Are privacy concerns a turn-off? Engagement and privacy in social networks

Jessica Staddon
staddon@google.com

David Huffaker
huffaker@google.com

Larkin Brown
larkinbrown@google.com

Aaron Sedley
asedley@google.com

ABSTRACT

We describe the survey results from a representative sample of 1,075 U.S. social network users who use Facebook as their primary network. Our results show a strong association between low engagement and privacy concern. Specifically, users who report concerns around sharing control, comprehension of sharing practices or general Facebook privacy concern, also report consistently less time spent as well as less (self-reported) posting, commenting and "Like"ing of content. The limited evidence of other significant differences between engaged users and others suggests that privacy-related concerns may be an important gate to engagement. Indeed, privacy concern and network size are the only malleable attributes that we find to have significant association with engagement. We manually categorize the privacy concerns finding that many are nonspecific and not associated with negative personal experiences. Finally, we identify some education and utility issues associated with low social network activity, suggesting avenues for increasing engagement amongst current users.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Security, Human Factors, Measurement

Keywords

privacy, social networks, control, transparency

1. INTRODUCTION

The challenges of privacy in social networks are well-known. Surveys help us gauge the rate of privacy concern by demographic segments (e.g. [7, 13]) and deep qualitative and quantitative research captures the negative privacy experiences (e.g. [30]) and nuanced privacy attitudes (e.g. [4]) of users. Others approach social network privacy from the behavioral side, finding significant relationships between privacy-related actions (e.g. [25]).

Less is understood about the link between the two: privacy concern and experience *and* the behaviors of the users reporting concerns. Clearly, each is very difficult to measure through self-report as we do here. Privacy concern is subject to self-report bias and hence, measuring it is an active area of research (e.g. [17]). In addition, the challenges of measuring behavior through self-report are well-known (e.g. [29, 9]). We attempt to moderate bias in the former by asking about several different aspects of privacy (e.g. comprehension, control, perception of others) and argue that even the *perception* on the part of a user that they are less engaged is important and suggests there is room for improvement in the user's experience with the service.

Our contribution toward understanding this link is a survey of $1,363$ users selected randomly from a representative panel of social network users residing in the U. S. [11]. Because our sample is dominated by $1,075$ users who regard Facebook as their primary social network, we focus on this subset, thus controlling for answer variations due to social network differences. However, we emphasize that the survey is *not* specific to Facebook and the findings may apply to other networks.

We consider both overall social network privacy concern and aspects of concern related to transparency and control, specifically, comprehension of information sharing in the network, control over information sharing in the network, and sharing practices of the user in relation to their friends in the network (all survey questions are in the Appendix). We find that each aspect of privacy concern is strongly associated with self-reported engagement across several measures, including visit frequency, comment frequency and frequency of "Like"ing content. Specifically, users who report higher concern are less engaged. Similarly, users who perceive their friends as sharing more *personal* information are less engaged (an interesting counterpart to [5] in which sharing by friends increases engagement). In contrast, those who report more control and comprehension over sharing of their information in the network, are more engaged.[1]

Importantly, these relationships with engagement gener-

Symposium on Usable Privacy and Security (SOUPS) 2012, July 11-13, 2012, Washington, DC, USA.

[1]For ease of exposition, we do not repeat the term "report" throughout this paper. That is, for example, we say "users are concerned" as shorthand for "users report to be concerned".

ally persist when we control for time spent on the network. For example, 21% of the 379 moderately concerned users who visit a few times a week or more never post personal photos, versus 12% of the 426 slightly concerned users who visit a few times a week or more.

In addition, we find few demographic differences between high engagement and low engagement users. Privacy concern and network size (with low engagement-users tending to have smaller networks) are the main significant differences, suggesting privacy is an important hurdle to social network engagement.

One opportunity for surmounting the privacy hurdle appears to be education. In some contexts, privacy-concerned users may not be aware of features that may address their concerns. For example, more than 40% of users reporting little control over the sharing of their information in the network, do not limit the visibility of any of the profile fields surveyed (picture, birth date, phone number, home address, residence city, email address, gender, relationship status, and interests/hobbies). Since we find privacy-concerned users are less engaged, this lack of use of visibility controls is compatible with the finding of boyd-Hargittai [4] that engaged users are more likely to adjust privacy settings.

In addition, we find evidence of the importance of "service-sanctioned" controls. That is, work-arounds may serve the immediate privacy goal, but do not contribute significantly to perceptions of control. For example, those who use nicknames or fake names in Facebook do not report significantly more control over the sharing of their information in the network, and on average they report less control, than other users.

1.1 Related work

There are a number of research themes related to our survey. The foundational underpinnings to privacy provided by Westin (e.g. [12]) and Altman (e.g. [2]) inform our goal of understanding the nature of reported social network privacy concern both in terms of issues to address and survey design. In addition, there have been numerous overlapping surveys (e.g. the Pew Internet series) and novel approaches to modeling social network privacy. We highlight some of the most related efforts in each theme in the following.

Models of privacy in social networks. Our work is perhaps closest to efforts to model privacy-related behaviors and attitudes toward understanding cause and effect. We highlight several contributions that are closely related to ours; each examines privacy settings and their connection to attitudes and behaviors.

Limiting visibility of profiles is found to be a boundary management tool for weak ties in [26]. Building on this [25] finds that users who report to have customized privacy settings tend to disclose more and users who have read more of a site's privacy policy tend to disclose less.

Privacy settings modifications are used as a proxy for privacy concern in [4]. boyd and Hargittai [4] provide a longitudinal study of privacy practices and attitudes of teenagers (specifically, 18-19 year olds). They find significant behavioral evidence of privacy concern in this age group, in contrast to popular opinion, and find that engaged users are more likely to change privacy settings. Homophily-like drivers for privacy settings modifications are found in [14] in addition to gender-based differences for privacy settings (with women limiting profile visibility more).

Table 1: Demographic background of survey respondents.

Gender	
Female	58%
Male	42%
Age	
18-29	23.3%
30-44	26.8%
45-59	30.5%
60+	20.4%
Ethnicity	
White, Non-Hispanic	75.4%
Black, Non-Hispanic	8.8%
Other, Non-Hispanic	3.4%
Hispanic	9.7%
2+ Races, Non-Hispanic	2.6%
Education	
Less than high school	7.3%
High school	27.6%
Some college	33.9%
Bachelor's degree or higher	31.3%

Acquisti and Gross [1] find discrepancies between reported privacy concerns and privacy-related behaviors amongst Facebook users.

In contrast to each of these, we focus more on associations between *reported* concerns and broad engagement metrics, as our goal is to understand the users who report privacy and how they are interacting with the social network service in both privacy and non-privacy related ways. We show these users have consistently low engagement with the system, and do not differ from engaged users in many other respects, thus suggesting that addressing privacy concern may be necessary to activity.

Our work is perhaps closest to efforts to predict privacy (e.g. [31]) however in contrast to previous work we focus specifically on social network privacy and look at several (related) aspects of social network privacy concern.

Privacy Surveys. A number of very valuable surveys related to social networks and privacy already exist (e.g. [13, 15, 21, 28]. These surveys gauge percentages of specific social network behaviors, user attributes and attitudes. Our survey differs from these in that we try to understand *associations* between attitudes and behaviors/attributes; that is, we look at what actions, attributes and beliefs are associated with f report privacy concerns to better understand the nature of those concerns.

That said, it is worth summarizing how our percentages compare with the most closely related numbers in existing surveys. Our findings are consistent with many of the Pew Internet surveys. In particular, [13] finds that 58% of social network users have restricted access to their entire profile and [15] finds that 58% of social network users have restricted access to parts of their profile. We find that the same percentage of Facebook users have restricted access to at least one profile field.

Also, [13] reports that 17% of users have more than one account on a social networking site, and we find that 10% of Facebook users have more than one account; we do not find

significant differences by age or gender.

Regret and other negative experiences are also reported on in [15], in particular, 11% of SNS users have posted content they regret. We find that 6% of Facebook users have posted to a wider audience than intended, with no significant differences by gender or age. In addition, [21] looks at negative outcomes from social network use, and reports that 26% have experience "bad outcomes". We look at the narrower question of bad outcomes due to profile fields and find lower numbers overall; the percentages are less than 10% with the exception of negative experiences from phone number (.17%), home address (16%) and email (12%).

Table 2: Pearson correlations between questions representing various control and transparency aspects of privacy.

	$Q22$	$Q22a$	$Q21$	$Q20$	$Q19$
Facebook Privacy Concern ($Q22$)	—	$.68^{***}$	$-.09^{**}$	$-.12^{***}$	$-.14^{***}$
InternetPrivacy Concern ($Q22a$)		—	$-.07^{*}$	$-.06$	$-.14^{***}$
Sharing Comprehension ($Q21$)			—	$.43^{***}$	$.14^{***}$
Sharing Control ($Q20$)				—	$.15^{***}$
Relative Sharing ($Q19$)					—

Note. *p<.05, **p<.01, ***p<.001.

Privacy Theory. Starting with the work of Alan Westin [12] (and continuing more recently with Palen-Dourish [19] and Margulis [16] among others) there has been a substantial effort to characterize, or categorize, users according to their privacy concerns. Irwin Altman [2] also made significant contibutions by initiating a theory of privacy processes, with a focus on social interactions. Our work overlaps with these lines of research in that we are studying specific aspects of privacy (overall concerns, control and sharing) and how they are connected to difference behaviors and attributes, in the specific context of social networking and Facebook. We are interested in more complete characterizations of users who report privacy concerns in the social network domain.

Overview. This paper is organized as follows. We begin in Section 2 with a discussion of our methodology. Section 3 summarizes the survey results, describes our content analysis and presents significant associations between survey variables. We discuss our results and conclude in Section 4. Survey questions are in the appendix.

2. METHODOLOGY

2.1 Survey

In order to examine attitudes regarding privacy and social networks, we rely on survey methodology deployed by Knowledge Networks (KN). KN uses a probability-based panel to create a representative sample of the United States population. Survey respondents are screened based on their social media use, and complete the survey through a web site (97.7% completion rate). Given the selection process, we estimate the margin of error at under 3% with 95% confidence. The complete questionnaire is listed in the Appendix.

Our sample consists of 1,075 respondents who report that Facebook is their primary social network. Table 1 shows the demographic make-up of the sample.

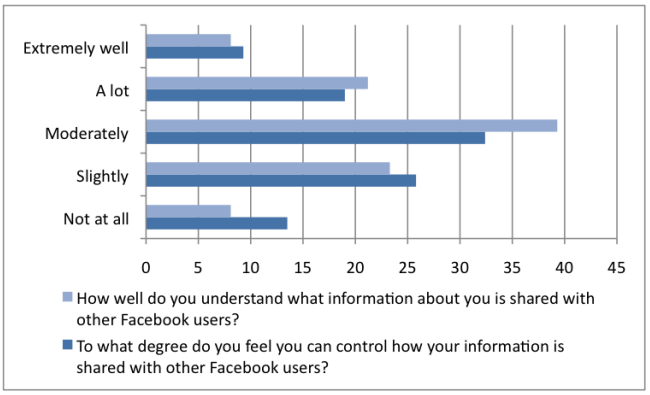

Figure 1: Degree of understanding and control of how information is shared to other Facebook users as a percentage of the responding participants (1060 for $Q20$ and 1061 for $Q21$). Differences are significant at the .01 level.

We look at associations with self-reported Facebook privacy concern ($Q22$) and Internet privacy concern ($Q22a$) as well as 3 questions related to information transparency and control: control over information sharing within Facebook ($Q20$), sharing of personal information in Facebook ($Q19$), and comprehension of information sharing in Facebook ($Q21$). The direct privacy questions, $Q22$ and $Q22a$, were coded so that high levels of concern correspond to larger integers. Similarly, the coded numbers increase with higher levels of comprehension in $Q22$, higher levels of control in $Q20$, and higher levels of sharing in $Q19$.

The complete survey is in the appendix, but we repeat these questions here since they are key to the analysis:

- ($Q22$) How do you feel about your privacy with regard to Facebook?
  - Not at all concerned
  - Slightly concerned
  - Moderately concerned
  - Very concerned
  - Extremely concerned
- ($Q22a$) How do you feel about your privacy with regard to the Internet overall?
  - *Same answer options as* $Q22$
- ($Q20$) To what degree do you feel you can control how your information is shared with other Facebook users?
  - Cannot control at all
  - Can control a little
  - Can control a moderate amount
  - Can control a lot

– Can control a great deal

- ($Q19$) Compared to the rest of the people in my network, I share personal information:
  - Much less
  - Somewhat less
  - About as much
  - Somewhat more
  - Much more
- ($Q21$) How well do you understand what information about you is shared with other Facebook users?
  - Don't understand at all
  - Understand slightly
  - Understand moderately well
  - Understand very well
  - Understand extremely well

The aspects of privacy represented by these questions are intuitively related, for example, as feelings of control and comprehension increase one would expect overall privacy concern to decrease. We find that this is the case; all the correlations are provided in Table 2.

2.2 Content Analysis

To gather valid data on user social network privacy concerns and consequences, our survey includes several open-ended text boxes for respondents to tell us directly about their fears and perceived negative outcomes [$Q23$: "What are your main privacy concerns online?", $Q24$: "What potential negative consequences are there from the concerns you mentioned in your answers to question $Q23$?"]. This method avoids aiding or biasing respondents with predetermined answer choices, and generates top-of-mind, salient responses. Two questions are used in order to both invite top of mind online privacy concerns and probe the specific consequences stemming from the reported concerns. This allows us to better understand how respondents reason through the logical ramifications of their privacy concerns and how their expressed fears might evolve when urged to reflect on the consequences.

We rely on hand-coded content analysis to complement our quantitative analysis. The content analysis utilizes a grounded theory approach [23] to develop common themes in the open response. Our codes reflect the most specific level of comment provided by each respondent, without assuming any details that are not explicitly mentioned. For example, when a respondent lists identity theft in their answer to [$Q23$: "What is your main privacy concern online"], we do not presume the probable repercussion of financial loss unless it is specifically mentioned by the respondent. Similarly, when a respondent voices concerns about access to their personal information, we code it as Access to Personal Data, while we reserve the Misuse of Personal Data coding for responses that mention this conventional implication. After coding the responses as narrowly as possible, we cluster the categories and identify a theme for each cluster. We utilize a second coder to measure agreement. The inter-coder agreement rate is 71.4%.

3. RESULTS

Our survey begins with questions regarding user engagement on Facebook, including how often they engage in certain activities. Table 3 shows that most respondents report a high level of using Facebook (61% visit at least once a day). Although respondents report a high level of usage, they also report low levels of creating content via posts, photos and comments. This points to the prevalence of 'lurker' style consumption patterns in online communities [18].

Table 5: Privacy concerns associated with specific consequences.

	Main Privacy Concern	Potential Negative Consequences?
Financial		
Indentity Theft	40%	35%
Financial Loss	11%	23%
TOTAL	51%	58%
Digital World		
Access to Personal Data	14%	8%
Account Hacking	11%	3%
Misuse of Personal Data	5%	2%
Unwanted Solicitations/Spam	3%	6%
Social Ramifications	3%	3%
Computer Virus	2%	2%
Unwanted Ad Targeting	1%	2%
TOTAL	42%	26%
Physical World		
Offline Threats	6%	5%
Harm to Family	2%	2%
Stalkers	1%	3%
Employment Risks	0.3%	2%
Hassle to Recover	0%	4%
TOTAL	9%	15%

We also focus on privacy concern and understanding. As a starting point, we ask how often users change their account security or privacy settings. Most respondents (72.9%) report changing their settings at least once, but these changes appear to occur sparingly (45% reported making a change less than once a month).

We ask about the kinds of information respondents provide on Facebook. As Table 4 shows, the more personal the information gets (e.g., home address), the lower the probability of being shared. Interestingly, very few users provide fake information instead.

Similarly, we ask respondents how much personal information they share compared to the rest of the people in their network. 44.8% report sharing much less, 28.5% report sharing somewhat less, and 23.6% report sharing about as much. Only about 3% report sharing more.

Concerning real name and pseudonym usage. 67.1% report using their full name (e.g. Bob Kawalski); 11.6% report using a first name only (e.g. Bob); 16.3% reported using a pseudonym or Nickname (e..g DreamWeaver21). 4.8% reported using a fake or made-up name (e.g. Joe Smith). 10.2% of the respondents also report having more than one account on Facebook.

We ask several questions about a respondent's degree of

Table 3: Self-reported Facebook usage and activities.

How Often...	Visit Facebook	Post a Status Update	Post a Photo	Comment on a Post	Like or +1 a Post
Multiple times a day	35.4%	3.7%	1.4%	10.3%	10.5%
About once a day	25.6%	6.0%	1.4%	10.3%	10.5%
A few times a week	15.1%	13.9%	4.7%	20.3%	20.7%
About once a week	8.1%	9.9%	4.9%	12.1%	8.8%
A few times a month	7.1%	15.2%	13.2%	16.9%	13.8%
About once a month	4.7%	11.5%	13.2%	8.4%	6.5%
Less than once a month	3.8%	22.5%	37.8%	14.6%	12.7%
Never		17.2%	23.4%	7.1%	16.5%

Table 4: Types of information provided to Facebook.

	My Picture	Birth Date	Phone Number	Home Address	Email	Gender	Relationship Status	Interests or Hobbies
Do provide.	62.7%	50.5%	7.3%	5.1%	25.2%	69.2%	54.3%	39.1%
Do provide but not everyone can see it.	22.5%	29.8%	14.4%	7.8%	36.6%	23.6%	24.2%	30.0%
Do provide but information is fake.	<1%	1.8%	<1%	1.0%	<1%	<1%	<1%	<1%
Do not provide.	14.0%	17.9%	77.4%	85.9%	37.2%	6.5%	20.8%	30.1%

Table 6: Privacy concerns associated with broad fears and meta attitudes.

	Main Privacy Concern	Potential Negative Consequences?
General Fears		
Unspecified Anxiety	11%	10%
Individuals with Ill Intent	2%	1%
Digital and Physical Theft	2%	6%
Fear of Big Brother Govt́	1%	1%
Potential Harrassment	1%	1%
TOTAL	17%	21%
Meta Attitudes		
No Concerns	5%	6%
Resigned to Minimal Privacy	2%	1%
Privacy Oriented	0.3%	2%
TOTAL	8%	9%

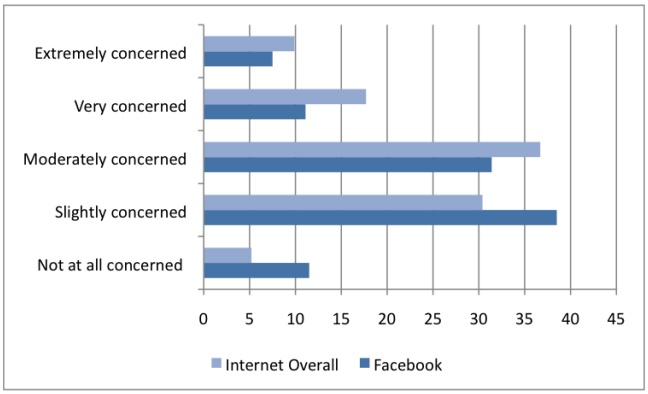

Figure 2: How do you feel about your privacy with regard to Facebook and the Internet Overall? Answers as a percentage of the 1061 respondents. Differences are significant at the .001 level.

control and understanding with regards to how their information is shared with other Facebook users. As shown in Figure 1, we see a normal distribution across all levels for both questions.

Finally, we ask how respondents feel about their privacy with regard to Facebook and the Internet overall. As shown in Figure 2, we see similar patterns in the answers for both, with most respondents slightly or moderately concerned.

We find little difference between the demographics of users with at least moderate privacy concern and all others. In particular, there are no statistically significant differences based on education level, gender, income level, marital status, work status, household size and home ownership status.

Significant differences do exist by age where most of the variation occurs between the 18-24 and 55-64 demographics, with the older users reporting more concern (Pearson correlation between age and privacy concern is .12***). Similarly, there is a significant negative association between feelings of control and age ($r = -.24$, $p < .001$) as there is with reported comprehension levels ($r = -.3$, $p < .001$). Finally, older users are more likely to think their friends share more personal information ($r = -.18$, $p < .001$) as do less experienced Facebook users ($r = -.21$, $p < .001$). Strongly significant differences also exist for different ethnicities. We find heightened concern levels of Hispanics and Blacks in com-

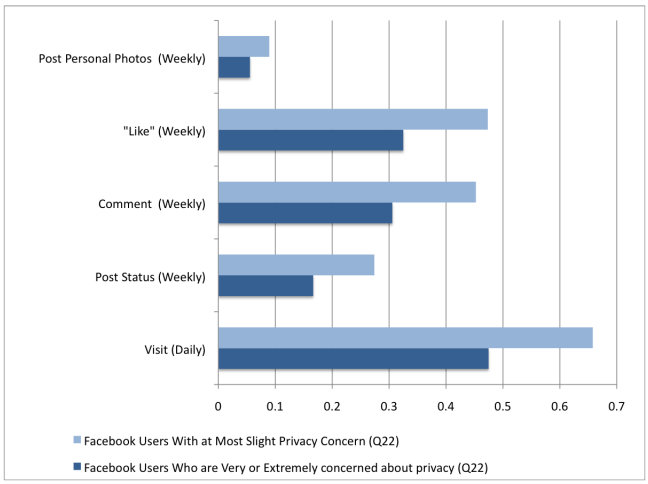

Figure 3: Percentages of users reporting certain behaviors, grouped by level of privacy concern. Users who report to be very or extrememly concerned about Facebook privacy also report less engagement across several metrics. The differences are all significant at the .001 level. There are 198 users reporting to be very or extremely concerned, and 530 who report to be at most slightly concerned. Most of the users in each group answered each of the behavioral questions shown.

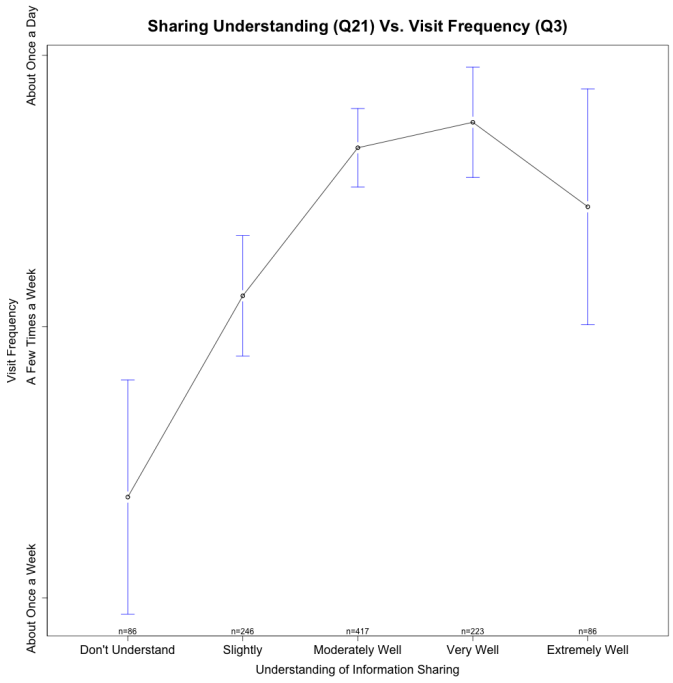

Figure 4: Users who understand how their information is shared in Facebook ($Q21$) tend to visit Facebook more. 95% confidence intervals for the means are shown.

parison to Whites. Differences between Blacks and Whites are significant at the .001 level, and differences between Hispanics and Whites are weakly significant ($p = .053$).

Reported comprehension also varies significantly by ethnicity. For example, Whites and Hispanics differ at the .05 level ($p = .009$ with TukeyHSD); with .41 of Hispanics reporting to "Understand very well" or "Understand extremely well", versus .26 of Whites.

3.1 The Nature of Privacy Concern

Based on manual coding of the 8924 responses to $Q23$ and the 841 responses to $Q24$, identity theft was the most commonly cited speciïňĄc concern overall. 40% of respondents express concern about identity theft; for example, "I'm concerned somebody could find my personal information and steal my identity." In comparison, the next most cited category was access to personal data, indicated by 14%, as expressed by one respondent, "I am concerned that people have access to info about me that I would not knowingly like them to have." Three categories followed with 11% respondents each: unspecified anxieties (fear without a specific concern, e.g. "It is just not good to provide such personal information about yourself. You never know what could happen."), hacking (unauthorized digital account break-ins, e.g. "That hackers can use my personal information for evil purposes"), and financial loss (e.g. "people getting money out of my accounts").

From our set of 23 codes, we manually cluster 5 meta-categories of concerns. These are financial, digi- tal world, physical world, meta-attitudes, and general fears. Financial and digital world concerns were each speciïňĄed by about half of respondents, while other meta categories had much lower incidence.

With a follow-up to $Q23$, we probe the potential consequences of the concerns mentioned by participants in $Q23$. Identity theft is the most-cited potential consequence, 35% of respondents. Fnancial loss is mentioned by 23% and unspeciïňĄed anxieties by 10%.

In the meta-category distribution, fewer respondents note digital world consequences than digital world concerns (38% less frequently), thus more respondents cite concerns across all other categories. The fact that digital world themes are the only meta-category to diminish may suggest a framework that people use to assess their privacy concerns. It reflects that while the online space is a catalyst for privacy concerns, the ultimate consequences are more often realized outside this milieu and are gateway to malfeasance in their real lives that people most fear. By addressing the catalyst, we may simultaneously lessen the fear of the consequences brought on by perceived weaknesses in digital privacy.

We did not try to correlate particular consequences with specific concerns, but did notice some qualitative trends. Not all respondents who initially express concrete concerns can describe concrete ramifications. Instead their concerns augment into unspecified anxiety. For example one respondent concretely stated their primary concern online as "ID theft" yet broadly articulated the potential consequence of identity theft as "Never know what people will do with your personal information." In a similar pattern, a respondent noted tangible concern of "Just plain old theft using your info to obtain items or money etc." yet described the consequence with the dramatically abstract response of "Destroying your life".

This pattern of response suggests that some people can express a fear more definitively than they can imagine the

plight of the victim. To accurately predict the potential risks of a privacy breach the user would likely need first hand experience or substantial domain knowledge. Lacking both of these traits, the user can remain in an anxiety loop. Despite having specific concerns, without an understanding of the potential consequences of those concerns, they are unable to identify tools and actions to mitigate them. Therefore to successfully build privacy understanding and trust, our education and policies would benefit by addressing both levels of fears: online privacy concerns and their widely interpreted consequences.

The sentiments in the free text responses are compatible with other survey responses. For example, only 6% of users report sharing accidents stemming from one of their posts in Facebook ($Q14\_5$). There is also a low rate of negative experiences connected to profile fields amongst those who provide content in those fields, but also a high rate of uncertainty around whether the fields could have led to positive or negative experiences. Table 7 summarizes the results.

Compatible with Tables 5 and 6, most reported privacy concern with sharing profile information is focused on Facebook's access to the information and the potential for identity theft. Users have little concern with being embarrassed in front of other users by profile information. Table 8 presents results for birth date and city of residence results; both are typical of profile demographics.

3.2 The Relationship between Privacy Concern and User Engagement

For each of the social network-specific privacy aspects, $Q22$, $Q21$, $Q20$ and $Q19$, we find significant associations with several engagement metrics including frequency of visiting, posting, commenting and "Like"ing content. The frequency of each generally increases as privacy concern decreases. Similarly, the engagement metrics generally increase with sharing comprehension, sharing control and sharing of personal information relative to others. We present all the Pearson correlations in Table 9. In this section we look more closely at each privacy aspect and its association with engagement, providing numerical and visual evidence for each.

Figure 3 provides a concrete example of the relationship between privacy and engagement for the general privacy question, $Q22$. Activity levels are shown for users very or extremely concerned about privacy ($n = 198$) and those with at most "slight" privacy concern ($n = 530$).

Figure 4 provides another illustration of the relationship between privacy and engagement, for $Q21$. We see that users who report to understand better how their information is shared in Facebook are more likely visit Facebook frequently.

Importantly, these associations with engagement generally persist when we control for visit frequency. That is, if we look at users who report to visit at about the same rate, then their frequency of posting, commenting, "Like"ing, etc., varies according to their reported privacy concerns as described above. We find users who visit daily and report to share personal information about as much as their friends show consistently more engagement than those who also visit several times a week but share somewhat less personal information than their friends. We see similar results for Q20, Q21, and Q22.

A linear model also shows that the significant association between the privacy aspects and posting ($Q7\_01$) persists when we control for visit frequency. Table 10 shows how frequency of posting by weekly visitors can be predicted from the privacy aspects $Q19$, $Q20$, $Q21$ and $Q22$ of weekly visitors.

Finally, Table 11 shows how the low engagement and high engagement users tend to cluster according to the privacy aspects. In particular, those who do not post status announcements report low control and low relative sharing, whereas the active posters report greater control and relative sharing.

Privacy concerns are also closely associated with an *indirect* indication of engagement, Facebook network size. Reported privacy concern ($Q22$) is negatively associated with Facebook network size ($r = -.1$, $p < .001$), feelings of control ($Q20$) are positively associated ($.12$, $p < .001$), sharing comprehension ($Q21$) is positively associated ($r = .17$, $p < .001$) and the perception that the user is sharing personal information at a rate on par or exceeding their friends is positively associated ($r = .15$, $p < .001$, $Q19$).

3.3 Profile Visibility and Restriction

Facebook and other social networks provide controls over the visibility of profile data. We find that increased feelings of control are associated with increased use of these features. We find differences between users reporting moderate and little control are significant at the .001 level in terms of visibility restriction for various profile fields such as phone, email and interests; at the .01 level for address and relationship status; and at the .05 level for picture. Differences for birth date and city are only weakly significant.

Users who perceive their friends share more personal information are less likely to use their real names in FB. For example 34% of those who report they share "somewhat less" or "much less" than their friends, do not use their real name. In contrast 26% of those who share "about as much", "somewhat more" or "much more" than their friends, use their real name in Facebook.

Interestingly, users who choose to use nicknames or fake names in Facebook do *not* report to have significantly more control in Facebook than others, and on average they report less control. This is in contrast to the profile visibility controls and may reflect the fact that nicknames and fake names are not an officially sanctioned control option, but rather, are discouraged by Facebook.

Similarly, *withholding* profile information (i.e. not providing it) is also associated with significantly less feelings of control except for gender (which more than 69% of users make publicly visible) and home address (which 86% of users withhold). In addition, users who report to have low control, often do not use visibility controls – more than 40% of the users reporting little control do not limit the visibility of any of the profile fields in $Q15$. We also note that of the users who choose to either withhold, make public or provide fake information, the most popular approach is a combination of public and withheld (86%)

4. DISCUSSION AND CONCLUSION

We have shown that self-reported activity in Facebook is significantly different for users who do and users who do not express privacy concerns. Users with privacy concern report to be less engaged across a variety of metrics, even when controlling for time spent using Facebook. While such a result may be intuitively reasonable, we are not aware of

Table 7: Percentage of users reporting a positive or negative consequence related to providing profile information.

Profile Field	Always or Sometimes Positive	Sometimes or Always Negative	Don't Know
My Picture (N = 901)	66.0%	9.2%	24.8%
My Birth Date (N = 853)	66.7%	5.7%	2.8%
My Phone Number (N = 227)	48.5%	3.9%	47.6%
My Home Address (N = 137)	48.2%	4.3%	47.5%
My City of Residence (N = 820)	52.9%	1.7%	45.4%
My Email (N = 648)	43.4%	12.2%	44.4%
My Gender (N = 978)	44.5%	8.1%	47.8%
My Relationship Status (N = 828)	50.4%	9.8%	39.9%
My Interests/Hobbies (N = 725)	53.2%	9.1%	37.7%

Table 8: Reported reasons for limiting access to birthdate in profiles ($Q18\_2$) and for withholding birthdate entirely ($Q17\_2$).

Reason	Birthdate (n = 317)	City of Residence
Limiting Access		
The information is only interesting to particular people	42%	24%
The information will only make sense to certain people	19%	6%
The information could be embarrassing if shared more broadly	5%	1%
There is a privacy/security risk to sharing the information	37%	58%
As a rule I restrict access to such information unless it's required	42%	50%
Withholding Entirely		
It was a hassle to enter it	2%	0%
I had technical problems entering the information	1%	1%
I didn't think the information would make for a better experience on the network	17%	21%
I withheld it because it could be embarrassing	3%	2%
I withheld it because of concerns about impersonation/identity theft	36%	37%
I felt uncomfortable providing it to the service provider	14%	26%
As a rule I withhold such information unless it's required to provide it	50%	64%

any large-scale studies demonstrating this association.

Of course, our study can only show a strong negative *association* between privacy concern and engagement; it is not designed to determine whether the relationship is causal. In the following subsections we discuss some of the potential explanations for this association and ways to investigate causality. We conclude with specific design implications stemming from our work.

4.1 Limitations

The relationship between privacy concern and engagement is likely complex. Users with privacy concern may interact with Facebook in sufficiently different ways than other users to impact their knowledge of the service, which itself may hinder engagement. Compatible with this, we find a negative association between self-reported comprehension and data control, and privacy concern.

Fully understanding the complexity of the relationship may not be possible within the constraints of a survey as the potential mediating factors are difficult to enumerate. In addition, causality tests benefit from longitudinal study, in which changes in concerns/attitudes in specific individuals can be tracked along with the behaviors of the individuals.

While additional qualitative and longitudinal research is needed to flesh out the relationship between privacy concern and engagement, we do not believe it is needed to determine the *importance* of the association. The strength of the association is itself evidence that privacy-concerned users are worth paying attention to as a group, regardless of the intricacies of the connection between the privacy-concerned state and low engagement.

An additional important attribute of our research is that it is entirely based on self-reports. We make no claims about associations between privacy concern and actual behavior. That said, the simple fact that a user *reports* to be weakly engaged is noteworthy and clearly not a desired user state from the point of view of a service provider.

4.2 Design Implications

Our work suggests three opportunities for improving privacy and engagement in social networks.

Education. As mentioned earlier, although we did not explicitly ask about awareness of profile visibility controls, there is evidence suggesting that many users who would prefer such controls are not aware of them. For example, almost 39% of users do not employ any profile visibility controls and 44% percent of these report having at most little control over how their information is shared in Facebook. Indeed, use of profile visibility controls is associated with significantly higher engagement for all the metrics ($Q7\_02$-$Q7\_07$, $p < .0001$) even when restricting to those users who visit Facebook at least a few times a week. We see no evidence

Table 9: A summary of the associations between various aspects of privacy concern and engagement. "NA" indicates that the association is not significant, with the exception of the association between $Q22$ and $Q7\_04$ which is weakly significant ($\mathbf{r = -.06}$, $p = .07$,).

	Visiting ($Q3$)	Posting ($Q7\_01$)	Commenting ($Q7\_02$)	"Like"ing ($Q7\_03$)	Personal Photo Posting ($Q7\_05$)	Video Posting ($Q7\_06$)	News/Web page ($Q7\_07$)	Resharing ($Q7\_04$)
$Q22$	$-.14^{***}$	$-.11^{***}$	$-.13^{***}$	$-.12^{***}$	$-.14^{***}$	NA	NA	NA
$Q21$	$.18^{***}$	$.18^{***}$	$.25^{***}$	$.2^{***}$	$.25^{***}$	$.22^{***}$	$.22^{***}$	$.19^{***}$
$Q20$	$.19^{***}$	$.22^{***}$	$.17^{***}$	$.17^{***}$	$.21^{***}$	$.15^{***}$	$.15^{***}$	$.15^{***}$
$Q19$	$.3^{***}$	$.33^{***}$	$.32^{***}$	$.35^{***}$	$.28^{***}$	$.23^{***}$	$.25^{***}$	$.3^{***}$

Table 10: Linear model with dependent variable $Q7\_01$ restricted to weekly visitors and independent variables $Q19$, $Q20$, $Q21$, and $Q22$, also all restricted to weekly visitors.

	Estimate	Std. Error	t value	$Pr(>\|t\|)$
(Intercept)	1.6	.32	5.00	$< 7.05e-07^{***}$
Share Less or More than Others on FB (Q19)	.52	.07	7.26	$< 8.72e-13^{***}$
Control Your Information on FB (Q20)	.21	.06	3.4	$.0007^{**}$
Understand What Information is Shared on FB (Q21)	.32	.07	4.55	$6.19e-06^{***}$
General Feeling of Privacy on FB (Q22)	$-.09$	.06	-1.47	.14
Res. std. error: 1.9, 873 DF	Multiple R^2: .13 Adj. R^2:.13	F-Statistic 33.52 on 4 873 DF	p-value $<2.2e-16$	Residuals Min: -4.7 Median: -.13, Max: 4.9

that this increased engagement is associated with the use of "unofficial" controls. For example, in the case of pseudonym-use mentioned earlier, we do not see increased engagement amongst users with pseudonyms, rather engagement is lower on average (although generally not significantly). The official nature of the controls offered, or even encouragement by the service provider to use the controls, may be important given the substantial amount of discomfort in sharing information with service providers (e.g. Table 8).

UTILITY. Profile information is often withheld because users do not perceive the benefits of providing it (see, Table 8. This perception may also contribute to the large number of users reporting to withhold or restrict information "as a rule" unless it is required.

TRUST. 30% of the respondents expressed general fears and attitudes as explanations for their privacy concerns. For example, one respondent said, "I am concerned with people knowing thing about me that I do not want them to know. Or people knowing things that could be used against me". Others expressed government-related fears, e.g., "The government can use my opinion against me later on." These responses were not tied to specific experiences of the respondents or their connections and rather suggest additional trust needs to be built between these users and service providers as responsible stewards of user data. Trust is a personalized concept and so difficult to broadly facilitate, but responsiveness of the service provider appears to contribute (e.g., [20, 22])

In conclusion, our work contributes to the ongoing discussion of social network privacy by offering additional insights into the common privacy concerns and perceived consequences, and the impact on user engagement and information restriction. Specifically, our research offers a characterization of privacy attitudes and user behaviors in Facebook (with findings relevant to any online social network) and argues for the importance of privacy concerns as a hurdle to engagement that can be overcome through education and controls..

Acknowledgments

We are grateful to Larry Osborn at Knowledge Networks for help launching the survey, Alma Whitten for supporting this research, Allison Woodruff and Robert Ing for useful discussions about the survey content, and Alessandro Acquisti for useful conversations about the analysis.

5. REFERENCES

[1] A. Acquisti and R. Gross. Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. Proceedings of Privacy Enhancing Technologies Workshop (PET), Lecture Notes in Computer Science 4258, Springer, 36-58, 2006.

[2] I. Altman. Privacy regulation: culturally universal or culturally specific? Journal of Social Issues, 33 (3), 66-84.

[3] L. Brandimarte, A. Acquisti and G. Lowenstein. Misplaced confidences: Privacy and the control paradox. WEIS 2010.

[4] D. boyd and E. Hargittai. Facebook privacy settings: Who cares? First Monday, Volume 15, Numer 8 - August 2, 2010.

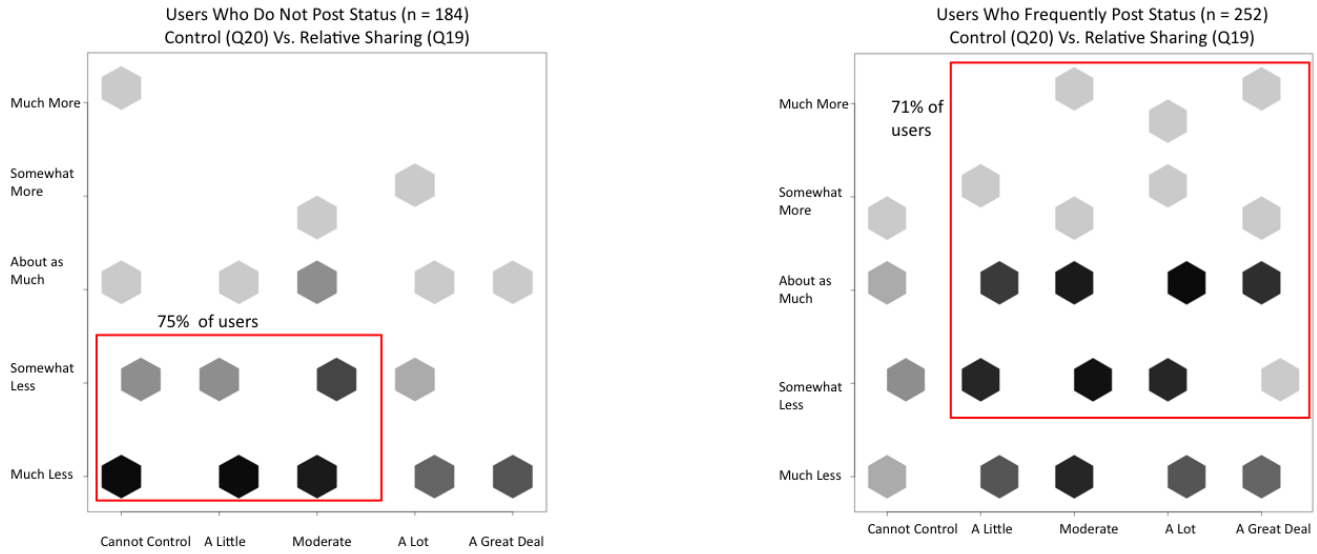

Table 11: A comparison of reported feelings of control (Q20) and relative sharing of personal information (Q19). The left image shows those who report to never post status ($n = 184$, "NeverStatus") and the right shows those who report status a few times a week ($n = 252$). The darker the hexagon, the more users there are who report the associated levels of control and relative sharing. Note the clustering of those who do not report status in the lower left, where reported feeling of control are the lowest and others are perceived to be sharing more personal information. In contrast, those who post status weekly are more in the center and to the right, indicating higher feelings of control and more perceived similarity with friends in personal information sharing.

[5] M. Burke, C. Marlow, C., and T. Lento. (2009). Feed me: Motivating newcomer contribution in social network sites. ACM CHI 2009: Conference on Human Factors in Computing Systems, 945-954.

[6] A. Chaudhuri. Randomized Response and Indirect Questioning Techniques in Surveys. Psychology Press, 2010.

[7] G. Cluley. 60% of Facebook users consider quitting over privacy. Naked Security. May 19,2010. `http://nakedsecurity.sophos.com/2010/05/19/60-facebook-users-quitting-privacy/`

[8] B. Debatin, J. Lovejoy, A. Horn and B. Hughes. Facebook and online privacy: attitudes, behaviors, and unintended consequences. Journal of Computer-Mediated Communication 15 (2009) pp.83-108.

[9] S. Gosling, O. John, K. Craik and R. Robins. Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. J. of Personality and Social Psychology, 1998, Vol. 74, No. 5, 1337-1349.

[10] R. Gross and A. Acquisti. Information revelation and privacy in online social networks (The Facebook case). WPES 2005.

[11] Knowledge Networks. `http://www.knowledgenetworks.com`

[12] P. Kumaraguru and L. Cranor. Privacy indexes: A survey of Westin's studies. Technical Report CMU-ISRI=5=138. December 2005.

[13] A. Lenhart. Adults and social network websites. Pew Internet Report. January 14, 2012.

[14] K. Lewis, J. Kaufman and N. Christakis. The taste for privacy: An analysis of college student privacy settings in an online social network. Journal of Computer-mediated Communication 14 (2008), pp 79-100.

[15] M. Madden. Privacy management on social media sites. Pew Internet Report. February 24, 2012.

[16] S. Margulis. Privacy and psychology. Contours of Privacy: Normative, Psychological and Social Perspectives. November 5, 2005.

[17] D. Nishioka, Y. Murayama and Y. Fujihara. Producing a Questionnaire for a User Survey on Anshin with Information Security for Users without Technical Knowledge. HICSS, January 2012.

[18] Nonnecke, B. and Preece, J. (2001). Why lurkers lurk. In Proceedings of the Americas Conference on Information Systems, Boston.

[19] L. Palen, P. Dourish. Unpacking "privacy" for a networked world. CHI 2003.

[20] E. Porter and N. Donthu. Cultivating trust and harvesting value in virtual communities. J. or the Institute for Operations Research and the Management Sciences, March, 2007.

[21] L. Rainie, A. Lenhart, A. Smith. The tone of life on social networking sites. Pew Internet Report. February 9, 2012.

[22] C. Ridings, D. Gefen and B. Arinze. Some antecedents and effects of trust in virtual communities. J. of Strategic Information Systems 11 (2002) 271-295.

[23] K. Sheehan. Toward a typology of Internet users and online privacy concerns. The Information Society, 18:21-32, 2002.

[24] Strauss, A. and Corbin, J. (1994). Grounded theory methodology: An overview. Sage Publications.

[25] F. Stutzman, R. Capra and J. Thompson. Factors mediating disclosure in social network sites. Computers in Human Behavior 27 (2011), pp 590-598.

[26] F. Stutzman and J. Kramer-Duffield. Friends only: examining a privacy-enhacing behavior in Facebook. CHI 2010.

[27] J. Turow, J. King, C. Hoofnagle and M. Hennessy. Contrary to what marketers say, Americans Reject Tailored Advertising and Three Activities that Enable It. September 2009. `http://www.ftc.gov/bcp/workshops/privacyroundtables/Turow.pdf`

[28] uSamp. Social Media Habits and Privacy Concerns Survey. January 30, 2012.

[29] S. Vazire and M. Mehl. Knowing me, knowing you: the accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. J. Personality and Social Psychology, 2008, November; 95(5): 1202-16.

[30] Y. Wang, S. Komanduri, P. leon, G. Norcie, A. Acquisti and L. Cranor. "I regretted the minute I pressed share": A Qualitative Study of Regrets on Facebook. SOUPS 2011.

[31] M. Yao, R. Rice and K. Wallis. Predicting user concerns about online privacy. J. of the American Society for Information Science and Technology, Volume 58, Issue 5, pages 710âĂŞ722, March 2007.

APPENDIX

A. SURVEY QUESTIONS

Users were first asked to state the select their primary social network from a list including Facebook, Twitter, LinkedIn, Google+, Foursquare, Photo-sharing sites and other options (including free-text). We analyzed the 1075 responses for which "your primary social network" is Facebook. All questions were optional.

- *Q*3: How often do you visit [primary social network]?
  - Never
  - Less than once a month
  - About once a month
  - A few times a month
  - About once a week
  - A few times a week
  - About once a day
  - Multiple times a day
- *Q*4: Approximately, when did you first start using [primary social network]?
  - More than Five Years Ago
  - 3-4 Years Ago
  - 1-2 Years Ago
  - 6-12 Months Ago
  - Less than 6 Months Ago
- *Q*5: Currently in your life, how many close friends would you say you have?
  - Less than 5
  - 5-10
  - 11-25
  - 25-50
  - More than 50
- *Q*6: Approximately, how many connections do you have in [primary social network]? [Free text response]
- *Q*7: How often do you:
  - *Q*7_01: Post a status update, broadcast message or question to [primary social network]?
    - *Same answer options as Q*3
  - *Q*7_02: React to someone elses post via a comment or reply in [primary social network]?
    - *Same answer options as Q*3
  - *Q*7_03: React to someone elses post via a Like or +1 a post in [primary social network]?
    - *Same answer options as Q*3
  - *Q*7_04: Share someone else's post via Reshare, Retweet or other syndication method in [primary social network]?
    - *Same answer options as Q*3
  - *Q*7_05: Post photos of myself or people I know in [primary social network]?
    - *Same answer options as Q*3
  - *Q*7_06: Post a link to a video in [primary social network]?
    - *Same answer options as Q*3
  - *Q*7_07: Post a link to a news story or web page in [primary social network]?

    - * *Same answer options as Q3*
  - Q7_08: View help pages within [primary social network]?
    - * *Same answer options as Q3*
  - Q7_09: Change your account security/privacy settings within [primary social network]?
    - * *Same answer options as Q3*
  - Q7_10: Close any social network service account?
    - * *Same answer options as Q3*
- Q8_01: What do you use as your user name on Facebook?
  - My full name (e.g. Bob Kawalski)
  - My first name only (e.g. Bob)
  - A pseudonym or Nickname (e..g DreamWeaver21, angryBobintheCity)
  - A fake, or made-up name (e.g. Joe Smith)
- Q9_1: Do you have more than 1 account on Facebook? (Y/N)
- Q10: Please explain why you have more than 1 account. (Free text response)
- Q11: How often do you visit other peoples online profiles in [primary social network]?
  - *Same answer options as Q7_01*
- Q12: How often have you clicked on a photo or photo collection of a friend of a friend or someone you do not know on [primary social network], to see more photos?
  - *Same answer options as Q7_01*
- Q13: How often do you view the following? [*All answer options are the same as Q7_01*]
  - *Q13_a*: The profiles or pictures of [primary social network] users with whom you have lost touch
  - *Q13_b*: The profiles or pictures of [primary social network] users who are acquaintances
  - *Q13_c*: The profiles or pictures of [primary social network] users whom you do not know
  - *Q13_d*: The profiles or pictures of celebrities or famous people
- Q14: Which of the following have happened to you on a social network service?
  - Q14_1: I have heard about a social event through my network (Y/N)
  - Q14_2: I have reconnected with an old friend (Y/N)
  - Q14_3: I have made progress with a job search (Y/N)
  - Q14_4: I found a buyer for something I was selling (Y/N)
  - Q14_5: I have accidentally posted something to more people than I intended (Y/N)
  - *Q14_none*: None of the above (Y/N)
- Q15: Please indicate whether you do or do not provide this information on your profile in your most used online social network.
  - Q15_01: My picture
    - * Do provide
    - * Do provide but not everyone can see it
    - * Do provide but information is fake
    - * Do not provide
  - Q15_02: My birth date (month and day)
    - * *Same answer options as Q15_01*
  - Q15_03: My phone number
    - * *Same answer options as Q15_01*
  - Q15_04: My home address
    - * *Same answer options as Q15_01*
  - Q15_05: My city of residence
    - * *Same answer options as Q15_01*
  - Q15_06: My email
    - * *Same answer options as Q15_01*
  - Q15_07: My gender
    - * *Same answer options as Q15_01*
  - Q15_08: My relationship status
    - * *Same answer options as Q15_01*
  - Q15_09: My interests/hobbies
    - * *Same answer options as Q15_01*
- Q16 Considering the information you do provide, has the information led to positive or negative experiences?
  - Q16_01: My picture
    - * Always positive
    - * Sometimes positive
    - * Sometimes positive, sometimes negative
    - * Sometimes negative
    - * Always negative
    - * I do not know
  - Q16_02: My birth date (month and day)
    - * *Same answer options as Q16_01*
  - Q16_03: My phone number
    - * *Same answer options as Q16_01*
  - Q16_04: My home address
    - * *Same answer options as Q16_01*
  - Q16_05: My city of residence
    - * *Same answer options as Q16_01*
  - Q16_06: My email
    - * *Same answer options as Q16_01*
  - Q16_07: My gender
    - * *Same answer options as Q16_01*
  - Q16_08: My relationship status
    - * *Same answer options as Q16_01*
  - Q16_09: My interests/hobbies
    - * *Same answer options as Q16_01*
- Q17_1: You said you do not provide your picture. Please tell us why.
  - Q17_1_1: It was a hassle to enter it
  - Q17_1_2: I had technical problems entering the information (e.g. the service did not recognize my city)
  - Q17_1_3: I did not think the information would make for a better experience on the network
  - Q17_1_4: I withheld it because it could be embarrassing
  - Q17_1_5: I withheld it because of concerns about impersonation/identity theft

  - $Q17\_1\_6$: I felt uncomfortable providing it to the service provider
  - $Q17\_1\_7$: As a rule I withhold such information unless it is required to provide it
  - $Q17\_1\_other$: Other/Free-text answer
- $Q17\_2$: You said you do not provide your birth date. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_3$: You said you do not provide your phone number. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_4$: You said you do not provide your home address. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_5$: You said you do not provide your city of residence. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_6$: You said you do not provide your email. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_7$: You said you do not provide your gender. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_8$: You said you do not provide your relationship status. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q17\_9$: You said you do not provide your interests/hobbies. Please tell us why.
  - *Same answer options as* $Q17\_1$
- $Q18\_1$: You said you limit access to your picture. Please tell us why.
  - $Q18\_1\_1$: The information is only interesting to particular people
  - $Q18\_1\_2$: The information will only make sense to certain people (e.g. rare hobbies or interests)
  - $Q18\_1\_3$: The information could be embarrassing if shared more broadly
  - $Q18\_1\_4$: There is a privacy/security risk to sharing the information
  - $Q18\_1\_5$: As a rule I restrict access to such information unless it is required to provide it more broadly
  - $Q18\_1\_other$: Other/Free-text answer
- $Q18\_2$: You said you limit access to your birth date. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_3$: You said you limit access to your phone number. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_4$: You said you limit access to your home address. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_5$: You said you limit access to your city of residence. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_6$: You said you limit access to your email. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_7$: You said you limit access to your gender. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_8$: You said you limit access to your relationship status. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q18\_9$: You said you limit access to your interests/hobbies. Please tell us why.
  - *Same answer options as* $Q18\_1$
- $Q19$: Compared to the rest of the people in my network, I share personal information
  - Much less
  - Somewhat less
  - About as much
  - Somewhat more
  - Much more
- $Q20$: To what degree do you feel you can control how your information is shared with other Facebook users?
  - Cannot control at all
  - Can control a little
  - Can control a moderate amount
  - Can control a lot
  - Can control a great deal
- $Q21$: How well do you understand what information about you is shared with other Facebook users?
  - Do not understand at all
  - Understand slightly
  - Understand moderately well
  - Understand very well
  - Understand extremely well
- $Q22$: How do you feel about your privacy with regard to Facebook?
  - Not at all concerned
  - Slightly concerned
  - Moderately concerned
  - Very concerned
  - Extremely concerned
- $Q22a$: How do you feel about your privacy with regard to the Internet overall?
  - *Same answer options as* $Q18\_1$
- $Q23$: What are your main privacy concerns online? (Free-text response)
- $Q24$: What potential negative consequences are there from the concerns you mentioned in the previous question? (Free-text response)
- $Q25$: What company or organization do you think sponsored this survey?

Knowledge Networks, who administered the survey, also has demographic information for the resondents including age, education level, race/ethnicity, gender, household information, marital status, geographic information and work status.