Cross-national privacy concerns on data collection by government agencies

Rebecca Cooper
Carleton University
rebecca_cooper@live.com

Hala Assal
Carleton University
HalaAssal@scs.carleton.ca

Sonia Chiasson
Carleton University
chiasson@scs.carleton.ca

***Abstract*—We conducted an online survey with 366 participants from Canada, India, the UK, and the US to compare privacy concerns and opinions about the collection of personal data by law enforcement and government agencies. We investigated what data participants were willing to share, in what circumstances participants were willing to allow data collection, what procedures companies should follow when they receive requests for customer information, and participants' general concern about their privacy. Statistical analysis showed that nationality and gender had significant impacts on participants' trust and perceptions of their governments, while nationality also impacted participants' willingness to share data under various circumstances. While participants were, on the whole, moderately amendable to data collection by government agencies given a court-ordered warrant, they also indicated a strong desire for increased transparency, and reported a lacklustre knowledge about privacy legislation.**

I. Introduction

Many researchers [1], [2] have investigated users' privacy concerns as they relate to data collection by corporate entities. In recent years, the documents leaked by Snowden revealed massive government surveillance programs, including the NSA's PRISM program, which allowed the NSA to directly access email content, browsing history, and other data on private companies' servers [3]. The documents also referenced GCHQ's Tempora program, which reportedly collected email content, Facebook entries, and other data [4].

The 2001 US Patriot Act contains numerous provisions for domestic and international surveillance [5] including provisions for legal mechanisms known as National Security Letters, which allow the FBI (and similar agencies) to acquire user data from telecommunications companies without a court-issued warrant to support investigations [5]. In 2016, the Email Privacy Act [6] was unanimously passed by the House of Representatives, but has stalled in the Senate.

In 2006, the UK passed the Data Retention Directive, which required telecommunications companies to store user data for 6 to 24 months to support police investigations but this law was struck down in 2014 due to privacy concerns [7]. The Data Retention and Investigatory Power Act followed but was replaced in 2016 by the Investigatory Powers Bill. It requires service providers to store data for 12 months to be disclosed with a warrant, and allows government security services to hack into digital networks and devices with a warrant [8].

India lacks comprehensive data privacy laws, despite multiple proposals since 2011. A 2014 draft of The Right to Privacy Bill includes provisions for data collection by government and law enforcement agencies, including interception of communications, and covert surveillance [9]. The Bill includes mandatory deletion after the purpose of collection is completed and data breach notification [9]; however, law enforcement and intelligence agencies would be exempted if data is collected for national security or crime prevention [9].

In 2015, Canada passed Bill C-51 [10], also known as the Anti-Terrorism Act, allowing the government to share Canadians' personal information with security and intelligence services, to add an individual to a no-fly list based on information from security and intelligence services (or other sources), and to disrupt or obtain information about activities that would undermine Canada's security. Critics suggest that the bill infringes Canada's Charter of Rights and Freedoms. The current government promises changes to C-51, but these have not yet materialized.

The controversy surrounding these laws highlights a riff between governments' desires to collect data and citizens' desires for privacy. In this study, we investigate participants' opinions and preferences about the collection of personal data by law enforcement and government agencies, and determine whether these opinions and preferences are influenced by participants' nationality, gender, or age. In particular, we examine which privacy issues people are most concerned about when it comes to government data collection, what data they are willing to share, and in what circumstances or through what processes such data collection should be allowed.

II. Related Work

Several studies relate to users' preferences about privacy and the collection of their personal data. Milberg *et al.* [1] found that individuals were particularly concerned about data collection, unauthorized secondary use of data, improper access to data, and errors in collected data. Leon *et al.* [2] investigated how privacy policies affect users' attitudes toward sharing data with advertising companies, and found that approximately half of survey participants were unwilling to share any data in exchange for targeted ads. Kang *et al.* [11] found that those with better understanding of the internet had heightened privacy concerns about corporate and government surveillance. Acquisti *et al.* [12] discuss how users' privacy

This paper was accepted for publication at the 15th Annual Conference on Privacy, Security and Trust (PST). This is the authors' copy for personal use. Additional appendices are included in this tech report.

concerns shift over time, can be affected by cultural dimensions, and can be manipulated by external influences.

In 1995, Milberg *et al.* [13] surveyed individuals from approximately 30 different countries, and found that privacy concerns varied significantly between different nationalities. This finding was supported by Bellman *et al.* [14], who investigated internet privacy concerns in 2004. In 2015, Kugler [15] discussed approaches to data protection in the private sector in the US, Europe, and Japan, and suggested that differences exist because Europe and Japan view privacy as a fundamental human right and encode it in legislation, while the US tends to rely on self-regulation.

Cavoukian [16] considered the implications of large-scale data collection by government entities, and highlighted the importance of maintaining individuals' privacy. The Office of the Privacy Commissioner of Canada (OPC) released a 2013 report [17] discussing their efforts to protect Canadians' privacy in the face of Big Data and demands for public safety. The OPC also conducted a survey [18] on privacy-related issues, including concerns about government and law enforcement agencies collecting personal data without warrants.

Furthermore, Marthews and Tucker [19] identified a chilling effect, both in the US and internationally, in online searches for terms that may trigger interest by government agencies after the Snowden revelations in 2013. Our user survey explores current opinions and concerns about government data collection, and investigate how these opinions and concerns differ due to nationality, gender, and age.

III. Methodology

Our study was cleared by our institution's Research Ethics Board. We conducted a crowd-sourced survey using CrowdFlower[1] to collect data from four countries. Data collection occurred in two rounds (November 2014 and February 2015). Participants received $0.50 (US) in compensation. A total of 400 surveys were equally distributed to participants in the United States (US), the United Kingdom (UK), Canada, and India. These groups were chosen to represent some variety in culture and governmental structure, ensure that English-speaking individuals would be able to complete the survey, and take advantage of the most active demographics on CrowdFlower. The survey was directed at CrowdFlower members ranked as Level 2 contributors or higher.

The survey was divided into three sections: (1) basic demographic information; (2) opinions about personal data collection by law enforcement and government agencies; (3) general levels of concern about the protection of the participants' privacy. To avoid biasing participants' reported opinions about data collection, there were no references to 'privacy' before the last section of the survey.

Table I summarizes the statistical tests used to evaluate the effect of nationality, gender, and age on participant responses. Participants' responses to Likert scale questions constitute ordinal data, while responses to multiple choice questions are categorical data. When considering the effect of age on participants' categorical responses, participants were grouped into four age ranges (18-29, 30-39, 40-49, and 50 and above) to allow Pearson's Chi-Squared test to be used.

[1] http://www.crowdflower.com/

TABLE I. Statistical tests used for each type of comparison

	Nationality	Gender	Age
Ordinal data	Kruskal-Wallis (KW) & Mann-Whitney (MW)	Mann-Whitney	Kendall's tau-b correlation metric (K)
Categorical data	Chi-Squared (Chi)	Chi-Squared	Chi-Squared

Mann-Whitney U (MW) tests were done with Bonferroni corrections as needed

IV. Results

After filtering, a total of 366 valid responses were collected (UK: 88, US: 92, India: 88, Canada: 96). There was an approximately even gender distribution among UK (48% M) and US (48% M) respondents, but responses from India (81% M), and Canada (39% M) were skewed.

Participants had a mean age of 37 years (UK: 41, US: 38, India: 30, Canada: 40) and 55% had some type of post-secondary degree. Participants held various occupations. Fourteen participants currently or have previously worked for a law enforcement agency (UK: 4, US: 3, India: 6, and Canada: 1), and 54 participants currently or have previously worked for a government agency (UK: 13, US: 15, India: 14, Canada: 10).

This user survey investigates whether nationality, gender, or age significantly influence opinions about government data collection; types of information participants most and least willing to share with law enforcement and government agencies; circumstances in which participants might be willing to allow data collection without a warrant; participants' willingness to share data with government or law enforcement agencies compared to their willingness to share data with private companies; and protocols that companies should follow when sharing customers' data with government or law enforcement agencies. The full survey is included in the appendix.

A. Perceptions of Government Agencies

Eight questions on the user survey related to participants' trust of government agencies and their perceptions of the benefits of government data collection. Participants' responses to the questions are illustrated in Figure 1. Results of our statistical analysis are shown in Table II.

These responses show that participants from India tend to respond more positively than participants from the US, the UK, or Canada. For instance, India's mean and median responses (which are above 3) suggest that participants from India believe data collection can lead to a reduction in crime, a stronger economy, and increased levels of safety. While UK respondents also tended to agree with the idea that data collection could increase safety and reduce crime, Canadian and US respondents tended to be more sceptical. Similarly, respondents from India were more likely to believe the government is capable of keeping data secure and collecting accurate data. Participants tended to have reservations about data being shared between multiple government agencies, and respondents from Canada, the UK, and the US indicated that they do not believe their governments are honest and transparent about the data they collect. Correspondingly, participants from all countries indicated (with mean scores of at least 4.11) that they would like their governments to be more transparent about data collection.

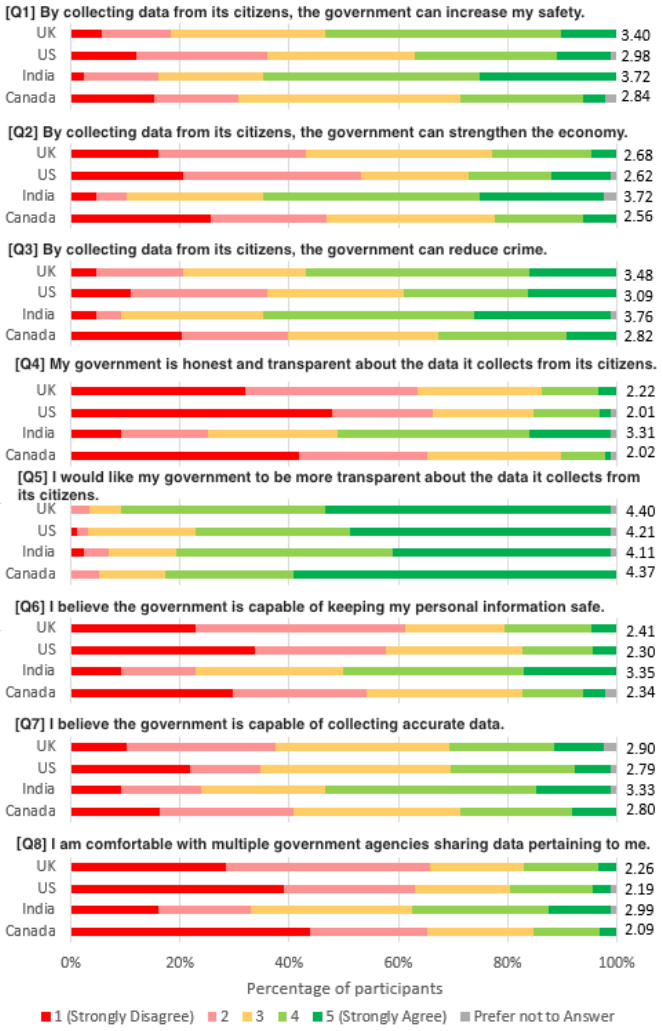

Fig. 1. Participant responses to questions 1 – 8 with means on the right

B. Data Collection by Governments

The bulk of the survey dealt with what, when, and how data should be collected by government and law enforcement agencies. Participants were asked how comfortable they were with data collection in specific circumstances (D1) and for specific purposes (D2), and what processes should be followed to gain access to personal data. We also investigated how participants' level of comfort with data collection varied depending on the type of data being collected (D3). The sums of each participant's responses in each category (D1, D2, D3) were considered in statistical analysis. Results of our statistical analysis are shown in Table III.

D1: Circumstances for Data Collection. Participants rated how comfortable they were with companies being compelled to disclose customers' personal information in various circumstances[2]. Means of the Likert scale responses are shown in Figure 2. Nationality had a significant effect on responses. Overall, Canadian respondents were least comfortable with data disclosure, regardless of the circumstances, and expressed significantly lower comfort levels than respondents from India and the UK; US respondents expressed only slightly higher levels of comfort than Canadians, and were significantly less comfortable than respondents from India. Respondents from all countries tended to be more amendable to data disclosure when the data was being used to protect a vulnerable person (i.e, a kidnapping victim, or a person who may be a danger to

[2]This question was adapted from the OPC's survey [18].

TABLE II. STATISTICAL ANALYSIS OF QUESTIONS 1 – 8

	Nationality	Gender	Age
Q1	* $\chi^2(3) = 34.21$ * **UK** – Canada: r = 0.26 * **India** – US: r = 0.31 * **India** – Canada: r = 0.38	ns	ns
Q2	* $\chi^2(3) = 52.73$ * **India** – UK: r = 0.45 * **India** – US: r = 0.42 * **India** – Canada: r = 0.46	**Male**-Female: * r = 0.16	**Young**-Old: * $r_\tau = -0.19$
Q3	* $\chi^2(3) = 31.33$ * **UK** – Canada: r = 0.27 * **India** – US: r = 0.28 * **India** – Canada: r = 0.37	ns	ns
Q4	* $\chi^2(3) = 62.62$ * **India** – UK: r = 0.43 * **India** – US: r = 0.48 * **India** – Canada: r = 0.50	**Male**-Female: * r = 0.15	**Young**-Old: * $r_\tau = -0.17$
Q5	ns	ns	Young-**Old**: * $r_\tau = 0.12$
Q6	* $\chi^2(3) = 42.45$ * **India** – UK: r = 0.38 * **India** – US: r = 0.41 * **India** – Canada: r = 0.40	ns	**Young**-Old: * $r_\tau = -0.21$
Q7	* $\chi^2(3) = 13.22$ * **India** – UK: r = 0.20 * **India** – US: r = 0.23 * **India** – Canada: r = 0.23	ns	**Young**-Old: * $r_\tau = -0.16$
Q8	* $\chi^2(3) = 28.74$ * **India** – UK: r = 0.29 * **India** – US: r = 0.31 * **India** – Canada: r = 0.35	**Male**-Female: ** r = 0.11	**Young**-Old: ** $r_\tau = -0.09$

* $p < 0.01$, ** $p < 0.05$, ns = non-significant, bold item has higher mean (indicating stronger agreement), green background indicates significant differences

TABLE III. STATISTICAL ANALYSIS OF OPINIONS ABOUT CIRCUMSTANCES, PURPOSES, AND DATA TYPES

	Nationality	Gender	Age
D1	* $\chi^2(3) = 25.51$ * **UK** – Canada: r = 0.24 * **India** – US: r = 0.26 * **India** – Canada: r = 0.33	ns	ns
D2	* $\chi^2(3) = 41.63$ * **India** – UK: r = 0.32 * **India** – US: r = 0.42 * **India** – Canada: r = 0.39	ns	ns
D3	* $\chi^2(3) = 38.78$ * **UK** – Canada: r = 0.20 * **India** – UK: r = 0.26 * **India** – US: r = 0.34 * **India** – Canada: r = 0.43	ns	**Young**-Old: ** $r_\tau = -0.08$

* $p < 0.01$, ** $p < 0.05$, ns = non-significant, bold item has higher mean (indicating greater comfort), green background = significant differences

	Canada	India	US	UK
[a] to locate a young person who has run away	2.76	3.44	3.01	3.40
[b] to locate a person who has gone missing or been kidnapped	3.59	3.70	3.87	4.02
[c] to locate a person who may be a danger to themselves or others	3.47	3.86	3.58	4.03
[d] to gather evidence in support of an investigation into a serious offence	3.15	3.87	3.33	3.82
[e] to gather evidence in support of an investigation into a minor offence	2.20	3.38	2.35	2.59
[f] to gather information about a person's online activities	1.88	3.07	2.04	2.09

Fig. 2. D1: Mean comfort level with government data collection, without a warrant (1 = Very uncomfortable, 5 = Very comfortable)

	Canada	India	US	UK
[g] preventing child exploitation	3.76	4.02	3.86	4.24
[h] combating the drug trade	3.31	3.95	3.39	3.90
[i] preventing crime within the country	3.37	4.03	3.41	3.83
[j] determining the desires of the country's populace	2.56	3.70	2.43	2.47
[k] preventing illegal streaming or downloading of copyright content	2.29	3.42	2.27	2.49
[l] monitoring political dissidence	2.26	3.10	2.10	2.28
[m] preventing political dissidence	2.11	3.20	2.14	2.28
[n] maintaining national security	3.22	4.03	3.24	3.65

Fig. 3. D2: Mean comfort level with government data collection for various purposes (1 = Very uncomfortable, 5 = Very comfortable)

themselves or others). Participants (particularly those from the UK and India) were also relatively comfortable when data was being used to support an investigation into a serious offence. Similar trends were observed by the OPC [18].

D2: Purposes of Data Collection. Participants were also asked how comfortable they were with their government collecting data for specific purposes. Means of the Likert scale responses are shown in Figure 3. Respondents from India were significantly more receptive to data collection than were respondents from other countries. However, participants from all four countries were generally comfortable (with mean responses above 3.0) with data being collected for preventing various crimes or protecting national security.

D3: Types of Collected Data. Participants indicated their level of comfort with government or law enforcement agencies collecting different types of data (Figure 4). Participants from the US, the UK, and Canada report average comfort levels below 3.0 for all data types (except for UK participants' current country), which suggests participants are moderately reluctant and/or divided when it comes to sharing specific types of information with government and law enforcement agencies. Participants from India were significantly more comfortable with sharing than other countries. Respondents from the UK also indicated significantly higher comfort levels than Canadians. We further found that older participants expressed lower levels of comfort than younger participants.

For comparison, we repeated the D3 questions, changing the collector to a private company instead. Participants from India were more comfortable with the collection of data by

	Canada	India	US	UK
Country	2.76	3.62	2.96	3.41
Religion	2.35	3.60	2.56	2.87
Town/city	2.32	3.49	2.52	2.88
IP address	2.30	3.15	2.62	2.86
Phone subscriber info	2.41	3.22	2.52	2.77
Sexual orientation	2.35	3.20	2.43	2.92
Internet subscriber info	2.29	3.24	2.55	2.71
Email metadata	2.37	3.24	2.40	2.67
Medical information	1.97	3.51	2.51	2.53
Text message metadata	2.23	3.03	2.38	2.71
Phone call metadata	2.26	2.86	2.51	2.73
Activity logs	2.16	3.13	2.41	2.58
OSN content	2.19	3.05	2.37	2.59
OSN friends list	2.15	3.07	2.42	2.54
Political views	2.10	3.26	2.32	2.33
Search terms	2.17	2.85	2.31	2.52
Exact current location	1.98	3.24	2.19	2.31
Online purchases	1.99	3.08	2.30	2.34
Email content	2.18	3.09	2.12	2.24
Browsing history	2.09	2.72	2.15	2.45
Text message content	2.02	2.57	1.98	2.03
Phone call content	1.80	2.55	1.92	2.00
Financial info	1.73	2.77	1.86	1.87
Credit card number	1.65	2.24	1.78	1.69

Fig. 4. D3: Mean comfort level with government collection of different data (1 = Very uncomfortable, 5 = Very comfortable)

private companies than were participants from other countries; however, average responses were less than 3 for all types of data across all countries. A Wilcoxon signed rank test with continuity correction showed that the collector (i.e., a company or a government agency) had a significant effect on responses ($p < 0.01, r = 0.44$). Participants were more willing to share data with government agencies than with private companies.

Processes for Data Collection. 50% of respondents indicated that government and law enforcement agencies should require a warrant, while 10% believed approval should be granted by a designated government member. Another 22% of respondents believed these agencies should be able to access or collect data at any time for policing or to maintain national security. Pearson's Chi-squared test revealed that nationality had a significant effect on responses ($\chi^2(15, N = 366) = 63.12, p < 0.01$). Participants from India were more likely to feel that government and law enforcement agencies should be able to collect or access data for any reason. Participants from Canada, the UK, and the US preferred that government and law enforcement agencies require a warrant.

C. Interactions with Service Providers

Corporate entities (such as Facebook, Google, and ISPs) often serve as access points to large amounts of personal data. We asked participants how companies should respond to and facilitate data collection and access by government and law enforcement agencies.

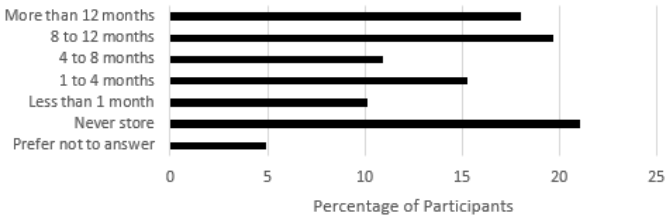

Fig. 5. Acceptable duration of data retention

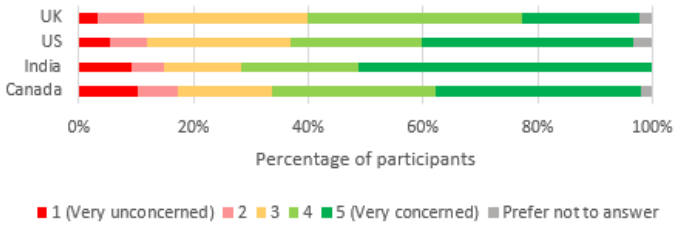

Fig. 6. P1: Concern about the protection of privacy

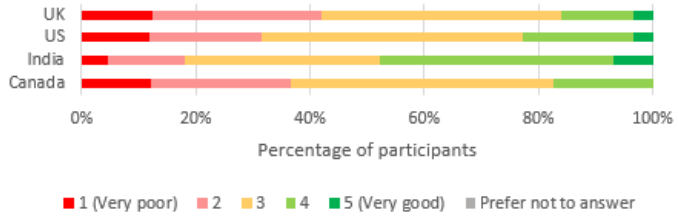

Fig. 7. P2: Self-reported knowledge about privacy laws

Compliance. We presented four statements describing when a company should comply with requests for personal data from law enforcement or government agencies (never, only with a warrant, in emergency situations, always), and asked participants to select with which they most agreed. 61% of respondents thought that companies should comply with requests for customer data only when presented with a warrant (or similar legal document), or in emergency situations. Pearson's Chi-squared tests revealed a significant effect of gender ($\chi^2(4, N = 364) = 21.32$, $p < 0.01$). Female respondents are more likely than males to feel that companies should release data in emergency situations (71% compared to 52%), while male respondents were more likely to select 'Never' (9% compared to 2%) or 'Whenever the company receives a request, without a warrant or other legal document' (10% compared to 3%). Nationality and age did not have statistically significant effects on participant responses.

Customer Notification. When asked if companies should notify a user when a government or law enforcement agency requests that user's data, 33% of respondents selected 'Always', while 53% selected 'Generally, yes, but there are some situations in which [companies] should not'. Neither nationality, gender, or age had a significant effect on responses.

Storing Data. Respondents were divided when it came to how long companies should store user data for use by government or law enforcement agencies. While 21% said companies should never store user data for the purpose of providing it to law enforcement or government agencies, another 18% said companies should store data for more than 12 months (Figure 5). Neither nationality, nor gender, nor age had a significant effect on responses.

D. Concern and Knowledge about Privacy

P1: Concern about privacy. Participants indicated how concerned they are about the protection of their privacy (P1, Figure 6)[3]. Despite being more amenable to data collection than other countries, participants from India indicated somewhat higher levels of concern about the protection of their privacy than other countries, particularly (and significantly) the UK ($\chi^2(3) = 8.12, India - UK : r = 0.22$). Overall, 65% of all participants were 'Concerned' or 'Very concerned' about the protection of their privacy, 21% responded neutrally, and only 14% were 'Unconcerned' or 'Very unconcerned'.

P2: Knowledge of Privacy Laws. Participants rated their knowledge of the laws pertaining to privacy protection (P2, Figure 7).[3] Our statistical analysis showed effects of nationality and gender. Indian participants rated their knowledge more highly than other countries ($\chi^2(3) = 27.96, India-UK : r = 0.35, India - US : r = 0.26, India - Canada : r = 0.33$), and overall, male respondents rated their knowledge more highly than females ($r = 0.10$). Overall, 42% of all participants responded neutrally when asked about their knowledge of privacy-related legislation, while 32% felt their knowledge was 'Poor' or 'Very poor', 22% felt it was 'Good', and only 3% felt it was 'Very good'. We note that this is self-reported; we did not assess participants' actual knowledge.

[3]This question was adopted from the OPC's 2013 survey [18]

V. Discussion

Indian respondents were most receptive to government data collection. They tended to agree with statements regarding the benefits of data collection, while respondents from other countries (particularly Canada and the US) were more sceptical. Respondents from India also seemed to have a better opinion of their government, overall, since they were more willing to believe that the government could protect and accurately collect data. These differences may be due, in part, to the recent (negative) media attention government data collection has received in the UK, the US, and Canada [3], [4], [7].

Age also had a significant effect. Older participants were less trusting of governments' abilities to accurately collect and safely store data. They were less comfortable with data being collected by government agencies and private companies. However, the effects of age and nationality are likely not independent. Respondents from India were significantly younger, thus, responses from India may have differed due, in part, to different age distributions. Likewise, younger participants may have appeared to respond differently than older participants due to different nationalities. We believe the latter possibility is more likely, here. We used Kendall's tau rank correlation metric (post-hoc) to investigate the effect of age on responses from individual countries. We found that age had far fewer significant effects within individual countries than were observed when responses from all countries were considered. This suggests that nationality may have amplified the effects of age. However, where significant age effects were observed within countries, the trends described above

held; older participants were less amendable to data collection. Gender had fewer significant effects, but where an effect was apparent, women had more reservations about data collection.

Participants were not wholeheartedly against data collection, but generally wanted a warrant. Participants from all countries exhibited willingness to allow data collection (even without a warrant) when data was used to protect a vulnerable person, or prevent serious crimes (e.g., child exploitation), but were less willing when data was used to thwart minor crimes.

When asked about sharing their own data, participants generally indicated low comfort levels for disclosure. Across the four countries, credit card numbers, financial information, and phone call content ranked among the five most discomforting types of data for government or law enforcement agencies to collect. In contrast, religion and country were consistently ranked among the least discomforting.

While participants indicated a certain willingness to allow data collection, they overwhelmingly want more transparency. They also indicated that individuals should be informed if and when a government or law enforcement agency requests their personal information, but largely recognized that this may not always be appropriate. There also seems to be a disconnect between the general sense of acceptance of data collection and what data participants themselves are willing to share.

Limitations. There are occasional concerns [20] raised about crowd-sourcing studies, but they are now commonly used in usable security research (e.g., [2]) and can lead to good-quality data [21]. Thus, the general trends observed in our research are likely to hold in the general population, even if exact levels of concern/comfort may vary; similar trends were noted in Canada [18]. The survey did not explicitly address data collection methods. However, this was done by design. We avoided drawing explicit attention to terms such as "interception" and potentially clandestine, controversial surveillance programs that might inaccurately prime respondents towards artificially high levels of concern. We thought that the type of data being collected was more important than *how* it was being collected, especially when *how* might invoke fear based primarily on media and recent events rather than encourage reflection on whether that data was particularly sensitive.

VI. Conclusion

The subject of data collection by government agencies and law enforcement is at the foreground as society struggles with balancing individuals' right to privacy with issues of protection and security. Given the amount of data digitally available, and advancements in processing capabilities and abilities to make inferences from these datasets, the consequences can be much more far-reaching than for previous generations. This user survey was conducted to examine cross-national privacy concerns as they relate to personal data collection by law enforcement and government agencies. Nationality, age, and, to a lesser extent, gender had significant impacts on participants' responses. Participants indicated willingness to allow data collection/access in emergency situations and to protect vulnerable people. However, they felt that a court-issued warrant should be obtained, and that companies should generally notify customers if their data is accessed. They also seemed hesitant to disclose their own data. It seems that approaches will need to be customized on a per country basis, to address the specific privacy-related attitudes of their citizens.

Acknowledgment

S. Chiasson acknowledges funding from the NSERC Discovery Grant and Canada Research Chair programs.

References

[1] S. J. Milberg, H. J. Smith, and S. J. Burke, "Information privacy: Corporate management and national regulation," *Organization science*, vol. 11, no. 1, 2000.

[2] P. G. Leon, B. Ur, Y. Wang, M. Sleeper, R. Balebako, R. Shay, L. Bauer, M. Christodorescu, and L. F. Cranor, "What matters to users?: factors that affect users' willingness to share information with online advertisers," in *SOUPS*. ACM, 2013.

[3] G. Greenwald and E. MacAskill, "NSA Prism program taps in to user data of Apple, Google and others," *The Guardian*, vol. 7, no. 6, 2013.

[4] E. MacAskill, J. Borger, N. Hopkins, N. Davies, and J. Ball, "Mastering the internet: how GCHQ set out to spy on the world wide web," *The Guardian*, vol. 21, no. 6, 2013.

[5] U. Gorham-Oscilowski and P. T. Jaeger, "National Security Letters, the USA Patriot act, and the constitution: The tensions between national security and civil rights," *Government Information Quarterly*, vol. 25, no. 4, 2008.

[6] K. Yoder, "H.R. 699 Email Privacy Act," 2016. [Online]. Available: www.congress.gov/bill/114th-congress/house-bill/699

[7] BBC, "Top EU court rejects EU-wide data retention law," 2014. [Online]. Available: www.bbc.com/news/world-europe-26935096

[8] M. Burgess, "It's official, the Snooper's Charter is becoming law: how the IP bill will affect you," 2016. [Online]. Available: www.wired.co.uk/article/ip-bill-law-details-passed

[9] G. Greenleaf, "India's draft the right to privacy bill 2014-will Modi's BJP enact it?" *129 Privacy Laws and Business International Report*, 2014.

[10] Minister of Public Safety and Emergency Preparedness, "Bill C-51: Anti-terrorism act," 2015. [Online]. Available: www.documentcloud.org/documents/1513457-bill-c-51.html\#document/p2

[11] 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 *USENIX SOUPS*, 2015.

[12] A. Acquisti, L. Brandimarte, and G. Loewenstein, "Privacy and human behavior in the age of information," *Science*, vol. 347, no. 6221, 2015.

[13] S. Milberg, S. Burke, H. Smith, and E. Kallman, "Values, personal information privacy, and regulatory approaches," *Communications of the ACM*, vol. 38, no. 12, 1995.

[14] S. Bellman, E. J. Johnson, S. J. Kobrin, and G. L. Lohse, "International differences in information privacy concerns: A global survey of consumers," *The Information Society*, vol. 20, no. 5, 2004.

[15] L. Kugler, "Online privacy: regional differences," *Communications of the ACM*, vol. 58, no. 2, 2015.

[16] A. Cavoukian, *Privacy and government 2.0: the implications of an open world*. Information and Privacy Commissioner of Ontario, 2009.

[17] J. Stoddart, *Privacy Priorities: Reflections on the Office of the Privacy Commissioner of Canada's Strategic Priority Issues*. OPC, 2013.

[18] "Survey of Canadians on privacy-related issues," Office of the Privacy Commissioner of Canada, 2013.

[19] A. Marthews and C. Tucker, "Government surveillance and internet search behavior," *Available at SSRN 2412564*, 2015.

[20] R. Kang, S. Brown, L. Dabbish, and S. B. Kiesler, "Privacy attitudes of mechanical turk workers and the us public." in *SOUPS*. ACM, 2014.

[21] A. J. Berinsky, G. A. Huber, and G. S. Lenz, "Evaluating online labor markets for experimental research: Amazon. com's mechanical turk," *Political Analysis*, vol. 20, no. 3, 2012.

Appendix: Survey of Opinions and Preferences Relating to Data Collection by Law Enforcement and Government Agencies

Thank you for participating in this survey. The questions in this survey are designed to collect basic demographic information, as well as your opinions about personal data collection by government and law enforcement agencies.

The survey should take approximately 20 minutes to complete and is rewarded with $0.50 (US).

Section 1: Demographic Information

We would like to begin by collecting some basic demographic information.

1. What is your age in years? (You may decline to answer by leaving this field blank.) *[Textfield]*
2. What is the highest level of education you have completed? *[Select one]*
    - [List of options – see Figure 1]
3. In what country do you currently reside? *[Select from list of countries in dropdown menu]*
4. What is your gender? *[Select one]*
    - Male
    - Female
    - Other
    - Prefer not to answer
5. Which of the following best describes your current occupation? *[Select one]*
    - [List of occupations – see Figure 2]
6. On average, how many hours a day do you spend online? *[Textfield]*
7. Do you use a smart phone (for browsing the internet, getting directions, downloading apps, storing photos/videos, etc.)? *[Yes/No/Prefer not to answer]*
8. Do you currently or have you ever worked in a field related to computer science or IT? (e.g., programmer, system administrator, computer helpdesk attendant, website developer) *[Yes/No/Prefer not to answer]*
9. Are you currently or have you ever been employed by a government agency? *[Yes/No/Prefer not to answer]*
10. Are you currently or have you ever been employed by a law enforcement agency? *[Yes/No/Prefer not to answer]*

Section 2: Data Collection Concerns, Preferences, and Perceptions

Many everyday activities are moving online, including shopping, banking, socializing, and entertainment. These activities generate large amounts of personal data, including phone records, emails, and website browsing activity. We would like to learn about your opinions on the collection of personal data by law enforcement and government agencies.

1. Please indicate your agreement with the following statements:
   [5-point Likert scales ranging from 'Strongly Agree' to 'Strongly Disagree', and 'Prefer not to answer']
    - By collecting data from its citizens, the government can increase my safety.
    - By collecting data from its citizens, the government can strengthen the economy.
    - By collecting data from its citizens, the government can reduce crime.

2. How comfortable are you with law enforcement agencies being able to require organizations to disclose personal information, without a warrant, in the following circumstances? (Personal information might include a service subscriber's name, address, phone number, phone records, or similar data.)
   [5-point Likert scales ranging from 'Very Comfortable' to 'Very Uncomfortable', and 'Prefer not to answer']
   - To locate a young person who has run away
   - To locate a person who has gone missing or been kidnapped
   - To locate a person who may be a danger to themselves or others
   - To gather evidence in support of an investigation into a serious offence
   - To gather evidence in support of an investigation into a minor offence
   - To gather information about a person's online activities
3. How comfortable would you be with government or law enforcement agencies collecting data from individuals if you knew the data was being used only for each of the following purposes?
   [5-point Likert scales ranging from 'Very Comfortable' to 'Very Uncomfortable', and 'Prefer not to answer']
   - Preventing child exploitation
   - Combating the drug trade
   - Preventing crime within the country
   - Determining the desires of the country's populace (e.g., to develop policies, legislature, etc.)
   - Preventing illegal streaming/downloading of copyright-protected content
   - Monitoring political dissidence, or protest against government policy
   - Preventing political dissidence, or protest against government policy
   - Maintaining national security
4. With which of the following statements do you most agree? Companies (such as Google, Facebook, Microsoft, or telecommunications service providers) should comply with requests from law enforcement or government agencies…
   [Select one]
   - Whenever the company receives a request, without a warrant or other legal document
   - Only when presented with a warrant or similar legal document, or in emergency situations (such as tracing the location of a kidnapped person's cell phone)
   - Only when presented with a warrant or similar legal document
   - Never
   - Prefer not to answer
5. Companies should notify a user when a government or law enforcement agency makes a request for that user's data.
   [Select one]
   - Always
   - Generally, yes, but there are some situations in which they shouldn't
   - Generally, no, but there are some situations in which they should
   - Never
   - Prefer not to answer

6. For how long should companies store user data (such as emails, phone records, or browsing activity) in case it's needed by law enforcement or government agencies?
   [Select one]
   - Less than 1 month
   - 1 to 4 months
   - 4 to 8 months
   - 8 to 12 months
   - More than 12 months
   - Companies should **never** store user data for the purpose of providing that data to law enforcement or government agencies.
   - Prefer not to answer
7. With which of the following statements do you most agree?
   [Select one]
   - Government or law enforcement agencies should be able to access or collect my data at any time for any reason.
   - Government or law enforcement agencies should be able to access or collect my data at any time for the purposes of policing or maintaining national security.
   - Government or law enforcement agencies should require approval (e.g., a warrant) from a court of law before accessing or collecting my data.
   - Government or law enforcement agencies should require approval from a designated member of the government before accessing or collecting my data.
   - Government and law enforcement agencies should **not** be able to access or collect my data at any time or for any reason.
   - Prefer not to answer
8. My government is honest and transparent about the data it collects from its citizens.
   [5-point Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree', and 'Prefer not to answer']
9. I would like my government to be more transparent about the data it collects from its citizens.
   [5-point Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree', and 'Prefer not to answer']
10. I believe the government is capable of keeping my personal information safe.
    [5-point Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree', and 'Prefer not to answer']
11. I believe the government is capable of collecting accurate data.
    [5-point Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree', and 'Prefer not to answer']
12. I am comfortable with multiple government agencies sharing data pertaining to me.
    [5-point Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree', and 'Prefer not to answer']
13. If government or law enforcement agencies were to collect data pertaining to you, would you be concerned by any of the following issues?
    [5-point Likert scales ranging from 'Very Concerned' to 'Very Unconcerned', and 'Prefer not to answer']
    - I would be unable to know what data they had collected.
    - There may be errors or inaccuracies in the collected data.

    - Government agencies could use collected data to verify information related to income taxes or eligibility for social welfare programs.

14. If government or law enforcement agencies were to collect data pertaining to you, how important or unimportant would you find each of the following abilities?
    [5-point Likert scales ranging from 'Very Important' to 'Very Unimportant', and 'Prefer not to answer']
    - The ability to view the data that had been collected about me.
    - The ability to rectify any errors in the collected data.
15. How comfortable are you with **government or law enforcement agencies** collecting each of the following types of data?
    [5-point Likert scales ranging from 'Very Comfortable' to 'Very Uncomfortable', and 'Prefer not to answer']
    - **Online activity**
        - Email content (e.g., email messages, attachments)
        - Email metadata (e.g., email sender, recipient, date and time email was sent)
        - Internet search terms
        - Internet browsing history (e.g., websites visited)
        - IP address
        - Content from social networking sites (e.g., posts, videos, photographs)
        - List of friends on social networking sites
        - Activity logs from social networking sites (e.g., time of login, IP address from which login occurred)
        - Internet service subscriber information (e.g., name and address of the individual paying for internet service)
    - **Mobile phones**
        - Content of phone calls
        - Phone call metadata (e.g., phone number called, time and date of call, duration of call, but not the actual content of the phone call)
        - Text message content
        - Text message metadata (e.g., message sender, recipient, date and time message was sent, but not the actual content of the text message)
        - Phone service subscriber information (e.g., name and address of the individual paying for phone service)
    - **Finances**
        - Financial information (e.g., online banking transactions, account balances)
        - Credit card number
        - Online purchases (e.g., items bought)
    - **Location**
        - Exact current location
        - Current town/city without pinpointing exact location
        - Current country without pinpointing exact location
    - **Other Information**
        - Medical information (e.g., medications taken, illnesses)
        - Political views
        - Religion
        - Sexual orientation

16. How comfortable are you with **private corporations** collecting each of the following types of data?
[5-point Likert scales ranging from 'Very Comfortable' to 'Very Uncomfortable', and 'Prefer not to answer']
    - **Online activity**
        - Email content (e.g., email messages, attachments)
        - Email metadata (e.g., email sender, recipient, date and time email was sent)
        - Internet search terms
        - Internet browsing history (e.g., websites visited)
        - IP address
        - Content from social networking sites (e.g., posts, videos, photographs)
        - List of friends on social networking sites
        - Activity logs from social networking sites (e.g., time of login, IP address from which login occurred)
        - Internet service subscriber information (e.g., name and address of the individual paying for internet service)
    - **Mobile phones**
        - Content of phone calls
        - Phone call metadata (e.g., phone number called, time and date of call, duration of call, but not the actual content of the phone call)
        - Text message content
        - Text message metadata (e.g., message sender, recipient, date and time message was sent, but not the actual content of the text message)
        - Phone service subscriber information (e.g., name and address of the individual paying for phone service)
    - **Finances**
        - Financial information (e.g., online banking transactions, account balances)
        - Credit card number
        - Online purchases (e.g., items bought)
    - **Location**
        - Exact current location
        - Current town/city without pinpointing exact location
        - Current country without pinpointing exact location
    - **Other Information**
        - Medical information (e.g., medications taken, illnesses)
        - Political views
        - Religion
        - Sexual orientation

Section 3: Concern and Knowledge about Privacy

1. In general, how concerned are you about the protection of your privacy?
[5-point Likert scale ranging from 'Very Concerned' to 'Very Unconcerned', and 'Prefer not to answer']
2. How would you rate your knowledge of the laws pertaining to privacy in your country?
[5-point Likert scale ranging from 'Very Good' to 'Very Poor', and 'Prefer not to answer']