

# Keeping Up Appearances: Understanding the Dimensions of Incidental Information Privacy

Kirstie Hawkey and Kori M. Inkpen

Faculty of Computer Science, Dalhousie University  
Halifax, NS B3H 1W5  
{hawkey, inkpen}@cs.dal.ca

## ABSTRACT

We conducted a survey of 155 participants to examine privacy concerns relating to the viewing of incidental information (i.e. traces of previous activity unrelated to the task at hand) in web browsers. We have identified several dimensions of privacy for this domain. Results revealed the scope of this problem and how location and device affect web browsing activity and contribute to the types of incidental information that may be visible. We found that there are different privacy comfort levels inherent to the participant and dependent on the context of subsequent viewing of incidental information, including the sensitivity of the content, their relationship to the viewer and the level of control retained over input devices.

## Author Keywords

Privacy, survey, incidental information, web browsing, collaboration

## ACM Classification Keywords

H.5.3 [Information Interfaces and Presentation] Group and Organization Interfaces: Collaborative Computing; Web-based Interaction

## INTRODUCTION

Colleagues often gather on an ad hoc basis around a display to collaborate on a project. Incidental information about past activities on the computer is then visible with casual inspection (as in Figure 1). This information may or may not be appropriate for the current viewing context. Typically with their personal displays, people position the screens so that they are not visible to others. They may also rely on social protocols that restrain others from intentionally viewing private displays [20]. However, during collaboration around a display, normative privacy does not apply as the display itself becomes an object in the collaboration. Any incidental information displayed will not only be visible, but likely be viewed.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2006, April 22-27, 2006, Montréal, Québec, Canada.  
Copyright 2006 ACM 1-59593-178-3/06/0004...\$5.00.

Privacy management of incidental information can be difficult for computer users. It is not always clear exactly which traces of activities are being created and stored and which can subsequently be viewed by others during normal computer usage [30]. Nor is it clear whom all the future viewers will be and the context under which material will be viewed, particularly when devices are mobile and used in both personal and business settings [23]. Privacy concerns can increase when displays are viewable by many people in a group and members aren't clear which information is being viewed, by whom, and how often [14].

Currently, users must make tradeoffs to manage the privacy of this information: they can either work efficiently in a familiar environment, with access to convenience features and usual layout, or work awkwardly in a sterile environment. Our overall goal is to provide users with tools to manage the privacy of their incidental information, only revealing information appropriate for the current context.

Research in the domain of incidental information privacy is just beginning. Previous research in other privacy domains has found that privacy concerns are highly nuanced and individual. The intersection of privacy management [3, 20] and personal information management [6] results in a challenging problem due to the complexity and volume of



Figure 1. Incidental information privacy example. Previous search terms are revealed to a collaborator when the user begins to type “privacy research” in the form.

information [12]. Before developing privacy management solutions for incidental information, we must fully understand the dimensions of the problem.

Web browsers were selected as the representative application for this research as they are often used during co-located collaboration to find information or share previously found web sites. In addition, web browsers are used for a wide variety of tasks, both personal and work related. Web browsers have many convenience features, such as History, Auto Complete, and Favorites, that assist users when navigating to previously visited pages, but also display traces of prior activity that users may prefer to remain private. The nature of these traces often leads to their unintentional viewing. For example, Auto Complete will reveal search terms previously entered; during a search for “privacy research” a previous search for “personal bankruptcy laws” may be revealed (as in Figure 1).

This paper presents a survey examining privacy concerns related to the incidental viewing of web browsing traces. The survey objectives were threefold: 1) to understand the scope of the privacy issues in this domain, 2) to examine how browsing behaviours affect the visible content, and 3) to investigate the role of content sensitivity, level of control, and viewer on privacy comfort levels. We first present related privacy work including tools, theory, research in other domains, and modeling. We next describe the dimensions of incidental information privacy that may affect the comfort of users in a given situation. We then present the survey methodology and our results. Following a discussion of the results, we conclude with future work.

## RELATED WORK

**Web Browsing Privacy Tools for Incidental Information**  
COLLABCLIO [17] was a research system developed to support automated sharing of web browsing histories in a workplace environment. It provided users with a binary classification scheme (public/private). Users expressed a wish for finer-grained classification to reflect differing privacy needs for sub-groups of collaborators.

While there are commercial products that allow the erasure of traces of browsing activities, those traces are often valuable during navigation and may decrease productivity if removed. As an example, Window Washer [2] allows a user to delete traces such as Auto Complete, History, and Recent Documents. However, with the exception of the ability to save selected Cookies, the decision to erase a class of traces erases all instances indiscriminately.

### Privacy Theory

Privacy theories vary according to the domain in which privacy issues occur. We discuss those that fit most closely for the domain of incidental information privacy. Palen and Dourish [23] describe three interrelated boundaries for privacy management: the disclosure boundary, the temporal boundary, and the identity boundary. These boundaries between what is considered public or private are

continuously refined depending on the context. This model of privacy fits incidental information privacy well. Users would like to be able to control an appropriate level of content sensitivity given the context of viewing (disclosure boundary). The persistence of traces of previous activity (temporal boundary) makes it difficult for users to ensure that they are presenting themselves appropriately for their current role (identity boundary).

Goffman [10] first introduced the need to project different personas or faces in our social interactions. The face that we present in any given situation depends not only on the current audience but also on the current conditions. The combination determines how much and what information will be disclosed. Lederer et al. [18] discuss how activities convey the essence of a persona. Knowledge of activities is more sensitive when identity is known as the activities can reveal hidden personae. With traces of incidental information, a person's actions in one area (e.g. personal browsing) may later be viewed in another area (e.g. workplace). Information that is appropriate for a friend to see may not be appropriate if viewed by an acquaintance or an authority figure with whom one would prefer to present a more formal or otherwise restricted face.

Moor [20] uses a “control/restricted access” theory of privacy; users can fine-tune the privacy of their information by both recipient and information type via zones of privacy. However, with incidental information, not all potential viewers of the information may be apparent at the time the traces are created.

## Privacy Research

### *Information Privacy*

Online privacy concerns have been examined in great detail and the Platform for Privacy Preferences Project (P3P) [1] has developed standards that allow users greater control over the use of their personal information at participating websites. A 1998 survey [3] examining privacy preferences for Internet users found differing levels of sensitivity about personal data, ranging from little concern about providing such information as their favourite television show to great concern over credit card and medical information. Interestingly, 18-20% expressed concern over even the most innocuous data. The authors suggest that an individualized approach is necessary given the large variance in reactions. However, online privacy research focuses on consumer privacy and not interpersonal privacy as occurs during the viewing of incidental information.

Recent information sharing research has looked at privacy comfort for various types of information and recipients of that information. Cadiz and Gupta [7] found that, in general, people were open to sharing information except with strangers; their results were also highly nuanced. A similar study by Olson et al. [21] found that, in terms of comfort of sharing information, participants clustered recipients into four groups: public, work relationships,

family, and spouse. Their results suggest that the types of incidental information that may be revealed during web browsing (e.g. personal activities like viewing non-work related websites, transgressions like viewing erotic material) are considered more sensitive than other content such as contact and availability information. Incidental information privacy when web browsing is less clear-cut than for static information (e.g. contact information) that may be shared electronically. There are likely several levels of sensitivity of content within the traces, the amount of highly sensitive content may fluctuate over time, and the user may be less aware of what is actually saved.

#### Computer Supported Collaborative Work (CSCW)

Privacy issues have been addressed extensively in distributed CSCW research as these systems are capable of capturing and displaying a great deal of awareness information in an attempt to replicate some of the benefits of co-located collaboration. Strategies for maintaining privacy include only storing and presenting aggregate data where possible [5] and adjusting the level of detail of information depending on the size and public nature of the display [14]. Awareness information privacy is also addressed in the ubiquitous computing community. Lederer et al. [19] examined the relative importance of the inquirer (spouse, employer, stranger, merchant) and the situation for the preferred accuracy of personal information disclosed. Preferred accuracy varied by inquirer, but not by situation (except when the inquirer was the employer).

Privacy issues raised in co-located CSCW research have been limited to privacy of data within an application (e.g. [13]) or on specialized devices dedicated for collaboration (e.g. [25]). However, this view of private information assumes that all information viewed is task-related; during ad hoc collaboration, this is often not the case.

#### Modeling Privacy

The Westin-Harris [22] privacy segmentation model explores consumers confidence in how their personal data is collected and used by companies. The model partitions consumers into three privacy categories: *fundamentalists*, *pragmatists*, and *unconcerned*. Privacy fundamentalists feel strongly that current information practices are a threat to their privacy, while those classified as privacy unconcerned have the opposite viewpoint. Privacy pragmatists tend to weigh the privacy risks of releasing information against the potential benefits (e.g. personalization of a web site). Spiekermann et al. [26] further divided pragmatists into the *identity concerned* and the *profiling averse*. Identity concerned participants were most concerned about revealing contact information such as their name, email, and address to corporate sites while the profiling averse were more concerned about information such as health status, hobbies, and other interests.

The consumer-based privacy segmentation model has been applied to other privacy domains with limited success.

Consolvo et al. [9] did not find the model to be a good predictor for disclosure of awareness information. Patil and Lai [24] used an extended questionnaire with a trust component to model their participants, but did not find correlations between the settings participants would choose and their questionnaire scores. Olson et al. [21] used questions from a trust scale and demographic information (e.g. age, gender) in an attempt to find a small number of questions that indicate privacy preferences for information sharing. Although they state that interesting patterns emerged, none were statistically significant.

Recent research (e.g. [3]) cautions that actual behaviour with respect to privacy practices often does not follow stated privacy concerns. Acquisti [4] proposed enriching privacy models by including psychological models of personal behavior such as immediate gratification and self-control. Jensen et al. [15] studied whether attitudinal information about on-line privacy practices gathered in a survey matched behaviour during an experimental purchasing scenario. The authors suggest that surveys may be best suited to evaluate attitudes and can be used as a baseline with which to compare actual behaviour.

#### INCIDENTAL INFORMATION PRIVACY

Through an examination of related work and our research results to date, several dimensions of incidental information privacy that impact a user's comfort level have been identified. Four dimensions that directly impact the *privacy comfort level* in a given situation include the user's *inherent privacy concerns*, their level of *control*, their *relationship to the viewer* of the display, and the sensitivity of potentially *visible content* (see Figure 2 for an influence diagram). Furthermore, the visible content may depend upon recent *browsing activity*, *browser settings*, and any *preventative actions* taken. Browsing activity itself may vary depending on the *location* of the activity and the type of *computer*.

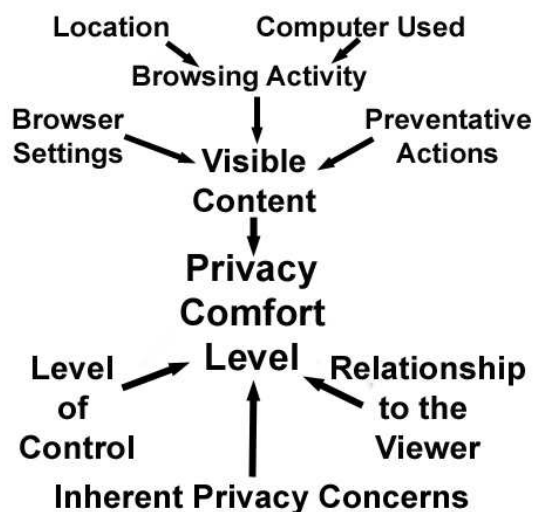


Figure 2. Dimensions that affect the comfort level of users during incidental viewing of traces of prior web activity.

While Figure 2 shows the major influences on privacy comfort levels, these dimensions are often inter-related. For example, advance knowledge of a specific viewer may trigger preventative actions to limit what is visible. The dimensions presented are specific to traces of web browsing activity; however, while the nature of the visible content will change for other types of incidental information, the impact of level of control, viewer, and inherent privacy concerns will likely remain consistent.

The *inherent privacy concerns* of an individual will have a large effect on his privacy comfort level in a given situation. By partitioning participants into a classification scheme such as the Westin-Harris segmentation model (e.g. privacy fundamentalists, unconcerned, and pragmatists) [22], an indication of inherent privacy concerns may be found. Such classifications could also be used as an initial predictor of privacy preferences. Privacy unconcerned participants should have relatively high comfort levels regardless of context; similarly, fundamentalists will have relatively low comfort levels. Privacy pragmatists will likely have varying comfort levels depending on visible content, level of control, and viewers.

Another dimension influencing privacy comfort is the *level of control* a person retains over the information viewed. A high amount of control (e.g. control over input devices) should lessen privacy concerns, while lower levels of control should increase concerns. Incidental information can be hard to control due to its dynamic and temporal nature. Furthermore, users are often uncertain what traces of activity are saved and may be subsequently revealed.

The *relationship to the viewer*, more accurately the persona that a user maintains with a viewer, also impacts privacy comfort levels. Traces of previous web browsing that are out of character for the persona being presented will likely increase discomfort in a situation.

The sensitivity of potentially *visible content* has an effect on privacy comfort levels. Traces of activity that are in character with the persona and setting should cause little concern (e.g. non-confidential work-related browsing activity in the workplace). However, activities that are perceived as transgressions (e.g. personal browsing if company policy does not allow it) may cause greater discomfort. Visible content depends on the *browsing activities* of the user. Browsing activities may depend on the *device* being used and also the *location* of the browsing. For example, someone with both a home and a work computer may refrain from conducting many personal activities while at work, while someone with access only at work may conduct a broader range of activities. Those using a shared computer without a separate login may not conduct the same activities as those with their own PC or own login. A laptop user may perform the majority of their browsing activities on their laptop and move between locations. *Browser settings* (e.g. saving 0 days history) can reduce the visible content as can *preventative actions* such

as erasing all traces. Recent techniques to increase the recognition of information stored in convenience features (e.g. thumbnails of web pages in history files) may help users more easily find a desired page [16], but are also a concern as they increase visibility of incidental information.

## **SURVEY**

We conducted a survey to explore the dimensions of incidental information privacy that arise when web browsers are used during face-to-face collaboration or sequentially. The three main objectives of the survey were to 1) determine the scope of the problem, 2) gain an understanding of the content that may be visible as a result of browsing activity, and 3) measure privacy comfort levels for different contexts of web browsing.

### **Methodology**

The survey was available on-line which allowed participants to complete the survey on their own time, in a place of their choosing which may have promoted more honest responses for questions of a sensitive nature [29]. Mode effects (e.g. elevated responses on ratings scales) between paper- and web-based surveys are generally minimal except for items that deal with computing and information technology [8]. We therefore elected to make the survey available only on-line in order to keep conditions as similar as possible for all participants. Access was controlled through unique personal identification numbers. The survey took about 20 minutes to complete and participants received no compensation. It was refined through several iterations of pilot testing and critiques.

### **Population**

Through email lists and hand-distribution of notices, we recruited 155 participants (57% male) from businesses, the Dalhousie University community, and the public. As participants were not randomly sampled, we cannot claim that our population is representative of all web users. In particular, our population has a high level of education (median Bachelor's degree) and computer experience (avg. 12 years, ranging from 2 to 35). Participants were frequent computer users (median 29-35 hours per week) and web browser users (median 15-21 hours per week). Our participants were diverse with respect to age (avg. 31.5, ranging from 17 to 59) and occupation. Participants' occupations ranged from homemakers to professionals, but students are over-represented at 42.6% of the sample.

### **Questionnaire**

#### *Scope of Privacy Concerns*

We asked a series of questions designed to learn about the general scope of privacy issues participants have related to the incidental information in web browsers. We asked participants to think about their overall computer use during an average day. We enquired about the *frequency* with which ten different types of people (both interpersonal and business/school relationships) might *view* or *use* a participant's computer. Participants were asked to think

about both who can clearly *see* the contents of their screen as they use it and who may *use* their computer and approximately how often they may be in those situations. We used a 5 point scale for frequency (daily, weekly, monthly, rarely, never).

Participants were also asked to reflect upon how they currently handle the tradeoff between convenience and privacy. We asked them to indicate their current settings for their History, AutoComplete, and Favorites browser features. In addition to typical settings, choices included *unsure*, *default*, and *don't use*. We also asked them to indicate all the actions they would take on each of their computers if given advanced warning that somebody else would be working closely with them as they used their web browser and could see all areas of their screen. In addition to actions such as checking and clearing or erasing data in Favorites, History, and Auto Completes, choices included taking *no action* and *limiting control* of input devices.

#### Dimensions Affecting Visible Content

Questions probed web browsing and computer usage both at home and away. We investigated the types of browsing activities in which participants engage, where those activities take place, and on which types of computers.

#### Effect of Context on Privacy Comfort Levels

This section of the survey was designed to learn how context affects privacy concerns. Rather than examining privacy comfort for all types and sensitivities of traces, we examined privacy comfort for overall context: viewer, level of control, and general content sensitivity. Participants were asked to think about "*how comfortable [a situation] makes you feel in terms of privacy.*" Participants rated comfort with a seven point Likert scale (extremely uncomfortable (1), to neutral (4), to extremely comfortable (7)).

We investigated privacy comfort levels for a subset of potential viewer categories (close friend, supervisor, parent, spouse/significant other, colleague/fellow student) for three scenarios that varied in the sensitivity of information. For each category of viewer, participants indicated a comfort level for three different circumstances that varied the level of their control: if the participant was the one in control of the web browser (you), if the viewer was in control of the web browser with the participant sitting right there (other), or if the viewer was in control of the web browser and the participant left the room (away).

The three scenarios were explicit descriptions of recent web browsing behaviour and their order of presentation was counter-balanced. The scenarios were designed to discover the range of comfort a participant had for information of varying sensitivity. All scenarios discussed a situation that led to information seeking behaviour on a web browser and described a set of search topics and web page visits that might be revealed during a future web browsing episode. The scenarios were contrived to be universally 1) *embarrassing* (genital shingles), 2) *neutral* (buying a car),

*You have been experiencing itching and pain in your groin area. You go see the doctor who unfortunately diagnosed you with shingles on the genitals. Shingles can occur in people who have previously had chicken pox. It is a very painful disease. You have been experiencing uncomfortable symptoms and have been looking for relief. You use your web browser to search for such topics as "burning genitals" and "itching groin" and have visited such web pages as [www.yoursexualhealth.com/stoptheburning.html](http://www.yoursexualhealth.com/stoptheburning.html) and [www.genitalhealthcare.com/topics/infectiousdiseases](http://www.genitalhealthcare.com/topics/infectiousdiseases) (which you add to your favorites for future reference).*

**Figure 3. The embarrassing web browsing scenario.**

and 3) *positive* (winning a trip). The embarrassing scenario (Figure 3) was designed to be extremely sensitive in content, but with no judgment about the activity's morality. We also asked participants to indicate their privacy comfort levels according to viewer and amount of control, after first reflecting on their *usual web browsing behaviour*.

## RESULTS

### Scope of the Incidental Information Problem

#### Frequency of Viewers and Users

We collapsed the frequency responses for potential *viewers* and *users* of participants' computers into the categories 'regularly' (daily, weekly), 'occasionally' (monthly, rarely) and 'never'. All participants reported at least one category of viewer that could sometimes *see* their display and only 10 participants reported no categories of potential *users*. The viewing frequency and using frequency (see Table 1) both vary depending on the category of the viewer/user. The most regular potential viewers are close friends, colleagues, spouse/significant other, and supervisor. Acquaintances and employees are more likely to be

Viewers	Frequency of viewing (%)			Frequency of using (%)		
	Reg.	Occ.	Never	Reg.	Occ.	Never
Close friends	36.5	49.6	13.9	16.4	41.8	41.8
Colleagues	56.0	29.9	14.1	16.7	35.6	47.7
Acquaintances	20.6	58.1	21.3	2.2	34.3	63.4
Spouse/S.O.	49.6	20.7	29.6	38.1	24.6	37.3
Tech. support	9.7	59.7	30.6	7.4	48.9	43.7
Supervisor	37.1	28.1	34.1	5.4	22.3	72.3
Audience	3.0	47.0	50.0	--	--	--
Employees	22.9	19.1	58.0	4.6	18.5	76.9
Parents	11.7	24.1	64.2	3.6	18.2	78.1
Clients	9.8	19.6	70.7	0.8	6.1	93.2

**Table 1. Percentage of participants at each frequency (regularly, occasionally, never) for each category of potential viewers and users. Most common response is highlighted.**

occasional viewers as are parents, clients, technical support staff, and audiences at presentations. The frequency of others *using* participants' computers was less than that for *viewing*, with spouses being the most regular users.

### Current Privacy Management Strategies

Participants reported a wide range of uses of their web browsing convenience features that often varied across their computers. For example, 31.8% of the 132 participants with responses for multiple computers indicated they used Favorites differently among those computers. Participants tended to be less likely to use convenience features on their desktop PCs away from home than on their home desktops or laptop computers (e.g. 31.2% don't use Favorites on away PCs, 6.0% don't use on home PCs, 14.7% don't use on laptops). The majority of participants reported that they would take some action if given advanced warning that someone could view their display (66.4% of home PC users, 76.6% of laptop users, and 60.2% of away PC users).

### Dimensions Affecting Visible Content

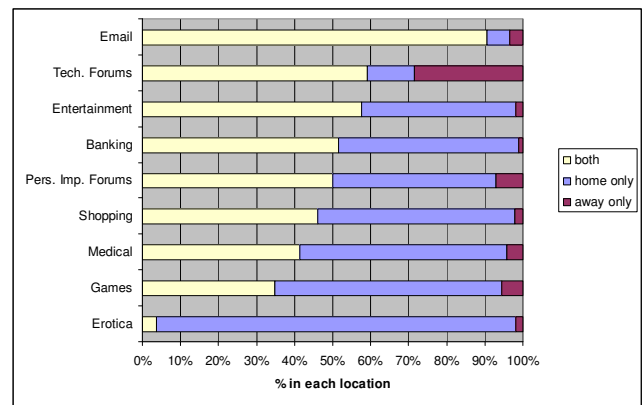
#### Browsing Activities

Almost all participants used their web browsers for email (99.4%) and accessing entertainment information (94.2%). Banking (82.5%), viewing medical information (81.3%), technical support forums (78.9%), shopping (75.5%), and playing games (57.9%) were also popular on-line activities. Fewer participants said they used their web browsers to view erotic material (43.0%) or visit personal improvement forums (37.7%). These activities were reported in a higher frequency than in a recent Stats Canada survey [27, 28]. In 2003, of the 64% of households had Internet access, 81% used it for email, 48% for banking, 56% for medical information, 29% for shopping, and 44% for games. Our participants may therefore be more frequent and experienced Internet users than other Canadians.

#### Location of Browsing Activities

Participants did vary their activities depending on the location of their computer (regardless of type of computer). Most participants that use their web browsers for a given activity will do that activity while at home (73.8% for technical support forums, 91.3-98.5% for remainder). However, only technical support forums and email are accessed similarly at home and away. The remaining activities, which are more personal in nature, are much less likely to occur when participants are away from home (6.2% for erotica, 40.9-55.2% for remainder).

We were interested whether participants partitioned their web browsing activities according to location or conducted the browsing in both locations. In order to reflect participants' choice rather than circumstance, we omitted participants that only indicated browsing activity for one of the locations. While activities such as email, technical support forums, and entertainment browsing often occurred in both locations, the more personal the type of activity, the



**Figure 4. For those participants that do each activity, the proportion who conduct the activity both at home and away, only at home, or only away from home.**

more likely the activity was conducted only at home (see Figure 4). With the exception of technical support forums, few users only conducted browsing activities while away from home. Interestingly, all participants who indicated viewing erotica while away from home were laptop users.

#### Computers Used for Web Browsing

Participants browsed the web on a wide variety of computer types including single user PCs, shared PCs, laptops, and other devices (Table 2). Few participants used web browsers in a single location; only 2.9% of participants reported never using one at home and only 6.5% reported never using one when away from home. Participants generally used multiple computers for web activity: only 7.2% used 1 device, while 52.9% used 2 devices, 23.2% used 3 devices, and 16.7% used 4 or more devices.

Overall, 65.9% of participants answered questions about laptop use. Of these participants, 83.5% used laptops for web browsing in multiple locations. Most also used other computers at times (87.9%). For web browsing of a personal nature, most would use their laptop (86.3%), although home PCs and away PCs would also be used by some participants (55.0% and 33.8% respectively).

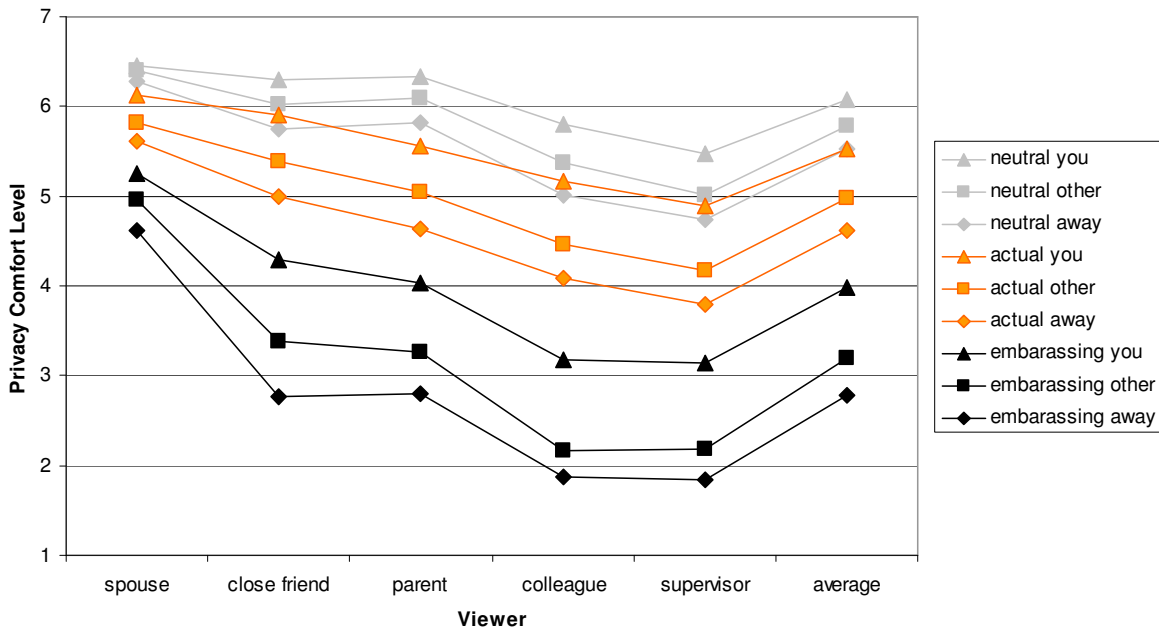
### Effect of Context on Privacy Comfort Levels

#### Overall

Privacy comfort levels (PCLs) were highly contextual with the viewer, the level of control, and the sensitivity of the content affecting the level of comfort (see Figure 5). The PCLs for the positive and neutral scenarios did not differ significantly, so the positive scenario has been omitted.

	Total	Single User PC	Shared PC	Laptop	Other
Home	97.1%	33.6%	41.8%	50.0%	7.5%
Away	93.5%	51.2%	38.8%	38.0%	3.1%

**Table 2. Percent of participants web browsing in each location (home, away) with each computer type in the location.**



**Figure 5. Comparison of privacy comfort levels (y-axis) according to the context of potential viewer (x-axis), scenario (colour of series), and level of control (you in control, other person in control with you there, other person in control and you leave the room) the participant has (shape of series points).**

*Privacy Comfort by Scenario*

Privacy comfort levels when participants reflected on their *usual web browsing* (across all computers) were lower than for the positive/neutral scenarios, but far higher than for the embarrassing scenario (as seen in Figure 5). This gives us some indication of how sensitive participants feel their typical browsing habits are. On average, 66.2% of participants rated their level of comfort higher when reflecting on their actual web browsing than when reflecting on the embarrassing scenario, 27.6% rated it the same, and 6.2% rated it lower.

We examined the average PCLs assigned by participants for the *usual web browsing scenario* using a two-way ANOVA for the variables of primary location for web browsing (home, away from home) and primary computer in use (single user desktop, shared desktop, laptop). A main effect was found for the primary location of use. Participants who performed the majority of their browsing at home gave significantly lower average PCL ratings for the usual browsing scenario than those performing the majority of their browsing away from home ( $F(1,148) = 7.45, p=0.007$ ). The difference may be due to the wider range of personal activities that participants stated they engage in on their home computers. The main effect for the type of computer used ( $F(2,148) = 0.85, N.S.$ ) and the interaction effect ( $F(2,148) = 1.76, N.S.$ ) did not reach statistical significance. The diversity of computers in use may have impacted our ability to detect changes in PCL by primary computer.

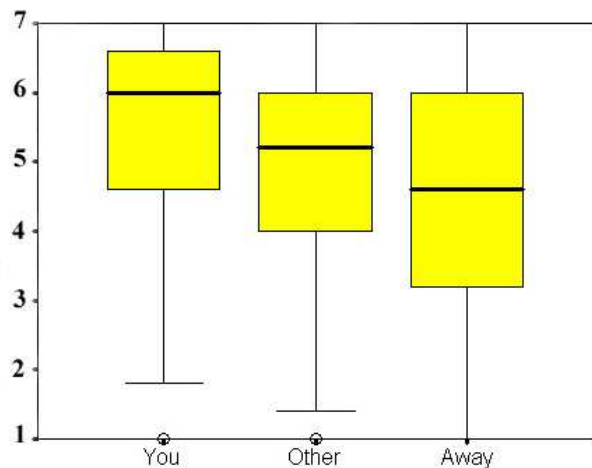
*Privacy Comfort by Level of Control*

Throughout, the impact of control was consistent: the higher the level of control retained by the user, the more

comfortable they were in terms of privacy. Figure 6 shows the variability of participants' PCLs according to level of control (for actual web browsing, averaged across viewers). The differences between the mean PCLs is statistically significant ( $\chi^2=134.74, p<.001$ ); however, both viewer and scenario impact the magnitude of the change (see Figure 5).

*Privacy Comfort by Viewer*

Participants had a relatively high comfort level for their spouse/significant other, and the amount of control and sensitivity of the scenario did not greatly change their comfort (see Figure 5). However, other viewer groups were much more sensitive to context. For example, spouse and



**Figure 6. Box plots showing the variability of average privacy comfort levels (y-axis) for the 3 levels of control (for actual web browsing scenario, PCLs averaged across viewers).**

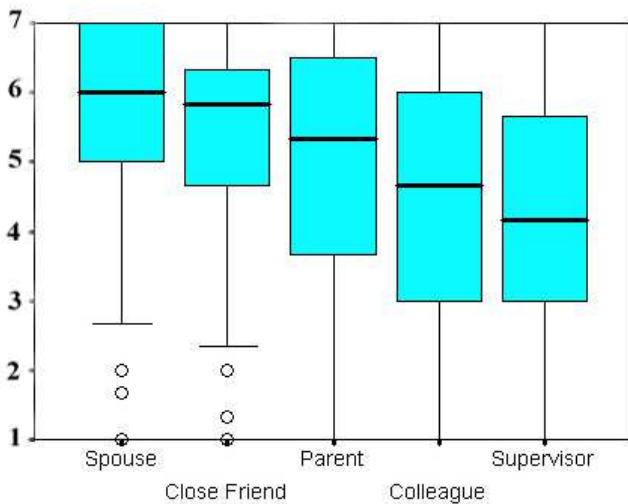
close friend have similar comfort levels with a high level of control; however, as control is lost, there is less comfort with the close friend than with the spouse.

Figure 7 shows the variability of participants' PCLs according to type of viewer (for actual web browsing, averaged across level of control). Results showed that the differences between the mean PCLs for viewer categories is statistically significant ( $\chi^2=206.30$ ,  $p<.001$ ). Participants were least comfortable with supervisors or colleagues as viewers.

#### *Inherent Privacy Concerns*

While the survey did not directly elicit participants' inherent privacy concerns, an examination of the privacy comfort levels according to context can give us a sense of participants' underlying concerns. For an initial estimate of inherent privacy concerns, we examined the PCLs that participants gave for the embarrassing scenario presented. We selected this scenario as it was most likely to provoke discomfort in participants and exhibited large comfort differences by context (level of control, type of viewer). We would expect privacy fundamentalists to have a low PCL regardless of context and the privacy unconcerned to have a relatively high PCL. However, privacy pragmatists might have differences in their PCL depending on the context.

K-means cluster analyses (by level of control and by viewer) were performed on participants' median comfort levels and the magnitude of the differences between their minimum and maximum PCL for the embarrassing scenario. Participants clustered into three groups: 28% as fundamentalists, 64% as pragmatists and 8% as unconcerned. When clustered solely by the magnitude of differences in PCL, only 40% of the pragmatists were concerned along both dimensions. The remaining 60% were concerned along only one dimension, level of control or viewer. Of those concerned along only one dimension, 82%



**Figure 7. Box plots showing the variability of average privacy comfort levels (y-axis) for the 5 types of viewers (for actual web browsing scenario, PCLs averaged across level of control).**

(49 of the 100 pragmatists) had high differences for viewers and low differences for level of control.

## **DISCUSSION**

In order to provide semi-automated management support for the domain of incidental information privacy, we must be able to build a rich model of its components. The main dimensions of incidental information privacy will frame discussion of our results.

### **Level of Control**

There was an overall effect of control on privacy comfort levels. When participants envisioned themselves in control of the keyboard and mouse, they have the least amount of concern across the viewing audience. As control is lost, the amount of concern grows. However, as seen in the inherent privacy concerns analyses, many participants (i.e. viewer concerned pragmatists, fundamentalists, and unconcerned) had relatively few differences by level of control. Clearly, control is a highly individual dimension.

### **Viewing Audience**

All participants had people view their display at least occasionally. Trusted viewers such as spouses and close friends were regular viewers; however, some of the most frequent viewers were colleagues and supervisors, both of whom have lower overall comfort levels. Our results are consistent with previous information privacy research such as [21] with respect to the relative comfort levels between categories of information receivers. However, the categories we used were relatively broad. Even within a viewer category there may be several levels of trust and sharing which may fluctuate depending on recent interpersonal interactions. Similar to level of control, the impact of potential viewers was highly individual.

### **Visible Content**

Comfort in a given situation, particularly for privacy pragmatists, will depend on the sensitivity of traces of previous activity that may become visible. Results showed that the activities that participants engaged in may depend on their location at the time and the device in use. While some participants conduct activities both at home and away, others partitioned their activities so that the more personal browsing was only conducted at home. Traces of differing sensitivity may be generated in each location of use. This impacts laptop users who do the majority of their browsing on a single device that moves between settings. It is also a consideration for users that consolidate their History files and Favorites for use in multiple locations.

People use a variety of computers regularly: laptops, single user PCs, and shared PCs, both at home and away from home. One problem with managing the privacy of traces of previous web browsing activity is that it is not always clear what traces will be revealed. With multiple devices, there may be increased uncertainty, particularly for those users that don't partition their browsing activities between

locations and devices. Additionally, many participants indicated that they used their web browser convenience features differently for each computer. This lack of standardized settings across computers could add to the uncertainty about what will be revealed for each computer.

Laptop users have increased privacy concerns and had a high likelihood (76.6%) of taking actions to protect their privacy if given advanced warning. This is understandable as participants used their laptop for browsing of personal nature and moved between multiple locations. The relatively low level of actions taken by away PC users may be due to their reduced convenience feature usage on those computers and their reduced likelihood of engaging in personal activities when away from home.

Our results showed that participants' actual web browsing appears to be more sensitive than the neutral and positive scenarios, but less sensitive than the embarrassing scenario given. The embarrassing scenario was designed to give us an indication of the upper bound of participants' discomfort for traces of their web browsing activity. For the 33.8% of participants who indicated they would have the same comfort or less than if traces of their usual web browsing were viewed, it was not the most embarrassing scenario imaginable. The medical nature of the risqué sites in the scenario might have mitigated any morality concerns. Further investigation showed that more participants indicated a higher level of discomfort for family than for co-workers in the usual browsing scenario as compared to the embarrassing scenario. The personal nature of the embarrassing scenario may have violated the persona kept for co-workers, thereby provoking a stronger response; however, participants may have envisioned sharing medical concerns with family, but not other private activities such as erotica.

### **Inherent Privacy Concerns**

Our initial examination of inherent privacy concerns looked only at the embarrassing scenario. However, some participants are also content sensitive and do not exhibit the same high concerns with other content scenarios so the fundamentalist group may be inflated. Similarly, solely looking at the neutral or positive scenarios may over-classify participants as privacy unconcerned. Using the *usual browsing scenario* as the basis for classification is problematic as the sensitivity of the content will vary according to browsing practices (e.g. is a participant concerned because they have recently conducted some sensitive browsing or because they have inherently high privacy concerns). Analysis is still underway to determine how best to partition the participants while accounting for content, control, and viewer sensitivities.

In a similar fashion to the subdivision of pragmatists in the consumer privacy domain into identity concerned and profiling averse [26], we may find it useful when modeling incidental information privacy to consider pragmatists in subcategories such as *control concerned*, *viewer concerned*,

*content concerned*, and *generally context concerned*. If we can determine an individual's inherent privacy concerns we can simplify configuration of a privacy management scheme. If a user is classified as a privacy fundamentalist, then the system should provide maximum privacy protection without requiring ongoing interaction. If a user is classified as a pragmatist, then knowing along which dimensions a user is concerned may allow the interface to be tailored to those concerns. Those that are privacy unconcerned would have little use for such a system.

### **CONCLUSIONS AND FUTURE WORK**

Privacy of incidental information is indeed a concern for many. The results from our research clearly show that not only did participants have regular occasions when others could view their displays, most were also concerned enough to take some steps to maintain the privacy of the incidental information that may be displayed. We found that a user's privacy comfort level is impacted by who is viewing their screen, the amount of control the user retains over the mouse and keyboard, and the sensitivity of the information being viewed.

However, it is also important to recognize that privacy is very individual; different users tend to have different underlying sensitivities regarding which aspects of privacy make them comfortable or uncomfortable. For each dimension of incidental information privacy, we need to know the extent and variability of user behaviour and concerns. If behaviour and concerns are consistent across users, we can use a standard approach in a privacy management solution. If participants cluster into groups, we can try to determine best management practice for those instances. However, we will also need methods of determining to which group an individual belongs so that the appropriate automated approach is taken. Individualized privacy management systems may be able to simplify privacy preference configuration by only presenting options along those aspects of privacy pertinent to the individual.

A limitation of surveys is that participants must reflect upon their attitudes and experiences while not in the context of those experiences. However, in the incidental information domain, current privacy management is largely a matter of speculation: What traces of my past activities will be visible? Who will be able to view it? Should I clear my History files? How will others regard these traces of past activity? In this regard, the survey was a good choice to explore attitudes and get self-reported data about typical web browsing behaviour and current privacy management practices. However, we are also grounding our research in actual behaviours. We have conducted field studies (e.g. [11]) and will synthesize the results to build a more complete model of incidental information privacy with respect to web browsers. Patterns have emerged in how participants apply privacy levels to traces of their actual web browsing. Our current work is exploring how the context of location of web browser use and page content

correlate with the privacy levels applied to actual web browsing. It is clear that we must utilize inherent privacy concerns, the current browsing context, and web browsing behaviour patterns to relieve the burden of the user classifying all incidental information in a privacy management system.

#### ACKNOWLEDGMENTS

Thanks to the members of the EDGE Lab for their support. Funding provided in part by NSERC and NECTAR.

#### REFERENCES

1. Platform for Privacy Preferences (P3P) Project. <http://www.w3.org/P3P/>.
2. WebRoot Software | Window Washer. <http://www.webroot.com/consumer/products/windowwasher/>.
3. Ackerman, M., Cranor, L. and Reagle, J. Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences. In *Proc. of EC '99*, ACM Press (1999), 1-8.
4. Acquisti, A. Privacy in Electronic Commerce and the Economics of Immediate Gratification. In *Proc. of EC '04*, ACM Press (2004), 21-29.
5. Begole, J., Tang, J.C. and Hill, R. Rhythm Modeling, Visualizations and Applications. In *Proc. of UIST 2003*, ACM Press (2003), 11-20.
6. Boardman, R. and Sasse, M.A. "Stuff Goes Into the Computer and Doesn't Come Out": A Cross-tool Study of Personal Information Management. In *Proc. of CHI '04*, ACM Press (2004), 583-590.
7. Cadiz, J. and Gupta, A. Privacy Interfaces for Collaboration. Microsoft Research. MSR-TR-2001-82 (2001).
8. Carini, R.M., Hayek, J.C., Kuh, G.D., Kennedy, J.M. and Ouimet, J.A. College Student Responses to Web and Paper Surveys: Does Mode Matter? *Research in Higher Education* 44, 1 (2003), 1-19.
9. Consolvo, S., Smith, I.E., Matthews, T., LaMarca, A., Tabert, J. and Powledge, P. Location Disclosure to Social Relations: Why, When, & What People Want to Share. In *Proc. of CHI '05*, ACM Press (2005), 81-90.
10. Goffman, E. *The Presentation of Self in Everyday Life*. Doubleday Anchor Books, Garden City, New York, 1959.
11. Hawkey, K. and Inkpen, K. Privacy Gradients: Exploring ways to manage incidental information during co-located collaboration. *Ext. Abstracts CHI 2005*, ACM Press (2005), 1431-1434.
12. Hawkey, K. and Inkpen, K. Web Browsing Today: The impact of changing contexts on user activity. *Ext. Abstracts CHI 2005*, ACM Press (2005), 1443-1446.
13. Hinckley, K. Distributed and Local Sensing Techniques for Face-to-Face Collaboration. In *Proc. of ICMI '03*, ACM Press (2003), 81-84.
14. Huang, E.M. and Mynatt, E.D. Semi-Public Displays for Small, Co-located Groups. In *Proc. of CHI '03*, ACM Press (2003), 49-56.
15. Jensen, C., Potts, C. and Jensen, C. Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies* 63, (2005), 203-227.
16. Kaasten, S., Greenberg, S. and Edwards, C. How People Recognize Previously Seen WWW Pages from Titles, URLs and Thumbnails. In *Proc. of Human Computer Interaction 2002*, Springer Verlag (2002), 247-265.
17. Lau, T., Etzioni, O. and Weld, D.S. Privacy Interfaces for Information Management. *Communications of the ACM* 42, 10 (1999), 89-94.
18. Lederer, S., Mankoff, J. and Dey, A.K. Towards a Deconstruction of the Privacy Space. *Workshop on Ubicomp Communities: Privacy as Boundary Negotiation, UBICOMP 2003*, <http://guir.berkeley.edu/pubs/ubicomp2003/privacyworkshop/papers/lederer-privacyspace.pdf> (2003),
19. Lederer, S., Mankoff, J. and Dey, A.K. Who Wants to Know What When? Privacy Preference Determinants in Ubiquitous Computing. *Ext. Abstracts CHI 2003*, ACM Press (2003), 724-725.
20. Moor, J.H. Towards a theory of privacy in the information age. *ACM SIGCAS Computers and Society* 27, 3 (1997), 27-32.
21. Olson, J.S., Grudin, J. and Horvitz, E. A Study of Preferences for Sharing and Privacy. *Ext. Abstracts of CHI '05*, ACM Press (2005), 1985-1988.
22. P&AB Consumer Privacy Attitudes: A Major Shift Since 2000 and Why. *Privacy & American Business Newsletter* 10, 6 (2003), 1,3-5.
23. Palen, L. and Dourish, P. Unpacking "Privacy" for a Networked World. In *Proc. of CHI '03*, ACM Press (2003), 129-136.
24. Patil, S. and Lai, J. Who Gets to Know What When: Configuring Privacy Permissions in an Awareness Application. In *Proc. of CHI '05*, ACM Press (2005), 101-110.
25. Shoemaker, G.B.D. and Inkpen, K.M. Single Display Privacyware: Augmenting Public Displays with Private Information. In *Proc. of CHI '01*, ACM Press (2001), 522-529.
26. Spiekermann, S., Grossklags, J. and Berendt, B. E-privacy in 2nd Generation E-Commerce: Privacy Preferences versus actual Behavior. In *Proc. of EC '01*, ACM Press (2001), 38-47.
27. Household Internet use at home by Internet activity. <http://www40.statcan.ca/101/cst01/comm09a.htm>.
28. Household Internet use by location of access. <http://www40.statcan.ca/101/cst01/comm12a.htm>.
29. Turner, C.F., Ku, L., Rogers, S.M., Lindberg, L.D., Pleck, J.H. and Sonenstein, F.L. Adolescent sexual behaviour, drug use, & violence: Increased reporting with computer survey technology. *Science* 280, (1998), 867-873.
30. Weisband, S.P. and Reinig, B.A. Managing User Perceptions of Email Privacy. *Communications of the ACM* 38, 12 (1995), 40-47.