Investigating Country Differences in Mobile App User Behavior and Challenges for Software Engineering

Soo Ling Lim, Peter J. Bentley, Natalie Kanakam, Fuyuki Ishikawa, and Shinichi Honiden

Abstract—Mobile applications (apps) are software developed for use on mobile devices and made available through app stores. App stores are highly competitive markets where developers need to cater to a large number of users spanning multiple countries. This work hypothesizes that there exist country differences in mobile app user behavior and conducts one of the largest surveys to date of app users across the world, in order to identify the precise nature of those differences. The survey investigated user adoption of the app store concept, app needs, and rationale for selecting or abandoning an app. We collected data from more than 15 countries, including USA, China, Japan, Germany, France, Brazil, United Kingdom, Italy, Russia, India, Canada, Spain, Australia, Mexico, and South Korea. Analysis of data provided by 4,824 participants showed significant differences in app user behaviors across countries, for example users from USA are more likely to download medical apps, users from the United Kingdom and Canada are more likely to be influenced by price, users from Japan and Australia are less likely to rate apps. Analysis of the results revealed new challenges to market-driven software engineering related to packaging requirements, feature space, quality expectations, app store dependency, price sensitivity, and ecosystem effect.

Index Terms—Requirements/specifications, market-driven software engineering, mobile application development, user requirements, survey research, app user behavior, software product lines, software ecosystems

✦

1   INTRODUCTION

MOBILE apps are software applications developed for use on mobile devices such as smartphones and tablets. Once developed, an app is sold via an application distribution platform, commonly known as an app store. App development is market-driven. Similar to traditional market-driven software [1], [2], the requirements for an app are usually derived from strategic business goals or from market opportunities. During the development of an app, developers have limited contact with potential users. Success is measured by the number of downloads and revenues generated from the app. The app store concept has democratized the software industry—almost anyone can build and sell apps to a worldwide population of users via app stores.

The benefits of app stores come with significant challenges. App developers face a crowded and highly competitive app market, and as a result, an app can fail (receive little or no downloads) due to features unrelated to its functionality and usability, such as app name, app icon or level of exposure. As the profit margins from app sales are small (Section 1.2), an app should ideally appeal to a large number of users worldwide in order to be successful. However, many developers are unaware that users from different countries have different behavior and needs, and that these factors affect app downloads.[1] There is also a lack of awareness about the importance of features such as app description, screenshots, pricing, and user feedback. These challenges have caused many apps to fail. Studies have found that 400,000 out of 600,000 apps in the iOS App Store have no downloads,[2] and 80 percent of paid Android apps received less than 100 downloads [3].

Despite these failures, app development continues to accelerate worldwide. Market-driven software engineering has been studied in the past [4], [5], [6], but today researchers are increasingly focusing on the new opportunities and challenges of app development. Recent studies have made advances in our understanding of app user behaviors through mining app store data, gathering user activity logs and surveys (e.g., [7], [8], [9]). These provide useful data relating to specific smartphones, app stores, apps, app categories (e.g., medical apps), countries, or age groups. However to date there has been little research that studies global user behaviors in different app stores and mobile devices, comparing

- S.L. Lim is with the Department of Computer Science, University College London, The National Institute of Informatics, Japan, and the Software Systems Research Centre, Bournemouth University.
  E-mail: s.lim@cs.ucl.ac.uk.
- P.J. Bentley is with the Department of Computer Science, University College London. E-mail: p.bentley@cs.ucl.ac.uk.
- N. Kanakam is with the Department of Clinical, Education and Health Psychology, University College London. E-mail: n.kanakam@ucl.ac.uk.
- F. Ishikawa is with the Digital Content and Media Sciences Research Division, National Institute of Informatics, Japan.
  E-mail: f-ishikawa@nii.ac.jp.
- S. Honiden is with the National Institute of Informatics, Japan.
  E-mail: honiden@nii.ac.jp.

1. http://www.guardian.co.uk/technology/appsblog/2012/dec/04/ios-android-revenues-downloads-country
2. http://www.pcadvisor.co.uk/news/mobile-phone/3373458/400000-ios-app-store-apps-have-no-downloads/

across countries. In this work we complement previous research by focusing on this important area.

1.1 Contributions

This work makes the following contributions:

- We conducted one of the largest surveys to date of mobile app users worldwide, in terms of questionnaire extent, participant number, and country coverage. Our questionnaire investigated user adoption of the app store concept, their app needs, and their rationale for selecting or abandoning an app, as well as the differences in user behaviors across countries. We surveyed 10,208 participants from more than 15 countries, including the United States of America, China, Japan, Germany, France, Brazil, the United Kingdom, Italy, Russian Federation, India, Canada, Spain, Australia, Mexico, and Republic of Korea. We anticipate that this extensive dataset will form a valuable resource for the fields of application software development, human-computer interaction, and mobile computing and we regard this as a major contribution of our work.

- We analyzed the data and identified clear evidence that there exist country differences in user app behavior, where some, but not all, of these differences can be correlated with known cultural differences between countries. The analysis was conducted using well-established statistical measures such as the Pearson correlation coefficient, linear regression, Pearson's chi-square test, and odds ratio. The large dataset enables our findings to be statistically significant.

- From analysis of our results and comparison with the market-driven software engineering literature, we identified new challenges and their corresponding implications for software engineering research.

1.2 Motivation

App development is now a mainstream form of software engineering. Just as the growth of web development resulted in every organization requiring its own webpages, today every organization requires its own apps. Major software companies such as IBM, Oracle and Accenture are providing mobile application development services and support.[3,4,5] The result is unprecedented growth and competition. For example, in January 2013, Apple's iOS (mobile operating system) App Store had more than 200,000 app developers,[6] 700,000 apps, and 1,000 new apps per day. A keyword search for "to do list" on 18 Jan 2013 returned more than 1,000 apps offering the feature. With so much competition, developers may lose downloads due to "packaging" features such as the app's icon, name, or description in the app store [10].

Apps often cost between $35,000 and $200,000 to develop,[7,8,9] and one study reported that almost 70 percent of developers earned on average a total revenue of $5,000 to date or less due to small margins (e.g., the profit of an app priced at $0.99 has to be shared between the app store and the developer).[10] It is not surprising that 80 percent of developers reported generating insufficient revenue to support their business.[10] Some failures are very costly. For example, a $41 million project to develop an app that allows users to share live video broadcasts and photos with their friends was abandoned due to insufficient users and a high churn rate.[11,12] Media attention received by the app attracted downloads, but users found the app did not meet their needs and was difficult to use, and therefore abandoned the app.[11, 12]

Some developers who have success in one country find difficulty repeating the success in others.[13] As developers have limited contact with their users, it is difficult for them to identify target users and their needs. Although developers can receive feedback or feature requests from users via ratings and reviews, review rates are very low with many developers reporting a rate of less than 1 percent.[14,15] For example a developer reported 81 reviews out of 91,534 downloads (i.e., averaging one review per 1,130 downloads).[14] Subsequently, only successful apps that have been downloaded thousands of times have a chance of obtaining useful user feedback. Previous research has found cultural differences in organizations and information systems (e.g., Hofstede et al. [11], Straub et al. [12], van Everdingen [13]) between countries. Findings such as these have led us to form the hypothesis that differences may also exist in mobile app user behavior between countries. However, cultural and country differences in the context of mobile apps have yet to be investigated. Our research aims to provide evidence to support the hypothesis and also to identify the precise differences in app user behavior across countries.

The remainder of the paper is organized as follows. Section 2 provides a review of related literature. Section 3 describes the research questions, Section 4 describes the methodology used, and Section 5 provides the results. Section 6 analyzes the country differences using Hofstede's cultural index [11], and discusses the new challenges and their implications for software engineering research. Section 7 discusses threats to validity, and Section 8 concludes.

2 BACKGROUND

Existing research into understanding the needs of a large population of app users and their app user behavior can be

3. http://www-935.ibm.com/services/us/gbs/mobileconsulting/
4. http://www.accenture.com/us-en/Pages/service-custom-mobile-application-development.aspx
5. http://www.oracle.com/technetwork/developer-tools/adf-mobile/overview/index.html
6. An app developer can be a sole developer or a team of developers.

7. http://www.bluecloudsolutions.com/blog/cost-develop-app/
8. http://answers.oreilly.com/topic/2292-how-much-does-it-cost-to-develop-an-app/
9. http://www.padgadget.com/2010/10/17/the-cost-of-building-an-ipad-app/
10. http://app-promo.com/wp-content/uploads/2012/04/App-Promo-TheNecessityofMobileAppMarketing.pdf
11. http://www.digitaltrends.com/mobile/where-did-color-go-wrong/
12. http://mashable.com/2012/10/17/color-shuts-down/
13. Personal communication with a global app analytics company.
14. http://iphonedevsdk.com/forum/business-legal-app-store/54678-how-many-reviews-per-download-do-you-get-on-average.html
15. http://www.cocos2d-iphone.org/forum/topic/1231

categorized into those that mine app store data, those that collect activity logs from mobile devices, and those that conduct surveys and elicit feedback from users.

2.1 Mining App Store Data

App stores have accumulated a large amount of data, such as app descriptions, user ratings, and reviews. As such, an increasing number of studies to understand user needs are conducted by mining data from the app stores themselves. For example, Pagano and Maalej collected data on user ratings and reviews for the top 25 free and paid apps of one country on 16 September 2012 from each app category in the Apple iOS App Store [7]. They used various statistical measures to investigate how and when users provide feedback, as well as analyze the content of the reviews. Their results showed that most user reviews were provided shortly after new releases, with a quickly decreasing frequency over time. In addition, user reviews typically contain multiple topics, such as user experience, bug reports, and feature requests. The quality and constructiveness of user reviews vary widely, from helpful advices and innovative ideas to offensive comments [7].

Harman et al. mined the Blackberry app store for information such as app description, app category, user ratings, price and the rank of the app based on downloads [14]. The authors found a strong correlation between user ratings and app ranking, but no correlation seemed to be present between price and number of downloads. Their study focused on priced apps, further work may be necessary in order to corroborate the findings by taking free apps into consideration [14]. Chen and Liu mined the Apple iOS App Store and collected app information such as name, developer, category, current ranking, average rating, and number of ratings [15]. Their analysis revealed that the top-ranked paid apps are not necessarily closely correlated with user ratings, and their finding was consistent with that of Pagano and Maalej [7].

2.2 Activity Logs

A large number of studies about mobile app users have collected activity logs from mobile devices. For example, Do et al. collected data about app access, location, and Bluetooth from 77 Nokia Smartphone users over a duration of nine months [16]. They found that app usage depends on the users' location. For example, utility apps such as clocks are used most frequently at home, while camera and map apps are used most frequently on holiday. Participants who spend more time at a friend's home also use communication apps more [16]. Their study highlighted the need for developers to recognize the physical and social usage context of the apps they build. Xu et al. studied network traffic created by apps [17]. Their results indicated that news and weather apps are often used daily and at a certain time and suggested that developers could implement prefetching mechanisms in their apps to reduce latency perceived by users.

Falaki et al. collected app usage data from 255 Android and Windows Mobile users [18]. They found immense diversity among users, for example, the average number of smartphone interactions per user per day ranged from 10 to 200, and suggested that apps should adapt to different user groups. Bohmer et al. collected data related to the status

information of apps, such as installing, uninstalling, opening, and closing, from 4,125 Android users [8]. Their study revealed many interesting app usage patterns, for example, new applications are most popular in the morning and games are most popular at night. However, the participants in Bohmer et al.'s study were biased towards early adopters and frequent app users [19]. Although these studies collected considerable data about app usage, they have limited information about the participants themselves [8], and as a result, have difficulty achieving statistical control over potentially confounding variables [19].

A number of studies focus on gathering requirements for specific apps. For example, Henze et al. published five game apps in the Android market and monitored how the apps were used [20]. Their most popular app collected data from 6,907 users. Their data showed that many users abandoned the apps after a short period and they suggested that developers should focus on app quality and providing incentives to users in order to motivate long-term use of an app [20]. Henze et al. also found that most of their participants were English-speaking users from the United States, hence limiting their ability to derive conclusions about a global population [20].

In another study, McMillan et al. collected usage data of their iPhone app from 8,676 users over five months [21]. Data logging seemed to be a cost effective way to collect data from a large number of geographically dispersed users. However, activity logs were unable to provide an in-depth understanding of user behavior, and log analysis failed to reveal the users' needs and rationale behind their behavior [21]. In addition, the data was biased towards users who enjoyed the app because users who did not enjoy the app, stopped using it and were unavailable for data logging [21]. The researchers supported the activity logs with questionnaires to elicit feedback on app features and user demographics (e.g., age, gender, country of residence). They also interviewed users from a range of countries, but due to language barriers and difficulty engaging the users, they could only interview 10 users [21].

To provide a richer set of data about users, Rahmati et al. collected demographic information such as age and household income in addition to activity logs [19]. Their study was longitudinal over the period of a year, involving iPhone 3GS usage among 34 university students. Their study revealed the importance of understanding target users of an app. For example, participants with a lower household income used social networking apps such as Facebook and YouTube more than their peers. They also downloaded more apps, used them more frequently, but found them more difficult to use. In another study, Rahmati and Zhong conducted a four-month study of HTC Wizard phone usage from 14 teenagers in the United States [22]. Recreational applications were the most popular, and boredom caused gaming apps to loose popularity.

2.3 Surveys and User Feedback Elicitation

Surveys are one of the best tools to learn about large groups of users, their interests and their preferences [23]. When conducted effectively, surveys can produce a higher degree of certainty about the user's profile compared to indirect analysis of user behavior via activity logs [23]. For example, in

addition to activity logs from 117 users of Nokia N95 smartphones in Switzerland, Chittaranjan et al. also used a questionnaire to collect the users' demographic information (e.g., gender, age, nationality) and self-reported personality traits [24]. They found that extraverted participants are more likely to use office and calendar apps, and receive more calls on their smartphone [24]. Male participants were more likely to use game apps, while female participants who were introverted were more likely to use Internet apps [24].

Franko and Tirrell conducted an online survey to examine the app needs of 3,306 medical practitioners in the United States [9]. They collected and analyzed data related to the app store adoption by physicians (e.g., use of smartphones, use of apps in clinical practice), app needs (e.g., commonly used apps, desired app features), and demographics (e.g., medical specialty, level of training). Their results indicated that more than 85 percent of the participants owned a smartphone and 56 percent used apps in their clinical practice. They also found that the most useful features are drug guides, followed by medical calculators, coding and billing apps, and pregnancy wheels. Most importantly, there was a mismatch between physician needs and app availabilities. For example, although a large number of reference materials apps already exist in app stores, they remained the most requested types of apps by physicians since the existing apps were of insufficient quality. Merely importing all information from a textbook into an app does not provide the optimal user experience due to screen size or other restrictions. Many reference apps cost nearly as much as equivalent print versions. In order for an app to be successful in being commonly used by physicians, it must be easy to use and reasonably priced. Finally, information contained within those apps may not be based on validated or peer-reviewed information [9].

In order to gain a better understanding of development practices for mobile apps, Agrawal and Wasserman conducted a survey on app developers, using existing mobile developer forums to solicit respondents [25]. Their survey revealed that developers adhered quite well to recommended sets of "best practices" but rarely used any formal development processes. In addition, developers rarely tracked their development efforts in an organized manner and gathered few metrics. As mobile apps move from inexpensive recreational uses to complex business-critical applications, it will be essential to apply software engineering processes to assure the development of secure, high-quality software [25]. Wasserman proposed that while many software engineering techniques will transfer easily to the mobile apps domain, there are other areas for new research and development such as user experience, non-functional requirements, processes, tools, and architecture [25].

In the field of requirements engineering, Seyff et al. proposed using mobile devices to elicit end-user needs [26]. Using their proposed method, mobile phone users can document their needs and requirements using text entry, audio recordings, and images captured using their phone. Their evaluation revealed that end-users are able to document their needs without being facilitated by requirements analysts [26].

2.4 Summary

To summarize, existing research into app user behavior focus on a specific smartphone, app store, app, app category (e.g., medical apps), country, or age group. Large-scale studies using activity logging and data mining can reveal interesting usage patterns but not the rationale behind the patterns. In addition, they lack information related to user demographics (e.g., age, country of residence), which can be useful to understand the usage patterns. User studies collect detailed data and can reveal interesting insights but they often involve insufficient number of participants for the results to be generalizable. Most importantly, the data is derived from highly focused studies, which are not able to elucidate the usage of many types of app at an international scale. There is a need for more comprehensive data that is representative of app user needs in many countries, which may help improve user experience and improve software development practice for mobile apps.

3 RESEARCH QUESTIONS

Our research questions first establish a baseline in order to enable the discovery of country differences. This baseline focuses on user adoption of the app store concept, their app needs, and their rationale for selecting or abandoning an app. We then focus on the differences of these findings between countries. The research questions are listed as follows.

RQ1. How are users adopting the app store concept?
It is important to understand how best to develop apps and app stores such that users can find apps. In this research question we investigate user behavior relating to seeking apps, in terms of the platform used, frequency of use of that platform, frequency of downloads, and methods used to search for apps.

- RQ1.1 What is the distribution of users across mobile app platforms?
- RQ1.2 How frequently do users visit their app stores to look for apps?
- RQ1.3 On average, how many apps do users download per month?
- RQ1.4 How do users find apps?

RQ2. What needs are users trying to meet with apps?
In addition to the mechanics of finding apps, there are the fundamental needs of the users. In this question we aim to understand what might prompt a user to consider looking for an app in the first place, why they download apps, and which types of apps they prefer.

- RQ2.1 What triggers users to start looking for apps?
- RQ2.2 Why do users download apps?[16]
- RQ2.3 What types of apps do they download?

RQ3. What are the features of an app that influence its selection or abandonment?
Apps must be advertised through app stores, potentially making non-functional and packaging requirements as important as functional requirements. In this research

---

16. RQ2.1 focuses on the stimulus to launch an app store and look for apps; RQ2.2 focuses on triggers to download apps once browsing within the store. It is possible that some answers to RQ2.2 may apply to RQ2.1.

question we investigate the importance of app features versus descriptions, ratings, price, and perceived quality.

- RQ3.1 What are the factors that influence users' choices of apps?
- RQ3.2 Given that ratings influence app selection, why do users rate apps?
- RQ3.3 Why do users pay for apps?
- RQ3.4 Why do users stop using an app?

RQ4. How do the behaviors above vary across countries?

Here we revisit all the previous research questions with the aim of detecting differences across countries. Do users in different countries have different approaches to finding apps, or needs; are they influenced by different factors when they choose or abandon apps?

4 METHODOLOGY

This study used a survey to investigate the research questions. We constructed a questionnaire in order to collect quantitative data from app users. In order to provide a representative and generalizable view of mobile app user behavior, we targeted a large number of participants with varied demographics. Our survey focused on the top 15 GDP[17] countries. The targeted countries were the United States of America, China, Japan, Germany, France, Brazil, the United Kingdom, Italy, Russian Federation, India, Canada, Spain, Australia, Mexico, and Republic of Korea, sorted by decreasing GDP.[18] Due to the large coverage of participants, we employed an online survey in order make the survey more accessible. To understand the participants' background, we also used questions to elicit information about their demographics and personality.

4.1 Questionnaire Construction

The objective of this work is to understand user adoption of the app store concept, their app needs, and their rationale for selecting or abandoning an app and the differences across countries. To achieve the objective, we formulated survey questions to correspond to each of the research questions in Section 3. For example, for RQ1.1 (user distribution across mobile app platforms), we asked participants to specify the make, model name and number of the mobile device they use, as well as the app store they use. We used close-ended questions whenever possible because open-ended questions require much more effort from the respondents [23].

For each closed-ended question, we assembled a list of options gathered from the literature, our previous research, and our experiences as app users and app developers.[19] For example, for RQ1.1 (user distribution across mobile app platforms), we compiled a list of popular app platforms

including Apple, Google Play, Blackberry, Windows Phone. For RQ3.1 (factors that influence the choice of apps), we compiled a list of items the user can see in the screen of purchase, such as app icon, app description, star ratings, and screen shots. (Previous research has shown that quality of the icon influences the user's perception of app quality and their decision of whether to download [10].) We attempted to capture the full variety of human behavior including those that were previously unknown. Therefore we included an "Other (please specify)" option where applicable [27].

We worded our survey carefully in order to avoid any misunderstanding of the questionnaire. We used language that can be easily understood by participants from ages 12 and above, and used unambiguous words [27]. For example, as "developer" is not a common word, we substituted it with "person who developed the app." Technical or uncommon words were followed by examples. For instance, for the app category "Utilities" we provided examples of apps belonging to the category such as Calculator and Alarm Clock. When asking about how frequently users visit the app store, we provided quantifiable options such as, "once a day" or "once a month", rather than "frequently" or "rarely", which are subjective words.

We arranged the questions so as to engage the participants in the survey because participants who are interested are more likely to complete the survey and provide better quality data [23], [27]. For example, we grouped the questions thematically and arranged questions to have a natural progression [23], e.g., start from how users find apps, to what influences them when downloading apps, the amount they spend on apps, to why they rate apps, and why they stop using apps. We put demographics questions at the end because they are considered boring and could be construed as intrusive at the start of the survey [23].

To reduce response bias, we randomized the ordering of the answer choices for choices that do not need to be sorted in order (e.g., answers for the app store questions). This method reduces bias that may occur when respondents choose answers without reading all of the options [27]. In doing so, some options (such as "I don't rate apps" and "I do not pay for apps") remain the first option so that participants who do not do those things can quickly move on to the next question, and some options (such as "Other") remain the final option where people usually find them.

To ensure participants do not miss out any questions, the online questionnaire highlights missing answers and respondents cannot proceed until the missing answers are completed. We also used skip logic so that respondents do not see questions that are not relevant to them and respondents who indicate that they do not own a mobile device or their mobile device cannot run apps were screened out. Finally, we tested the questionnaire on common browsers, including Internet Explorer (v6 and above), Apple Safari (v3 and above), Mozilla Firefox (v4 and above), and Google Chrome (v2 and above).

4.2 Pilot Study

We recruited eight participants to pre-test the questionnaire in order to identify potential problems [28]. We selected the participants to reflect, as much as possible, the varied demographics of our target audience in terms of age (M = 31.75,

---

17. Gross domestic product (GDP) is the market value of all officially recognized final goods and services produced within a country in a given period of time. GDP per capita is often used an indicator of a country's standard of living, so people from higher GDP countries are more likely to be spending money on apps.

18. Data retrieved from the World Bank http://data.worldbank.org/data-catalog/GDP-ranking-table.

19. The second author developed an app that has received more than three million downloads (http://www.bbc.co.uk/news/11145583). The app was the number 1 top downloaded app in the UK iOS App Store in August 2010.

TABLE 1
Participant Feedback and Questionnaire Modification

Feedback	Modifications
RQ3.4 App abandonment. A participant reported that he stopped using many of his apps because he does not need them anymore, but this option was not available. RQ1.4 Finding apps. A participant reported that he finds apps using search engines, but the option was not available.	The options were added to the questionnaire.
The checklist questions constrained the number of options a participant can select. Several participants found the constraint very restrictive and counter-intuitive because many of the options applied equally well to them. Forcing them to choose between the options may result in frustration as well as missing data.	The constraints were removed such that participants can select all options that applied to them.
The draft questionnaire started with the personality question to draw the participants into the survey. However, two participants commented being confused about the focus of the survey.	The personality questions were moved to the end of the questionnaire.
Some participants did not understand the rationale behind the demographics and personality questions for a mobile app user survey. They felt that understanding the purpose behind the questions would motivate participants to provide better quality responses.	The purposes of our questions were explained in the questionnaire. For example, we explained that the demographic questions were "to ensure that we have surveyed a varied mix of people in society."
Some participants were uncomfortable providing sensitive information such as their ethnicity and household income.	"Prefer not to say" options were added to sensitive questions.

SD = 10.17), gender (Female = 3, Male = 5), and countries (the United Kingdom, Germany, Japan, China, and Australia). We asked the participants to complete the questionnaire and point out any problems they encountered. In particular, we asked them to (1) highlight ambiguous instructions, questions, and options, (2) identify missing questions and options to the survey questions, and (3) point out improvements we can make to the questionnaire in order to motivate potential respondents. Based on feedback from participants, we revised the questionnaire as summarized in Table 1.

We evaluated our revised questionnaire on four new participants. Feedback from all the participants was positive. Participants reported that the survey "was very engaging" and "very well designed."

4.3 Questionnaire Translation

The survey targets individuals from a variety of countries, ages, and background. As such, the questionnaire was translated into the first languages of the target countries in order to avoid misunderstanding and increase the accuracy of responses. The questionnaire was translated from English into nine other languages: Spanish, Korean, French, German, Japanese, Italian, Mandarin, Russian, and Portuguese. We selected our translators from native speakers of the language who were also proficient in English. Each translator was asked to use words that can be easily understood by an audience from ages 12 and above, and to ensure that the translated questionnaire matches the English questionnaire. Finally, we validated the translated questionnaires by asking a separate set of native speakers to trial the survey in each language.

4.4 Final Questionnaire

The final questionnaire had three sections and had 31 questions in total. The first section asked respondents about their user behavior in terms of mobile app usage, including the app stores they use, what triggers them to

look for apps, why they download apps, why they abandon apps, and the types of apps they download. The second section consisted of demographic questions in order to understand the types of people who responded to the survey. These questions asked about the respondent's gender, age, marital status, nationality, country of residence, first language, ethnicity, education level, occupation, and household income. The final section asked the respondents about their personality, using the Big-Five personality traits [29]. Finally, participants were asked to provide us with optional comments and their email addresses if they were interested to know the results. We also collected their browser and operation system information. An excerpt of the questionnaire can be seen in Fig. 1, and the complete questionnaire is available in the supplementary material of the paper, which can be found on the Computer Society Digital Library at: http://doi.ieeecomputersociety.org/10.1109/TSE.2014.2360674, and at: http://www.cs.ucl.ac.uk/research/app_user_survey/.

The online questionnaire was set to automatically default to the respondents' browser language, so that participants

Fig. 1. Excerpt of questionnaire (second page).

could answer the survey in the language that they were most comfortable with. Participants could also select their preferred language on each page of the questionnaire (Fig. 1).

4.5 Data Collection

Two methods were used for data collection: snowballing and online panels. The survey was conducted from the 26th of September 2012 to the 26th of November 2012. In the first method, we used the snowballing method (used in our previous research [30], [31]) to recruit participants. Specifically, we invited individuals in our social networks to complete the survey, and then asked them to invite individuals in their social networks to complete the survey, and so on. The following methods were used: emails to specific colleagues or friends, emails to mailing lists, posting the survey link on Twitter, Facebook, and LinkedIn.

The second method comprised the distribution of our survey to a panel of international participants provided by Cint,[20] an ISO certified panels company for conducting opinion and social research.[21] To achieve a representative sample of the target population, the panels used a random and stratified sampling technique, and enabled the recruitment of participants that is census representative.[22] Within the required targets, sample is randomly generated as well as being stratified by high, medium and low responders. A total of 32,491 panel members were recruited to participate in the survey.

4.6 Data Cleaning Approach

We used the following approach to clean our data. We focused on questions with an "Other (please specify)" option where participants provided textual answers, in order to codify their answers. We first translated each textual answer to English, and then coded all the translated responses into categories [32]. For example, for the question "Why do you rate apps?" The Spanish answer "para que los creadores las hagan funcionar mejor" was translated to English as "for creators to make them work better," and coded as "feedback to developers." We assigned the same code to other answers that when translated have the same meaning, e.g., "to provide feedback to the developers" and "to inform creators of defects in the app".

We then parsed the codes as follows. If the code duplicated an existing option in the same question, we merged it with the existing option, and removed the participants' selection of the "Other" option. (We found the majority of codes to fall in this category.) If the code duplicated an existing option in another question, we selected the option in the other question, and maintained the participants' selection of the "Other" option in the original question. If the code was new, but the number of answers sharing the same code was more than 5 percent, we created a new option for the

question, and participants were recoded to select the new option rather than "Other." If the code was new, but the number of answers sharing the same code was less than 5 percent, the participants remained selecting the "Other" option. This approach was used so that the "Other" option was the one with the fewest answers among all options [33]. Only the question "Which app store do you use?" had more than 5 percent with the same code. The original questionnaire and the questionnaire with the coded options are available in the supplementary material of the paper, available online, and at: http://www.cs.ucl.ac.uk/research/app_user_survey/.

Finally, for respondents who did not know their app store, we used the mobile phone specifications they provided in order to derive their app stores. For example, if their mobile phone is iPhone, we recoded their app store as Apple iOS App Store, because the iOS App Store is the most common and the only official app store used by iPhone users.

4.7 Data Analysis Techniques

We analyzed RQ1−3 using descriptive statistics. We also used parametric statistics to analyze the relationship between variables as follows. We used the Pearson correlation coefficient to analyze the relationship between users' age and other variables, such as whether they use search engines to find apps, or whether price influences their app choice, as well as frequency of app store visits and the average number of apps downloaded. Moderate sized correlations ($r > .5$) were followed up with linear regressions in order to assess whether one variable was a significant predictor of the other variable.

In RQ4 we revisited all previous research questions, analyzing them across countries. Direct comparisons were made for multiple-choice, single-answer questions (RQ1.1 to RQ1.3). We analyzed the data using Pearson's chi-square test (χ^2) for multiple-choice, multiple-answer questions (RQ1.4 onwards). Specifically, we used Pearson's chi-square test to analyze whether there were significant differences across countries for the categorical variables such as "compare several apps" or "browse randomly." A p value of less than 0.001 was used to determine variables that differed significantly across countries [34]. We measured the magnitude of the difference between each country and the other countries in the dataset combined using odds ratios [34]. For example, if country C has an odds ratio of R for behavior B, it means that users from country C are R times more likely to exhibit behavior B compared to users from the other countries.

All quantitative analyses were conducted using SPSS.[23] The results are presented using the APA standard [34].

5 RESULTS

Out of the 32,491 participants recruited from the panel, a total of 9,818 participants responded, and a further 390 participants responded from our snowballing method, resulting in a total of 10,208 participants who responded to our survey (96 percent panel, 4 percent snowballing method). The overall response rate was approximately 30 percent. This is similar to the highest response rate achieved for

---

TABLE 2
Countries and Response Rates from Panel

Panel Country	Recruited	Responded	Response Rate (%)
Australia	968	622	64.3
Brazil	5,350	707	13.2
Canada	3,650	1,075	29.5
China	4,507	811	18
France	965	715	74.1
Germany	760	612	80.5
India	1,388	479	34.5
Italy	810	362	44.7
Japan	2,350	1,439	61.2
Mexico	900	453	50.3
Rep. of Korea	5,350	371	6.9
Rep. of Russia	1,521	553	36.4
Spain	650	430	66.2
United Kingdom	810	518	64
United States	2,512	671	26.7
Total	**32,491**	**9,818**	**30.2**

A further 390 participants responded through the snowballing method.[26] For some participants, the panel country differed from the country of residence. In our analysis of different countries, we used country of residence provided by the participant in the demographics section of the questionnaire.

online surveys reported in Deutskens et al.[24] [35]. Table 2 provides the response rate for each country.

A total of 8,082 participants completed the survey (panel = 7,831, snowballing = 251). (We excluded incomplete surveys in our analysis.) A total of 3,258 participants were screened out because they did not use apps.[25] Only three participants provided bad data (e.g., garbage or obscenities) and were excluded from the analysis. Thus the final total comprised 4,824 participants (Male = 2,346 (49 percent), Female = 2,478 (51 percent), aged 11-87, average age = 34.51, standard deviation = 15.19). Fig. 2 shows the country of residence of the participants at the time of the survey. A total of 1,805 participants (37.4 percent) were interested to learn about the results of the survey and volunteered their contact details. The complete dataset is available in the supplementary material of the paper, available online, and at: http://www.cs.ucl.ac.uk/research/app_user_survey/.

24. In Deutskens et al.'s study of the response rate of online surveys with different configurations (e.g., short vs. long, donation to charity versus lottery incentive, early vs. late reminder), they found that the response rate ranged from 9.4 to 31.4 percent.

25. A total of 3,258 responses were screened out because the respondent did not own a mobile device (N = 1,208), the respondent's mobile device could not run apps (N = 1,653), and the respondent did not use apps (N = 394). These 394 participants selected "Other" for all app related questions and provided the explanation that they "do not use apps". This small percentage of participants may have completed the survey even though they believed that they did not use apps because panel users were rewarded only if they completed the survey, and their responses were valid and did not contain bad data. Among the 3,258 responses that were screened out, 48 were from our snowballing method (respondent did not own a mobile device (N = 18), respondent's mobile device could not run apps (N = 25), and respondent did not use apps (N = 5)).

26. The response rate for these participants could not be calculated because there is no way of knowing how many of our contacts saw our posts on Facebook or Twitter, and no way of knowing which of the contacts forward the link to their contacts and so on. In general, we found that asking individuals to complete the survey (e.g., via personal email) was more effective than posting the survey on Facebook or Twitter.
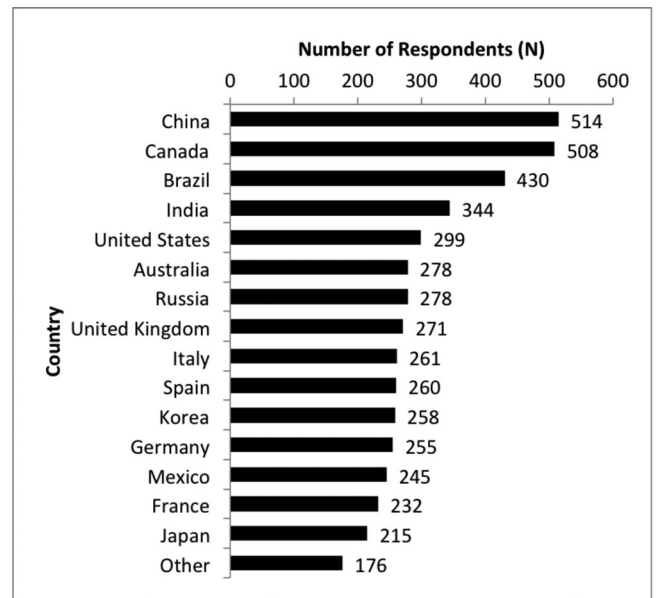

Fig. 2. Number of respondents per country after screening (N = 4,824). Countries in the "Other" category included, in decreasing number of participants, Cyprus, Malaysia, Belarus, Ukraine, Colombia, Costa Rica, Indonesia, Vietnam, Sweden, Guatemala, Kazakhstan, Singapore, Chile, Puerto Rico, Thailand, Argentina, El Salvador, Peru, Philippines, Croatia, Ecuador, Greece, Norway, Panama, Paraguay, Romania, Austria, Belgium, Bolivia, Caribbean, Dominican Republic, Fiji, Ghana, Honduras, Ireland, Ivory Coast, Kyrgyzstan, Mauritius, Netherlands, Pakistan, Poland, Portugal, St. Vincent, Switzerland, Taiwan, Turkey, Uruguay, and Venezuela.[27]

The following sections describe our results for each research question. The results consider all users regardless of how long they have used apps, and include both paid and free apps. For the purposes of brevity, we report the results for correlation that are > 0.2 or < -0.2 and significant. The complete correlation results are available in the supplementary material of the paper, available online, and at: http://www.cs.ucl.ac.uk/research/app_user_survey/.

5.1 App Store Adoption (RQ1)

This section reports the results for RQ1: How are users adopting the app store concept?

5.1.1 User Distribution (RQ1.1)

The app store that was most used was Google Play/Android Market (39 percent), followed by Apple iOS App Store (22 percent), Nokia Ovi Store (15 percent), Samsung Application Store (13 percent), Blackberry App World (6 percent), and Windows Phone Marketplace (3 percent) (Fig. 3). This distribution was consistent with the market share of smartphone operating systems in Q1 2012: Android had the highest market share, followed by Apple, Symbian,[28] Blackberry, and Windows.[29] This

27. Among the 4,824 responses, 203 are from our snowballing method (Australia (N = 13), Brazil (N = 1), Canada (N = 3), China (N = 1), France (N = 1), Germany (N = 2), Italy (N = 30), Japan (N = 19), UK (N = 50), USA (N = 7), and Other (N = 76)). Countries in the "Other" category included, in decreasing number of participants, Cyprus, Malaysia, Sweden, Vietnam, Indonesia, Singapore, Thailand, Taiwan, Norway, Netherlands, Greece, Ireland, Mauritius, Austria, Portugal, and Switzerland.
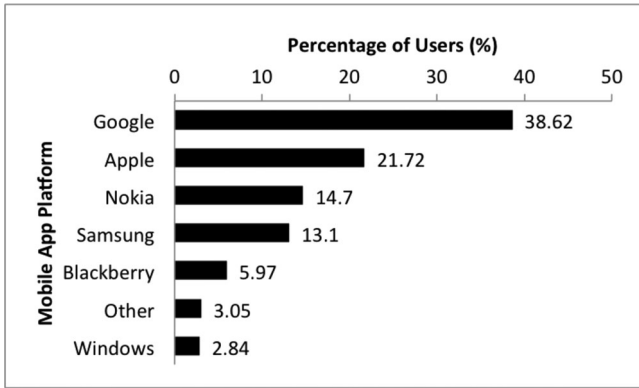
28. Nokia phones run on Symbian OS.

29. http://www.idc.com/getdoc.jsp?containerId=prUS23503312

Fig. 3. User distribution across mobile app platforms.
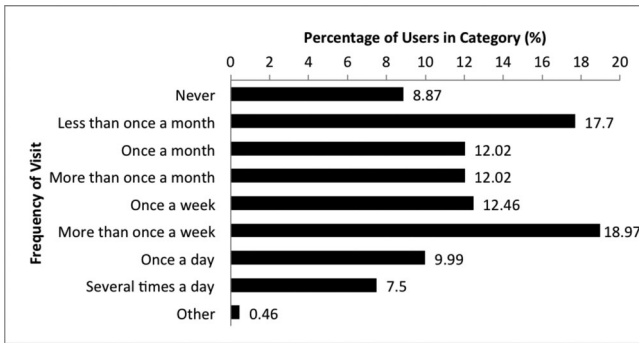
Fig. 4. Frequency of visiting app stores to look for apps.

Fig. 5. Average number of app downloads per month.

Fig. 6. Methods used to find apps.

result differed from that of Franko and Tirrell, which found that the majority of practitioners used Apple iOS (48 percent), followed by Android (19 percent) and Black-Berry (13 percent). This could be due to their participants being only medical practitioners in the United States of America, which was a subset of the whole population.

Approximately 15 percent of users did not know what their app store was, despite visiting the store to download apps. This might be due to some smartphone providers supporting a number of operating systems (e.g., some Samsung smartphones supporting Android, some Windows, and others Samsung Bada), some app stores being rebranded (the Android Market has been rebranded as Google Play,[30] Ovi has been rebranded as Nokia[31]), and in Japan some app stores are "wrapped" within local mobile communication carrier stores. In the survey, some Apple iOS users reported iTunes, Apple's media player and media library application, as their app store.

5.1.2   Frequency of Visit (RQ1.2)

More than once a week was the most common frequency that users visited their app store (19 percent) (Fig. 4). This was followed by less than once a month (18 percent) and once a week (12 percent). The least common frequency of visiting the app store was several times a day (8 percent). Approximately 9 percent of users reported not visiting the
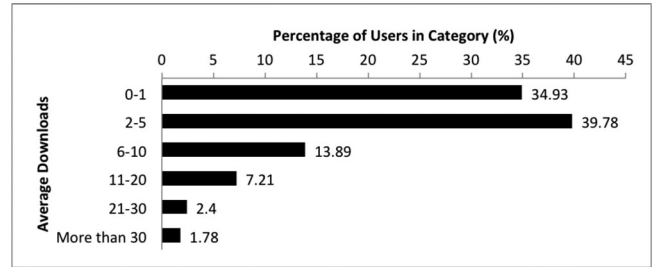
30. http://techcrunch.com/2012/03/06/goodbye-android-market-hello-google-play/
31. http://conversations.nokia.com/2011/05/16/the-evolution-of-nokia-and-ovi/
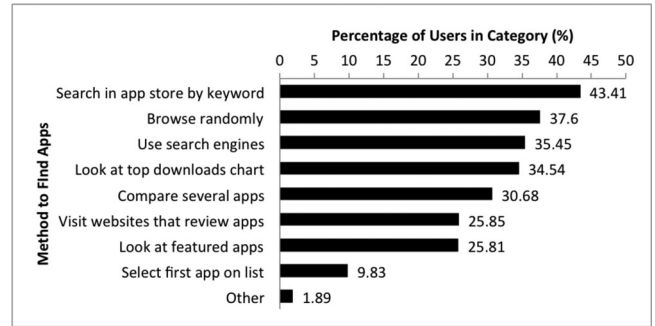
app stores to look for apps. Correlation analysis revealed that as age increased, the frequency of visiting the app store decreased significantly, $r = -.292, p = .000$.

5.1.3   Average Downloads (RQ1.3)

The highest proportion of users downloaded 2-5 apps per month (40 percent) (Fig. 5). This was followed by 0-1 apps (35 percent), 6-10 apps (14 percent), 11-20 apps (7 percent), and 21-30 apps (2 percent). Only 2 percent of users downloaded more than 30 apps per month.

The frequency of visits to the app store was significantly correlated with the average number of apps downloaded per month, $r = .662, p = .000$. A linear regression revealed that the frequency of app store visits accounted for 43.9 percent of the variation in the average number of apps downloaded per month ($R^2 = .439, p = .000$). Correlation analysis showed that with increasing age the average number of apps downloaded per month decreased significantly, $r = -.233, p = .000$.

5.1.4   Finding Apps (RQ1.4)

The majority of people found apps by keyword search in the app store (43 percent) (Fig. 6). This was followed by browsing randomly (38 percent), using search engines such as Google (35 percent), looking at top downloads chart (35 percent), and comparing several apps (31 percent). The least number of users reported downloading the first app they found (10 percent), suggesting that users tend to spend some time choosing apps, even if the apps were free. Correlation analysis showed that as age increased, the likelihood of users finding apps by looking at top downloads chart decreased significantly, $r = -.209, p = .000$.

5.2   User Needs (RQ2)

This section reports the results for RQ2: What needs are users trying to meet with apps?
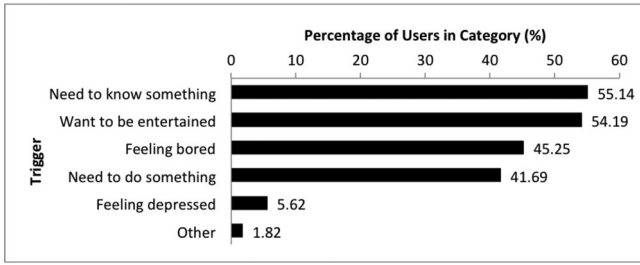
Fig. 7. Triggers to start looking for apps.

5.2.1 Triggers (RQ2.1)

The most popular situation that triggered users to look for apps was when they needed to know something (55 percent), followed by when they wanted to be entertained (54 percent), and when they were feeling bored (45 percent) (Fig. 7). The least popular reason to look for apps was when users were depressed (6 percent). However, the respondents' willingness to specify this option might have been influenced by social desirability bias.

With increasing age, users were significantly less likely to be triggered by boredom ($r = -.331$, $p = .000$), and the need for entertainment ($r = -.305$, $p = .000$).

5.2.2 Reasons for Download (RQ2.2)

The most popular reason for users to download an app was to be entertained (58 percent), followed by to carry out a task (51 percent) (Fig. 8). The third most popular reason for users to download an app was because the app was recommended by friends or family (36 percent). This shows the importance of viral marketing and social networks on app downloads. Curiosity was also an important reason (35 percent), which meant that novel or quirky apps have the potential to attract downloads in the app store.

With increasing age, users were significantly less likely to download apps for entertainment, $r = -.269$, $p = .000$.
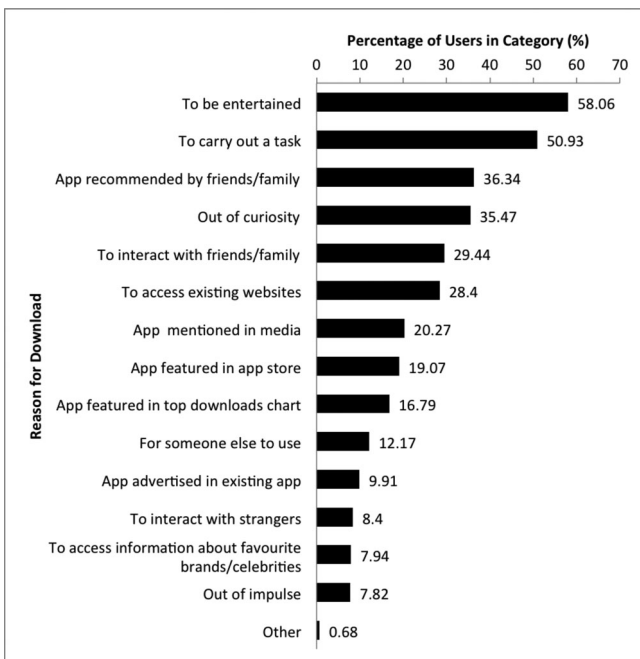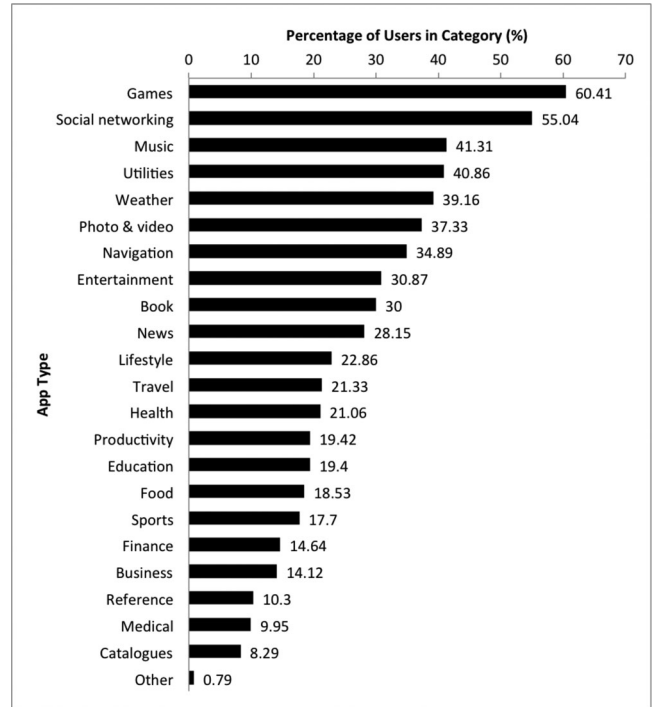

Fig. 8. Reasons for downloading apps.

Fig. 9. Types of apps that users download.

5.2.3 App Types (RQ2.3)

The most popular app category was games (60 percent) followed by social networking (55 percent) and music apps (41 percent) (Fig. 9), which was consistent with the fact that the most common reason to download apps was to be entertained (Section 5.2.2). Utility apps and weather apps were very popular too (41 and 39 percent respectively), indicating that apps play an important role in supporting very specific tasks and providing specific information.

As age increased, users were significantly less likely to download entertainment apps ($r = -.231$, $p = .000$), games apps ($r = -.332$, $p = .000$), social networking apps ($r = -.228$, $p = .000$), and music apps ($r = -.221$, $p = .000$). Learning and empowerment may also be factors that can reduce boredom. However, the likelihood of downloading apps that can provide learning and empowerment is not correlated with age: education apps ($r = -.149$, $p = .000$), productivity apps ($r = -.075$, $p = .000$) and reference apps ($r = -.025$, $p = .078$).

5.3 Influencing Features (RQ3)

This section reports the results for RQ3: What are the features of an app that influence its selection or abandonment?

5.3.1 Choice (RQ3.1)

The most important factors that people consider when choosing apps were: price (57 percent), app features (49 percent), app description (49 percent), reviews by other users (48 percent), and star ratings (46 percent) (Fig. 10). Sadly, the least important factor that influenced a user's choice of apps was the developer (11 percent). This meant that developers would find it difficult to use the success of their previous apps to promote future apps. This finding was
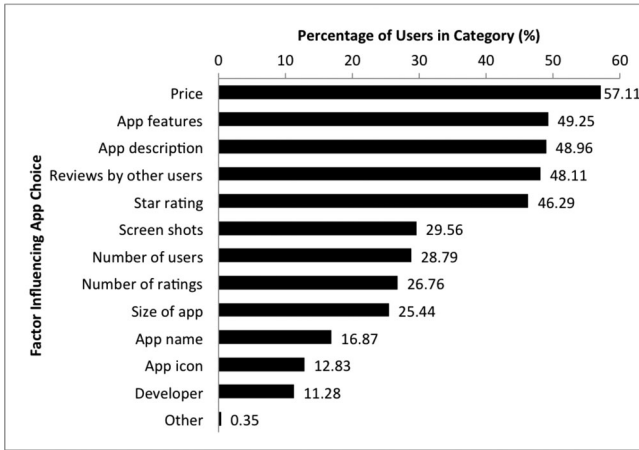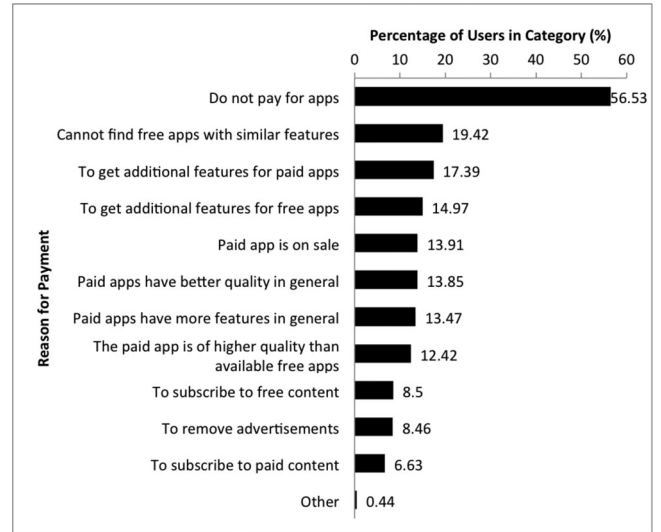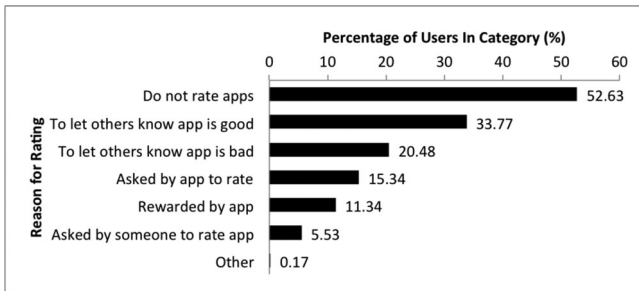
Fig. 10. Factors that influence app choice.

Fig. 11. Reasons for rating apps.

consistent with our experience.[32] As age increased, screen shots became significantly less likely to influence the users' app choice, $r = -.238$, $p = .000$.

5.3.2   Rating (RQ3.2)

Approximately 53 percent of users did not rate apps. The most popular reasons for rating apps was to let other users know that the app was good (34 percent), followed by to let other users know that the app was bad (20 percent) (Fig. 11). Interestingly, the app rewarding users to rate it (11 percent) was a less popular reason compared to the app simply reminding the users to rate it (15 percent). The least common reason for users to rate apps was because they were asked by someone else to do so (6 percent).

5.3.3   Payment (RQ3.3)

Most app users did not pay for apps (57 percent). The most popular reasons to pay for apps were that users could not find free apps with similar features (19 percent). This was followed by the need to get additional features for paid apps (17 percent) and for free apps (15 percent), and that the apps were on sale (14 percent) (Fig. 12). However, a similar number of users selected each reason (M = 13%, SD = 4%). The least common reason people paid for apps was to subscribe for paid content (7 percent). This might be that when the content had to be paid for, users expected the app to be free.
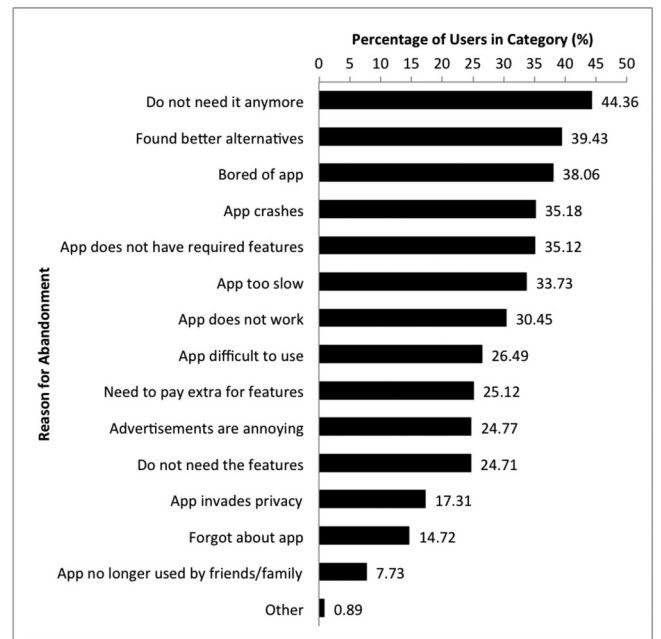
32. When the iStethoscope Pro app by the second author was number 1 in the United Kingdom App Store, other apps by the same developer received no increase in downloads.

Fig. 12. Reasons for paying for apps.

Fig. 13. Reasons for abandoning apps.

5.3.4   Abandonment (RQ3.4)

The most common reason for app users to abandon an app was because they did not need the app anymore (44 percent) (Fig. 13). This was followed by finding better alternatives (39 percent) and getting bored of the app (38 percent). This finding suggested that many apps served temporary functions, unlike desktop software. Correlation analysis showed that with decreasing age, users were significantly more likely to abandon apps because they were bored of the app, $r = -.261$, $p = .000$.

Non-functional requirements such as performance, reliability and usability, were important for app users. Reasons such as the app crashed, the app did not have the required features, the app was too slow, the app was difficult to use, the app did not work, were, on average, adequate reasons for more than 30 percent of users for abandoning an app (Fig. 13). This result showed that the quality of an app was

crucial to encourage continued usage. This is consistent with the more recent study by Khalid et al. that functional errors and app crashes are among the most frequent complaints by users in their app reviews [36].

Only 17 percent of users stopped using an app because it invaded their privacy. However, this might be due to app users being largely unaware of their privacy being invaded and the implications [37].

5.4 Differences between Countries (RQ4)

The results for RQ1 to RQ3 established the baseline mean user behaviors across all countries in our study. We now focus on the main aim of the paper: to investigate the differences in app user behavior between countries.

When comparing the results for the first research question (RQ1.1 to RQ1.3) between countries, some clear differences were evident. Respondents in different countries used some app stores more frequently than others less frequently than the global trend (RQ1.1, Section 5.1.1). At the time of the survey, Google Play was the app store used by the highest number of respondents in all countries (RQ1.1, Section 5.1.1). However, in Australia the highest number of respondents (41 percent) used Apple, likewise in Canada the highest number of respondents (33 percent) used Apple; in India the highest number of respondents (44 percent) used Nokia, and in Japan 50 percent of the respondents selected "Other" as their responses to the app store question and specified Japanese communication carriers such as Docomo and AU as their app stores. Until recently, Japanese communication carriers such as Docomo and AU created their own app stores specific to feature phones. Even today, for Android devices, Japanese communication carriers have developed a wrapper around Google Play such that users can access Google Play apps via the app store of the communication carriers.[33, 34] This also results in fewer Japanese users knowing the name of their app store compared to any other country. A total of 49 percent of app users in Japan did not know their app stores while the average percentage per country was 16 percent and the standard deviation was 11 percent.

Although the global results showed that the most common frequency of visits to app stores was more than once a week (RQ 1.2, Section 5.1.2), in many countries the most common frequency of visits was less than once a month. Only Brazil (22 percent), China (34 percent), South Korea (32 percent), Spain (20 percent), and the United States (20 percent) had the most common frequency of visits as more than once a week. In India, the highest number of respondents visited the app store once a day (21 percent). Countries where respondents visited app stores more frequently also had a higher average number of downloads. This was consistent with our findings of a strong correlation between the frequency of app store visits and the average number downloads per month (RQ1.3, Section 5.1.3).

Fig. 14 shows a heat map visualization of the differences normalized per country so that the values of the odds ratio range from 0 to 1, where 0 is the lowest odds ratio and 1 is the highest odds ratio. Low odds ratio means
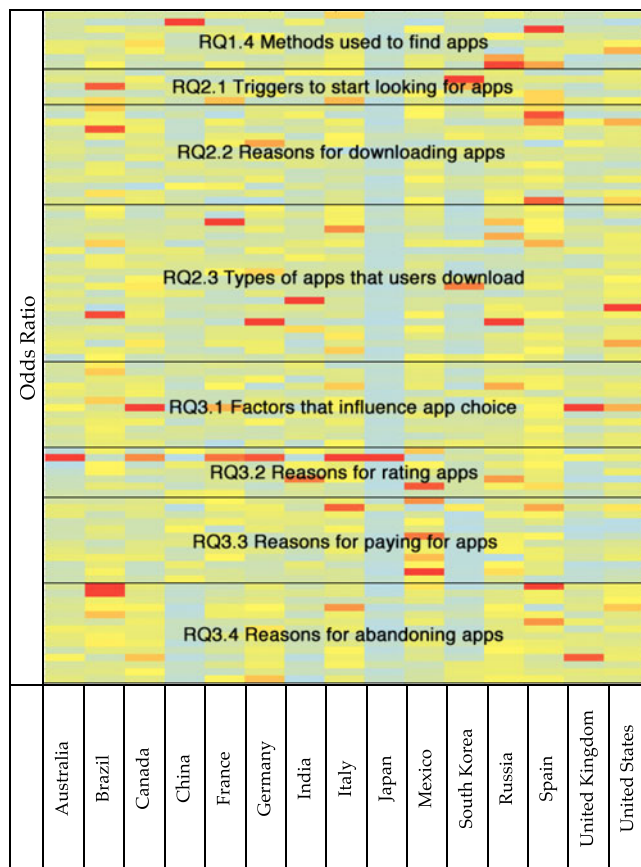
Fig. 14. Heat map of odds ratio per variable normalized per country. Blue to yellow shades indicate lower odds ratios (between 0 and 0.65 respectively), yellow to red shades indicate higher odds ratios (between 0.65 and 1 respectively). Each row of the heat map corresponds to each answer choice for the research question in the order depicted in Figs. 15 16, 17, 18, 19, 20, 21, and 22.

low differences in behavior and high odds ratio means high differences in behavior. (As described in Section 4.7, if country C has an odds ratio of R for behavior B, it means that users from country C are R times more likely to exhibit behavior B compared to users from the other countries.) It is clear from Fig. 14 that many countries have unique differences compared to other countries. The mostly blue stripe representing Japan shows that app users from Japan are indifferent for most answers apart from not rating apps (the only red box in the blue stripe)—Japanese users strongly prefer not to rate apps compared to users from the other countries.

Figs. 15, 16, 17, 18, 19, 20, 21, and 22 illustrate the odds ratio results per country for RQ1.4 onwards. Stacked bar charts are used in order to show cumulative odds ratio results (i.e., odds ratios for all answers to a given question are stacked in one bar per country). A longer bar corresponds to a higher cumulative odds ratio. For each question, the stacked bar charts are ordered by decreasing cumulative odds ratio, so that the country with highest cumulative odds ratio appears first. For example, China had the highest cumulative odds ratios for many questions, with Brazil, India and Mexico following behind. Japan had the lowest cumulative odds ratios for all questions except for reasons for rating apps where Germany had the lowest cumulative odds ratio (Fig. 20). The different colors within each bar
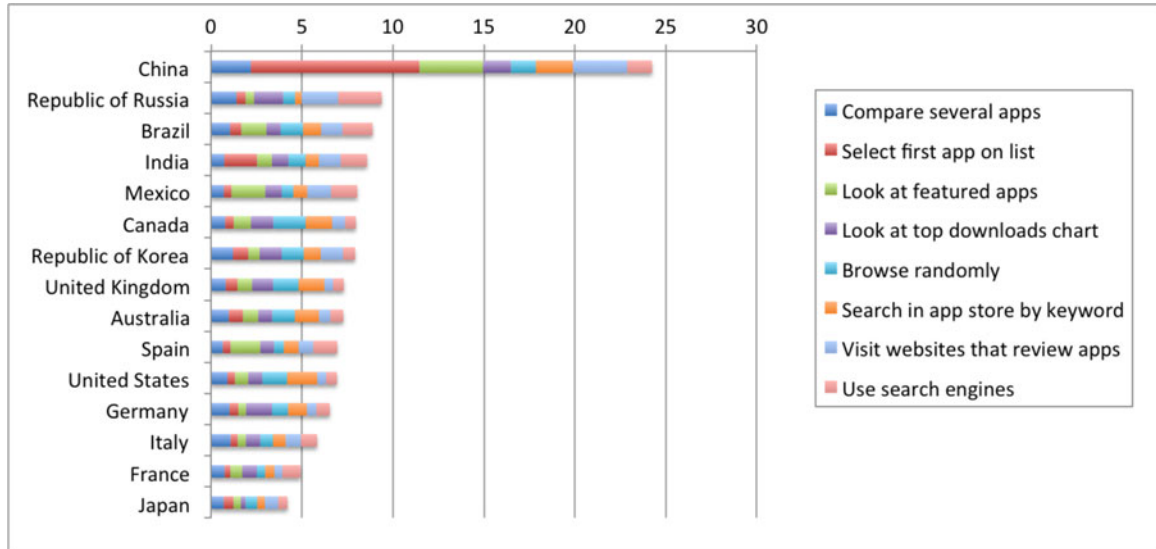
Fig 15. Stacked bar chart showing odds ratio for methods used to find apps per country (RQ4 country differences for RQ1.4).

shows the odds ratio for each answer to each question to enable a direct visual comparison across countries (each color corresponds to a specific answer). For example, in reasons for downloading apps (Fig. 17), Germany, United Kingdom and China are more likely to download out of impulse, compared to Spain, Mexico and Brazil. The option "Other" was not analyzed for odds ratio because for these research questions it comprised less than 5 percent of the responses per country. All countries showed similar odds ratios for reasons to abandon an app, with Brazil showing the largest deviation (Fig. 22).

Pearson's chi-square test on the countries and user behaviors provides a clear picture of the significant differences between countries of app user behaviors (RQ1.4 onwards). Table 3 reports the odds ratio results for each country in turn, highlighting the top three largest differences of that country for brevity. The complete odds ratio and Pearson's chi-square results are available in the

supplementary material of the paper, available online, and at: http://www.cs.ucl.ac.uk/research/app_user_survey/

Together, these results clearly indicate that significant differences exist in mobile app user behavior between countries, confirming our hypothesis. The findings presented here provide a crucial snapshot of the differences to enable future work to track their evolution over time.

6   ANALYSIS AND DISCUSSION

Previous research in cultural differences in organizations and technology usage by Hofstede et al. [11] led to our hypothesis that country differences may exist in app user behavior. The results in Section 5 confirm the hypothesis and in addition highlight specific differences for each country in terms of app user behavior. Section 6.1 analyzes the country difference results by comparing them with Hofstede's work [11]. Section 6.2 compares our
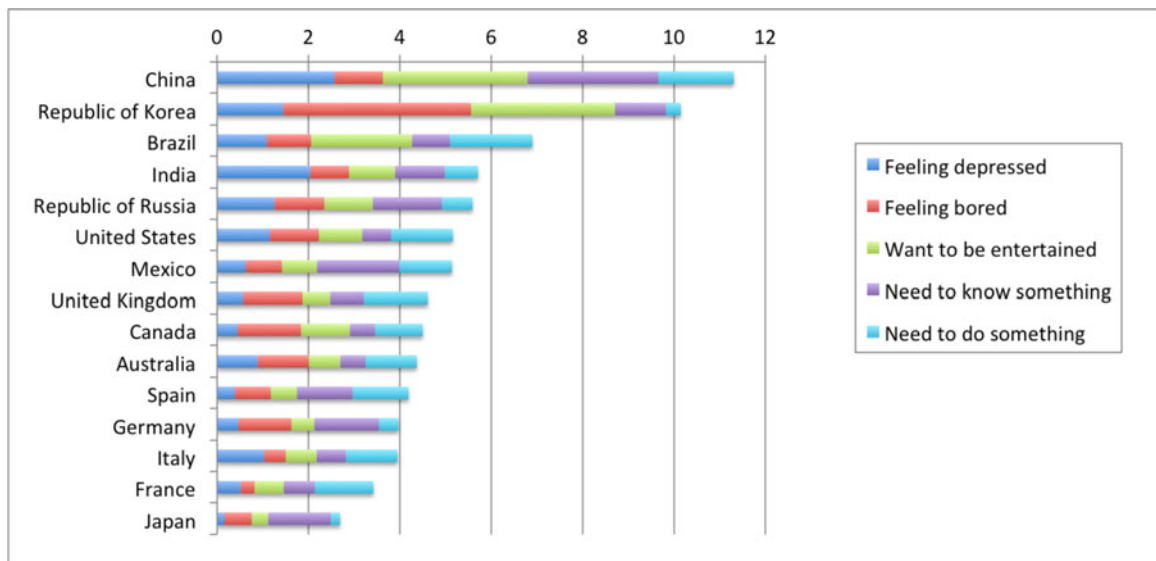

Fig 16. Stacked bar chart showing odds ratio for triggers to start looking for apps per country (RQ4 country differences for RQ2.1).
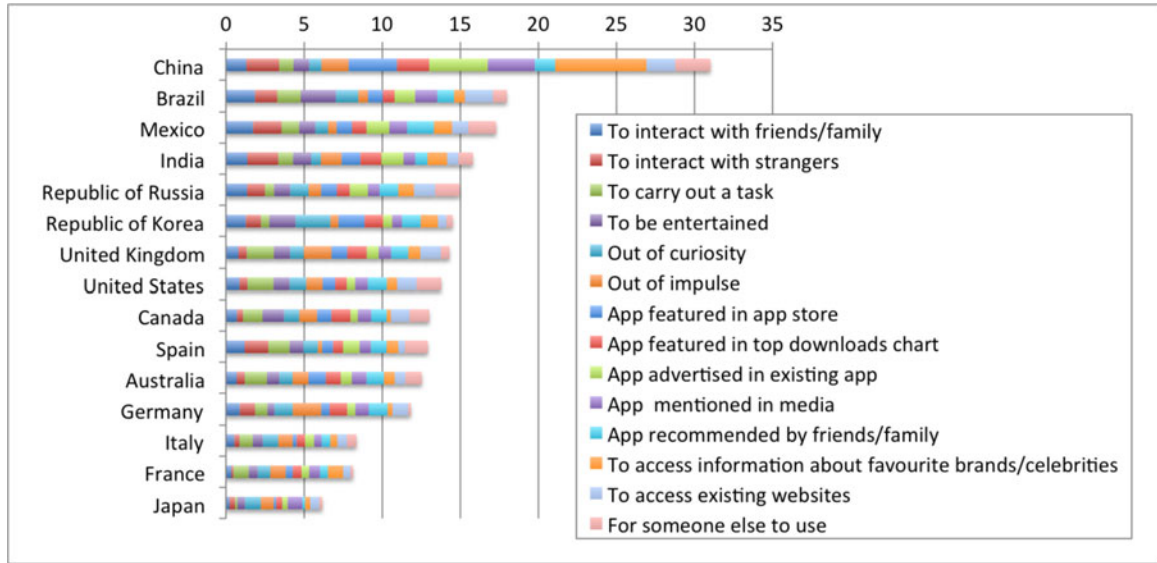
Fig 17. Stacked bar chart showing odds ratio for reasons for downloading apps per country (RQ4 country differences for RQ2.2).

findings with the literature in market-driven software engineering in order to identify new challenges and to inform our discussion of their implications for software engineering.

6.1 Country Differences

While some differences are related to historical or technological legacies as in the case of app store awareness in Japan (Section 5.4), the causes of other differences are perhaps more complex and difficult to track. The differences in user behaviors are largely independent of GDP—when ranked in order of differences, the rankings do not correspond to the relative wealth of those countries. Our results indicate that country-specific differences exist in almost all categories: users from the United Kingdom are most forgetful about their apps and most influenced by price, users from Japan prefer not to rate

apps, users from China are more likely to select the first app on the list more than any other, users from Mexico think that paid apps have more features, and users from Germany and Russia are more likely to download reference apps.

In order to understand the differences, we measured the correlation between app user behavior and Hofstede's cultural index as follows [11]:

1) Power Distance Index (the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally),
2) Individualism Index (the preference for a loosely-knit social framework in which individuals are expected to take care of themselves and their immediate families only),
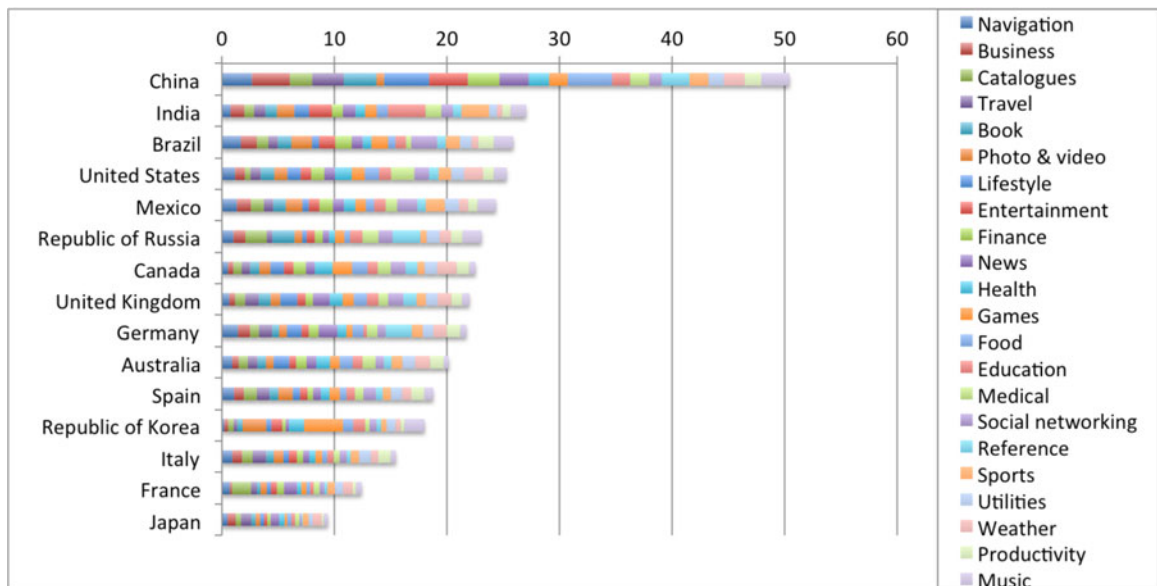

Fig 18. Stacked bar chart showing odds ratio for types of apps that users download per country (RQ4 country differences for RQ2.3).
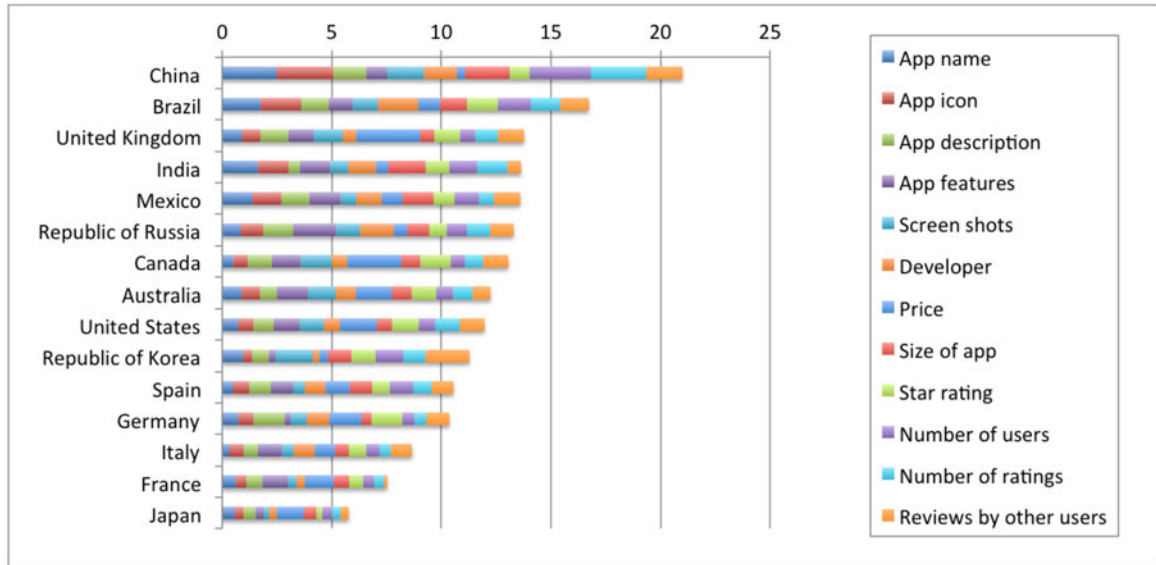
Fig 19. Stacked bar chart showing odds ratio for factors that influence app choice per country (RQ4 country differences for RQ3.1).

3) Masculinity Index (masculine societies have clearly distinct emotional gender roles: men are supposed to be assertive, tough, and focused on material success, whereas women are supposed to be more modest, tender, and concerned with the quality of life),

4) Uncertainty Avoidance Index (the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity),

5) Long-Term Orientation Index (the fostering of virtues oriented towards future such as persistence and personal adaptability), and

6) Indulgence Versus Restraint Index (indulgent societies have a tendency to allow relatively free gratification of basic and natural human desires related to enjoying life and having fun, restrained societies have a conviction that such gratification needs to be curbed and regulated by strict norms).

Our analysis indicates that Hofstede's cultural index helps to explain some, but not all, of the country differences we observed. Results with some correlation to the cultural index include:

- Users from strong power distance countries are less likely to be influenced by price when choosing apps ($r = -.219$, $p = .000$), more likely to spend money on apps because they believe paid apps have better quality in general ($r = .203$, $p = .000$), less likely not to rate apps ($r = -.275$, $p = .000$), more likely to rate an app to let others know that it is good ($r = .262$, $p = .000$). For example, app users in Russia, Mexico, China and India (high power distance) are more likely to spend money on apps because they believe paid apps have better quality in general than app users in Canada, Australia, Germany and the United Kingdom (low power distance) (Section 5.4, Fig. 21).
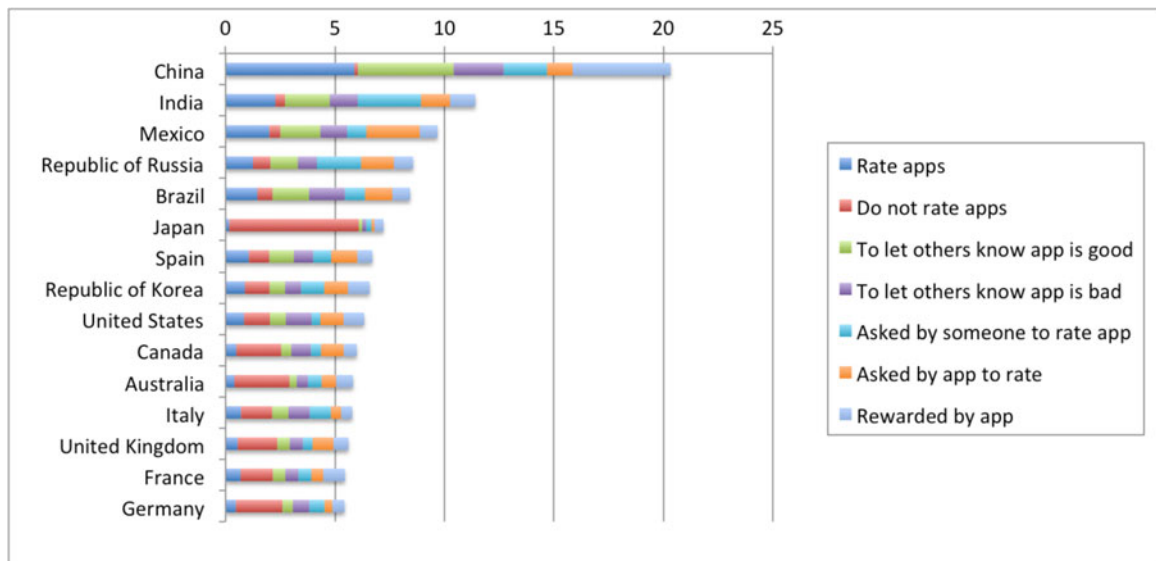

Fig 20. Stacked bar chart showing odds ratio for reasons for rating apps per country (RQ4 country differences for RQ3.2).
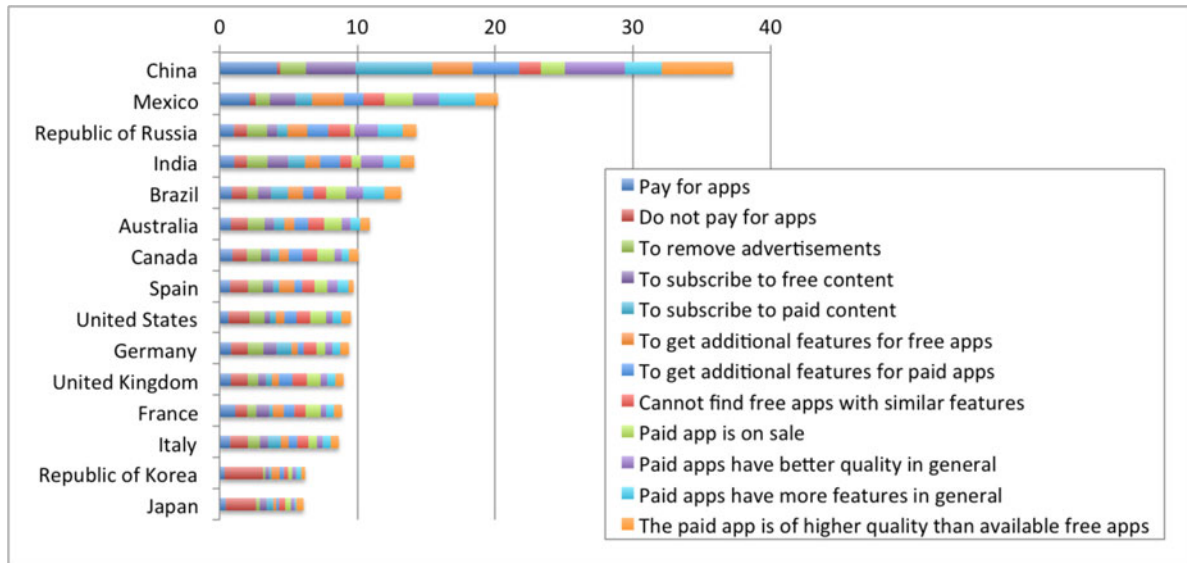
Fig 21. Stacked bar chart showing odds ratio for reasons for paying for apps per country (RQ4 country differences for RQ3.3).

• Users from strong individualism index countries are more likely to be influenced by price when choosing apps ($r = .240$, $p = .000$). They are also more likely not to rate apps ($r = .250$, $p = .000$) and less likely to rate an app in order to let others know that it is good ($r = −.241$, $p = .000$). For example, app users in the United States, Australia, the United Kingdom and Canada (high individualism index) are more likely to be influenced by price when choosing apps than app users in China and Mexico (low individualism index) (Section 5.4, Fig. 19). The former group of users is also less likely to rate an app in order to let others know that it is good compared to the latter (Section 5.4, Fig. 20). In previous work, individualist cultures are less likely to share information with their groups [21]. In individualist countries, media is primary source of information. In collectivist countries, social network is primary source of information.

• Users from strong uncertainty-avoidance countries are less likely to download the first app they see on the list ($r = −.211$, $p = .000$). They are also less likely to download lifestyle apps ($r = −.248$, $p = .000$). For example, app users from Russia, Japan and France (high uncertainty-avoidance index) are less likely to download the first app they see on the list and download lifestyle apps than app users from India, the United Kingdom and China (low uncertainty-avoidance index) (Section 5.4, Figs. 15 and 18). In previous work, lower uncertainty-avoidance index cultures are found to take fewer risks and exhibit hesitancy toward new products and technologies [11].

However, some correlations are not explained by cultural differences. For example, we find that users from strong
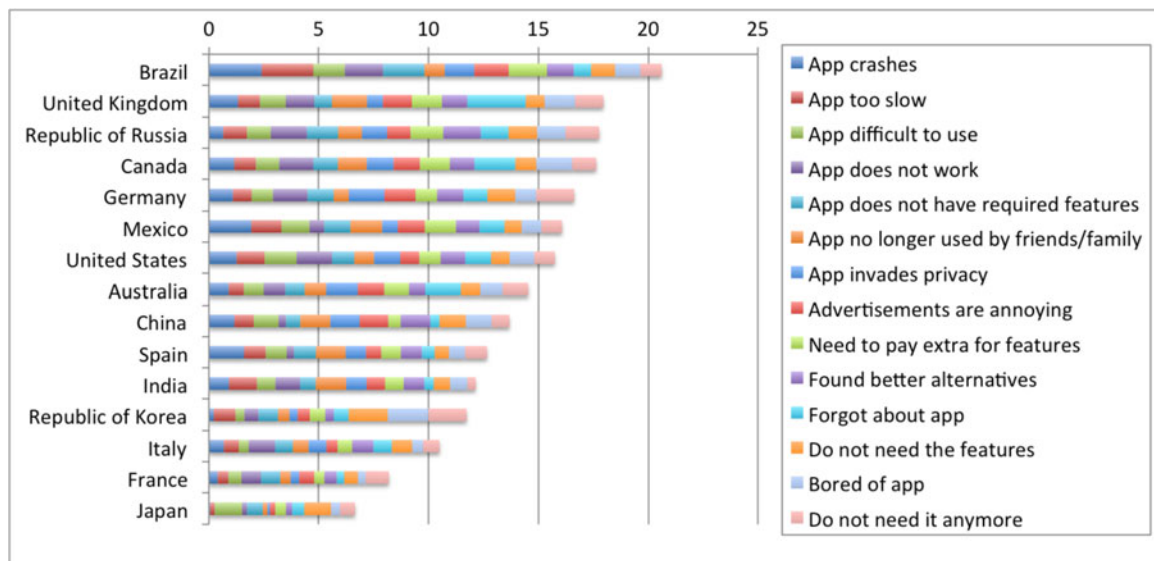

Fig 22. Stacked bar chart showing odds ratio for reasons for abandoning apps per country (RQ4 country differences for RQ3.4).

TABLE 3
Top Three Largest Differences in App User Behavior between Each Country and the Other Countries

Country	Top Three Differences in App User Behavior Compared to Other Countries	Pearson's Chi-square
Australia	1. App users are 2.51 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 47.47, p = .000$
	2. App users are 1.64 times more likely to be influenced by price when choosing apps, Fig. 19.	$\chi^2(1) = 14.24, p = .000$
	3. App users are 1.61 times more likely to abandon an app because they had forgotten about it, Fig. 22.	$\chi^2(1) = 9.95, p = .002$
Brazil	1. App users are 2.39 times more likely to stop using an app because it crashes, Fig. 22.	$\chi^2(1) = 76.64, p = .000$
	2. App users are 2.34 times more likely to stop using an app because it is slow, Fig. 22.	$\chi^2(1) = 73.06, p = .000$
	3. App users are 2.29 times more likely to download social networking apps, Fig. 18.	$\chi^2(1) = 57.02, p = .000$
Canada	1. App users are 2.45 times more likely to be influenced by price when choosing apps, Fig. 19.	$\chi^2(1) = 74.19, p = .000$
	2. App users are 2.05 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 53.18, p = .000$
	3. App users are 1.85 times more likely to stop using an app because they had forgotten about it, Fig. 22.	$\chi^2(1) = 29.8, p = .000$
China	1. App users are 9.27 times more likely to select the first app on the list presented to them, Fig. 15.	$\chi^2(1) = 541.92, p = .000$
	2. App users are 6.02 times more likely to rate apps, Fig. 20.	$\chi^2(1) = 278.4, p = .000$
	3. App users are 5.83 times more likely to download apps that feature their favorite brands or celebrities, Fig. 17.	$\chi^2(1) = 264.32, p = .000$
France	1. App users are 1.69 times more likely to download catalogue apps, Fig. 18.	$\chi^2(1) = 6.9, p = .009$
	2. App users are 1.47 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 7.93, p = .005$
	3. App users are 1.32 times more likely to be influenced by price when choosing apps, Fig. 19.	$\chi^2(1) = 3.89, p = .049$
Germany	1. App users are 2.31 times more likely to download reference apps, Fig. 18.	$\chi^2(1) = 27.4, p = .000$
	2. App users are 2.12 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 30.4, p = .000$
	3. App users are 1.83 times more likely to download apps out of impulse, Fig. 17.	$\chi^2(1) = 9.82, p = .002$
India	1. App users are 3.35 times more likely to download education apps, Fig. 18.	$\chi^2(1) = 119.46, p = .000$
	2. App users are 2.89 times more likely to rate apps because someone asked them to do so, Fig. 20.	$\chi^2(1) = 40.35, p = .000$
	3. App users are 2.43 times more likely to download sports apps, Fig. 18.	$\chi^2(1) = 56.11, p = .000$
Italy	1. App users are 1.43 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 7.6, p = .006$
	2. App users are 1.30 times more likely not to pay for apps, Fig. 21.	$\chi^2(1) = 3.94, p = .047$
	3. App users are 1.21 times more likely to download travel apps, Fig. 18.	$\chi^2(1) = 1.67, p = .196$
Japan	1. App users are 5.91 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 100.78, p = .000$
	2. App users are 2.2 times more likely not to pay for apps, Fig. 21.	$\chi^2(1) = 26.34, p = .000$
	3. App users are 1.36 times more likely to look for apps when they need to know something, Fig. 16.	$\chi^2(1) = 4.7, p = .03$
Mexico	1. App users are 2.64 times more likely to pay for apps because they believe that paid apps have more features in general, Fig. 21.	$\chi^2(1) = 45.15, p = .000$
	2. App users are 2.44 times more likely to rate an app because they were asked by the app to do so, Fig. 20.	$\chi^2(1) = 39.22, p = .000$
	3. App users are 2.31 times more likely to pay for an app to get additional features for free apps, Fig. 21.	$\chi^2(1) = 33.17, p = .000$
South Korea	1. App users are 4.1 times more likely to look for apps when feeling bored, Fig. 16.	$\chi^2(1) = 103.8, p = .000$
	2. App users are 3.46 times more likely to download game apps, Fig. 18.	$\chi^2(1) = 59.91, p = .000$
	3. App users are 3.15 times more likely to look for apps when they want to be entertained, Fig. 16.	$\chi^2(1) = 61.78, p = .000$
Russia	1. App users are 2.47 times more likely to download reference apps, Fig. 18.	$\chi^2(1) = 35.6, p = .000$
	2. App users are 2.39 times more likely to find apps using search engines, Fig. 15.	$\chi^2(1) = 51.3, p = .000$
	3. App users are 2.02 times more likely to rate apps because someone asked them to do so, Fig. 20.	$\chi^2(1) = 11.62, p = .000$
Spain	1. App users are 1.62 times more likely to find apps by looking at the featured apps section of the app store, Fig. 15.	$\chi^2(1) = 13.16, p = .000$
	2. App users are 1.6 times more likely to stop using an app because it crashes, Fig. 22.	$\chi^2(1) = 13.52, p = .000$
	3. App users are 1.52 times more likely to download apps to interact with people they don't know, Fig. 17.	$\chi^2(1) = 4.45, p = .035$
United Kingdom	1. App users are 2.91 times more likely to be influenced by price when choosing apps, Fig. 19.	$\chi^2(1) = 54.12, p = .000$
	2. App users are 2.66 times more likely to abandon an app because they had forgotten about it, Fig. 22.	$\chi^2(1) = 52.65, p = .000$
	3. App users are 1.81 times more likely not to rate apps, Fig. 20.	$\chi^2(1) = 20.74, p = .000$
United States	1. App users are 2.07 times more likely to download medical apps, Fig. 18.	$\chi^2(1) = 21.51, p = .000$
	2. App users are 1.68 times more likely to download weather apps, Fig. 18.	$\chi^2(1) = 19.31, p = .000$
	3. App users are 1.66 times more likely to be influenced by price when choosing apps, Fig. 19.	$\chi^2(1) = 16.08, p = .000$

Differences are measured using odds ratio. The results are statistically significant where p < .001.

power distance countries are more likely to download music apps ($r = .206, p = .000$) and users from strong individualism countries are less likely to download music apps ($r = -.214, p = .000$).

Some differences seem to be in contradiction to previous findings in cultural research. For example, according to Hofstede, countries with higher indulgence versus restraint index tend to be less thrifty. However, Australia, Canada and the United Kingdom, which are the three countries in our dataset with the highest indulgence versus restraint scores, are significantly more likely than other countries to be influenced by price when choosing apps (Section 5.4). Only Mexico appears to follow the trend predicted by Hofstede as the users are 2.64 times more likely than other countries to pay for apps to get more features (Section 5.4).

Some correlations are predicted by the cultural index but are missing. For example, in masculine countries, more nonfiction is read [11]. However, there is no correlation between masculinity index and app types that relate to nonfiction such as reference, business, and catalogues. Countries with high indulgence index are expected to put more emphasis in leisure enjoyment [11], however no correlation is found between indulgence versus restraint index with entertainment related answers. Countries with a low individualism index might have more correlation with influence from friends, to interact with friends or family, download app for someone else to use, apps recommended by friends or family [11], but this was not found to be the case.

Consequently, this analysis suggests that country differences in apps are significant but they are not entirely consistent with previous findings on cultural differences nor are they fully explained by those findings.

Many app user behaviors are different in different countries. However, one universal factor worldwide is app abandonment—all users are very likely to cease using apps of bad quality (e.g., crashes, too slow, difficult to use, does not work). It seems that only an effectively engineered app will stand the test of time and become a popular addition to the mobile device of users. Evidence for this can also be seen in the participants' responses when asked to name of the app they spent most money on and describe the best and/or worst feature of the app. One of the most common answers was *WhatsApp Messenger*[35] with very positive feedback on simplicity and ease of use. This successful app is an example of a well-engineered, cross-platform app that has been popular for most of the life of the app store itself. The app allows users to exchange messages without having to pay for SMS. Its user base is large and users are satisfied, evident by its consistently high ranking and a majority of favorable reviews from users saying it is easy to use and well developed. In this sense, app development is no different from other forms of software development: good software engineering practices matter. (Since the writing of this paper, *WhatsApp* was sold to Facebook for $19 Billion.[36])

6.2 Challenges for Software Engineering

Analysis of the survey results suggests that app-based software development brings new challenges to market-driven software engineering. In this section, we discuss the challenges and their implications for software engineering, in the context of our results and challenges suggested by previous research in market-driven software engineering listed in Table 4.

6.2.1 Addressing Packaging Requirements

Packaging requirements such as app description, title, keywords and screenshots play an important role in app discovery and download. For example, 43 percent of users find apps by searching for keywords and 38 percent browse randomly to find apps that catch their attention (Section 5.1.4). A number of factors that influence users' choice of apps are packaging related, such as app description (49 percent), screenshots (30 percent), app name (17 percent), and app icon (13 percent) (Section 5.3.1). Due to the rapidly increasing number of apps on the app store, packaging requirements have a large influence on the visibility of the app and hence its discoverability and download. As shown in Table 4, marketing influence and communication have been identified as challenges by other researchers. However the specific challenge of addressing the packaging requirements of apps in app stores has not been identified previously.

This challenge is complex, for the packaging requirements vary across different countries. Some countries are more influenced by the packaging of an app. For example, when choosing apps, users from China are 2.5 times more likely than other countries to be influenced by app name and 2.6 times more likely to be influenced by app icon (Fig. 19). Equally, the same packaging can be appealing in one country but not in another. For example, in Japanese app stores, many apps targeted at adults have elements of "cuteness" in their icon and interface, which is inline with the cute culture in Japan [44], but this is not found in app stores in other countries.

Traditionally packaging requirements were met by marketing teams. However, app stores have enabled individual developers and small developer teams to be involved in global market-driven software engineering. This brings additional responsibilities to developers that are not within their skill set.

To address this challenge, natural language processing tools can be used to mine descriptions of existing apps in the app store for each country and evaluate the developers' app description in terms of clarity and attractiveness as well as to suggest improvement using recommender systems. For example existing work in pattern analysis that uses natural language processing and statistics based machine learning to identify news popularity [45] could be adopted to evaluate app descriptions. Research has also been conducted to investigate the use of Latent Dirichlet Allocation (LDA) to evaluate app description against app behavior [46]. In addition, large-scale data mining of app stores and local media for a specific country can be used to automatically suggest popular locale-specific names, using machine learning and pattern analysis methods [47].

35. http://www.whatsapp.com/
36. http://www.bloomberg.com/news/2014-02-19/facebook-to-buy-mobile-messaging-app-whatsapp-for-16-billion.html

TABLE 4
Summary of Software Engineering Challenges from Market-Driven Software Engineering Literature

Challenge	References
Volatile requirements due to market changes, competitors and customers	[5], [38]
Communication gaps between marketing and development	[5], [38]
Balancing the influence between marketing and development on requirements decisions	[5], [38]
Limited value of monolithic requirements specifications	[5], [38]
Requirements overloading as requirements volume is potentially very large and continuously growing	[5], [38]
Dependencies among requirements make release planning difficult	[5], [38]
Balancing between elaborate and elementary development processes, and finding tools and solutions that are not too complex	[5], [38]
Requirements often overlap with design, it is difficult to draw a clear line between the phases	[5], [39]
Due to ad hoc processes and lack of documentation, companies rely on low staff turnover to succeed	[5], [40]
Difficulty completely satisfying the end user as the quality level that is considered acceptable is dependent on both the usage and the application domain	[2]
The developing organization makes all decisions but also takes all risks	[38], [41]
Difficulty managing and testing requirements that are often informally described	[38]
Pressure on short time-to-market on initial and subsequent releases, frequent releases	[1], [2], [39]
Large markets	[2]
Users are difficult to identify or initially unknown	[2], [4]
Limited contact with end users	[2], [39], [40], [42]
The need to constantly invent new, selling requirement	[2], [39]
Difficulty in portfolio and product line planning	[43]
Difficulty managing knowledge share with other participants in the software ecosystem	[43]
Challenges architecting for extensibility, portability, and variability	[43]
Difficulty interfacing functionality with other systems that may have differing organizational contexts	[43]

Research could also be conducted to develop tools that can automate or semi-automate app packaging design. For example, techniques in search-based software engineering, in particular, evolutionary computation techniques such as genetic algorithms that have been used to generate attractive art [48], can be adapted to generate country-specific attractive app icons and app graphics based on existing icons in the app store for each country. This tool would be particularly useful for countries such as China where users are highly influenced by visibility.

6.2.2 Managing Vast Feature Spaces

Traditional market-driven software tends to offer a large feature set in order to meet all of the users' anticipated needs, and the number of features increases as new versions are released, and such releases may be very frequent (these have previously been identified as challenges as shown in Table 4). Mobile apps tend to be highly specific with very few features, and developers release new updates frequently in order to engage with and retain their customer base.[37] Our study shows that users' preferences for features also differ across countries. For example, users in India are three times more likely to download education apps and users in Germany are two times more likely to download reference apps (Section 5.4). Developers face the challenge of selecting an optimal and small subset of features or combination of features in a very large feature space and the ability to tailor the features for each country. The challenge is as much about which features to omit as which to include. This is an interesting contrast to the "requirements overloading" challenge listed in the literature for more general market-driven software engineering (Table 4).

To address the challenge, insights from country and culture differences can be used to inform app feature selection and tailoring. For example, a medical app for personal use by an adult user in an individualist country can be tailored to include features for a high collectivism country that might enable a user to use it to help care for their elderly parents [11].

App users have a wide range of needs. For example, users look for apps when they need to know something (55 percent), want to be entertained (54 percent), feel bored (45 percent), and need to do something (42 percent) (Section 5.2.1). Although users download apps mainly to be entertained (58 percent) or to carry out a task (51 percent), a large proportion of users also download apps out of curiosity (35 percent)—curiosity is the fourth most popular reason for app downloads (Section 5.2.2). Techniques from creative requirements engineering can be used to invent features for apps that will catch a user's interest. (Creative requirements engineering is the use of creative thinking techniques including random idea combination, analogical reasoning and storyboarding as part of a requirements process [42].) Creative requirements engineering can be applied in all app types and is particularly useful for those with large demand and supply such as games (downloaded by 60 percent of users) and social networking apps (downloaded by 55 percent of users) (Section 5.2.3). User needs and trends change quickly: 38 percent of users abandon apps because they are bored of them and 44 percent users abandon apps because they are no longer needed (Section 5.3.4). Indeed, volatile requirements due to market changes is listed as a challenge

37. https://blog.kissmetrics.com/mistakes-in-app-marketing/

in the literature (Table 4). Techniques used in evolutionary computation to automate the creative process of producing design can be applied to the feature space to evaluate and suggest interesting combinations of app features [49].

The challenge of managing vast feature spaces is aided by their method of sale. In traditional market-driven software engineering some of the known challenges include limited contact with end users, difficult-to-identify users, and communication gaps between marketing and development (Table 4). These challenges are now much reduced because app stores provide an unprecedented opportunity for researchers to access a large amount of historical data about app features, user preferences, and download patterns. These data can be mined and used to support the requirements engineering process and is more cost effective and scalable compared to market research and focus groups. Recent research has investigated the use of data mining to extract user requirements from app reviews [7], [50]. Natural language processing can be used to automatically identify useful information such as bugs and feature requests from the large amount of textual reviews. In addition, data mining and pattern analysis can also be used to identify features that are popular, and can be used to predict trends and changing needs. Recommender systems have been applied to large-scale requirements elicitation (e.g., [31], [51]). The large amount of data in the app store suggests that requirements engineering researchers can build an extensive app user profile, and use recommender systems for requirements elicitation for apps. Recommender systems can also be used to identify features that are popular in one country and bring it to another country with similar profile interests. Some app stores have already begun implementing country specific recommender systems to suggest apps to users in the same region.

Finally, research in end-user programming [52] can investigate simple methods that can engage users to customize the content and features of the apps and compose different apps together to meet their goals. Apps that allow interfaces with other apps (e.g., document editor apps that can interface with cloud storage app) do better. Techniques should be developed to enable the development of highly customizable apps where users can "turn-on" and only pay for features or content that they need (today implemented via in-app purchases).

6.2.3 Meeting High Quality Expectations

App users have high expectations on the usability and performance of apps and tend to be unforgiving when an app fails to meet their expectations. For example, 34 percent of users stop using an app because it is too slow, 30 percent of users stop using an app because they cannot get it to work, 26 percent of users stop using an app because they found it difficult to use, and 25 percent of users stop using an app because they found the advertisements annoying (Section 5.3.4). However, users from different countries have differing concerns about app quality. For example, users from Brazil and Spain are two times more likely than other countries to stop using an app because it crashes and users from Brazil are also two times more likely than other countries to stop using an app because it is slow (Section 5.4). The difficulty of completely satisfying the end user has always been a challenge in market-driven

software engineering (Table 4), but due to the large number of competing mobile apps, the challenge of meeting high quality expectations has more severe consequences for app developers, i.e., their app may be abandoned.

With so many apps offering the same features, 39 percent of users abandon apps because they found better alternatives (Section 5.3.4). A large number of apps offering the same or similar features also means that non-functional requirements determine if an app will be downloaded and used. Users can assess non-functional requirements from app description, screen shots and ratings and reviews from other users. As a result, non-functional requirements have become, in some instances, more important than functional requirements.

To address this challenge, requirements engineering researchers need to develop effective techniques to capture non-functional requirements for apps, taking into account the country differences in priorities of the requirements. Requirements prioritization methods for non-functional requirements for commercial off the shelf software and the NFR framework can be adapted to prioritize country specific non-functional requirements [53], [54]. Techniques in data mining and recommender systems mentioned in the previous sections can also be used to identify and prioritize non-functional requirements. There is also the need to develop methods to quantitatively evaluate apps against their non-functional requirements.

6.2.4 Managing App Store Dependency

Traditional market-driven software can be sold via multiple channels such as directly through the software vendor or via other software vendors and resellers, and in soft or hard copies. The challenge of "dependencies among requirements making release planning difficult" has been noted by other researchers for market-driven software (Table 4). However mobile apps have a new and very specific dependency which may override all others. Mobile apps can only be sold via the app store of the platform they are developed for. Although "jail-broken" platforms[38] exist, less than 1 percent of users reported using such platforms (Section 5.1.1). Apps are governed by app store guidelines, which are frequently updated and vary across app stores. Apps that do not adhere to the guidelines will be removed from the store, which makes the success or failure (or even existence) of an app highly dependent on the app store. For example, *App-Gratis*, an app that recommends other apps to app users (which was used by 12 million iOS users and developed by a team of 45), was removed from the iOS app store because of a new app store guideline that stated that "Apps that display Apps other than your own for purchase or promotion in a manner similar to or confusing with the App Store will be rejected."[39] As such, developers need to consider app stores as important stakeholders during requirements elicitation and be alert and responsive towards changes in app store guidelines.

---

38. Jail-breaking allows the download of additional applications, extensions, and themes that are unavailable through the official platforms.
39. http://appgratis.com/blog/2013/04/09/appgratis-pulled-from-the-app-store-heres-the-full-story/

There are differences in app store uses across countries and those differences change rapidly. For example, Japan has its own app store system (Section 5.1.1). One app store in Japan provides a "smart pass" where users can access a selection of apps for free for a monthly fee.[40] There are even Japanese app stores specifically designed for girls.[41] Laws within different countries can cause app store rules and regulations to change (e.g., the need for FDA approval in the USA can affect medical apps). Some rules apply only in some countries and not others. It is possible for the functionality of an app to contravene the customs or laws of some countries, e.g., religious or freedom of speech, and be banned from the countries. Consequently, there may therefore be unanticipated costs and benefits of developing for each platform and country, which developers should consider when planning app projects.

Techniques to model app store guidelines such that app specifications can be verified to meet the guidelines before the app is developed would be very useful to keep developers from investing time developing apps that will be rejected from the store. As guidelines change frequently and are different across platforms and countries, and enforced at different levels of rigor, requirements traceability tools are needed to track different versions guidelines, guidelines for different countries, and app specifications to ensure continuous alignment between the features offered by an app and the app store guidelines.

Software developers often wish to make their apps available on multiple mobile app platforms due to the distribution of users across different app platforms. For example, 39 percent of users use Google Play, 22 percent use Apple iOS, 15 percent use Nokia, 13 percent use Samsung (Section 5.1.1). However, they face difficulty having to port an app from the source platform to the target platforms due to differences in hardware specifications (screen size, resolution, memory), software architecture, API, programming languages and app store guidelines.

To address this challenge, techniques from software product line engineering can be used to develop apps with very few features and need to be released and updated in quick succession across platforms. These techniques should be ultra lightweight in comparison with traditional techniques used by large companies that have long term return on investment [55] and should support short deadlines and the ability to be responsive to market pressure and trends [55], [56]. Feature maps can be used as a lightweight method for defining a feature space of options as well as assessing the value of a particular subset of those options [57]. Recent research in search-based software engineering has investigated the use of indicator-based evolutionary algorithm to maximize the use of user preference knowledge and optimize feature selection [57]. Techniques to model and visualize feature space for apps would be very useful to support optimization.

Research in product line software engineering should also be conducted to support variability in app platforms [58]. For example, recent work by Gokhale et al. has developed a technique to systematically infer likely mappings between the APIs of Java2 Platform Mobile Edition and Android graphics APIs [59]. Techniques from software product lines should be applied to enable strategic and systematic reuse to support the release of an app to different platforms, particularly platforms with high user distribution. Researchers should also investigate the development of a meta-language such that an app can be developed once and then deployed on different platforms, in different architectures, and in different configurations.

6.2.5 Addressing Price Sensitivity

App users are highly sensitive towards app prices. For example, 57 percent of users do not pay for apps and 19 percent pay for apps only if they cannot find free apps with similar features (Section 5.3.3). Price is the most important influence in app choice (Section 5.3.1), and users from some countries more likely than others to be influenced by price when choosing apps (e.g., United Kingdom three times more likely and Canada two times more likely) (Section 5.4).

In the past, it was difficult to attain accurate information about product prices and number of purchases and their variations over time, making pricing a challenging and rarely-studied topic in software engineering. However, in many app stores, the daily price and number of downloads for each app are publicly available, providing researchers with new opportunities to study pricing and develop predictive models on the effects of pricing changes to downloads. Such studies can help developers identify the optimal price point for their apps, which should vary according to the country of sale, as users from different countries are receptive to different price points. The large amount of data also enables the possibility to develop accurate predictive statistical models to model complex country-specific users behavior towards prices [60]. Previous researchers have identified the challenge that "requirements often overlap with design, it is difficult to draw a clear line between the phases" (Table 4). In this case price clearly impacts design decisions and requirements prioritization. For example, design process of a free app with incrementally added paid content differs from the design process of a free or paid app.

Finally, research in software product lines can develop methods to support common variability in function that are related to pricing, such as free version of an app with adverts, free version with limited features, free version with in-app purchases, paid version, and paid version with in-app purchases.

6.2.6 Balancing Ecosystem Effects

Traditionally, software vendors function as relatively independent units, where performances are largely dependent on product features, reputation, and marketing efforts [43]. For example, software houses involved in market-driven software engineering build reputation and the reputation influences users' buying decisions (e.g., Microsoft, Norton). In contrast, the developer's identity is the least important factor that influences a user's app choice (Section 5.3.1). App stores have created a software ecosystem where vendors have become networked and their success or failure highly dependent on one another and on app users who can influence the sale of their apps. For example, users are highly influenced by other users when choosing apps: other

users' reviews (48 percent), their ratings (46 percent), the number of existing users (29 percent), and the number of ratings (27 percent) (Section 5.3.1). As a result, an app that has received good reviews can receive more downloads, and in turn, receives even more ratings and reviews. Reviews can be positive or negative. For example, 34 percent of users rate an app in order to tell other users an app is good, and 20 percent do so to warn other users about a bad app (Section 5.3.2).

Recommendation by friends or family is one of the top reasons for downloading apps (36 percent), more so than other forms of publicity such as being mentioned in the media (20 percent), featured in the app store (19 percent) or in the top downloads chart (17 percent). This result is consistent with results from consumer research, which found that consumers trust "earned media" such as word-of-mouth and recommendations from friends or family, above all other forms of advertising.[42]

With the importance of "earned media", there is a need to develop techniques to effectively elicit feedback, reviews and ratings from users. The elicitation strategies need to be country specific because users' app rating behavior differs across countries. For example, users from Australia, Canada, Japan are more likely not to rate apps, users from China are 6 times more likely to rate apps, users from India are more likely to rate an app if someone asks them to do so, and users from Mexico are more likely to rate an app if asked by the app (Section 5.4). As such the app can be configured to elicit reviews from users more proactively and creative methods that leverage the user's social network should be developed. In addition, with so much data available, there is a need for tools to manage and analyze user feedback, identify unmet needs and prioritize the needs based on level of demand.

Previous challenges in the literature relating to software ecosystems have identified issues such as portfolio planning, knowledge management with other participants in the ecosystem, architecting sufficient flexibility, and integrating functionality with other systems (Table 4). These challenges reflect only part of our challenge of balancing ecosystem effects, for in addition to interactions between vendors and their apps, the significant interaction of the users and the app stores also has a significant impact. App vendors have to consider their strategic role in the software ecosystem to survive [43]. Addressing this challenge requires an understanding of complex app ecosystems and the network effects of all the players, which are themselves challenging research topics.

In our previous research, we have developed multi-agent systems and artificial life simulations to understand interactions between developers, users and apps, and specifically the effect of publicity, and developer strategies on app downloads and ecosystem health [61], [62], [63], [64]. The data collected from this study can be used to provide a more accurate model of mobile app ecosystems, in particular user profile differences across countries. Using our data in combination with historical data from the app stores, there is a potential to develop a tool that can estimate the performance of an app and explore pricing strategies for the app during the planning and development phases of an app. We can leverage knowledge from interdisciplinary research such as biology and artificial life where such predictions are often used to understand natural ecosystems.

7 THREATS TO VALIDITY

Considerable care and attention has been made to ensure the rigor of this work, but as with any chosen research methodology, it is not without limitations. One common issue in survey research is non-response bias. This is where the behaviors of users who responded differ from the behaviors of those who did not respond. Due to the scale of our survey, we were unable to follow up non-respondents and ask for their reason of non-response. However, we found that most respondents who did not complete the survey did not use apps. Thus, it was likely that people who did not respond to the survey did not use apps, hence the sample is unlikely to be subjected to systematic bias among non-respondents. In addition, our use of a panel helped to obtain a sample resembling the actual population.

Respondent demographics from the snowballing method differ from that of the panel. For instance, the majority of respondents from the snowballing method had completed their masters (34 percent), PhDs (28 percent), and undergraduate degrees (27 percent). Occupation sectors from the snowballing dataset were mainly Computer and Mathematical Occupations (30.54 percent), Students (21.18 percent), and Education, Training, and Library Occupations (15.27 percent). In contrast, the panel dataset had occupations spread across all categories, including "Building and Grounds Cleaning and Maintenance Occupations", "Farming, Fishing, and Forestry Occupations", and "Construction and Extraction Occupations." These results highlight the benefits of using a panel to collect a representative dataset and demonstrate one disadvantage of the snowballing method by using social group recommendations, that is, data may cluster on fewer education and occupation types. The inclusion of the snowballing dataset may influence the results. However, participants from this method were a very small percentage of the entire dataset (4 percent) and spread across multiple countries, and further analysis has shown that they have no significant effect on the overall findings.

A survey is a self-report, which means that the answers could be subjective. Although it is likely that participants are aware of the types of apps they download, they may be unaware of the exact number of apps they download per month. Ideally, we could confirm our findings by collecting activity logs from each respondent. However, due to the scale and coverage of the study, it is difficult to collect logs from so many users. In addition, it is infeasible to set up activity logging across all mobile phone operating systems given that some operating systems such as iOS restricts such functionality in apps. (However, it may be possible for app platforms to access the data and confirm our results.) Finally, cultural differences can influence the way respondents answer survey questions [65], which can in turn, influence the survey results. For example, respondents from China often selected many answers in the multiple-answer

42. http://www.nielsen.com/us/en/press-room/2012/nielsen-global-consumers-trust-in-earned-advertising-grows.html

questions. To counter such bias, we normalized the heat map per country (Fig. 14) to enable a clearer visualization of differences of each country.

This study focused on the relationship between app user behavior and the country demographic. Other demographics such as age, gender, education, occupation, and household income also influence user behavior and should be investigated. User behavior may also differ across app stores and app categories (e.g., game apps versus medical apps) and warrants further study. The dataset that was generated from this research contains extensive user demographic data, enabling future research to be conducted to understand the relationship between app user behavior and individual demographic as well as a combination of demographics. The dataset also enables future research to study the differences in app user behavior across app stores and app categories. In addition, this dataset comprises one of the largest and most international ten-item personality inventory to date, and the data can be used to study the relationship between app user behavior and user personality. Finally, future work should also study countries with high smartphone usage per capita such as Sweden and Saudi Arabia, as well as corroborate the findings in this paper through activity logs of a small subset of selected respondents.

8 CONCLUSION

Mobile apps are software developed for use on mobile devices and made available through app stores. App stores are highly competitive markets with a rapidly increasing number of apps, and developers need to cater to a large number of users due to low margins per sale. In this study, we conducted one of the largest surveys to date of mobile app users across the world. We demonstrated that app user behavior differs significantly across countries, a result that was shown in other domains but never before in app-based software engineering, indicating that app developers should carefully consider the countries of their target users. We also investigated user adoption of the app store concept, their app needs, and their rationale for selecting or abandoning an app. Through analysis of the survey results, we identified new challenges to market-driven software engineering related to packaging requirements, feature space, quality expectations, app store dependency, price sensitivity, and ecosystem effect, and their implications for software engineering research in terms of research directions and tool development.

We have released the results of our survey to the app developer community and received feedback that the insights are very useful. Some developers have requested for other countries to be studied as they are building apps for those countries.

We anticipate that the new challenges identified in this paper can guide software engineering researchers towards the development of tools and techniques to improve market-driven software engineering for mobile apps.

REFERENCES

[1] P. Sawyer, I. Sommerville, and G. Kotonya, "Improving market-driven RE processes," in *Proc. Int. Conf. Product Focused Softw. Process Improvement*, 1999, pp. 222–236.

[2] J. Natt Och Dag, "Elicitation and management of user requirements in market-driven software development," Dept. Commun. Syst. Lund Institute Technol., Licentiate Thesis, 2002.

[3] Distimo, "In-depth view on download volumes in the Google Android market," Distimo Company Rep., Utrecht, Netherlands, 2011.

[4] C. Potts, "Invented requirements and imagined customers: requirements engineering for off-the-shelf software," in *Proc. 2nd IEEE Int. Symp. Requirements Eng.*, 1995, pp. 128–130.

[5] L. Karlsson, N. G. Dahlstedt, B. Regnell, J. Natt Och Dag, and A. Persson, "Requirements engineering challenges in market-driven software development: An interview study with practitioners," *Inform. Softw. Technol.*, vol. 49, pp. 588–604, 2007.

[6] K. Wnuk, B. Regnell, and B. Berenbach, "Scaling up requirements engineering, exploring the challenges of increasing size and complexity in market-driven software development," in *Proc. 17th Int. Working Conf. Requirements Eng.: Found. Softw. Quality*, 2011, pp. 54–59.

[7] D. Pagano and W. Maalej, "User feedback in the appstore: An empirical study," in *Proc. 21st IEEE Int. Requirements Eng. Conf.*, 2013, pp. 125–134.

[8] M. Bohmer, B. Hecht, J. Schoning, A. Kruger, and G. Bauer, "Falling asleep with Angry Birds, Facebook, Kindle: A large scale study on mobile application usage," in *Proc. 13th Int. Conf. Human Comput. Interaction Mobile Devices Services*, 2011, pp. 47–56.

[9] O. I. Franko and T. F. Tirrell, "Smartphone app use among medical providers in ACGME training programs," *J. Med. Syst.*, vol. 36, pp. 3135–3139, 2012.

[10] B. Westerman, "How do people really use the iPhone," Presentation slides published in 2008 by Create with Context, Inc. URL: http://createwithcontext.com/insights-how-peoplereally-use-the-iphone.php.

[11] G. Hofstede, G. J. Hofstede, and M. Minkov, *Cultures and Organizations: Software of the Mind*, vol. 2. London, U.K.: McGraw-Hill, 1991.

[12] D. Straub, M. Keil, and W. Brenner, "Testing the technology acceptance model across cultures: A three country study," *Inf. Manage.*, vol. 33, pp. 1–11, 1997.

[13] Y. Van Everdingen, J. Van Hillegersberg, and E. Waarts, "Enterprise resource planning: ERP adoption by European mid-size companies," *Commun. ACM*, vol. 43, pp. 27–31, 2000.

[14] M. Harman, Y. Jia, and Y. Zhang, "App store mining and analysis: MSR for app stores," in *Proc. 9th IEEE Work. Conf. Mining Softw. Repositories*, 2012, pp. 108–111.

[15] M. Chen and X. Liu, "Predicting popularity of online distributed applications: iTunes app store case analysis," in *Proc. iConf.*, 2011, pp. 661–663.

[16] T. M. T. Do, J. Blom, and D. Gatica-Perez, "Smartphone usage in the wild: A large-scale analysis of applications and context," in *Proc. 13th Int. Conf. Multimodal Interfaces*, 2011, pp. 353–360.

[17] Q. Xu, J. Erman, A. Gerber, Z. Mao, J. Pang, and S. Venkataraman, "Identifying diverse usage behaviors of smartphone apps," in *Proc. ACM SIGCOMM Conf. Internet Meas. Conf.*, 2011, pp. 329–344.

[18] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, "Diversity in smartphone usage," in *Proc. 8th Int. Conf. Mobile Syst., Appl.Services*, 2010, pp. 179–194.

[19] A. Rahmati, C. Tossell, C. Shepard, P. Kortum, and L. Zhong, "Exploring iPhone usage: The influence of socioeconomic differences on smartphone adoption, usage and usability," in *Proc. ACM Int. Conf. Human Comput. Interaction Mobile Devices Services*, 2012, pp. 11–20.

[20] N. Henze, M. Pielot, B. Poppinga, T. Schinke, and S. Boll, "My app is an experiment: Experience from user studies in mobile app stores," *Int. J. Mobile Human Comput. Interaction*, vol. 3, pp. 71–91, 2011.

[21] D. McMillan, A. Morrison, O. Brown, M. Hall, and M. Chalmers, "Further into the wild: Running worldwide trials of mobile systems," *Pervasive Comput.*, vol. 6030, pp. 210–227, 2010.

[22] A. Rahmati and L. Zhong, "Studying smartphone usage: Lessons from a four-month field study," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1417–1427, Jul. 2012.

[23] M. Kuniavsky, *Observing User Experience: A Practitioner's Guide User Res.* San Mateo, CA, USA: Morgan Kaufmann, 2003.

[24] G. Chittaranjan, J. Blom, and D. Gatica-Perez, "Mining large-scale smartphone data for personality studies," *Pers. Ubiquitous Comput.*, vol. 17, pp. 433–450, 2012.

[25] T. Wasserman, "Software engineering issues for mobile application development," in *Proc. Workshop Future Softw. Eng. Res.*, 2010, pp. 397–400.

[26] N. Seyff, F. Graf, and N. Maiden, "Using mobile re tools to give end-users their own voice," in *Proc. 18th IEEE Int. Requirements Eng. Conf.*, 2010, pp. 37–46.

[27] E. R. Babbie, *Survey Research Methods*. Belmont, CA, USA: Wadsworth, 1990.

[28] B. A. Kitchenham and S. L. Pfleeger, "Principles of survey research: Part 4: Questionnaire evaluation," *ACM SIGSOFT Softw. Eng. Notes*, vol. 27, pp. 20–24, 2002.

[29] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the big-five personality domains," *J. Res. Personality*, vol. 37, pp. 504–528, 2003.

[30] S. L. Lim, D. Quercia, and A. Finkelstein, "StakeNet: Using social networks to analyse the stakeholders of large-scale software projects," in *Proc. ACM/IEEE 32nd Int. Conf. Softw. Eng.*, 2010, pp. 295–304.

[31] S. L. Lim and A. Finkelstein, "StakeRare: Using social networks and collaborative filtering for large-scale requirements elicitation," *IEEE Trans. Softw. Eng.*, vol. 38, no. 3, pp. 707–735, May 2012.

[32] J. M. Hellerstein, "Quantitative data cleaning for large databases," United Nations Economic Commission Eur. (UNECE), Geneva, Switzerland, 2008.

[33] D. List, *Know Your Audience: A Practical Guide to Media Research*. 3rd ed., Original Books, Wellington, New Zealand, 2007.

[34] A. Field, *Discovering Statistics Using SPSS*. Newbury Park, CA, USA: Sage, 2009.

[35] E. Deutskens, K. De Ruyter, M. Wetzels, and P. Oosterveld, "Response rate and response quality of internet-based surveys: An experimental study," *Marketing Lett.*, vol. 15, pp. 21–36, 2004.

[36] H. Khalid, E. Shihab, M. Nagappan, and A. Hassan, "What do mobile app users complain about? a study on free iOS apps," *IEEE Softw.*, no. 1, p. 1, PrePrints, 2014.

[37] L. F. Cranor, J. Reagle, and M. S. Ackerman, "Beyond concern: Understanding net users' attitudes about online privacy," in *Internet Upheaval: Raising Questions*, I. Vogelsang and B. M. Compaine, Eds. Cambridge, MA, USA: MIT Press, 2000, pp. 47–70.

[38] B. R. Regnell and S. Brinkkemper, "Market-driven requirements engineering for software products," in *Engineering and Managing Software Requirements*. New York, NY, USA: Springer, 2005, pp. 287–308.

[39] M. Lubars, C. Potts, and C. Richter, "A review of the state of the practice in requirements modeling," in *Proc. IEEE Int. Symp. Requirements Eng.*, 1993, pp. 2–14.

[40] E. Kamsties, K. Hormann, and M. Schlich, "Requirements engineering in small and medium enterprises," *Requirements Eng.*, vol. 3, pp. 84–90, 1998.

[41] L. Lehtola and M. Kauppinen, "Suitability of requirements prioritization methods for market Aedriven software product development," *Softw. Process: Improvement Practice*, vol. 11, pp. 7–19, 2006.

[42] N. Maiden, S. Jones, K. Karlsen, R. Neill, K. Zachos, and A. Milne, "Requirements engineering as creative problem solving: A research agenda for idea finding," in *Proc. 18th IEEE Int. Requirements Eng. Conf.*, 2010, pp. 57–66.

[43] S. Jansen, A. Finkelstein, and S. Brinkkemper, "A sense of community: A research agenda for software ecosystems," in *Proc. 31st Int. Conf. Softw. Eng.-Companion Volume, ICSE-Companion*, 2009, pp. 187–190.

[44] L. Hjorth, "Odours of mobility: Mobile phones and Japanese cute culture in the Asia-Pacific," *J. Int. Stud.*, vol. 26, pp. 39–55, 2005.

[45] E. Hensinger, I. Flaounas, and N. Cristianini, "Modelling and predicting news popularity," *Pattern Anal. Appl.*, vol. 16, no. 4, pp. 623–635, 2013.

[46] A. Gorla, I. Tavecchia, F. Gross, and A. Zeller, "Checking app behavior against app descriptions," in *Proc. 6th Int. Conf. Softw. Eng.*, 2013, pp. 1025–1035.

[47] T. Snowsill, I. Flaounas, T. De Bie, and N. Cristianini, "Detecting events in a million New York Times articles," in *Proc. Eur. Conf. Mach. Learning Knowl. Discovery Databases*, 2010, pp. 615–618.

[48] J. Secretan, N. Beato, D. B. D Ambrosio, A. Rodriguez, A. Campbell, and K. O. Stanley, "Picbreeder: Evolving pictures collaboratively online," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2008, pp. 1759–1768.

[49] J. S. Gero and V. A. Kazakov, "Evolving design genes in space layout planning problems," *Artif. Intell. Eng.*, vol. 12, pp. 163–176, 1998.

[50] C. Iacob and R. Harrison, "Retrieving and analyzing mobile apps feature requests from online reviews," in *Proc. 10th Int. Workshop Mining Softw. Repositories*, 2013, pp. 41–44.

[51] C. Castro-Herrera, C. Duan, J. Cleland-Huang, and B. Mobasher, "Using data mining and recommender systems to facilitate large-scale, open, and inclusive requirements elicitation processes," in *Proc. 16th IEEE Int. Requirements Eng*, 2008, pp. 165–168.

[52] J. Wong and J. I. Hong, "Making mashups with Marmite: Towards end-user programming for the web," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2007, pp. 1435–1444.

[53] L. Beus-Dukic, "Non-functional requirements for COTS software components," in *Proc. ICSE Workshop COTS Softw.*, 2000, pp. 4–5.

[54] L. Chung, B. Nixon, E. Yu, and J. Mylopoulos, "Non-functional requirements in Software Engineering," *Kluwer Academic Publishers Int. Series Softw. Eng.*, 2000.

[55] J.-C. Trigaux and P. Heymans, "Software product lines: State of the art," Product Line ENgineering of food TraceabilitY software, FUNDP - Equipe LIEL, Belgique, pp. 9–39, 2003.

[56] P. Knauber, D. Muthig, K. Schmid, and T. Widen, "Applying product line concepts in small and medium-sized companies," *IEEE Softw.*, vol. 17, no. 5, pp. 88–95, Sep./Oct. 2000.

[57] A. S. Sayyad, T. Menzies, and H. Ammar, "On the value of user preferences in search-based software engineering: A case study in software product lines," in *Proc. Int. Conf. Softw. Eng.*, 2013, pp. 492–501.

[58] F. Bachmann and L. Bass, "Managing variability in software architectures," *ACM SIGSOFT Softw. Eng. Notes*, vol. 26, pp. 126–132, 2001.

[59] A. Gokhale, V. Ganapathy, and Y. Padmanaban, "Inferring likely mappings between APIs," in *Proc. Int. Conf. Softw. Eng.*, 2013, pp. 82–91.

[60] I. Zukerman and D. W. Albrecht, "Predictive statistical models for user modeling," *User Modeling User-Adapted Interaction*, vol. 11, pp. 5–18, 2001.

[61] S. L. Lim and P. J. Bentley, "From natural to artificial ecosystems," in *Proc. Frontiers Natural Comput.*, 2012, p. 15.

[62] S. L. Lim and P. J. Bentley, "App epidemics: Modelling the effects of publicity in a mobile app ecosystem," in *Proc. 13th Int. Conf. Synthesis Simul. Living Syst.*, 2012, pp. 202–209.

[63] S. L. Lim and P. J. Bentley, "How to become a successful app developer? Lessons from the simulation of an app ecosystem," in *Proc. Genetic Evol. Comput. Conf.*, 2012, pp. 129–136.

[64] S. L. Lim and P. J. Bentley, "Investigating app store ranking algorithms using a simulation of mobile app ecosystems," in *Proc. IEEE Congr. Evol. Comput.*, 2013, pp. 2672–2679.

[65] S. L. Lim, D. Damian, F. Ishikawa, and A. Finkelstein, "Using Web 2.0 for stakeholder analysis: StakeSource and its application in ten industrial projects," in *Managing Requirements Knowledge*. New York, NY, USA: Springer, 2013.

Soo Ling Lim received the bachelor's of software engineering degree with First Class Honors from the Australian National University in 2005, and the PhD degree from the School of Computer Science and Engineering, the University of New South Wales in Sydney, Australia in 2011. She is a research associate at the Department of Computer Science in University College London, an assistant professor at the Honiden Laboratory in the National Institute of Informatics, Japan, and a research associate at the Software Systems Research Centre in Bournemouth University. She received the Dean's prize from Australian National University, and the Australian Computer Society prize. Before the PhD degree, she was a SAP consultant at the Computer Sciences Corporation and as a software engineer at CIC Secure in Australia. Her research interests are in software engineering (in particular stakeholder analysis, requirements elicitation and prioritization, and mobile app development), survey research, social network analysis, recommender systems, and agent-based modeling. She is a member of the British Computer Society.

Peter J. Bentley received the BSc (Hons.) degree in computer science (artificial intelligence) in 1993 and the PhD degree in evolutionary computation applied to design in 1996. He is an honorary reader and senior college teacher at the Department of Computer Science, University College London (UCL), a collaborating professor at the Korean Advanced Institute for Science and Technology (KAIST), a freelance writer, and an app developer. He has published more than 200 scientific papers and is editor of the books *Evolutionary Design by Computers*, *Creative Evolutionary Systems,* and O*n Growth, Form and Computers*, and author of *The PhD Application Handbook* and the popular science books *Digital Biology*, *The Book of Numbers, The Undercover Scientist*, and *Digitized*. He runs the Digital Biology Group at UCL. His research focuses on modeling and bioinspired computation such as agent-based modeling, evolutionary algorithms, computational development, and other complex systems, applied to diverse applications including software engineering, design, mobile wireless devices, security, and novel computation.

Natalie Kanakam received the bachelor's of science degree in psychology with First Class Honors from Brunel University in 2007 and the PhD degree from the Eating Disorders Unit at King's College London in 2012, by Medical Research Council scholarship. She is a research associate in the Department of Clinical, Educational & Health Psychology at University College London. During the course of her degree and doctorate studies, she used in-depth quantitative and qualitative methods to investigate thinking styles and behaviors. She has authored three journal articles and two book chapters. She is a member of the British Psychological Society.

Fuyuki Ishikawa received the bachelor's of science in 2002 and masters in information science and technology in 2004, both from the University of Tokyo. He received the PhD degree in information science and technology from the University of Tokyo in Japan, in 2007. He is an associate professor at the National Institute of Informatics (NII) in Tokyo, Japan. He is also a visiting associate professor at the University of Electro-Communications in Japan, an adjunct lecturer at the University of Tokyo, Waseda University, Tokyo Institute of Technology, and Japan Advanced Institute of Science and Technology. His research interests include software engineering (e.g., requirements analysis, formal methods, self-adaptation), and services computing (e.g., service description, selection and composition), and he has published more than 70 papers in these areas, some of which are in high-ranking conferences such as the International World Wide Web Conference (WWW), International Conference on Web Services (ICWS), and International Conference on Service-oriented Computing (ICSoC).

Shinichi Honiden received the PhD degree from Waseda University in 1986. He was at Toshiba Corporation from 1978 to 2000. Since 2000, he has joined the National Institute of Informatics (NII) as a professor. His research interests include software engineering, agent technology, and pervasive computing, and he has published over 200 scientific papers in these areas. He is a professor and the deputy director general at NII in Tokyo, Japan, a professor in the Graduate School of Information Science and Technology at the University of Tokyo, and Director of the Center for Global Research in Advanced Software Science and Engineering (GRACE Center). He is also a visiting professor at the University College London, United Kingdom, and Waseda University in Japan. He is a steering committee member of the International Conference on Automated Software Engineering since 2006. In 2006, he received the ACM Recognition of Service Award from the Association for Computing Machinery. In 2012, he received The Commendation for Science and Technology from the Japanese Minister of Education, Culture, Sports, Science and Technology.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.