SOCIAL MEDIA PRIVACY CONCERNS AMONG COLLEGE STUDENTS

Alex Koohang, Middle Georgia State University, alex.koohang@mga.edu
Joanna Paliszkiewicz, Warsaw University of Life Sciences, joanna_paliszkiewicz@sggw.pl
Jeretta Horn Nord, Oklahoma State University, jeretta.nord@okstate.edu

ABSTRACT

Social media privacy concerns have grown as the number of users and the amount of data collected over the years has increased. The purpose of this study was to investigate significant differences between four selected independent variables (likeliness of using a social media site on a typical day, the amount of time spent using a social media site on a typical day, age, and gender) and the six dependent variables of social media privacy concerns SMPC (SMPC: Collection, SMPC: secondary usage, SMPC: errors, SMPC: improper access, SMPC: control, and SMPC: awareness). Collected data from a sample of 120 college students were analyzed using ANOVA. Results indicated several significant differences between independent variables and dependent variables. These results are discussed, and recommendations are made.

Keywords: Social media; privacy; privacy concerns; social media privacy concerns, usage, age; gender

INTRODUCTION

Social media has become a frequent research topic (Alfaro, Watson-Manheim, 2015; Han, 2016; Teng, et al. 2017). Boyd and Ellison (2008, p. 211) defined social media as “… web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections, and those made by others within the system.” Ahlqvist, et al. (2008) referred to social media as web-based platforms that stimulate communication and sharing of information/content between individuals in virtual communities. Within public and private entities, social media is an effective tool for advertisements (Cavinder, et al. 2017). It builds strong customer relationships (Grewal and Levy, 2016), facilitates communication between marketers and customers (Han, Kim, 2016), and is used as a significant strategic tool among organizations (Parveen, et al. 2016).

Social media is also used as an effective tool in governments (Hong, 2013; Madestam and Falkman, 2017) and education (Paliszkiewicz and Koohang 2016). Despite its many useful features, social media tends to pose serious privacy concerns for users. Rainie (2018) reported that 91% of Americans believe they have no control over their personal information collected by social media and that they are not aware of how their personal information is being collected and used by social media.

In general, privacy concerns are “… the degree to which an Internet user is concerned about website practices related to the collection and use of his or her personal information’ (Hong and Thong, 2013, p. 276). In their research, Smith et al. (2011) describe privacy as the ability of an individual to control his/her personal information. Malhotra et al. (2004) and Smith et al. (1996) described privacy concerns as the individuals’ inherent worries about possible loss of information privacy. Xu et al. (2011) defined it as consumers’ concerns about potential loss of privacy because of information disclosure to an external agent (e.g., a specific website).

Privacy concerns have been researched in conjunction with data mining (White, 2012); user behavior (Chen and Sharma, 2012); trust and risk behavior (Fogel and Nehmad, 2009); Internet of Things (Hsu and Lin, 2016); and control of personal information (Bansal, 2017). In the literature, concerns for information privacy have been considered as obstacles to individuals’ acceptance of e-commerce (Dinev and Hart, 2006), their purchasing behavior (Eastlick et al., 2006), and their willingness to expose personal information online (Bansal et al., 2010). Toch (2014) believes that users should know that their information is collected and used consistent with their expectations. According to Palen and Dourish (2003), users’ sense of privacy is correlated with the control they have over their information.
Hong and Thong (2013) identified six dimensions of the Internet privacy concerns (i.e., collection, secondary usage, errors, improper access, control, and awareness) based on an extensive prior research. Koohang (2017) adapted these six dimensions of the Internet privacy concerns to explicitly define the users' social media privacy concerns (SMPC):

“Collection - the amount of specific user data absorbed by the social media sites.
Secondary usage – [users’] personal information collected by the social media sites for one purpose, but used, without authorization/permission from the user, for another secondary purpose.
Errors - inadequate protections against deliberate and/or accidental errors in user personal data collected by the social media sites.
Improper access - [users’] personal information held by the social media sites that is readily available to others and/or not properly authorized to be viewed or accessed by others.
Control - inadequate control over personal information held by the social media sites.
Awareness - [users are] not being made aware of information privacy practices by the social media sites.”
(Koohang, 2017, p.17)

PURPOSE OF THE STUDY

Much of the recent research in the area of social media privacy concerns has focused on defining social media privacy concerns (Koohang 2017), users’ motivation to control personal information (Compañó and Lusoli, 2010); users’ understanding of privacy settings (Liu et al., 2011); and users’ disclosure of information (Madejski et al., 2012). Little research has been done to examine variables that may be significant in determining users’ social media privacy concerns.

The purpose of this study was to determine whether there are significant differences between the four selected independent variables (likeliness of using a social media site on a typical day, the amount of time spent using a social media site on a typical day, age, and gender) and the six dependent variables of SMPC (SMPC: Collection, SMPC: secondary usage, SMPC: errors, SMPC: improper access, SMPC: control, and SMPC: awareness). The four independent variables were selected because they have been of interest in social media research in general (e.g., Wells and Link, 2014; Lenhart et al. 2010; Paliszkiewicz and Koohang, 2016).

RESEARCH QUESTIONS

Consistent with the purpose of the study, we ask four research questions. They are:

RQ1: Are there significant mean differences between the independent variable of likeliness of using a social media site on a typical day and each of the six dependent variables of SMPC (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?

RQ2: Are there significant mean differences between the independent variable of the amount of time spent using a social media site on a typical day and each of the the six dependent variables of (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?

RQ3: Are there significant mean differences between the independent variable of age and each of the six dependent variables of SMPC (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?

RQ4: Are there significant mean differences between the independent variable of gender and each of the six dependent variables of SMPC (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?
METHODOLOGY

Instrument
This study uses a Likert-type instrument with six constructs, i.e., SMPC: Collection, SMPC: secondary usage, SMPC: errors, SMPC: improper access, SMPC: control, and SMPC: awareness. This instrument was empirically validated to measure privacy concerns among social media users (Koohang, 2017). The items of each construct are as follows:

“Social media privacy concerns (Collection) Construct
1. It bothers me when social media sites ask me to provide personal information.
2. When social media sites ask me for personal information, I sometimes think twice before providing it.
3. I am concerned that social media sites are collecting personal information about me.

Social media privacy concerns (Secondary Usage) Construct
4. I am concerned that social media sites would use my stored personal information for their own advantage/profit.
5. I am concerned that social media sites would sell my stored personal information in their databases to other companies.
6. I am concerned that social media sites would share my stored personal information in their databases with other companies without my authorization.

Social media privacy concerns (Errors) Construct
7. I am concerned that social media sites do not take enough steps to make sure that my personal information in their files is accurate.
8. I am concerned that social media sites do not have adequate procedures to correct errors in my personal information.
9. I am concerned that social media sites do not devote enough time and effort to verifying the accuracy of my personal information in their databases.

Social media privacy concerns (Improper Access) Construct
10. I am concerned that social media site databases that contain my personal information are not protected from unauthorized access.
11. I am concerned that social media sites do not devote enough time and effort to preventing unauthorized access to my personal information.
12. I am concerned that social media sites do not take enough steps to make sure that unauthorized people cannot access my personal information on their computers.

Social media privacy concerns (Control) Construct
13. It usually bothers me when I do not have control of personal information that I provide to social media sites.
14. It usually bothers me when I do not have control or autonomy over decisions about how my personal information is collected, used, and shared by social media sites.
15. I am concerned when control of my personal information on a social media site is lost or unwillingly reduced because of marketing transactions with other companies.

Social media privacy concerns (Awareness) Construct
16. I am concerned when a clear and visible disclosure is missing in online privacy policies of social media sites.
17. It usually bothers me when I am not aware or knowledgeable about how my personal information will be used by social media sites.
18. It usually bothers me when social media sites seeking my information online do not disclose the way the data are collected, processed, and used.” (Koohang, 2017, pp. 17-18)

The instrument used the following scale: 7 = completely agree, 6 = mostly agree, 5 = somewhat agree, 4 = neither agree nor disagree, 3 = somewhat disagree, 2 = mostly disagree, and 1 = completely disagree.
Sample and Data Collection
Via an Internet survey site, data were collected electronically from 127 undergraduate students from a medium-sized university in the Southeast USA. Of the 127, seven were removed because of incomplete data. This yielded a final sample of 120 for the study. The participants in this study were undergraduate students who were using Facebook daily. In addition to using Facebook, 90 indicated using Snapchat, 87 indicated using Instagram, 42 indicated using Pinterest, and 26 indicated using Twitter. Participants were females (N = 67) and males (N = 53). The participants’ age categories were 18 - 20 (N = 29), 21 - 29 (N = 35), 30 – 39 (N = 37), and 40 or older (N = 19).

Data Analysis
SPSS™ version 25 was used to conduct four one-way analysis of variance (ANOVA) to test the significance of group differences. The ANOVA test shows the effect that one independent variable has on one dependent variable. The F test establishes the significance of the groups. Post-hoc analysis was conducted for groups with more than two levels. Descriptive analysis was conducted to show the means and standard deviation of each dependent variable with independent variables. This study included four independent variables (likeliness of using a social media site on a typical day, the amount of time spent using a social media site on a typical day, age, and gender) and six dependent variables (SMPC: Control, SMPC: secondary usage, SMPC: errors, SMPC: improper access, SMPC: control, and SMPC: awareness).

RESULTS

Descriptive Analysis
As revealed in Figure 1, significant concerns were expressed by the respondents regarding social media privacy. Respondents were most concerned about what personal information social media sites collect (COL – 5.87) and about improper access (ACC – 5.87) to this information. Also of concern were the accuracy (ACC – 5.64) of the user information and transparency or awareness (AWA – 5.12) of how their personal information may be used. Finally, secondary usage (SEC), or the concern of social media sites using personal information for profit, and lack of control (CON) of the usage of the respondents’ information averaged 4.91 and 3.49 on the Likert scale, respectively.

![Social Media Privacy Concerns](chart)

**Figure 1.** Means of the six dimensions of Social Media Privacy Concerns

**RQ1:** Are there significant mean differences between the independent variable of likeliness of using a social media site on a typical day and each of the six dependent variables of SMPC (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?

The results of ANOVA for RQ1 showed a significant mean difference between the independent variable of likeliness of using a social media site on a typical day and the dependent variable of **SMPC: Collection** ($F_2, 117 = 3.101, .049$). The post-hoc results showed a significant difference between group 2 “very likely” and group 3 “Moderately likely.” Table 1 shows the means and standard deviation of SMPC: Collection and likeliness of using a social media site on a typical day.
typical day. Those subjects who moderately likely used a social media in a typical day were more concerned about their privacy in terms of the amount of specific data being collected by the social media sites.

**Table 1.** Means & Std. Deviation for SMPC & Likeliness of Using a Social Media on a Typical Day (N = 120)

<table>
<thead>
<tr>
<th>SMPC: Collection</th>
<th>Likeliness of Use</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Extremely likely</td>
<td>55</td>
<td>5.8545</td>
<td>1.18420</td>
</tr>
<tr>
<td></td>
<td>Very likely</td>
<td>34</td>
<td>5.5882</td>
<td>1.13724</td>
</tr>
<tr>
<td></td>
<td>Moderately likely</td>
<td>31</td>
<td>6.2688</td>
<td>0.91659</td>
</tr>
</tbody>
</table>

Note: There were no answers for Not at all likely level.

No significant mean differences were reported between the independent variable of likeliness of using a social media site on a typical day and the following dependent variables: SMPC: Secondary usage ($F_{2,117} = .353, p = .703$), SMPC: Errors ($F_{2,117} = .298, p = .743$), SMPC: Improper Access ($F_{2,117} = 2.363, p = .099$), SMPC: Control ($F_{2,117} = 2.034, p = .135$), and SMPC: Awareness ($F_{2,117} = 1.910, p = .153$).

**RQ2:** Are there significant mean differences between the independent variable of the amount of time spent using a social media site on a typical day and each of the six dependent variables of SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness separately?

The results of ANOVA for RQ2 showed a significant difference between the independent variable of the amount of time spent using a social media site on a typical day and the dependent variables of SMPC: Errors ($F_{1,118} = 6.907, p = .010$), SMPC: Control ($F_{1,118} = 4.183, p = .043$), and SMPC: Awareness ($F_{1,118} = 4.154, p = .044$).

Table 2 shows the means and standard deviation of SMPCs (Errors, Control, and Awareness) and time spent using a social media on a typical day. As regards SMPC: Errors, subjects who spent more time (4 – 5 hours) using a social media site in a typical day were more concerned about their privacy. As regards SMPC: Control, subjects who spent less time (1 – 3 hours) using a social media site in a typical day were more concerned about their privacy. As regards SMPC: Awareness, subjects who spent more time (4 – 5 hours) using a social media site in a typical day were more concerned about their privacy.

**Table 2.** Means & Std. Deviation for SMPC & Time Spent Using a Social Media on a Typical Day (N = 120)

<table>
<thead>
<tr>
<th>SMPC: Errors</th>
<th>Time Spent</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 - 3 Hours</td>
<td>57</td>
<td>5.2807</td>
<td>1.63031</td>
</tr>
<tr>
<td></td>
<td>4 - 5 Hours</td>
<td>63</td>
<td>5.9577</td>
<td>1.17393</td>
</tr>
<tr>
<td>SMPC: Control</td>
<td>Time Spent</td>
<td>N</td>
<td>Mean</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>1 - 3 Hours</td>
<td>57</td>
<td>3.7719</td>
<td>1.47890</td>
</tr>
<tr>
<td></td>
<td>4 - 5 Hours</td>
<td>63</td>
<td>3.2275</td>
<td>1.43538</td>
</tr>
<tr>
<td>SMPC: Awareness</td>
<td>Time Spent</td>
<td>N</td>
<td>Mean</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td></td>
<td>1 - 3 Hours</td>
<td>57</td>
<td>4.8538</td>
<td>1.40721</td>
</tr>
<tr>
<td></td>
<td>4 - 5 Hours</td>
<td>63</td>
<td>5.3598</td>
<td>1.31206</td>
</tr>
</tbody>
</table>

Note: There were no answers for 6 - 8 Hours and Over 8 Hours levels.

No significant differences were reported between the independent variable of the amount of time spent using a social media site on a typical day and the dependent variables of SMPC: Collection ($F_{1,118} = 1.233, p = .269$), SMPC: Secondary usage ($F_{1,118} = .453, p = .502$), and SMPC: Improper Access ($F_{1,118} = 1.135, p = .289$).

**RQ3:** Are there significant mean differences between the independent variable of age and each of the six dependent variables of SMPC (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?

The results of ANOVA for RQ3 showed no significant difference between the independent variable of age and the dependent variables of SMPC: Collection ($F_{3,116} = 936, p = .426$), SMPC: Secondary usage ($F_{3,116} = 1.357, p = .259$), SMPC: Errors ($F_{3,116} = 1.238, p = .299$), SMPC: Improper Access ($F_{3,116} = 1.651, p = .182$), SMPC: Control ($F_{3,116} = 0.36, p = .991$), and SMPC: Awareness ($F_{3,116} = 1.118, p = .345$). The age categories were 18 - 20 (N = 29), 21 - 29 (N
= 35), 30 - 39 (N = 37), and 40 or older (N = 19).

RQ4: Are there significant mean differences between the independent variable of gender and each of the six dependent variables of SMPC (SMPC: Collection, SMPC: Secondary usage, SMPC: Errors, SMPC: Improper access, SMPC: Control, and SMPC: Awareness) separately?

The results of ANOVA for RQ4 showed a significant difference between the independent variable of gender and the dependent variable of SMPC: Improper Access ($F_{1,118} = 5.278, p = .023$). Table 3 shows the means and standard deviation for SMPC: Improper Access and gender. Female subjects were more concerned about their privacy.

<table>
<thead>
<tr>
<th>Table 4. Summary of the Significant Findings</th>
</tr>
</thead>
</table>
| Likelihood of using a social media site on a typical day and SMPC: Collection. There was an overall significant mean difference reported between likelihood of using a social media on a typical day and SMPC: collection. Post-hoc analysis showed that subjects who “moderately likely” used a social media in a typical day were more concerned about their personal information collected by social media sites followed by those who “very likely” used a social media in a typical day. These results may suggest that users who are the most concerned about the collection of their personal information by social networking sites, simply don’t use social media as often.

Amount of time spent using a social media site on a typical day and SMPC: Awareness. Significant mean differences were reported between the amount of time spent using a social media site on a typical day and the dependent variables of SMPC: Errors, SMPC: Control, and SMPC: Awareness. For SMPC: Errors, subjects who spent more time (4 – 5 hours) using a social media site in a typical day were more concerned about the insufficient protections against errors in their personal data being collected by the social media sites. For SMPC: Control, subjects who spent less time (1 – 3 hours) using a social media site in a typical day were more concerned about the inadequacy of control over their personal information being held by the social media sites. For SMPC: Awareness, subjects who spent more time (4 – 5 hours) using a social media site in a typical day were more concerned about the lack of awareness of privacy practices by the social media sites. The amount of time respondents spent using social media sites had an impact on the greatest number of dependent variables. It is interesting to note that...
that users spending the least amount of time in a day (1 – 3 hours) were more concerned about the lack of control over their personal information. However, regardless of the number of hours respondents used social media in a typical day, significant privacy concerns were revealed.

Gender and SMPC: Improper access. There was a significant difference between gender and SMPC: Improper access. Female subjects were more concerned about the improper access of their personal information held by the social media sites. One explanation for these findings is that it is widely known that predators often use social media to form relationships under a false pretense. These warnings may likely leave more females than males fearful about improper access to their personal information.

CONCLUSION

Several points of significance between the independent and dependent variables were revealed from the results of this study. A total of three of the four independent variables were found to have significant mean differences between a total of five of the six dependent variables. Equally as important were the significant privacy concerns indicated by the social media users responding in this study. These findings suggest that social media platforms must ramp up efforts to secure user trust, including better security, user transparency, and ethical policies while users gain knowledge on securing their own data.

As described by Smith et al. (2011), privacy is the ability of a person to control and manage his or her personal information. We assert that the ability to control and manage personal information by users, to a large extent, depends on the privacy design of social media sites. Cavoukian (2010) outlined seven foundational principles of privacy by design. These principles are “1) proactive not reactive - preventative not remedial 2) privacy as the default, 3) privacy embedded into design, 4) full functionality – positive-sum, not zero-sum, 5) end-to-end lifecycle protection, 6) visibility and transparency, and 7) respect for user privacy” (Cavoukian, 2010, para. 18). To protect user’s privacy, social media sites should be mindful of the privacy by design principles. Social media sites should make these principles a prerequisite to the design of their sites. The following seven questions, based on the concept of privacy by design, merit attention for ensuring privacy safeguards, defense, and protection of personal information on social media sites:

1. Are social media sites foreseeing and stopping privacy breach of personal information of their users?
2. Are social media sites guaranteeing that the personal information of their users is naturally protected with a) their technology infrastructure and b) their business practices of collecting and sharing personal information?
3. Are social media sites making privacy a vital element of the principal functionality of their business?
4. Are social media sites accommodating all reasonable interests and avoiding unnecessary trade-offs to create a positive outcome for all parties?
5. Are social media sites securing users’ personal information from the collection all the way until the information is wiped out?
6. Are social media sites making all business practices known to users, e.g., collecting/sharing of personal information?
7. Are social media sites having the interest of the users in mind when collecting/sharing their information?

Social media is big business, and responsibility falls on not only the social media sites but also involves businesses and the user. This study is limited to college students in a specific geographic area of the United States. For comparison purposes, recommendations for future related research using the same instrument could include the following targeted respondents: 1) college students throughout the USA; 2) a random sample of the general population throughout the USA; 3) college students in different countries; and 4) a random sample of the general population in different countries.

REFERENCES


