Predicting Mobile App Privacy Preferences with Psychographics

Andrew McNamara
N. C. State University
ajmcnama@ncsu.edu

Akash Verma
N. C. State University
averma3@ncsu.edu

Jon Stallings
N. C. State University
jwstalli@ncsu.edu

Jessica Staddon
N. C. State University
jessica.staddon@ncsu.edu

ABSTRACT

Using a multi-country data set of over 600 survey participants, we show that psychographics and various attributes of the mobile app context are predictive of user privacy preferences. In particular, we find that a user's decision-making style is predictive of their privacy preferences, with users who are more independent decision makers tending to be more conservative about sharing personal information. We also provide additional insights into several of the privacy preference clusters identified in Lin et al. 2014 by finding that users who are more privacy-concerned tend to be more knowledgeable of app privacy risks and more attentive to app permissions requests, and those individuals with less concern are more likely to trust well-known app companies.

1. INTRODUCTION

Privacy-related decision-making has been actively studied for some time, with much of the research focusing on the "privacy paradox" in which behavioral decisions seem at odds with reported preferences (see, for example, [6, 2]). Mobile apps are an area in which decision making is particularly important to understand given that such apps are very popular in the ever-growing smartphone market (expected to be 2 billion by the end of 2016 [1]).

The mobile app context is also important given that apps often come with substantial privacy risks. For example, several studies find a strong potential for the leakage of private smartphone information and the unexpected ability to track smartphone user activities (e.g., [18, 21, 22, 38]). In addition, Android malware/grayware that maliciously accesses and misuses user personal data by exploiting certain features of smartphones, such as premium messages, has been found in app stores [39]. The situation is made more challenging by the fact that the Android permissions model, requiring every app to declare and seek user approval before accessing device resources, has not been found to be an effective mechanism for meeting user privacy preferences [8, 20, 34].

Significant progress has been made in broadly understanding privacy-related decision making through research connecting psychographics with attitudes and behavior. In particular, in [15] it is shown that decision-making style and risk tolerance are predictive of privacy-related attitudes and security-related behavioral intent. In addition, others have found personality attributes to be predictive of privacy concern (e.g., [23]).

Important progress has also been made in understanding privacy decision-making in the mobile app context. Lin et al, [30], quantitatively link app privacy-related behaviors to user privacy preferences. They identify four user "privacy clusters" that can be used to craft app permissions profiles meeting the preferences of the clusters.

In this paper we connect the psychographic models with the privacy clusters of [30], by showing how psychographics and other user attributes yield a deeper understanding of privacy preferences and the app installation decision process. Through a multi-country survey with over 600 participants (using both Amazon Mechanical Turk and Prolific.ac), we segment participants into the 3 privacy clusters in [30]: unconcerned users, fence-sitters and conservative users. We show that decision-making style is predictive of a user's cluster, while risk tolerance is not. In particular, those who are most concerned about sharing personal information ("conservative" users) tend to have a more independent decision-making style, whereas those with less concern ("unconcerned") are more likely to seek advice from others when making decisions.

In addition, we find that specific user attitudes towards permissions, app companies and the user's own knowledge of the app are also predictive of their cluster, thus providing a more complete portrait of the 3 privacy clusters. The fence-sitter cluster, the largest and most enigmatic cluster in [30] ([30] poses understanding the decision process of the fence-sitters as an open problem) is characterized by significantly increased trust in app developers whereas conservatives are significantly more attentive to, and more critical of, permissions.

In summary, we make the following contributions:

- Multi-country data characterizing user app experiences, attitudes and preferences

WPES'16, October 24 2016, Vienna, Austria

DOI: http://dx.doi.org/10.1145/2994620.2994631

- The first study linking psychographics with an established clustering of users by mobile app privacy preferences
- The identification of factors important to the app installation decision process by user privacy cluster

2. RELATED WORK

Our work is related to two areas of active research: 1. the privacy aspects of mobile apps and, 2. drivers of privacy and security-related behaviors. We also make use of established scales for measuring risk tolerance and decision-making style. In the following we highlight some of the previous research in each area and compare with our work.

Apps and Privacy. User perception of privacy in the mobile app context is an active research area. For example, Felt et al., [20], study various app privacy risks and rank user concerns. Similarly, Shklovski et al., [34], identify several privacy-related app behaviors and find that many users report feeling violated when aware of those behaviors. Our work is closest to [8], which studies smartphone privacy perceptions broadly, including a focus on the self-reported factors leading users to install apps. We also study app installation drivers, but do so through measuring associations between privacy preferences and various behavioral and attitudinal attributes.

App permissions requests are an important source of information for assessing the privacy risk of an app, and are well-studied. Felt et al., [21], show that user-awareness of app permissions is low and that users often have difficulty understanding app permissions descriptions. Understanding user preferences for app permissions is an active area of research. For example, Lin et al., [30], find that when the purpose of a permission is considered, users can often be grouped into a small number of clusters according to preferences. The authors use quantitative analysis to link app privacy-related behaviors to user privacy preferences and identify four distinct clusters (three that are privacy-focused). Our work builds on [30] by identifying factors that are predictive of these clusters. We are also broadly interested in understanding the drivers of app installation decisions, with app permissions being one potential influence on the decisions.

Finally, we note that technology and product guidelines are being developed to improve the privacy experience of app users. Kelley et al., [26], develop a display that concisely represents app permissions information prior to installation and test the effects on user behavior. Similarly, Felt et al., [19], develop guidelines to aid platform designers in determining appropriate permission-granting mechanism for a given permission request. Also, privacy and security "nudges"[1] (user alerts) have been found to be effective in helping users control their privacy experience (e.g., [4]).

Drivers of Privacy-Related Behaviors. There is a growing body of research on associations between experience and security and privacy behaviors. For example, Vaniea et al., [35], find that users who have had negative experiences with software updates are less willing to install security updates going forward. Technology familiarity and knowledge have also been found to be associated with privacy behaviors. In particular, Kang et al [25], find that Internet knowledge is associated with more perceived privacy risk when online, but that this may not translate into more privacy or security-related actions.

Under certain conditions, social information (e.g. technology adoption by friends) has been shown to be associated with increased adoption of security features [13], more sensitivity to security [12], and increased perceptions of informed consent [5].

Egelman et al., [14], study user willingness to pay a premium for apps that do not ask permissions for personal information and find that many smartphone users are willing to pay in order to preserve the privacy of their personal data.

Our study focuses on drivers of app privacy preferences. In particular, we explore whether app knowledge and experience are predictive of a user's privacy cluster (using the terminology of [30]).

Scales. Several scales exist for measuring the association between psychometrics and user privacy/security attitudes and behaviors. Quite a few of them measure associations with the "Big 5", [11], personality model (e.g., [23, 27]), one of the most widely used personality models. Egelman and Peer, [15, 17], study personality scales other than the Big 5 and examine if individual differences in the decision-making styles and risk tolerance are predictive of privacy attitudes. One of the goals of our paper is to expand this study into the mobile app domain and determine if such individual differences are strong predictors of the privacy preferences of mobile app users.

To understand associations between psychographics and privacy preference clusters, we use several well-established scales. We tested the Mobile Users Information Privacy Concerns (MUIPC) scale [37] with a small group of users and found that it was not predictive of privacy behaviors and appeared to bias responses. Instead, we use the Generalized Decision Making Style (GDMS) scale [33] and the Domain-Specific Risk-Taking (DoSpeRT) scale [7] (both found to have strong predictive power in [17]), and examine their predictive power for privacy-related mobile apps behaviors and attitudes. The GDMS scale was developed to measure decision-making styles in a broad sense (not context-specific). The breadth of GDMS makes it applicable to our context, since mobile app installation decisions may be based on several factors. The DoSpeRT scale measures risk-taking-attitudes in six commonly encountered content domains. In our study, we measure the risk attitudes of participants for the purpose of predicting their privacy preferences in the mobile app context. We note that [16] and [17] have been found to be predictive of security behavior intentions.

Finally, we do not use the Westin Privacy Segmentation Index [28] since it has been found to not be predictive in many privacy contexts (e.g., [36]).

3. DATA AND METHODOLOGY

In [30], Lin et al., develop a user clustering based on comfort with different app permissions. As mentioned in Section 1, this clustering classifies individuals as one of 3 privacy-related clusters: *unconcerned*, *fence-sitter*, *conservative*, and an additional cluster, termed *advanced*. Informally speaking (for more discussion, see Section 3.3), the *unconcerned* users have a high level of comfort with disclos-

[1]In this paper we use the term "nudge" in a different way to refer to additional information given to a survey participant to test the strength of their attitudes and opinions.

ing personal data, the *concerned* users are uncomfortable with disclosing personal data, and the *fence-sitters* appear to be largely "on the fence" about disclosing personal data. Our data set does not support the identification of *advanced* users as in [30], a category largely unrelated to privacy.

While [30] hypothesizes potential drivers for individuals based on their cluster and estimates the predictive power of the clusters, identifying these drivers is beyond the scope of their paper. We build on their work by developing hypotheses to explore how user characteristics, knowledge, and context are related to these privacy clusters. In particular, we explore the following four hypotheses for the unconcerned, fence-sitter and conservative clusters:

H1. *Decision-making style and risk tolerance are predictive of user privacy clusters.*

H2. *Knowledge about specific apps (i.e. ratings and reviews) is predictive of user privacy clusters.*

H3. *Previous bad experiences, warning fatigue, and other contextual knowledge are predictive of user privacy clusters.*

H4. *An app's search results rank, an individual's consideration of permissions when updating, and other immediate context when installing apps are each predictive of user privacy clusters.*

3.1 Pilot Interviews

In order to get a baseline for the mobile app intallation motivations, pilot interviews were conducted with six participants either in person or by phone. We recruited participants from multiple backgrounds (English, Computer Science, Business) with the goal of covering a range of motivations. In these interviews, we tried to understand why the individuals might have decided either not to install an app or reluctantly installed an app. In the latter part of the interviews, we also tried to understand how much they considered permissions requests when making their decisions. The interviews served as a good starting point for a more detailed survey.

3.2 Survey

The structure of the final survey is described below (see the complete survey here: https://goo.gl/PYrXmZ).

- *Demographics*: Seven questions to gather basic demographic information including gender, age, educational background and amount of time using a smartphone or tablet device.
- *General Motivations*: Factors considered when making decisions to install an app are measured through a list of 11 factors identified through the pilot interviews and a review of related work. Participants select the top three considered. Individuals were also given a free-response "other" factor if needed. After this, we presented the participant with several questions to further understand the extent to which these motivations influence their app installation behavior. Answer options are either Likert-scaled or yes/no.
- *Decision Making Style*: As mentioned in Section 2, we follow [17] and use GDMS to measure decision-making style. We include the dependent decision-making style since we are interested in how peers may impact user decisions. To manage survey length, we reduce the number of questions to the three questions with the highest loadings for each of the decision-making styles, as determined by Scott and Bruce [33]. These fifteen questions are each presented to participants as a 5-point Likert scale.
- *App Scenarios*: In addition to understanding what motivations individuals consider when installing apps, we study how these motivations affect app installation decisions. Hence, we present each participant with five hypothetical apps, specifying the app's name, category, and basic permissions requests. The participants use a 5-point Likert scale to report how likely they are to install each app. Based on their response, participants are shown follow up questions intended to nudge them in the opposite direction from their initial decision by presenting additional information from various factor categories. If the individual is not sure initially, they are presented with a combination of nudge categories.

  If an individual is initially unlikely to install an app and their response does not change after nudging, they are labeled as a privacy conservative for that app. Similarly, if an individual is initially likely to install an app and is never less likely to install the app after nudging, they are labeled as a privacy unconcerned for that app. These labels for all questions are then used to determine an individual's overall app privacy cluster (conservative, fence-sitter or unconcerned). Section 3.3 provides more details and an example.
- *Social Risk Score*: As mentioned in Section 2, we follow [17] and use DoSpeRT to measure risk taking attitudes. To manage survey length, we only ask participants the questions related to social risk as they had the most predictive power in [17]. This consists of six questions with a 7-point Likert scale from the DoSpeRT scale specified by Blais et al [7].
- *Knowledge of Malicious Capabilities*: This knowledge was measured by presenting participants with seven questions randomly selected from 15 of the user-reported risks in [20]. Participants are categorized into individuals who correctly identify 0-2, 3-5, and 6-7 of the malicious activities as being possible.
- *Follow up questions*: At the end of the survey, we ask the participants a couple of questions about the permissions model change in Android M and provided a free response comment box.

3.2.1 Sample

We launched the survey with two populations, Amazon Mechanical Turk (AMT)[2] and Prolific.ac[3], in order to collect responses from a range of demographics.

In total, 572 responses were collected from AMT and 141 responses from Prolific. Out of the 713 total responses, 62 were filtered out due to a failed attention check question (57

[2]https://www.mturk.com/mturk/welcome
[3]https://www.prolific.ac/

from AMT, 5 from Prolific) and 45 were filtered for completing the survey too fast (AMT). While the estimated completion time for the survey was 15 minutes, the filter threshold was set at 5 minutes to allow for faster completion times by professional survey takers. After filtering, 626 responses were left (490 AMT and 136 Prolific) with a completion time ranging from 5.03 to 49.38 minutes ($\bar{x} = 11.87$), 353(56.4%) male, 272(43.5%) female and 1(0.2%) undisclosed respondents.

The fraction of the population considered a fence sitter is similar to the numbers determined by Lin et al. [30], however, our results have fewer individuals categorized as unconcerned and more as conservative (Table 1).

Finally, while our survey does not measure privacy attitudes in a general sense, we note that the Amazon Mechanical Turk population has been found to be more privacy concerned than the general U.S. population [24].

3.3 Analysis

To test the hypotheses, we categorize individuals as either conservative, fence-sitters, or unconcerned as determined by their responses to all of the app scenario questions. If an individual answers conservatively on any question and never answers unconcerned, they are considered to be part of the *conservative* cluster. Similarly, if an individual answers unconcerned on any question and never answers conservatively, they are considered to be part of the *unconcerned* cluster. The rest of the individuals are part of the *fence-sitter* cluster, because they either respond to the app scenarios as a conservative and unconcerned participant, or neither. For example, if an individual answers conservatively on two questions, unconcerned on one, and neither on the last two, they would be clustered as a fence-sitter. If another individual answers unconcerned on two questions and does not answer conservatively on the other three, they would be clustered as unconcerned.

While we do not use a clustering analysis like Lin et al. [30], our approach to assigning privacy clusters is a natural adaptation of their privacy clusters to our survey context in that we assign clusters based on participant comfort with app privacy issues. We do not identify advanced users (as is done in [30]) as being "advanced" is not directly related to a user's privacy attitudes. While this means our findings are not exactly comparable to [30], we believe we are using a natural interpretation of three of their clusters (conservative, fence-sitter and unconcerned) that is consistent with the intent of the clusters.

H1. To test this hypothesis, we build a model with inputs: the rational, intuitive, and dependant General Decision Making Style (GDMS) subscores, the DoSpeRT social risk subscore, and a categorization based on the individual's knowledge of possible malicious activities *(Questions GDMS, soc_risk, and behav, respectively).*

The GDMS subscore ranges from 3 (an individual strongly disagrees with a decision-making style) to 15 (an individual strongly agrees with a decision-making style). Similarly, the DoSpeRT social risk ranges from 5-35 (ranging from extremely unlikely to extremely likely to take social risks). The variable was included in the model in two different forms, by calculating the geometric mean for each participant and by categorizing the individuals based on the answers over all five questions.

We bucket participant knowledge of potentially malicious apps into three categories based on the number of correctly identified capabilities. A higher malicious category corresponds to the correct identification of more of the malicious capabilities and is therefore considered a more advanced[4] user.

H2. To test this hypothesis, we build a model with inputs: the participant response to whether they will consider installing an app if it makes permission requests unrelated to its major functionality, the star rating of an app below which they have doubts about an app's quality, the trade-off between the number of ratings and stars received , and participant preference for downloading an app with a higher download count *(Questions fact_det05, fact_det08, fact_det11, and fact_det12, respectively).*

H3. To test this hypothesis, we build a model with inputs: whether or not the participant has reluctantly installed an app, had any negative experiences with an app, reports that security warnings are too intrusive, or reports that apps with the same functionality have the same permissions *(Questions reluctant, fact_det03, fact_det13, and fact_det14, respectively).*

H4. To test this hypothesis, we build a model with inputs including: whether or not the participant considers any additional permission requests by apps while updating them, whether they accept the additional permission requests by mobile apps while updating them, whether the location of an app in search results on the app store impacts participants decision to install an app and whether participants had ever reversed their decision to install an app *(Questions fact_det06, fact_det02, and fact_det15, respectively).*

Trust is an additional model input, measured as participant level of trust in well known and lesser known mobile app companies and whether participants believe that their data is handled properly by such companies *(Questions fact_det04, fact_det09, and fact_det10).* The last input to this model is whether participants would critically consider each permission before accepting them in the new runtime permissions model introduced in Android 6.0 *(Questions p_model1 and p_model2).*

3.3.1 Statistical methodology

Statistical analysis is conducted using the `R` software package [31]. Since the data are categorical and ordinal, we fit cumulative link models using the `ordinal` package [10] to determine if significant correlations exist. We use likelihood ratio tests to determine the variables for which we can reject the null hypothesis (namely, that the specific drivers considered are independent of an individual's privacy cluster) with $\alpha = 0.05$. Since the privacy cluster is an ordinal response, the proportional odds form of a cumulative logit model is used to determine the effect of the independent variables on the context. Each cumulative logit has its own intercept (μ_c) while the effects are the same for each logit (β_i) [3].

For the general form of the logit model we use, see equation 1 [9] where $P(behav \leq c)$ is the probability that the observed behavior is less than or equal to a specific privacy cluster (1 = unconcerned, 2 = fence-sitter, 3 = conservative), μ_c is the estimated threshold coefficient between clus-

[4]Note that this is similar to, but not necessarily the same as the"advanced" designation in [30])

ter c and $c+1$, β_i is the estimated coefficient for a particular effect, i. If an effect is modeled as a factor, every category after the first is included as its own effect which takes a value of 0 or 1 based on whether an individual expresses that factor. The first category of every factor is implicitly included in the threshold coefficient, μ_c.

$$logit(P(behav \leq c)) = \eta_{c,effects} \quad (1)$$

where

$$\eta_{c,effects} = \mu_c - \sum \beta_i effect_i \quad (2)$$

In order to determine the effect size of the models, the predicted probability for an individual to be in each of the privacy clusters is calculated with the inverse logit function (equation 3).

$$P(behav \leq c) = (1 + \exp(-\eta_{c,effects}))^{-1} \quad (3)$$

Much of the effect size analysis is in the Appendix due to space constraints.

To determine if specific attitudes and perceptions differ between the clusters we use nonparametric significance tests, Chi Square and Wilcoxon signed-rank.

4. RESULTS

In this section we describe how psychographics and other user attributes relate to the three privacy clusters (conservative, fence-sitters and unconcerned). We group these attributes into four categories: psychographics (e.g., decision-making style) in Section 4.1, app-specific information (e.g., ratings and reviews) in Section 4.2, prior contextual knowledge (e.g., a user's past experiences with apps) in Section 4.3 and current context (e.g., whether the individual will accept additional permissions when updating an app) in Section 4.4. Additional effects size analysis can be found in Appendix A, due to space constraints.

4.1 Psychographics

In this section we consider how user decision-making style and risk tolerance are related to their privacy cluster (H1). The likelihood ratio test indicates the rational and dependant GDMS subscore and the behavior categorization predict the privacy cluster with at least a 99% confidence ($p = 0.0195$, .000012, and 0.00322, respectively).

The effects fit for the model are shown in equation 4 with the coefficients as summarized in Table 2.

$$\begin{aligned}\eta_{c,effects} =& \mu_c - \beta_1 GDMS_{rat} - \beta_2 GDMS_{dep} \\ & - \beta_3 malicious_2 - \beta_4 malicious_3\end{aligned} \quad (4)$$

As can be seen in Figure 11 in Appendix A, with an increase in the rational GDMS subscore, an individual is more likely to be conservative in terms of installation behavior. As the dependant subscore increases, however, that trend decreases and the individual is more likely to be unconcerned.

The malicious knowledge categorization indicates that advanced users[5] are more likely to be conservative when it comes to making app installation decisions. Figures 11 and 12 in Appendix A show the effect size of these coefficients by taking the inverse of the logit model for all behavior categories.

Figures 1 and 2 show the distribution of answers across the privacy clusters. While the range of dependant subscores does not vary much with the privacy cluster, the median score is slightly higher for those labeled as unconcerned. The opposite relationship holds for the rational subscore, while at least 50% of individuals in each cluster record a rational subscore of at least 12, there is a larger fraction of unconcerned individuals that score less than 12 than in the other categories. It can be seen that unconcerned individuals have the highest fraction of novice users while conservatives have the largest fraction of advanced users in our population. Novice users make up a larger fraction of fence sitters than conservatives.

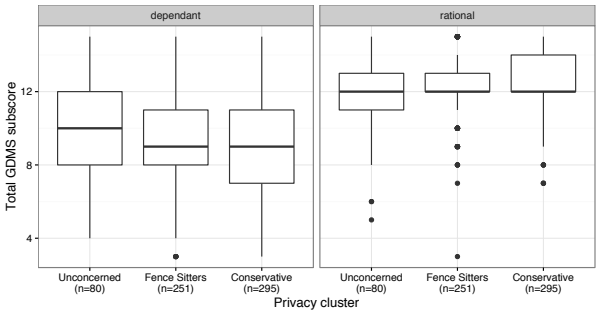

Figure 1: Distribution of GDMS subscores across privacy clusters. The dependant subscore range is larger than the rational subscore for all privacy clusters, but the distribution does not change much for either subscore. The differences are statistically significant.

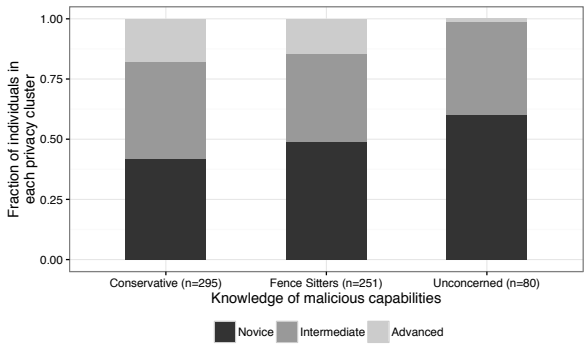

Figure 2: Distribution of knowledge about malicious capabilities across privacy clusters. The fraction of novices increases significantly as the privacy cluster becomes less conservative.

4.2 App-Specific Information

The likelihood ratio test indicates that whether users consider installing apps requesting permissions unrelated to the major functionality is predictive of the privacy cluster with 99.9% confidence ($p = 1.66e - 11$).

[5]See Section 3.3, H1 for a description of advanced users

Category	Number of individuals	Fraction of population	Males	Females	Average time (m)
Unconcerned	80	0.128	48	32	12.56
Fence sitters	251	0.401	140	111	14.01
Conservative	295	0.471	165	129	13.64

Table 1: Distribution of individuals into app installation behavior categories. Note that in figures corresponding to particular questions, the counts vary a bit from the above due to answers of "I'm not sure/Not sure".

Factor	Level	β	p
$GDMS_{rat}$	—	$0.1388(\beta_1)$	0.00241
$GDMS_{dep}$	—	$-0.1396(\beta_2)$	$8.78e-6$
malicious	Neither advanced nor novice (2)	$0.3185(\beta_3)$	0.0576
malicious	Advanced (3)	$0.8359(\beta_4)$	$4.52e-4$

Table 2: Coefficients of model for H1. Factors are 1) *Individual's rational GDMS subscore* 2) *Individual's dependant GDMS subscore*, and 3) *An individual's knowledge of potential malicious app capabilities*. Estimated coefficients for thresholds are -1.336 and 0.7975 for μ_1 and μ_2, respectively.

Factor	Level	β	p
perm	Neither (2)	$0.7884(\beta_1)$	$1.79e-4$
perm	Not consider (3)	$1.340(\beta_2)$	$1.18e-13$

Table 3: Coefficients of model for H2. Factor is *An individual will consider an app if it requests permissions unrelated to its major functionality.* Estimated coefficients for thresholds are -1.408 and 0.7699 for μ_1 and μ_2, respectively.

The effects fit for the model are in equation 5 in which a category of 1 indicates that an individual will consider an app requesting unrelated permissions permissions and a category of 3 indicates that an individual will not consider such an app. The coefficients of the model are summarized in Table 3.

$$\eta_{c,effects} = \mu_c - \beta_1 perm_2 - \beta_2 perm_3 \quad (5)$$

The predicted effects of this model are large; as an individual is less likely to consider installing an app that requests permissions unrelated to its major functionality, he or she is more likely to be in the conservative privacy cluster as seen in Figure 13 in Appendix A.

Participants identified as unconcerned are much more likely to consider these apps while conservatives are much more likely to not consider them (Figure 3).

4.3 Prior contextual app knowledge

A likelihood ratio test indicates that the belief that apps with the same functionality will also have the same permissions, bad experiences with apps, and the belief that security warnings are too intrusive, are predictive of privacy cluster with at least a 95% confidence ($p = 0.0011$, 0.0207, and $7.198e-5$, respectively). When the model was run with these parameters (see equation 6), not all of the coefficients were statistically significant (Table 4).

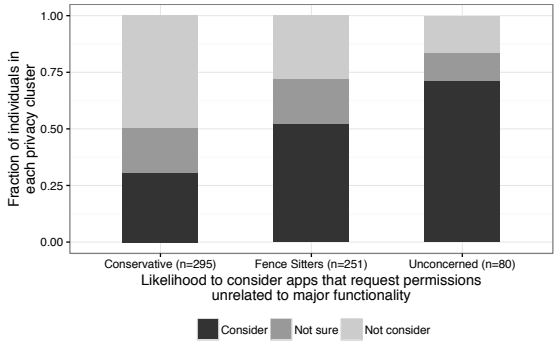

Figure 3: Distribution of the likelihood to consider installing an app that requests permission unrelated to its major functionality across privacy clusters. The fraction of individuals that will consider these apps increases significantly as the privacy cluster becomes less conservative.

Factor	Level	β	p
func	Neither (2)	$0.5104(\beta_1)$	0.0212
func	Disbelieve (3)	$0.6155(\beta_2)$	$7.87e-4$
neg	Not sure (2)	$-0.08269(\beta_5)$	0.825
neg	No (3)	$-0.4110(\beta_7)$	0.0134
warn	Neither (3)	$0.5477(\beta_{10})$	0.0217
warn	Disbelieve (4)	$0.9126(\beta_{11})$	$1.34e-5$

Table 4: Coefficients of model for H3. Factors are 1) *Different apps with the same functionality have the same permissions*, 2) *Whether individual has had a bad experience with an app*, and 3) *Security warnings are too intrusive.* Estimated coefficients for thresholds are -1.253 and 0.8884 for μ_1 and μ_2, respectively.

$$\begin{aligned}\eta_{c,effects} = &\mu_c - \beta_1 func_2 - \beta_2 func_3 \\ &- \beta_3 neg_2 - \beta_4 neg_3 \\ &- \beta_4 warn_2 - \beta_5 warn_3\end{aligned} \quad (6)$$

The effect size is small for fence-sitters when considering whether individuals believe that apps with the same functionality also have the same permissions, however, as individuals believe this statement, they are more likely to be unconcerned than conservatives (Figure 14, Appendix A). The effect size for whether individuals have had a negative experience with an app is also small, but individuals who have not had any bad experiences are less likely to be conservatives (Figure 15, Appendix A). The belief that security warnings are too intrusive has a larger effect size; as indi-

viduals believe this, they are more likely to be unconcerned or fence-sitters (Figure 16, Appendix A).

We find that most fence-sitters and unconcerned individuals believe that different apps with the same functionality should also have the same permissions (Figure 4) while unconcerned individuals are evenly split on this issue. As seen in Figure 5, there is little difference in the fractions of conservatives and fence-sitters who have had a negative experience with an app. A larger fraction of unconcerned individuals, however, have not had a negative app experience. Conservative and fence-sitter individuals are much more likely to believe that security warnings are not too intrusive while unconcerned individuals are roughly evenly split between believing and not believing this (Figure 6).

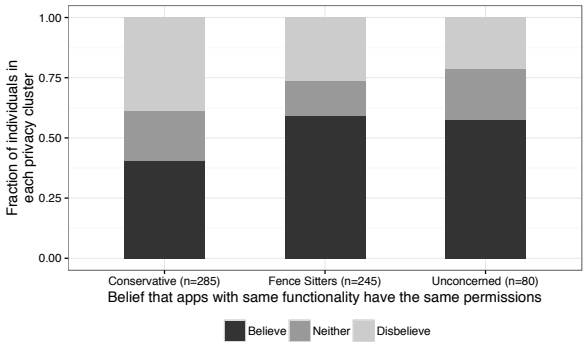

Figure 4: Distribution of the belief that apps with the same functionality should have similar permissions across privacy clusters. The fraction of individuals with this belief is significantly larger for the fence sitters and unconcerned.

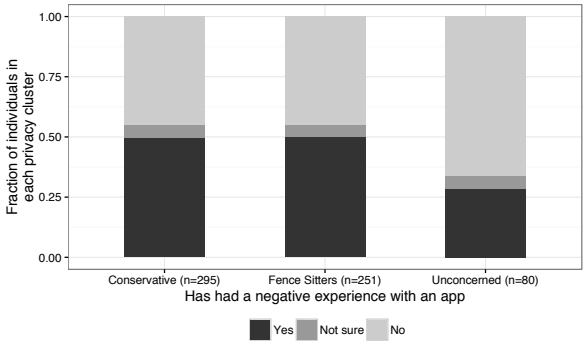

Figure 5: Distribution of participants who have had negative app experiences across privacy clusters. The fraction who have had bad experiences does not differ between conservatives and fence sitters but is significantly larger than the fraction of unconcerned who have had a bad experiences.

4.4 Current context

The likelihood ratio test indicates that participant likeliness to consider each permission request in the new runtime permissions model and to accept additional permissions when updating apps predict the privacy cluster with a 99.9% confidence ($p = 7.902e-4$ and $4.346e-4$, respectively). Additionally, the test also indicates that participant trust

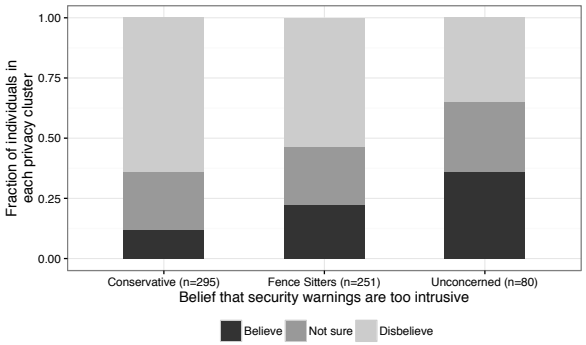

Figure 6: Distribution of the belief that security warnings are too intrusive across privacy clusters. The fraction of individuals that with this belief increases significantly as the privacy cluster becomes less conservative.

Factor	Level	β	p
trust	Neither (2)	$0.6815(\beta_1)$	$4.15e-4$
trust	Disbelieve (3)	$1.045(\beta_2)$	$2.40e-4$
perm	Neither (2)	$0.6632(\beta_3)$	$2.45e-3$
perm	Unlikely (3)	$1.123(\beta_4)$	$1.78e-4$
change	Not sure (2)	$-0.8093(\beta_5)$	$8.76e-4$
change	No (3)	$-0.3916(\beta_6)$	0.0275
droidM	Not sure (3)	$-0.3043(\beta_7)$	0.0409
droidM	Not consider (4)	$-1.062(\beta_8)$	$1.59e-4$

Table 5: Coefficients of model for H4. Factors are 1) *Individuals trust well known companies* 2) *Individuals will likely accept additional permissions when updating*, 3) *Whether an individual has ever changed their mind to install an app after initially considering*, and 4) *Individuals would critically consider each new permission under Android M's new permission model*. Estimated coefficients for thresholds are -1.906 and 0.3411 for μ_1 and μ_2, respectively.

in well known companies and whether they have changed their mind not to install an app, predict their privacy cluster with 95% confidence ($p = 0.01143$ and $p = 0.001685$, respectively).

The effects fit for the model are shown in equation 7.

$$\begin{aligned}\eta_{c,effects} =& \mu_c - \beta_1 trust_2 - \beta_2 trust_3 \\ & - \beta_3 perm_2 - \beta_4 perm_3 \\ & - \beta_4 change2 - \beta_5 change_3 \\ & - \beta_6 droidM_2 - \beta_7 droidM_3\end{aligned} \tag{7}$$

The coefficients for the independent variables are shown in Table 5. The effect size for *perm* and *trust* are similar, as the individual disagrees with the question more, they are more likely to be classified as a conservative (Figures 17 and 18 in Appendix A). The effect size for *droidM* is largest for those that will not consider the new permissions model as more individuals are predicted to be unconcerned (Figure 20). When considering whether individuals have changed their mind when installing an app, the largest effect size is for the only non-significant answer (*2 – Not Sure*). While the predicted probability of being a fence sitter does not change between the other answers for this question, individ-

uals are slightly more likely to be conservative if they have not changed their mind (Figure 19 in Appendix A).

As seen in Figure 7, participants in less conservative privacy clusters are more likely to believe that well known companies are trustworthy. Similarly, as the privacy cluster becomes less conservative, the fraction of individuals that will accept additional permissions when updating also increases (Figure 8). The survey also shows as the privacy cluster becomes less conservative, participants are less likely to have changed their mind after installing an app. Additionally, the highest fraction of unconcerned individuals are not sure whether they have changed their mind (Figure 9). While a large fraction of the respondents identifying as each of the privacy clusters will consider permissions in the Android M permission model (10), the fraction is slightly smaller for unconcerned individuals.

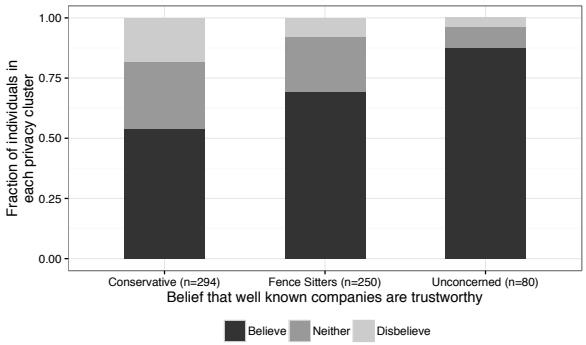

Figure 7: Distribution of belief that well known companies are trustworthy across privacy clusters. The fraction of individuals that trust well-known app companies increases significantly as the privacy cluster becomes less conservative.

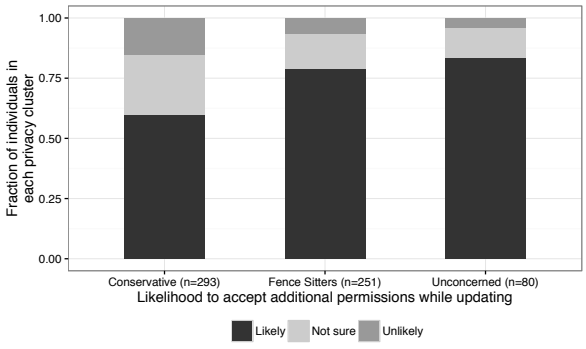

Figure 8: Distribution of the likelihood to accept additional permissions when updating across privacy clusters. The fraction of individuals likely to accept additional permissions requests is significantly lower in the conservative cluster than in either the unconcerned or fence-sitter clusters.

4.5 Other Results

Besides understanding the predictive power of various factors for user privacy clusters we also find several drivers of app installation decisions. In particular, app placement in search results in an app store is considered as a factor for

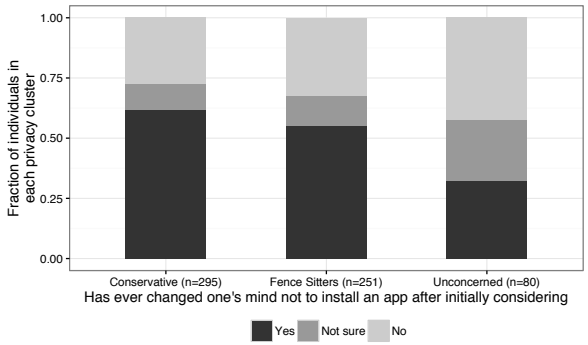

Figure 9: Distribution of participants who have changed their minds about installing an app across privacy clusters. The fraction of participants who have changed their minds is significantly less in the unconcerned cluster than in either the conservative or fence-sitter clusters.

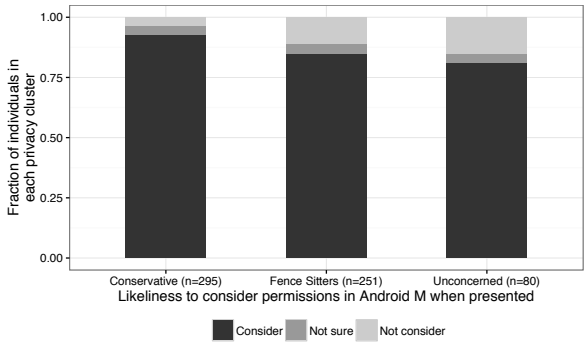

Figure 10: Distribution across the privacy clusters of the likelihood of considering permissions as presented in Android M's permission model. The fraction of participants likely to consider permissions is significantly less in the conservative cluster than in either the fence-sitter or unconcerned clusters.

installation by a large minority of our sample: 37% of people report to decide against installing an app if the app does not show on the first page of the search results; 32.6% of participants would still consider installing such apps.

We also find that a significant proportion (>40%) believe that similar apps request the same permissions, which may discourage attention to permission requests.

Trust in well-known companies is high across our sample; a majority of users (>60%) report that apps developed by well known companies are trustworthy and more than 55% of users trust well-known companies with their personal data as compared to just about 18% who trust lesser known companies with their data as well.

Finally, a large minority of our sample (>40%) report to be willing to install apps that request for one or more permissions that are unrelated to the app's functionality.

5. DISCUSSION

The over-arching goal of our study is to provide a deeper understanding of the privacy clusters in [30] (conservative, fence-sitters and unconcerned). As mentioned previously,

our sample differs from [30] in the distribution of the clusters. In particular, while our fence-sitter proportion is similar (0.4 in our sample, versus almost 0.5 in [30]), we differ significantly in the conservative proportion (0.47 in ours versus 0.12 in [30]) and unconcerned (0.13 in ours versus 0.23 in [30]). This may be due to the different survey content we employ (motivated by our goal of studying the connection between clusters and psychographics and other attributes) and due to the intervening time period. In particular, the data analyzed in [30] predates much of the Edward Snowden revelations (e.g., [29]) and other studies have shown that the Snowden events have had a strong impact on user privacy attitudes and behaviors (e.g., [32]).

Existing research may help interpret and gain confidence in some of our findings. For example, a large proportion of our sample is not attentive to app permissions requests which may be related to the user difficulty understanding permissions found in [21]. In addition, our results relating knowledge to privacy concern are similar to what has been found previously. In particular, [25] finds that Internet knowledge is associated with more perceived privacy risk and our conservative privacy cluster is more aware of malicious app capabilities.

That said, our findings differ from other studies in some important ways. In particular, while risk tolerance is found to be predictive of privacy attitudes in [17], we do not find it to be predictive of privacy clusters. However, in [17], privacy attitudes are measured by several general scales that are not specific to the mobile app context. Similarly, given that bad experiences with security updates have been shown to decrease willingness to install updates [35], it might be expected that bad app experience would predict a user's privacy cluster, which we do not find. However, the fact that more of the unconcerned users have *not* had a negative app experience, suggests it has some impact on privacy cluster.

We note that as with all survey-based studies, our reliance on self-reported data is a potential limitation. To mitigate this risk, our survey collects self-reported attitudes and behaviors *and* measures attitudes and behavioral intent through scenarios, a strategy that has proven successful in other studies (e.g. [36]).

Finally, our study does not take into account the usage of third party app stores or the number of apps users actively use. Understanding how these factors might affect the privacy clusters is an important open problem.

6. CONCLUSION

This paper adds to previous research on the connection between psychographics and security and privacy related behavior. We study this connection in the context of smartphone apps and identify personality aspects and user perceptions of app attributes that predict a user's privacy cluster. In particular, we find that decision-making style, knowledge of app privacy risk, attentiveness to app permissions and trust in app companies are all predictive of privacy cluster. Users in the conservative cluster are characterized by attentiveness to permissions and knowledge of privacy risks, whereas users in the other clusters are more likely to trust well-known app companies. Open questions include how stable the privacy clusters are in the presence of additional inputs (e.g., information about app privacy risks) and over time.

Therefore, if app developers are able to determine the psychographics of the users that might be drawn to their apps then they can infer the users' privacy cluster. These clusters can then be used in conjunction with other research ([30]) to aid in tailoring of apps towards the preferences of users.

Ethics

Our study was approved by North Carolina State University's Institutional Review Board (IRB).

7. REFERENCES

[1] 2 Billion Consumers Worldwide to Get Smart(phones) by 2016. http://www.emarketer.com/Article/2-Billion-Consumers-Worldwide-Smartphones-by-2016/1011694.

[2] A. Acquisti and J. Grossklags. Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 2(2005):24–30, 2005.

[3] A. Agresti. *Wiley Series in Probability and Statistics : Categorical Data Analysis*, chapter 8, pages 301–303. Wiley, Somerset, US, 3rd edition, 2013.

[4] H. Almuhimedi, F. Schaub, N. Sadeh, I. Adjerid, A. Acquisti, J. Gluck, L. F. Cranor, and Y. Agarwal. Your location has been shared 5,398 times!: A field study on mobile app privacy nudging. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 787–796. ACM, 2015.

[5] M. Balestra, O. Shaer, J. Okerlund, M. Ball, and O. Nov. The effect of exposure to social annotation on online informed consent beliefs and behavior. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2016.

[6] S. B. Barnes. A privacy paradox: Social networking in the united states. *First Monday*, 11(9), 2006.

[7] A.-R. Blais and E. U. Weber. A domain-specific risk-taking (dospert) scale for adult populations. *Judgment and Decision Making*, 1(1), 2006.

[8] E. Chin, A. P. Felt, V. Sekar, and D. Wagner. Measuring user confidence in smartphone security and privacy. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 1–16, 2012.

[9] R. H. B. Christensen. *Analysis of ordinal data with cumulative link models — estimation with the R-package ordinal*, June 2015.

[10] R. H. B. Christensen. *ordinal—Regression Models for Ordinal Data*, 2015. R package version 2015.6-28. http://www.cran.r-project.org/package=ordinal/.

[11] P. T. Costa and R. R. McCrae. The revised neo personality inventory (neo-pi-r). *The SAGE handbook of personality theory and assessment*, 2:179–198, 2008.

[12] S. Das, T. H.-J. Kim, L. A. Dabbish, and J. I. Hong. The effect of social influence on security sensitivity. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 143–157, 2014.

[13] S. Das, A. D. Kramer, L. A. Dabbish, and J. I. Hong. The role of social influence in security feature adoption. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1416–1426, 2015.

[14] S. Egelman, A. P. Felt, and D. Wagner. *The Economics of Information Security and Privacy*, chapter Choice architecture and smartphone privacy: there's a price for that, pages 211–236. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[15] S. Egelman and E. Peer. Predicting privacy and security attitudes. *ACM SIGCAS Computers and Society*, 45(1):22–28, 2015.

[16] S. Egelman and E. Peer. Scaling the security wall: Developing a security behavior intentions scale (sebis). In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2873–2882. ACM, 2015.

[17] S. Egelman and E. Peer. The myth of the average user: improving privacy and security systems through individualization. In *Proceedings of the 2015 New Security Paradigms Workshop*, pages 16–28, 2015.

[18] W. Enck, P. Gilbert, S. Han, V. Tendulkar, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems (TOCS)*, 32(2):5, 2014.

[19] A. P. Felt, S. Egelman, M. Finifter, D. Akhawe, D. Wagner, et al. How to ask for permission. In *HotSec*, 2012.

[20] A. P. Felt, S. Egelman, and D. Wagner. I've got 99 problems, but vibration ain't one: a suvey of smartphone users' concerns. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 33–44, 2012.

[21] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android permissions: User attention, comprehension, and behavior. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 1–14, 2012.

[22] C. Gibler, J. Crussell, J. Erickson, and H. Chen. Androidleaks: automatically detecting potential privacy leaks in android applications on a large scale. In *International Conference on Trust and Trustworthy Computing*, pages 291–307. Springer, 2012.

[23] I. A. Junglas, N. A. Johnson, and C. Spitzmüller. Personality traits and concern for privacy: an empirical study in the context of location-based services. *European Journal of Information Systems*, 17(4):387–402, 2008.

[24] R. Kang, S. Brown, L. Dabbish, and S. B. Kiesler. Privacy attitudes of mechanical turk workers and the us public. In *SOUPS*, pages 37–49, 2014.

[25] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler. "My data just goes everywhere:" user mental models of the internet and implications for privacy and security. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 39–52, 2015.

[26] P. G. Kelly, L. F. Cranor, and N. Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3393–3402, 2013.

[27] M. L. Korzaan and K. T. Boswell. The influence of personality traits and information privacy concerns on behavioral intentions. *Journal of Computer Information Systems*, 48(4):15–24, 2008.

[28] P. Kumaraguru and L. F. Cranor. Privacy indexes: a survey of westin's studies. *Technical report, Carnegie Mellon University*, 2005.

[29] S. Landau. Making sense from snowden. *IEEE Secur Priv*, 11(4):54–63, 2013.

[30] J. Lin, B. Liu, N. Sadeh, and J. I. Hong. Modeling users' mobile app privacy settings: restoring usability in a sea of permission settings. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 199–212, 2014.

[31] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.

[32] L. Rainie and M. Madden. Americans' privacy strategies post-snowden. *Pew Research Center*, 2015.

[33] S. G. Scott and R. A. Bruce. Decision-making style: The development and assessment of a new measure. *Educational and psychological measurement*, 55(5):818–831, 1995.

[34] I. Shklovski, S. D. Mainwaring, H. H. Skúladóttir, and H. Borgthorsson. Leakiness and creepiness in app space: perceptions of privacy and mobile app use. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI)*, pages 2347–2356, 2014.

[35] K. E. Vaniea, E. Rader, and R. Wash. Betrayed by updates: how negative experiences affect future security. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI)*, pages 2671–2674, 2014.

[36] A. Woodruff, V. Pihur, S. Consolvo, L. Schmidt, L. Brandimarte, and A. Acquisti. Would a Privacy Fundamentalist Sell Their DNA for $1000... If Nothing Bad Happened as a Result? The Westin Categories, Behavioral Intentions, and Consequences. In *Symposium on Usability of Privacy and Security (SOUPS)*, pages 1–18, 2014.

[37] H. Xu, S. Gupta, M. B. Rosson, and J. M. Carroll. Measuring mobile users' concerns for information privacy. *International Conference on Information Systems*, 2012.

[38] Z. Yang, M. Yang, Y. Zhang, G. Gu, P. Ning, and X. S. Wang. Appintent: Analyzing sensitive data transmission in android for privacy leakage detection. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pages 1043–1054. ACM, 2013.

[39] Y. Zhou and X. Jiang. Dissecting android malware: Characterization and evolution. In *2012 IEEE Symposium on Security and Privacy*, pages 95–109. IEEE, 2012.

Appendix A: Effects Size Analysis

This appendix contains bar plots showing the effect sizes of the estimated coefficients derived from each of the equations in section 4.

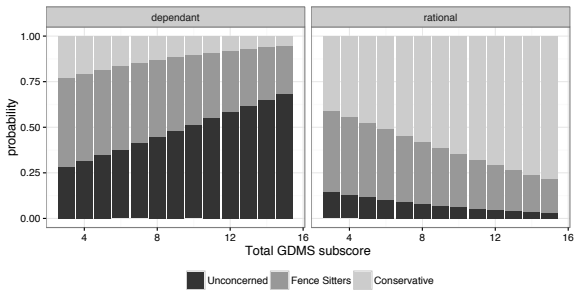

Figure 11: As measured by GDMS, users who are rational are significantly more likely to be conservative and users who are dependent are significantly more likely to be unconcerned.

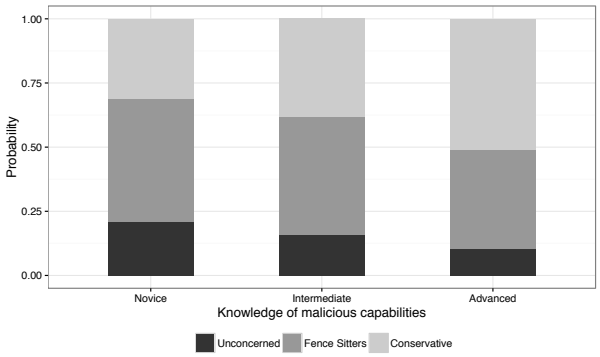

Figure 12: Users who can correctly identify more potentially malicious behaviors of apps are more likely to be conservative.

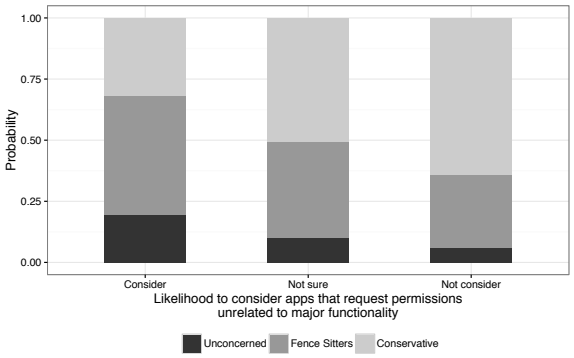

Figure 13: Individuals who are less likely to install an app if it makes permission requests unrelated to its major functionality are less likely to be unconcerned.

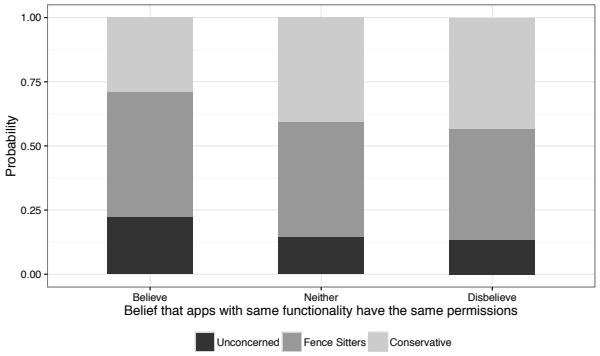

Figure 14: Individuals believing that apps with the same functionality will also have the same permissions are slightly more likely to be unconcerned and less likely to be conservative. The effect sizes with respect to fence sitters is small.

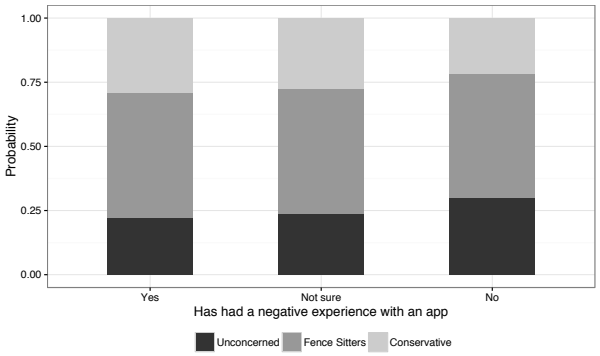

Figure 15: Individuals that have had negative experiences with apps are more likely to be conservative than those who have not had negative experiences..

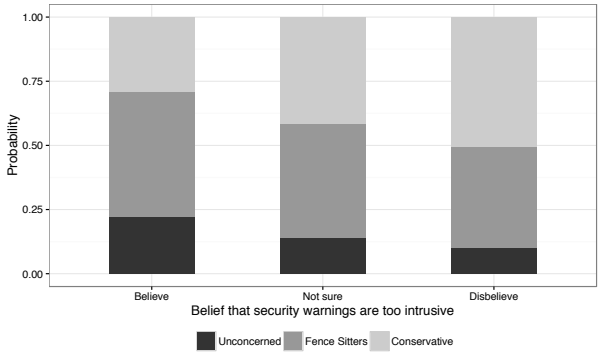

Figure 16: As individuals' belief that security warnings are too intrusive, their likelihood of being conservative is reduced while the likelihood of being unconcerned or a fence sitter increases.

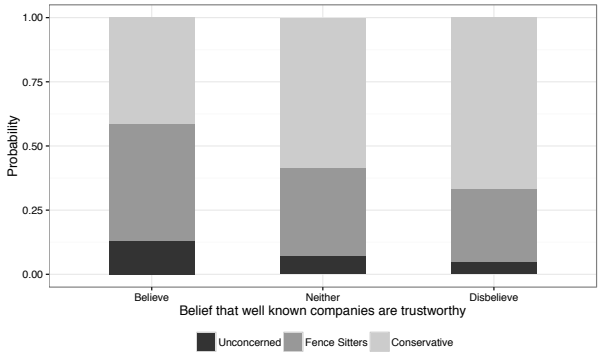

Figure 17: Individuals who trusted well known companies more were slightly more likely to be fence sitters. As the trust of these companies decreases, an individual is more likely to be conservative.

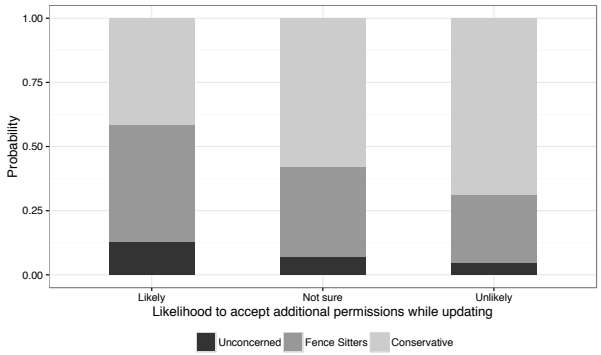

Figure 18: As individuals are less likely to accept additional permissions while updating apps, they are slightly less likely to be unconcerned and more likely to be conservative. Fence sitters represent the largest proportion of the individuals who are likely to accept additional permissions.

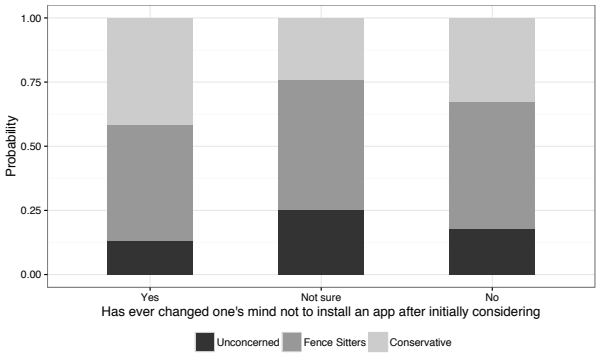

Figure 19: If an individual has changed their mind to install an app after initially considering it, they are more likely to be conservative as compared to those that have not changed their mind.

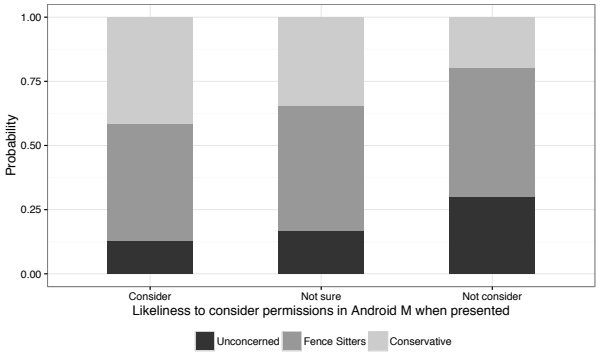

Figure 20: As individuals are less likely to consider each permission requested under Android M's permissions model, they are significantly more likely to be unconcerned.