# The Privacy Attitude Questionnaire (PAQ): Initial Development and Validation

Mark H. Chignell, Anabel Quan-Haase, and Jacek Gwizdka
Knowledge Media Design Institute
University of Toronto

Privacy has been identified as a key issue in a variety of domains, including electronic commerce and public policy. While there are many discussions of privacy issues from a legal and policy perspective, there is little information on the structure of privacy as a psychometric construct. Our goal is to develop a method of measuring attitudes towards privacy that can guide the design and personalization of services. This paper reports on research that we have been carrying out on the development of a Privacy Attitudes Questionnaire (PAQ). The development of an initial version of the PAQ is described. The factor structure of the PAQ is assessed, and cluster analysis is used to identify potential stereotypes with respect to attitudes towards privacy amongst different groups of people. The results of this study are then discussed in terms of a) the refinement and utility of the PAQ, and b) their implications for further research on attitudes towards privacy..

## Introduction

As technologies become more functional and complex there is increasing scope and need for personalization of interfaces and functionality to accommodate individual differences. Privacy has been identified as a key issue in a variety of domains, including electronic commerce and public policy associated with activities of police forces and other agencies charged with protecting the public good. While there are many discussions of privacy issues from a legal and policy perspective, there is little information on the structure of privacy as a psychometric construct. Thus there is a need for privacy measures that will support software design in general, and the personalization of services in particular. In view of increasing concern about erosions of privacy and rapid advances in technology, there is a need to assess what attitudes towards privacy exist, and how these attitudes differ amongst individuals, and between different groups of people. This paper reports on the development of an initial version of the Privacy Attitudes Questionnaire (PAQ) and the assessment of its factor structure and its potential for identifying stereotypes with respect to attitudes towards privacy amongst different groups of people.

## Background

Privacy has been identified as a key issue in a variety of domains, including electronic commerce and public policy associated with activities of police forces and other agencies charged with protecting the public good. Among a number of concerns, there is the issue of the role of privacy as a right (Thomson, 1984), and longstanding fears concerning growing restrictions on privacy as information technologies are harnessed by strong governments (Orwell, 1948), most recently exemplified by the opposition to the U.S. National Identity Program (Gartner, 2002).

Attitudes to what we think of as "privacy" may differ along multiple dimensions, relating to the type of information being dealt with, the level of trust we have in people or organizations with access to the information, and so on. For instance, Solove (2002), reviews a number of different conceptions of privacy, that may also be interpreted as having correlated attitudinal dimensions. His list included the following conceptions: The right to be left alone; limited access to the self; secrecy; control over personal information; personhood; intimacy.

Past empirical studies on privacy attitudes have generally been related to or stimulated by emerging public policy issues. For instance, Ekos (1993) carried out a telephone survey of 3,000 Canadian households between October 28 and November 4, 1992. They found a relatively high level of concern about privacy amongst the Canadian public with 92 percent of the respondents expressing at least a moderate level of concern, and 52 percent claiming to have "extreme" concern with personal privacy. This level of concern closely matched that for topics such as unemployment (56 percent) and the environment (52 percent). Studies in the United States have also found high levels of concern about privacy. For instance, UCLA's 2001 Internet Report found that 95% of respondents were "very concerned" about privacy, while the AOL/Roper Starch's 2000 Worldwide Adult Cyberstudy survey found that 94% of respondents were very concerned about privacy.

Westin (1967) identified three different types of technical threat to privacy: physical surveillance; psychological surveillance; data surveillance. Given the vast amounts of data that are now routinely collected, there has been considerable attention to the issue of data surveillance and the uses to which confidential personal information are put. Other technologies have created opportunities for increased physical surveillance at a distance, whether by organizations (e.g., using video cameras in public spaces) or

by other individuals. Video monitoring, instant messaging and other potentially disruptive technologies change the way that people communicate, and the ways in which various forms of personal data are disseminated (Isaacs et al. 2002, Tang, 1994).

There is relatively little research literature on attitudes towards privacy that deals specifically with how individuals view the construct of privacy. Thus designers have relatively little information to go on when design new technologies and interfaces that have privacy implications. Video monitoring, instant messaging and other potentially disruptive technologies change the way that people communicate, and the ways in which various forms of personal data are disseminated (Isaacs et al. 2002, Tang, 1994). Attitudes to privacy may also be influenced by stories and movies about such issues as stalking, identity theft, and the like. Since attitudes towards privacy may vary widely between individuals, a method for measuring privacy attitudes is needed as a basis for design and personalization.

## Development of the PAQ: Methodology

In order to identify some dimensions of privacy, we developed a set of questions in a workshop at the CASCON conference in Toronto in October 2002. There were about 20 participants in the workshop and we separated them in to two equal groups asked each group to generate questions which were thought to be relevant to the issue of privacy. We gave the groups about an hour to generate the questions. After this exercise about 150 questions were generated. We then swapped the questions/groups so that the questions generated by each group were reviewed and critiqued by the other group. Each group was asked to remove those questions that did not seem useful or relevant. In some cases, potentially good questions were re-worded so that they would make more useful items in a survey relating to privacy. It should be noted that the participants in the workshop were generally professionals working in the domain of information technology and were not selected for any prior interest or expertise with respect to the domain of privacy research or policy.

Once the questions/items were selected, an initial privacy survey was constructed out of them. Items in the survey were listed in the same order in which they were generated. No attempt was made to edit the items nor to eliminate redundancy. 79 items were chosen to be included in the draft version of the PAQ and each one of these items was rephrased into a statement with which people were asked about their level of agreement on a five point scale (1=strongly disagree; 2=disagree; 3=neither disagree nor agree; 4=agree; 5=strongly agree). Some of the statements were then identified as items that a privacy survey might try to predict (i.e., items that might be used to develop an initial validation of the survey). At the close of the workshop 13 of the participants filled out the 79 items. We then gave the same survey to 8 students working in a research laboratory that specializes in user interface design at the name-to-be-added. This first set of 21 responses (in total) was then characterized as the Information technology/user-centred design subsample. About a month later (November 2002) the survey was administered to 25 students in a graduate course on research methods within the Mechanical and Engineering Department at the University of Toronto. A third subsample was then collected in January 2003 by administering the initial version of the PAQ to 22 graduate students in the Faculty of Information Studies at the University of Toronto, resulting in a total initial sample of 68 people, which contained three subsamples. The sample data was then subjected to statistical analyses (factor analysis, item analysis and cluster analysis) that were designed to answer two questions:

- What are the underlying factors/scales of privacy based on similarities (as measured by statistical correlation) in how people answered groups of items?
- What subgroups of people (i.e., types of people) exist with respect to how they feel about the different privacy scales?

## Results

The survey responses were subjected to Principal Factor analysis with Varimax Rotation (using Kaiser Normalization). Nine of the items in the survey were removed from this analysis because they were chosen to be criteria to predicted in subsequent analyses (see Table 2). A number of initial analyses were run. After those initial analyses, items which did not seem to be participating in interpretable factors and which were judged to be either redundant with other items, unrelated to privacy issues, or were ambiguous or unclear, were removed from the analysis.

A factor analysis was then run with the edited data set. A five factor solution accounting for 33% of the variance was chosen as the best characterization of the data. The resulting factors were then subjected to item analysis using Cronbach's alpha as the criterion.

For each factor, corresponding scales were identified (linear sums of the highly loading items) that were interpretable, and that provided the highest level of alpha (internal validity) for the lowest number of items. The scales thus identified are shown in Table 1. Note that the first four scales had fairly good reliability (alpha greater than .7), while the fifth scale (labelled as "protection") showed only moderate reliability (alpha = 0.666). One of the scales (labelled as "NotPrivacy") seemed more related to low social desirability or even deviance, rather than privacy. This is indicated by the types of item in that scale, which included willingness to speed, liking gossip, interest in sharing confidences, and lying to one's doctor.

The "Personal Information" scale contains seven items that can be interpreted as relating to one's willingness to make one's personal information available to others. Items in this scale are: willingness to undergo DNA testing or retinal scanning (item 1); willingness to provide one's personal identification number (item 13); willingness to disclose one's credit rating (tem 15); willingness to wear a name tag (item

23), willingness to have one's email monitored (item 28), willingness to have one's messages tracked (item 31), and willingness to have strangers in one's house (item 58).

The "Exposure" scale contains four items reflecting uses of one's image or one's family image either in reality television (item 66), or as photos that are made public (item 19) or published on the Web (items 16 and 61). This scale may relate to the conceptions of "limited access to the self" and "personhood" noted by previous researchers.

| | **PersonalInfo** | **0.762** | | **Monitoring** | **0.696** |
|---|---|---|---|---|---|
| 1 | DNA/retina | | -4 | Showcalldisplay | |
| 13 | Personal ID Number | | 6 | phonesurveys | |
| 15 | creditrating | | -18 | apartmentbuzzer | |
| 23 | name tag | | 29 | PublicVideo | |
| 28 | monitor email | | 35 | HomePhoneforBiz | |
| 31 | messagetracking | | 49 | surveysOK | |
| 58 | StrangersInHouse | | 56 | RedLightCameras | |
| | | | 57 | SpeedingCameras | |
| | **Exposure** | **0.731** | | **NotPrivacy** | **0.703** |
| 16 | familyphotos | | 27 | SpeedIfPossible | |
| -19 | publicPhotoOk | | 34 | CellDiffPrivate | |
| 61 | PhotoWebPage | | 40 | gossip | |
| 66 | RealityTV | | 42 | confidences | |
| | | | 65 | lieToDoctor | |
| | **Protection** | **0.666** | | | |
| 9 | unlisted phone | | | | |
| 10 | clearCache | | | | |
| 11 | ChangePasswords | | | | |
| 12 | CloseCurtains | | | | |
| 14 | QuestionPersonalInfo | | | | |
| 55 | PersonalFirewall | | | | |

**Table 1. Five subscales identified**

The fourth privacy scale identified in this study has lower reliability (Cronbach's alpha of 0.666) and needs to be evaluated further to establish its usefulness (as do the other scales identified here). This scale is labelled "Protection" because it generally refers to actions one takes to Protect oneself against unwanted intrusion. Although willingness to protect oneself against invasions of privacy is not listed in Solove's conceptualisations, it seems to be a potentially important attitude relating to privacy, particularly with respect to predicting how people might respond to various informational policies or service offerings. The seven items in this scale refer to the use of an unlisted phone number (item 9), the clearing of cache when Web browsing (item 10), the frequency of changing passwords (item 11), closing curtains (item 12), questioning why personal information has to be provided (item 14), and using a personal firewall (item 55).

## Predictive Validity

Nine of the survey items were identified by participants in the original workshop (where the PAQ items were derived) as criteria that privacy attitude scales might have some predictive relationship with. These potential criteria were not used in developing the subscales. While these items are not likely to cover all aspects of privacy, they provide a way of assessing the predictive validity of the four privacy scales developed in this study. These items (criteria) are shown in Table 2.

| **Item** | **Item Text** | **Significant Predictors** | **$R^2$** |
|---|---|---|---|
| 24 | I use loyalty cards. | PersonalInfo, Monitoring | .213 |
| 32 | I would pay for things with a cell phone. | PersonalInfo, Exposure | .217 |
| 33 | I don't mind loaning out my credit card to people I trust. | | |
| 43 | I would rather use cash than credit or debit cards. | PersonalInfo | .157 |
| 44 | I use an air miles card. | | |
| 45 | I purchase things online. | PersonalInfo, Exposure | .187 |
| 69 | I give copies of my housekeys to other people | PersonalInfo | .199 |
| 70 | I would be comfortable having a housekeeper | | |
| 79 | I'd be interested in using digital cash. | PersonalInfo, Exposure | .187 |

**Table 2. Nine predictive criteria derived from the PAQ and predictive relationships involving the privacy scales**

Stepwise multiple regression analyses were used to predict scores on each of the criteria listed in Table 2, using the four privacy scales identified earlier. Backward entry was used, where all four scales were entered initially, with predictors being removed if they did not contribute significantly to the overall variance fitted. The results of these analyses are summarized in Table 3, which shows all the significant predictive relationships ($p<.01$) along with which privacy scales were the significant predictors. It can be seen that there were six items involved in predictive relationships, two of which (using cash not credit, loaning house keys) were predicted by personal information alone, three of which were predicted by a combination of personal information and exposure (cell phone payment, purchasing online, using digital cash), and one of which involved personal information with monitoring (loyalty cards). As shown by the corresponding r-squared values, roughly 20% of the variance in the criterion was accounted for in most cases.

As a second way of checking external validity we then used analysis of variance to see if scores on the privacy scales differed between the three subsamples that were used. As can be seen in Table 3, the exposure scale differed significantly between the subsamples ($p<.05$), while the difference in attitudes towards personal information was borderline significant ($p<.10$).

| | **F(2,65)** | **Sig.** |
|---|---|---|
| **PersonalInfo** | 2.556 | 0.085 |
| **Monitoring** | 1.583 | 0.213 |
| **Exposure** | 3.835 | 0.027 |
| **Protection** | 0.495 | 0.612 |

**Table 3. ANOVA results: Effect of subsample on Privacy Factors Scales**

Figure 1 shows a bar chart of the two significant (or borderline significant) privacy scales across the three subsamples. It can be seen that willingness to disclose personal information tended to be higher in the first subsample (information technologists), while willingness to expose one's image was lowest in the group of information studies graduate students

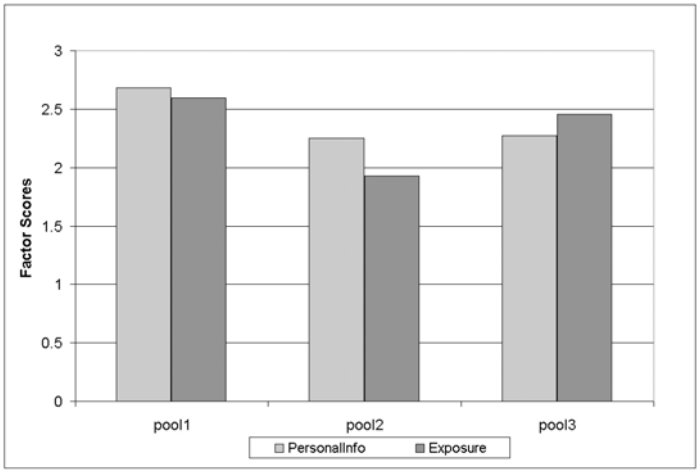


**Figure 1. Differences in privacy attitudes between the subsamples**

## Privacy Stereotypes

In order to explore the usefulness of the draft version of the PAQ, and the four privacy scales derived from it, we then attempted to identify privacy stereotypes, i.e., subgroups of people with similar attitudes towards privacy. K-means cluster analysis of the 68 respondents was carried out using the four privacy scales as the variables. Each scale was created by summing together its component items. Items that were reversed from the other items in the scale (e.g., item 19 in the "Exposure" scale) were inverted using a transformation (new value = 6 - Rating, where 1 = strongly disagree, 2 = disagree, etc). For instance, for an item with a value of strongly disagree, the reversed value of that item would be strongly agree. In order to improve the comparability of the scales, we then divided the scale value by the number of items in the scale, yielding a level of agreement (between 1 and 5) with the scale concept.

Three clusters were identified (numbered as clusters 1 through 3 below) with 25, 19, and 24 people, respectively in each. Average ratings on each of the four privacy scales differed significantly between these clusters (the corresponding ANOVA summary table is shown in Table 4.

| | F(2,65) | Sig. |
|---|---|---|
| **PersonalInfo** | 13.999 | 0 |
| **Monitoring** | 4.804 | 0.011 |
| **Exposure** | 42.326 | 0 |
| **Protection** | 31.97 | 0 |

**Table 4. ANOVA results: Four Privacy Scales across Three Clusters**

The differences in privacy attitudes between the three clusters are summarized in Figure 1. People in cluster one are less likely to give out personal information, but also less likely to adopt protective strategies (like clearing cache, changing passwords, etc.). They are relatively unwilling to expose themselves (e.g., by posting photos on the Web), but they generally are willing to be monitored. People in cluster two tend to be willing to provide personal information, be monitored, and expose their images to the public. They have a moderate level of interest in Protection. People in cluster three tended to be most concerned with privacy, being generally unwilling to divulge personal information, be monitored, or expose their images. They also had a high protection score (i.e., adoption of strategies to protect their privacy). In summary, people in cluster one seem most concerned about retaining personal information and their images, but being relatively unconcerned about being monitored or the need for protection. People in cluster two seem relatively unconcerned about privacy (at least in terms of the scales derived in this study). People in cluster three were the most concerned about privacy as measured here.

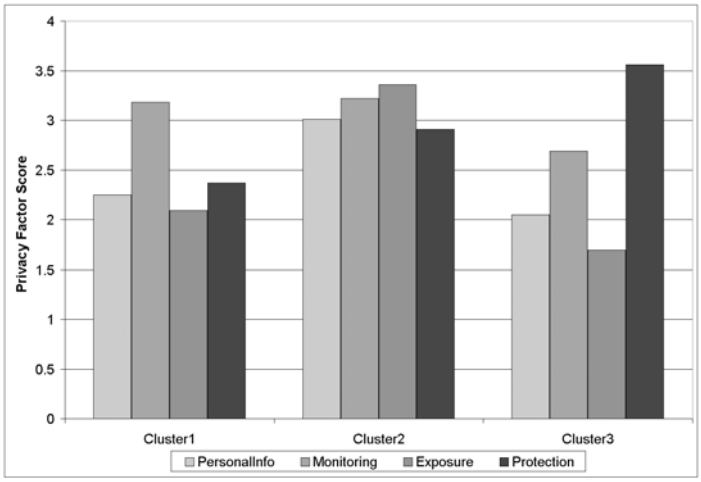


**Figure 2. Differences in Privacy Attitudes between the Clusters**

## Conclusions

From a methodological perspective this study has shown that people in a workshop are able to generate useful attitude instrument items, which can then form an effective first draft of a corresponding attitude questionnaire. In the present case, the resulting draft version of the PAQ yields interpretable attitude scales that are useful both in predicting privacy-related criteria and in differentiating subsamples of people. In the small sample of 68 people used in this study we also identified three clusters (roughly equal in size) that differed according to their scores on the four privacy scales.

It is our expectation that this type of empirical research can contribute greatly to discussions of attitudes towards privacy and their impact on behaviour. It is also highly relevant to human-computer interaction, since many new interfaces and services have implications for how personal information is handled and how people are portrayed or presented to others. Thus the PAQ may prove to be a useful tool for designers and human factors engineers seeking to

personalize services where privacy issues are relevant. A revised version of the PAQ is presented in the Appendix.

## References


Carroll, J.M. (1991). Confidential Information Sources: Public and Private, Second Edition. Boston: Butterworth-Heinemann.

Ekos (1993). Privacy Revealed: The Canadian Privacy Survey. Ottawa: Ekos Research Associates Inc.

Isaacs, E., Walendowski, A. & Ranganathan, D. (2002). Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions, Proceedings of CHI 2002, Minneapolis, MN.

Gardner N., CTI and evaluation, Br. J. Educ. Technol. 19, 225-226, 1988.

Kang, J. (1998). Information Privacy in Cyberspace Transactions, Stan. Law Rev., 50, 1202-1246.

Keisler S.B. and Sproull L. S. (eds), Computing and Change on Campus. Cambridge University Press, 1987.

Ng-Kruelle, G., Swatman, P.A., Rebne, D.S, and Hampe, J.F. (2002) Price of Convenience: Dynamics of Adoption Attitudes and Privacy Sensitivity Over Time, School of Information Systems, Deakin University, Working Paper 2002/33

Public Interest Advocacy Centre (1995). Surveying Boundaries: Canadians and their Personal Information. Ottawa: Public Interest Advocacy Centre.

Solove, D.J. (2002). Conceptualizing Privacy. California Law Review 1087

Tang, J., Isaacs, E., & Rua, M. (1994) Supporting distributed groups with a Montage of lightweight interactions, Proceedings of CSCW, Chapel Hill, NC, 23-34.

Thomson, J.J. (1984). The Right to Privacy. In F.D. Schoema (Ed.), Philosophical Dimensions of Privacy: An Anthology. Cambridge: Cambridge University Press.

Westin, A.F. (1967). Privacy and Freedom. N.Y.: Atheneum.


## Appendix - PAQ Version 1.0

The following items are arranged by subscale (E: Exposure; M Willingness to be Monitored; P: Interest in Protection; PI: Willingness to share Personal Information). Responses should be expressed as agreement with each of the statements (1=strongly disagree; 2=disagree; 3=neither agree nor disagree; 4=agree; 5=strongly agree). For items marked as negative (-E or -M) the scale should be reversed prior to scoring. To obtain the score on each subscale, sum the total score across each of the items in the subscale.

E I would like to keep photos of my family on the internet
-E I'd object to my photograph appearing in a public place without my permission
E I would put my photo on my personal web page.
E I would like to participate in reality TV.
E I'd like a high fence in my backyard
E No organization or person should disseminate personal information about me without my knowledge
E I tell some of my work colleagues about my personal life
E I like to tell stories about my family to my friends
E I would not mind appearing on television or being quoted in a newspaper
-M I frequently would like to block my phone number on call display
M I respond to telephone marketing surveys
-M I prefer not to have my name listed on a building directory
M Video cameras should be used in public places to improve public safety and security.
M I would give my home phone number to business clients?
M I like to fill out surveys and contests.
M Red light (intersection) cameras should be used.
M Speeding cameras should be used
M Insurance companies should not have access to people's health records
P I would prefer to have an unlisted phone number
P I like to clear my cache frequently for privacy reasons
P I like to change my passwords frequently
P I like to close my curtains at home at night
P I frequently question why I'm providing personal information
P I would prefer people to knock before coming into my office or bedroom
P I worry about the possibility that my conversations will be overheard
P I would use a personal firewall
PI I am comfortable with giving a DNA sample
PI I am comfortable giving out my personal identification number
PI I am comfortable in allowing others to check my credit

PI I am comfortable wearing a name tag
PI Employers should be able to monitor employee email
PI It is ok to use messaging services even if the messages could in principle be tracked.
PI I allow strangers to enter my house while I'm not there
PI I am comfortable with having my retina scanned
PI I do not mind using my real name in online discussions
PI My medical information should never be communicated to people or organizations without my permission