Men | Women |
| Praise | 41.38 | 50.00 | 20.39 | 10.81 | 38.30 | 30.77 |
| Report bug | 40.52 | 48.48 | 73.79 | 56.76 | 46.81 | 42.31 |
| Request feature | 26.72 | 18.18 | 32.04 | 21.62 | 27.66 | 38.46 |
| Ask question | 2.59 | 6.06 | 88.35 | 94.59 | 68.09 | 65.38 |
| Recommend to others | 12.96 | 16.67 | NA | NA | 36.17 | 19.23 |
| Dissuade others | 10.34 | 6.06 | NA | NA | 8.51 | 11.54 |
| Discuss shortcomings | 47.41 | 36.36 | NA | NA | 46.81 | 34.62 |
| Dispraise or criticise | 18.10 | 15.15 | 16.50 | 8.11 | NA | NA |
| Helpful situation | 36.21 | 27.27 | NA | NA | NA | NA |
| Discuss feature | 21.55 | 22.73 | NA | NA | NA | NA |
| Assist others | NA | NA | 55.34 | 21.62 | NA | NA |

stores, 20.3% of men and 14.5% of women reported to give feedback. On product forums, 18.0% of men and 8.1% of women reported to give feedback. On social media, the difference was the smallest, with 8.2% and 5.7% respectively reporting to give feedback. Chi-squared tests showed that the difference between men and women respondents was statistically significant for app stores ($p=0.02$) and product forums ($p<0.001$).

Men and women respondents reported some differences in the types of feedback they give on all three feedback channels, shown in Table VI. Differences of note are indicated in bold in the table. On app stores, more women feedback givers reported praising apps, than feedback giving men ($w: 50%$, $41.38%$) and also reported giving bug reports. More men reported describing a situation an app was helpful, reported an app short coming and request a new features.

On product forums, both genders were very likely to ask

TABLE VII
USER FEEDBACK WITH EMPLOYMENT TYPE

| | Number of Respondents | Forums (%) | Under 25 years (%) |
|---------------------------|-----------------------|------------|--------------------|
| Full-time | 215 | 21.40 | 20.93 |
| Part-time | 78 | 12.82 | 66.67 |
| Student | 644 | 10.87 | 78.57 |
| Full-time (no under 25's) | 170 | 22.94 | 0.00 |
| Part-time (no under 25's) | 26 | 11.54 | 0.00 |
| Students (no under 25's) | 138 | 18.84 | 0.00 |

TABLE VIII
FEEDBACK OF SOFTWARE PROFESSIONALS

| | Number of Respondents | App Store (%) | Forums (%) | Social Media (%) |
|------------------------|-----------------------|---------------|------------|------------------|
| Software Professionals | 171 | 27.49 | 19.88 | 12.87 |
| Other Respondents | 869 | 16.32 | 12.18 | 5.98 |

a question about software, with 88.35% of men feedback givers and 94.59% of women. Men feedback givers were more likely to give other types of feedback, including: report a problem, request a feature, give praise, give criticism and assist others. On social media, more men reported recommending software to others and discussing short comings. More women reported requesting new features.

Employment: Respondents working full time reported using product forums at a higher rate than those working part time and students (Table VII). However, there is a strong association between employment level and age as 78.57% of students are also under 25. In the bottom half of table VII, all under 25 year old respondents were removed from the analysis, showing the difference between employment levels is not as large when considering only older respondents. The feedback differences between employment groups were not found to be statistically significant, using chi-squared tests, after the exclusion of the under 25 year old respondents.

Software professionals: Respondents who work, or have worked in software (software professionals), reported to have given feedback at a higher rate than those who have not worked in software, on all channels (Table VIII). Chi-squared tests showed that the feedback rate difference between software professionals and other respondents was significant on all channels (app stores: $p=0.001$, product forums: $p=0.01$, social media: $p=0.002$).

Ethnicity: The majority of survey respondents were either Caucasian (790) or Asian (149), which limited our findings with respect to ethnicity. However, the ethnic demographics of the respondents are representative for a study based in New Zealand and Germany. Only the difference between Caucasian and Asian feedback rate could be investigated and this difference was not statistically significant on any channel.

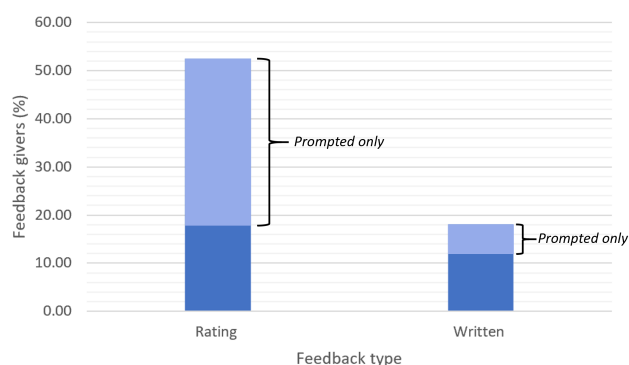


Fig. 3. Impact of in app prompts

Answer to RQ1: There are statistically significant differences in the amount of written feedback given by software users with respect to traditional demographics. For gender, men reported giving more feedback than women, on all three feedback channels. The types of feedback men and women reported to give also varied in notable ways. With age, distinct patterns emerged with respondents between 35 and 45 reporting to give the most feedback and under 18's reporting to give the least, on all channels. Additionally, software professionals reported giving significantly more feedback than other respondents.

B. Motivations

RQ2: What motivates software users to give online feedback and are there differences across demographics?

The section presents our findings with respect to what motivates users to give online feedback. The difference in motivations across the three channels and between groups are given.

Overall: The reported motivations to give feedback on app stores, product forums and social media are given in table IX. The motivations are given as a percentage of all users who give written feedback on each channel. Multiple motivations could be given by each respondent. As can be seen, the motivations vary across feedback channels. Show appreciation for software was the most common cited motivation on app stores (65.15%) and Social media (56.76%). Get help with software was the top motivation to post on product forums (70.37%). Influencing improvement was also a prominent motivation, being the third most cited on all channels.

Mobile app prompts (Fig 3): 52.45% of all survey respondents reported having previously given a star rating to an app. Of those who have given a star rating, 65.75% only gave the rating when prompted within the app, never directly on the app store. 18.16% of respondents reported having given a written review to an app. Of those who have given a written review, 31.75% only gave a written review when prompted to within the app.

Gender: Some differences in motivations to give feedback were reported between men and women. The percentage of men and woman feedback givers who cited

TABLE IX
MOTIVATIONS TO GIVE FEEDBACK

| App Store | (%) | Product Forum | (%) | Social Media | (%) |
|---------------------------|-------|---------------------------|-------|---------------------------|-------|
| 1. Show appreciation | 65.15 | 1. Get help | 70.37 | 1. Show appreciation | 56.76 |
| 2. Influence improvement | 52.02 | 2. Influence improvement | 44.29 | 2. Influence improvement | 51.35 |
| 3. Show dissatisfaction | 34.85 | 3. Show appreciation | 26.43 | 3. Show dissatisfaction | 37.84 |
| 4. Recommend to others | 29.80 | 4. Recommend to others | 17.86 | 4. Connect or socialise | 35.14 |
| 5. Discourage others | 12.63 | 5. Show dissatisfaction | 16.43 | 5. Recommend to others | 32.43 |
| 6. Get help | 9.20 | 6. Connect or socialise | 15.72 | 6. Get help | 22.73 |
| 7. No specific motivation | 5.05 | 7. No specific motivation | 7.86 | 7. Discourage others | 14.86 |
| 8. Connect or socialise | 1.52 | 8. Discourage others | 3.57 | 8. No specific motivation | 8.11 |

TABLE X
MOTIVATIONS TO GIVE FEEDBACK WITH GENDER

| Motivation | App Store (%) | | Product Forums (%) | | Social Media (%) | |
|------------------------|---------------|-------------|--------------------|--------------|------------------|--------------|
| | Men | Women | Men | Women | Men | Women |
| Show appreciation | 67.24 | 72.73 | 28.16 | 18.92 | 57.45 | 53.85 |
| Show dissatisfaction | 36.21 | 36.36 | 13.59 | 21.62 | 31.91 | 46.15 |
| Influence improvement | 57.76 | 50.00 | 49.51 | 27.03 | 55.32 | 42.31 |
| Recommend | 29.79 | 34.62 | 32.76 | 30.30 | 18.45 | 13.51 |
| Discourage | 16.38 | 6.06 | 1.94 | 5.41 | 10.64 | 19.23 |
| Connect/ socialise | 4.31 | 7.58 | 13.59 | 18.92 | 27.66 | 42.31 |
| Get help | 10.14 | 5.88 | 71.11 | 77.78 | 41.18 | 0.00 |
| No specific motivation | 0.86 | 3.03 | 9.71 | 0.00 | 4.26 | 11.54 |

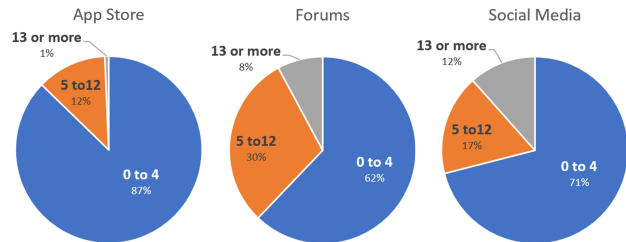


Fig. 4. Feedback given by individual users each year, on each channel

each motivation are shown in table X, where notable differences are indicated in bold. On app stores men were more motivated to discourage others from using a disliked app. On product forums, more men cited influencing an improvement in the software as a motivation. On social media, more women were motivated to show dissatisfaction and connect or socialise about a software product. Also on social media, more men cited influence improvement and get help. These results are bolded in Table X.

Feedback frequency: The majority of feedback givers reported having given feedback between 0 and 4 times in the last year, across all channels (Fig. 4). App stores had the least respondents reporting to give more than 4 pieces of feedback, product forums had the most respondents giving feedback more than 4 times.

Perception of influencing developers: Survey respondents who believed that software developers would *definitely not* be influenced by online feedback, were less likely to give feedback than those who believed influence

TABLE XI
USER FEEDBACK WITH PERCEPTION OF INFLUENCING DEVELOPERS

| | Number of Respondents | App Store (%) | Forums (%) | Social media (%) |
|---------------------|-----------------------|---------------|-------------|------------------|
| Definitely will | 83 | 14.46 | 18.07 | 7.23 |
| Probably will | 265 | 19.25 | 13.96 | 7.92 |
| Might or might not | 416 | 18.75 | 14.90 | 6.49 |
| Probably will not | 248 | 18.95 | 9.68 | 8.06 |
| Definitely will not | 27 | 3.70 | 7.41 | 0.00 |

was more likely, on all channels. However, chi-squared tests showed these differences weren't statistically significant. Feedback rates with perception of influencing developers are shown in table XI, the lower *definitely not* values are indicated in bold.

Answer to RQ2: Showing appreciation was the top motivation given to write feedback on app stores and social media. On product forums, getting help was the most commonly cited motivation (table IX). Differences in the motivations of men and women to give written feedback on each channel were also reported.

In-app prompts were reported to be very effective at motivating app users to give star ratings, but less effective at eliciting written feedback. Individual survey respondents reported engaging with each feedback channel at different frequencies, writing on product forums the most times a year and least on app stores.

C. Type of software and duration of use

RQ3: Does the likelihood of giving online written feedback vary based on the type of software used and the duration of software usage?

iPhone/Android: Android users reported giving feedback to the app store at a higher rate than iPhone users (Table XII). 13.48% of iPhone users reported having given written feedback on app stores compared to 21.84% of Android users. A chi-squared test showed this difference to be statistically significant, given in table XIII. The higher feedback rate of Android users on app stores has been indicated in bold in table XII.

Windows/Mac/Linux: Linux users reported giving written feedback on app stores and product forums at a higher

TABLE XII
USER FEEDBACK WITH DEVICE TYPE

| Device | Number of Respondents | App Store (%) | Product Forum (%) | Social Media (%) |
|---------|-----------------------|---------------|-------------------|------------------|
| Android | 618 | 21.84 | 13.75 | 6.80 |
| iPhone | 423 | 13.48 | 12.77 | 7.33 |
| Linux | 94 | 31.91 | 26.60 | 10.64 |
| Windows | 759 | 19.10 | 14.6 | 6.46 |
| Mac | 275 | 16.73 | 12.36 | 10.18 |

TABLE XIII
USER FEEDBACK WITH DEVICE TYPE, SIGNIFICANCE TESTS

| | App Store | | Product Forums | |
|------------------|---------------|--------------|----------------|--------------|
| | Chi2 | p | Chi2 | p |
| Android > iPhone | 11.144 | 0.001 | 0.087 | 0.769 |
| Linux > Windows | 7.651 | 0.006 | 8.073 | 0.004 |
| Linux > Mac | 8.974 | 0.003 | 9.531 | 0.002 |

Statistically significant results are bolded

TABLE XIV
USER FEEDBACK WITH DAILY COMPUTER USE

| Daily Computer Use | Number of Respondents | App Store (%) | Forums (%) | Social Media (%) |
|--------------------|-----------------------|---------------|--------------|------------------|
| Less than 1 hour | 109 | 18.35 | 10.09 | 6.42 |
| 1 - 4 hours | 436 | 15.14 | 9.40 | 5.96 |
| 4 - 8 hours | 363 | 20.66 | 17.08 | 7.99 |
| More than 8 hours | 110 | 21.82 | 23.64 | 9.09 |

rate than Windows and Mac users (Table XII). Chi-squared tests showed these differences to both be statistically significant (Table XIII). The difference between Windows and Mac users feedback was not statistically significant. The statistically significant higher feedback rates of Linux users are indicated in bold in table XII.

Hours of computer use: Respondents who reported a higher daily computer use (hours), were more likely to give feedback to product forums, indicated in bold in table XIV. The least forum feedback was given by respondents using their computer less than 1 hour or between 1 and 4 hours a day. Those using their computer between 4 and 8 hours gave more feedback and those using their computer more than 8 hours a day gave at the highest rate. Chi-squared tests showed that the feedback rate differences between 1 - 4 hours and 4 - 8 hours and between 1- 4 hours and over 8 hours were statistically significant (Table XV).

Hours of phone use: Respondents who reported a higher daily phone use (hours), were more likely to give feedback to social media, indicated in bold in table XVI. However, chi-squared tests showed these differences weren't statistically significant.

TABLE XV
COMPUTER DAILY USE, SIGNIFICANCE TESTS (PRODUCT FORUMS)

| Daily Computer Use | Chi2 | p |
|------------------------------|---------------|-------------------|
| Less than 1 hour < 1-4 hours | 0.001 | 0.971 |
| Less than 1 hour < 4-8 hours | 2.619 | 0.106 |
| 1 - 4 < Over 8 hours | 15.233 | < 0.001 |
| Less than 1 < Over 8 hours | 6.221 | 0.013 |
| 1-4 hours < 4-8 hours | 9.722 | 0.002 |
| 4-8 hours < Over 8 hours | 1.983 | 0.159 |

Statistically significant results are bolded

TABLE XVI
USER FEEDBACK WITH DAILY PHONE USE

| Daily Phone Use | Number of Respondents | App Store (%) | Forums (%) | Social Media (%) |
|-------------------|-----------------------|---------------|------------|------------------|
| Less than 1 hour | 52 | 15.38 | 15.38 | 3.85 |
| 1 - 4 hours | 664 | 16.57 | 13.70 | 6.63 |
| 4 - 8 hours | 266 | 22.93 | 11.65 | 7.89 |
| More than 8 hours | 51 | 17.65 | 17.65 | 13.73 |

Answer to RQ3: Statistically significant differences were reported in the amount of written feedback given based on the type of software used and the duration of daily use. Respondents who spend more hours each day on their computer reported giving more written feedback to product forums. Those using the Linux OS gave more written feedback to app stores and product forums than those using Windows and Mac. Android users reported giving more written feedback to app stores than iPhone users.

V. DISCUSSION

In this section, we first discuss the threats to validity of our study and then describe the implications and potential avenues for future work.

A. Threats to validity

Convenience sampling, used in this study to elicit survey participants, is a non-probabilistic sampling method and a possible source of bias [17]. The target population of this study are users of software and mobile applications. Survey participants were engaged via Facebook, Twitter, the Karlsruhe Institute of Technology survey pool and in Auckland cities public areas, therefore only a subset of the target population had the opportunity to participate. Additionally, all respondents who completed the survey were self-selected, and their feedback habits may not generalise to all software users.

To mitigate this bias, we collected data from a large number of software users (1040 participants). Recruitment was done through three channels; social media, public spaces, and the hroot survey pool to increase the odds of recruiting a diverse set of respondents. However, we cannot claim that our results generalise outside of our sample. Future studies can replicate our survey to validate our findings.

Our participants are not representative across all demographics. The demographics of our respondents are listed in Table II. The majority of participants are white/European and many are students. When presenting our results, we present proportions based on the total number of respondents in each demographic group. We also used chi-squared tests to determine significance between different demographic groups, which accounts for the sample size of each population being compared. Therefore, findings found to be statistically significant had a sufficient number of respondents in each demographic to satisfy the test. However, due to a low number of participants in some demographic groups, not all demographics could be analysed. Future studies should replicate this survey to enable analysis of additional demographics.

In addition, the majority of participants identified as men or women. We did give participants the option to self-specify gender, but very few participants chose to self-specify. Thus, our analysis was limited to only the differences between participants who identified as men and women. Again, future studies should replicate the survey to enable analysis beyond these binary genders.

Our results are based only on self-reported feedback habits. Demographic information of feedback givers is not readily available on the feedback channels we investigated (often the real name of the writer isn't even given). This data sparsity problem means our findings can't be directly validated against actual feedback data. One previous study, by Guzman et al. [5], approximated the gender of feedback givers on app stores from their usernames. Using these approximations, they found that men were more likely than women to provide feedback on the Apple app store, which is in line with what our respondents reported and supports our findings.

B. Implications and Future Work

Implication 1: *The findings presented in this paper suggest that, when leveraging online user feedback, to get the most representative user views and desires, feedback from multiple feedback channels should be considered.* We found statistically significant differences in the users who reported to give feedback on app stores, product forums, and social media with respect to traditional demographics and software usage habits. For example, older respondents prefer product forums to app stores, while younger respondents prefer app stores.

Importantly, a majority of feedback giving respondents reported only engaging with one of these three feedback channels. This indicates that considering multiple channels will enable feedback from a more diverse set of users.

We also found key differences in what motivates software users to engage with each of the three channels. The most cited motivation on app stores and social media was to show appreciation for the app/software. Whereas, on product forums showing appreciation was much less of a motivating factor, instead, getting help was the top cited motivation. Showing dissatisfaction, recommending and discouraging others were also significantly more cited on app stores and social media. On social media, connecting with other users was reported to be a more common motivation than on the other channels.

These motivation differences suggest that the feedback on each online channel is likely to contain different product development insights. For example, feedback on product forums contain users trying to get help and therefore likely describes ways the software is unintuitive or difficult to use. On app stores and social media, users are more motivated to communicate how they feel about the software/app to the developers and other users. These differences again emphasize the benefit to considering feedback from all channels, as each channel may provide unique insights.

Implication 2: *How to elicit feedback from under-represented groups needs further study.* We saw some demographics were less likely to give feedback than others. For example, respondents 35-44 years old report to provide the most feedback on all three feedback channels, while both older and younger respondents gave less feedback (Fig 2). Also, men reported giving feedback at a higher rate than women across all three channels. This is in line with the results of Guzman et al. which found that the Apple app store had more feedback from men [5]. Future work should investigate why some user groups give more feedback than others to understand how to better elicit feedback from these underrepresented groups. This work could survey underrepresented user groups on why they don't give feedback, so that these factors can be directly addressed. Changes to the feedback channels themselves may even be considered to make the tool more inclusive. Recent research found that most software has gender inclusivity issues [21], so it is possible that similar inclusivity issues exist in the software that collects online feedback.

Additionally, studies could be performed to see if other means of collecting feedback, that augments written feedback on the feedback channels we studied, could be used to help provide a more comprehensive set of feedback across demographics. For example, recent research proposed using a conversational agent (ladderbot) to conduct short requirements interviews to elicit user feedback [22]. Future research could evaluate whether some demographics would prefer other user feedback collection methods (like ladderbot).

Implication 3: *Feedback prompts are effective at eliciting feedback for app stores and may be effective if applied more widely in computer software. However, more work is needed to understand how to prompt users to give detailed feedback.* Mobile apps widely use prompts to elicit feedback. More survey respondents reported giving written feedback on app stores than on any other channel. Much of this feedback is prompted. The number of respondents who have provided unprompted app store feedback (12.39%) is very similar to the number who report to have written posts on product forums (13.45%). This suggests that the prompts are successful in eliciting additional feedback givers. The prompts are even more effective at eliciting app ratings, which take less time to provide than written feedback. Future research could study whether prompts could be successful in collecting other types of user feedback and how they could be integrated into other feedback channels.

Implication 4: *The types of software devices respondents use also has an association to feedback habits. Investigating why users of some devices give more feedback may give insights into how to motivate and facilitate feedback.*

On phones, more Android users give written feedback than iPhone users. It is not clear why there are differences in feedback across devices, but it may be influenced by differences in prompt rates, app quality, app store usability or even those who choose to use each phone type. iOS developers could benefit in understanding these factors in order to encourage more feedback from their users.

On computers, respondents who use the Linux OS more commonly had given written feedback to app stores and product forums than those who do not use Linux. The feedback habits of Mac and Windows users were relatively similar across all feedback channels. The higher feedback rates of Linux users may be related to the prevalence of software developers using it. In fact, 43% of respondents using Linux also reported working in the software industry, compared to only 16% of all respondents. Our results showed that software professionals are more likely to provide online feedback, possibly because they understand how that feedback will be used by development teams. Future research can investigate more thoroughly the reasons for differences across devices.

Other avenues for future work: *Investigate other feedback channels.* Our study was limited to app stores, product forums, and social media. Future work could perform a similar investigation considering other feedback channels like issue trackers.

Replicate survey in other countries. Our survey respondents were mostly from two countries; New Zealand and Germany. Future work could replicate our survey by eliciting responses in additional countries. This would also enable analysis at the ethnicity level if more ethnic diversity in the participants was achieved.

Understand gender differences in product forum engagement. In addition to men being more likely than women to post on product forums, men also reported using product forums for different reasons. While men and women both primarily used forums to ask software related questions, men also reported higher rates of giving other types of feedback on product forums, including: reporting problems, requesting features, praising and criticising the software, as well as assisting others. Further research is needed to understand the gender difference in engagement with product forums.

Making missing demographics more transparent. Currently, it is difficult for product development teams to know whether the feedback collected from online feedback channels is biased and misses the voices of some underrepresented groups. Future research could devise ways to make this more transparent to enable software development teams to more proactively consider the needs of the underrepresented groups and produce more inclusive software.

VI. CONCLUSION

The online user feedback written on app stores, product forums, and social media is a valuable source of requirements for software developers and has been a focus of requirements engineering researchers. However, limited studies have been done to understand which software users give this feedback and what motivates them. In this work, we directly surveyed 1040 software users about

their feedback habits, software use, and demographic information.

The responses indicate significant differences in the demographics of software users who give feedback on each online channel. For gender, men reported giving more feedback than women, and with age respondents between 35 and 45 reporting to give the most feedback across all channels. We also found strong evidence that younger software users (under 25), prefer to engage with app stores whereas older software users use product forums at equal (to app stores) and sometimes higher rates.

We identified key differences in what motivates software users to engage with each of the three channels. Comparing channels, respondents reported the top motivation to give feedback on app stores and social media was to show appreciation, whereas on forums the most cited motivation was to get help with software products. Differences between the motivations of men and women to give feedback were also reported for each of the channels. Respondents reported in app prompts to be significantly more effective in motivating them to give app ratings over written feedback. Additionally, individual feedback givers reported to engage more times a year on product forums than on with app stores.

Differences in feedback habits were also reported with the ways respondents use software. Those who spend more hours each day on their phone or computer reported giving more feedback about the software they're using. The software platform being used also presented a relationship to feedback rates, with more Linux (computer) and Android (phone) users reporting to give feedback than those who use the alternatives.

The findings presented in this paper give meaningful insights into which software users give online feedback and the motivations they have to give it. We found notable differences in those who give feedback to each online channel, which emphasises the need to mine all three feedback channels to get the most representative requirements from software users when leveraging online feedback.

REFERENCES

- [1] E. Guzman, R. Alkadhi, and N. Seyff, "A needle in a haystack: What do twitter users say about software?" in 2016 IEEE 24th International Requirements Engineering Conference (RE), Sep. 2016, pp. 96–105.
- [2] D. Pagano and W. Maalej, "User feedback in the appstore: An empirical study," in 2013 21st IEEE International Requirements Engineering Conference (RE), July 2013, pp. 125–134.
- [3] J. Tizard, H. Wang, L. Yohannes, and K. Blincoe, "Can a conversation paint a picture? mining requirements in software forums," in 2019 IEEE 27th International Requirements Engineering Conference (RE). IEEE, 2019, pp. 17–27.
- [4] E. Guzman, R. Alkadhi, and N. Seyff, "An exploratory study of twitter messages about software applications," Requirements Engineering, 07 2017.
- [5] E. Guzman and A. P. Rojas, "Gender and user feedback: An exploratory study," in 2019 IEEE 27th International Requirements Engineering Conference (RE). IEEE, 2019, pp. 381–385.
- [6] E. Guzman, M. Ibrahim, and M. Glinz, "A little bird told me: mining tweets for requirements and software evolution," in 2017 IEEE 25th International Requirements Engineering Conference (RE). IEEE, 2017, pp. 11–20.
- [7] W. Maalej and H. Nabil, "Bug report, feature request, or simply praise? on automatically classifying app reviews," in 2015 IEEE 23rd International Requirements Engineering Conference

- (RE), vol. 00, Aug. 2015, pp. 116–125. [Online]. Available: doi.ieeecomputersociety.org/10.1109/RE.2015.7320414
- [8] A. D. Sorbo, S. Panichella, C. V. Alexandru, C. A. Visaggio, and G. Canfora, “Surf: Summarizer of user reviews feedback,” in 2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C), May 2017, pp. 55–58.
- [9] E. Guzman, L. Oliveira, Y. Steiner, L. C. Wagner, and M. Glinz, “User feedback in the app store: a cross-cultural study,” in 2018 IEEE/ACM 40th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS). IEEE, 2018, pp. 13–22.
- [10] E. C. Groen, N. Seyff, R. Ali, F. Dalpiaz, J. Doerr, E. Guzman, M. Hosseini, J. Marco, M. Oriol, A. Perini et al., “The crowd in requirements engineering: The landscape and challenges,” IEEE software, vol. 34, no. 2, pp. 44–52, 2017.
- [11] S. Panichella, A. Di Sorbo, E. Guzman, C. A. Visaggio, G. Canfora, and H. C. Gall, “Ardoc: App reviews development oriented classifier,” in Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering, ser. FSE 2016. New York, NY, USA: ACM, 2016, pp. 1023–1027. [Online]. Available: <http://doi.acm.org/10.1145/2950290.2983938>
- [12] N. Chen, J. Lin, S. C. H. Hoi, X. Xiao, and B. Zhang, “Ar-miner: Mining informative reviews for developers from mobile app marketplace,” in Proceedings of the 36th International Conference on Software Engineering, ser. ICSE 2014. New York, NY, USA: ACM, 2014, pp. 767–778. [Online]. Available: <http://doi.acm.org/10.1145/2568225.2568263>
- [13] N. Papadopoulos, O. M. Martín, M. Cleveland, and M. Laroche, “Identity, demographics, and consumer behaviors,” International Marketing Review, 2011.
- [14] “New zealand census, 2018,” <https://www.stats.govt.nz/information-releases/2018-census-population-and-dwelling-counts>, accessed: December 2019.
- [15] R. Likert, “A technique for the measurement of attitudes,” Archives of psychology, 1932.
- [16] U. ISCED, “International standard classification of education 2011,” 2012.
- [17] I. Etikan, “Comparison of convenience sampling and purposive sampling,” American Journal of Theoretical and Applied Statistics, vol. 5, p. 1, 01 2016.
- [18] “Qualtrics survey platform,” <https://www.qualtrics.com>, accessed: December 2019.
- [19] O. Bock, A. Nicklisch, and I. Baetge, “hroot: Hamburg registration and organization online tool,” in WiSo-HH Working Paper Series, 2012.
- [20] M. L. McHugh, “The chi-square test of independence,” Biochemia medica: Biochemia medica, vol. 23, no. 2, pp. 143–149, 2013.
- [21] M. Burnett, S. Stumpf, J. Macbeth, S. Makri, L. Beckwith, I. Kwan, A. Peters, and W. Jernigan, “Gendermag: A method for evaluating software’s gender inclusiveness,” Interacting with Computers, vol. 28, no. 6, pp. 760–787, 2016.
- [22] T. Rietz and A. Maedche, “Ladderbot: A requirements self-elicitation system,” in 2019 IEEE 27th International Requirements Engineering Conference (RE). IEEE, 2019, pp. 357–362.