

## Examining the Usability of Web Privacy Policies

Robert W. Proctor

Department of Psychological Sciences,  
Purdue University, West Lafayette

M. Athar Ali

Department of Computer Technology,  
Purdue University, West Lafayette

Kim-Phuong L. Vu

Department of Psychology,  
California State University, Long Beach

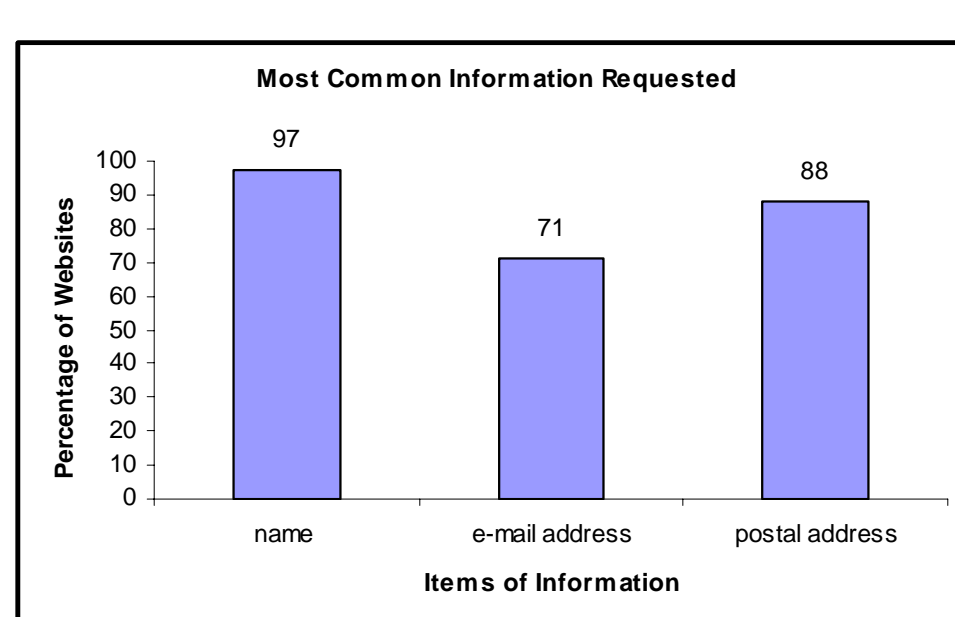
### Goals

- To understand what types of information are included in privacy policies
- To obtain measures of users' comprehension of privacy policies and attitudes toward the policies and their host sites

### Analysis of Web Sites

Last year, we reported an analysis of 7 Web site categories, with 6 sites in each, that examined the types of personal information that each site requested users complete for on-line transactions Ali, Proctor, & Vu, 2005).

| Website Category           | Number of items of information requested |
|----------------------------|--|
| Financial Institutions     | $M = 30; sd = 10$                        |
| Insurance                  | $M = 17; sd = 3$                         |
| Online Pharmacies          | $M = 17; sd = 3$                         |
| Travel agencies & Airlines | $M = 12; sd = 2$                         |
| Retail                     | $M = 13; sd = 2$                         |
| Technology                 | $M = 12; sd = 2$                         |
| Online Games               | $M = 9; sd = 3$                          |



We found that considerable differences exist between site categories. More importantly, there were significant differences within categories with respect to the type and amount of information that a user is asked to provide.

The within-category variability indicates that many sites are asking for information that is not necessary to conduct a transaction.

At least 80% of the Web sites in all categories except on-line gaming had links to the site's privacy policy. These links were typically located at the bottom of the page, centered or to the right.

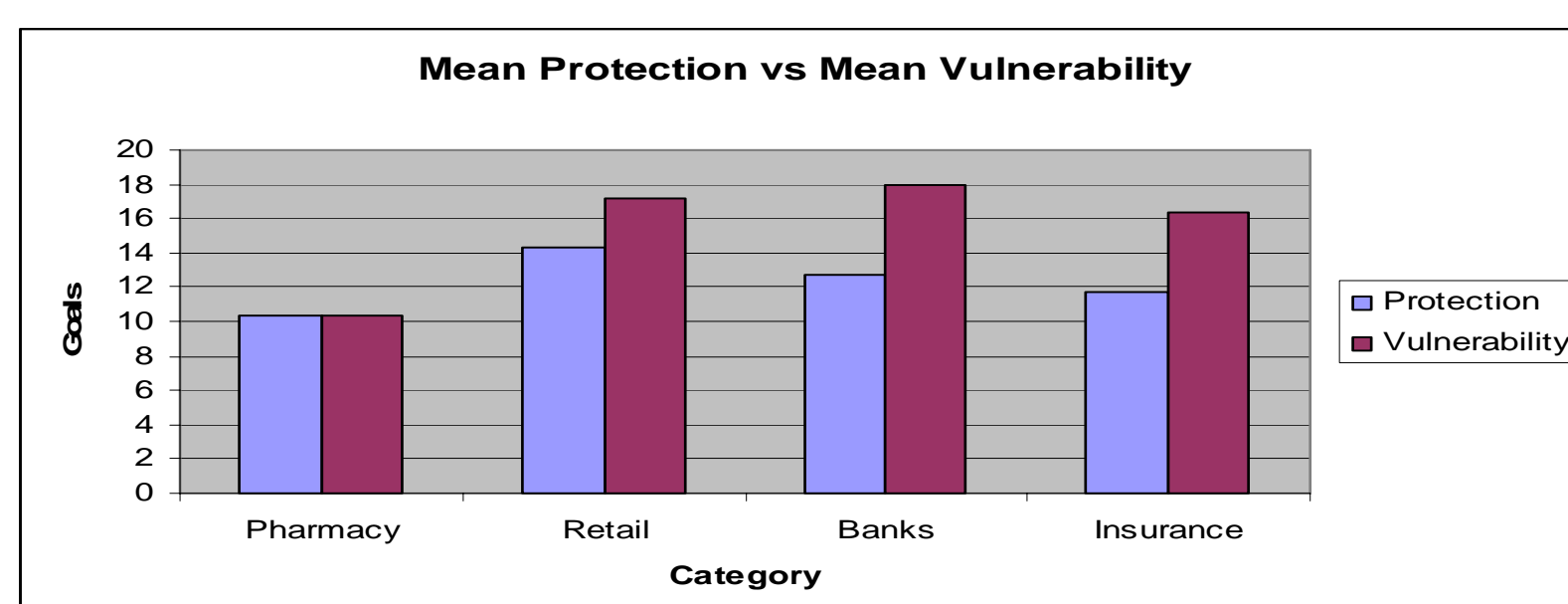
### Content Analysis

Privacy policies were analyzed to determine the number of protection and vulnerability goals included in each.

### Method

Privacy policies from 25 Web sites in each of the 4 most widely used categories (pharmacy, retail, banks, and insurance) from the first phase were selected for content analysis. Goal mining analyses were performed to determine the number of privacy goals contained in each policy and whether they consisted of mainly protection or vulnerability goals (Antón et al., 2004).

### Results



Across the four categories, we found that the privacy policies contained more vulnerability goals than protection goals when goals addressing both issues were not taken into account.

For the pharmacy category, protection goals were almost equal in number to vulnerability goals, but the number of goals of each type was lower for that category than for the other three categories. Also, online pharmacies, retailers, and banks were more likely to be low on both protection and vulnerability goals or high on both.

### Readability analysis

The same 25 Web sites in each category were also analyzed for readability. The readability analysis was performed using MS Word and ReadMe tool to determine the word count. Flesch Reading Ease Scores (FRES) and Flesch Grade Level (FGL) were obtained to score the readability of the policies.

### Results

| Category          | Mean FRES | Mean FGL |
|-------------------|-----------|----------|
| Online pharmacies | 29.10     | 14.84    |
| Retail            | 37.27     | 13.57    |
| Banks             | 35.59     | 13.56    |
| Insurance         | 37.84     | 12.99    |

The reading analysis showed that the mean grade level needed to understand these policies is at least 13 years. Given that not everyone using the Web has a college education, many users may not be able to understand the privacy policies.

Moreover, many privacy policies were too long to be read by most users. Overall the length of the policy did not affect its readability ( $r = 0.10$ ). So, making the policy more concise will increase the likelihood that the policy will be read without necessarily decreasing its readability.

### Reading Comprehension Experiment

We conducted an experiment to examine comprehension of privacy policies by college students, who should be at the 13-year reading level at the which the average policy is written.

### Method

20 students enrolled in Introductory Psychology, all experienced with Web-based transactions, participated. The students were at the educational level at which the average policy is written.

Privacy policies were selected from the bank and retail categories because college-age students are most likely to use Web sites within these categories. Four privacy policies from those used in the content analysis were selected that were of similar reading level, with two being LL and two being HH on goals. The company names were removed to eliminate any possible influence of previous experience.

Each person read all policies, with the order counterbalanced across subjects. Each policy was read one-at-a-time with no time limit. When the subject entered a command to continue, s/he was asked 9 questions about the content of the policy ("Does this website share information with other websites/outside companies?") and 9 questions about perceptions of the policy (e.g., "Do you think that the company collects more customer information than is required?").

### Results

| Website Type                             | % correct | Website Category |
|--|-----------|------------------|
| Short length (510 words) & LL on goals   | 59        | Policy A: Bank   |
| Long length (1628 words) & HH on goals   | 59        | Policy B: Bank   |
| Long length (2100 words) & HH on goals   | 54        | Policy C: Retail |
| Medium length (1306 words) & LL on goals | 37        | Policy D: Retail |

- Performance was least for the policy with low protection and vulnerability goals and of medium length (policy D).
- Privacy policies that had high protection and high vulnerability goals (HH) were perceived by users to provide better security measures than those that had low protection/ vulnerability (LL) goals.
- Privacy Policies that had low total goals (protection and vulnerability) were perceived by users to be less repetitive than policies that had high total goals.
- Privacy policies that had high total goals were perceived to be better at information protection.
- The privacy policy that was the shortest of the four was also considered to be the most unreliable.
- Users reported being least likely to provide confidential information like credit card number or SSN to sites with privacy policies of shorter length and low total goals.
- Overall, the research indicates that privacy policies as applied currently on Web sites do not satisfy users' needs. These policies are only slightly understandable to even more educated users, thus making the chances that an average user will comprehend them very low.

### Recommendations

- Web site privacy policies need to provide more protection goals. More research needs to be conducted to understand how this can be done without the policies sounding repetitive, while keeping the number of vulnerability goals low.
- Privacy policies should define the different categories of information, such as Personally Identified Information (PII) and non-PII. It was found that privacy policies use such terms but users are unaware of what they mean. Privacy policies should define these categories and what information should be considered as part of each category.

### References

- Ali; M. A., Vu, K.-P. L., & Proctor, R. W. (2005, March). An analysis of privacy and security information provided and elicited by different types of Web sites. Poster presented at the 6<sup>th</sup> Annual CERIAS Conference Information Security Symposium: Security in Motion. W. Lafayette, IN.
- Antón, A.I., et al. (2004a). The lack of clarity in financial privacy policies and the need for standardization. *IEEE Security and Privacy*, 2(2), Mar-Apr, 36-45.