

Understanding Compliance to Privacy Guidelines Using Text-and Video-Based Scenarios

Abdullah Al Mahmud¹, Maurits Kaptein¹, Oliver Moran¹,
Evelien van de Garde-Perik², and Panos Markopoulos²

¹User-System Interaction Program, Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{A.Al-Mahmud, M.C.Kaptein, O.P.Moran}@tm.tue.nl
²Department of Industrial Design, Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{e.m.v.d.garde, p.markopoulos}@tue.nl

Abstract. Privacy is a major concern for the design and user acceptance of pervasive technology. Investigating privacy poses several methodological challenges. A popular approach involves surveying reactions of people to scenarios that highlight privacy issues. This paper examines the validity of this approach. It reports an experiment that compared people's ability to correctly judge compliance to privacy principles when scenarios are presented in video versus textual form. It was found that such privacy-related concepts are hard to understand, leading to a large number of erroneous judgments regardless of medium and that interpretation varied across media. Comprehension in such studies can be improved, if a text scenario is preceded by a video-based version.

1 Introduction

Developments in the related fields of ubiquitous computing, pervasive computing and ambient intelligence have been paralleled by warnings for caution against invading people's privacy [25], considered as an interpersonal need or even an individual right. A growing body of research has been pursuing the development of privacy enhancing technologies, interaction design principles, and design methods that can help ensure people's privacy. Recently researchers into privacy and ubicomp have come to recognize the methodological difficulties for researching privacy in this domain and successive workshops at international conferences were organized on this topic [22, 23].

There are several serious methodological difficulties in studying privacy. Inquiry into user preferences is not sufficient to guide design, and privacy requirements tend to vary a lot depending on the technological and societal context in which an application is used, let alone its value to the individual. A well-known issue concerns the apparent discrepancy between privacy-related attitudes and behaviors, which has been reported in several different empirical investigations of privacy preferences, e.g., [2, 5, 12]. An important reason for this, which is well known in psychological studies concerning attitudes, is that attitudes expressed outside a specific context and without taking into account the context specific difficulty of a particular privacy-related behavior are very poor predictors of actual behavior [4]. In order to make reported

attitudes regarding privacy correspond better to actual privacy related behaviors, one possibility is to contextualize inquiry with the use of scenarios, i.e. to inquire regarding particular actions in the context of a specific situation described in a scenario.

Human-computer interaction research and interaction design practice have endorsed scenarios as a means to contextualize inquiry and design activities (e.g., see [7]). Scenarios for human-computer interaction are typically delivered as narrative texts, but are often also shown as storyboards or short videos. Indeed several recent empirical investigations concerning privacy and ubiquitous computing systems are based on participant surveys about their reactions to scenarios delivered to them in textual or video form [1, 13, 24]. Textual scenarios were used in [1] as a means to investigate how people would respond to situations where personal information is collected. Participants were asked about their concerns through specific scenarios involving online data collection. One scenario, for example, asked respondents whether they would be more or less likely to provide data to a Web site with a privacy policy that explained that their information would be removed from the site's database if they did not return to the site for three months. Video scenarios were used in [24] to illustrate RFID technology, its benefits and drawbacks in a neutral way. Two groups of participants were shown the same video, except for the PET available to consumers for controlling their privacy. Afterwards, participants were questioned about perceived control in the scenario presented to them.

In this study, it is examined whether the interpretation of scenarios relating to privacy are affected by the medium in which the scenario is delivered. In the following section, the context of this study and the reasons behind the hypothesis that a medium effect exists with respect to the ability of informants to discuss privacy-related issues are presented.

The study presented is useful to HCI in two ways. The results of this study provide insight into the reliability of text or video scenarios for surveying attitudinal responses to different system designs. Secondly, the results provide quantitative evidence regarding the difficulty of conveying privacy related functions and policies to users; earlier research has shown that users often do not read such policies [17, 19] or are not able to comprehend them [15, 19].

1.1 Privacy Guidelines for Embedded User Modeling

A growing body of work in the domain of ubiquitous networked applications concerns privacy, including analyses of fundamental concepts [21, 6] and discussions of different cases. Researchers have proposed collections of design principles and guidelines for ensuring privacy through interaction design. Jiang and Landay [16] proposed principles for regulating the flow of information in ubiquitous systems. Chung et al. [8] proposed a set of design patterns to solve privacy problems. Others have adopted principles originating from legislation for technical systems and privacy, such as the five principles of Fair Information Practices [9, 18].

For the purpose of this study the focus is on one such set of guidelines, namely those by the OECD [20]. These guidelines were originally intended for assisting governments, business and consumer representatives in shaping policies and practices that protect privacy and personal data while preventing unnecessary restrictions to

data flows across borders, both on- and off-line. OECD guidelines are the basis of the discussion by M. Langheinrich on designing for privacy in ubiquitous computing [18]. This study is often cited and has significant influence over work on privacy and ubiquitous computing/ambient intelligence. However, it did not attempt to provide empirical evidence to demonstrate that these principles do indeed lead to higher acceptance of such systems by end-users. The same holds for the other sets of principles cited above; currently they represent hypotheses that need to be validated empirically or through repeated application in practice.

The broader research aims to complement that work by providing empirical evidence regarding the importance of compliance to the OECD guidelines for users. In a recent study reported elsewhere [10], subjects were required to indicate their preferences regarding the relative importance of individual OECD guidelines by means of a scenario relating to tele-monitoring for healthcare. As a methodological precaution, a pre-study experiment was conducted, discussed in [11], where the ability of participants to answer correctly whether a particular system complied to a particular guideline or not was tested. Surprisingly, this pilot study revealed a very limited comprehension of the system's privacy policy. While this was improved substantially by rephrasing the text scenarios in the eventual experiment, many participants still could not recognize the system's privacy policy, even though the scenario contained the guideline text almost verbatim.

Based on this study it was suspected that the subtleties of the wording of OECD guidelines are not understood with common sense; people tend to make assumptions regarding the use of their personal information when they are told the purpose of its collection, or to assume policies regarding purpose by a description of the nature of information collected, etc. Further, it was suspected that a minimalist scenario lacking a narrative structure and a believable context is hard to understand. Hopefully, a video scenario, (which is typically used in informal evaluations with during interviews or focus groups to assess how a particular technology might be used), could provide a more compelling and easily comprehensible version of the scenario and would allow users to better express their preferences. (See for example, the study by Spiekermann [24] where privacy preferences were solicited after informants viewed two video scenarios for an RFID-based automated supermarket check-out). It was set to test experimentally whether providing a video-scenario would improve matters. A formal experiment was conducted to test these hypotheses. The main aim was to compare the degree of variance in interpretation across video- and text-based scenarios.

2 Method

The experiment involved comparing the medium (video or text) in which a scenario was presented to participants. A within-subject design was chosen, so participants were exposed to both the text-based and video-based scenario. Some participants were interviewed face-to-face and others over the internet. Participants were given judgment tasks, i.e. to indicate whether a given scenario complied or not with the guidelines presented to them in a shorthand manner. For each guideline, there was a correct answer. The dependent variable was the correctness of the response for each guideline and the overall correctness of the interpretation for each scenario.

2.1 Apparatus and Materials

The original text descriptions of the eight OECD guidelines [20] were adapted in such a way that they cover the essence of each guideline, without being too elaborate and especially omitting those parts of the guideline that are meaningful only as legal text rather than as design guidance. In this spirit, also the guideline of accountability was omitted, as it serves a legal purpose stipulating compliance with all other guidelines.

In this study Collection Limitation (CL) refers to the fact that data is collected with the knowledge of the user, or in other words, that the user is informed about the type of data that is collected. The part of the original guideline that requires collection by lawful and fair means is omitted.

Data Quality (DQ) pertains to the relevance of the collected data to the purposes for which they are to be used. The original OECD guideline requires also that this data be accurate, complete and kept up-to-date.

Purpose Specification (PS) requires that the purposes for which the data are collected be specified. The OECD guideline also prescribes that the timing of Purpose Specification, limitation of data use to the fulfillment of those purposes, and each occasion of change of purpose should be specified as well.

Use Limitation (UL) in this study refers to the fact that data is not used for purposes other than those specified. The OECD guideline also states that the data is not be disclosed, or made available for purposes other than those specified, and that there are exceptions in case of consent of the data subject and by authority of law.

The guideline for Security Safeguards (SS) refers to the protection of data by reasonable security safeguards. In the OECD guidelines it is specified that it offers protection against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data, but this addition was again considered more relevant in the legal context and is omitted for this study.

Openness (OP) requires that the user is informed about the other parties that have access to the collected data. This is somewhat different from the original OECD guideline that implies a general policy of openness about developments, practices and policies with respect to personal data, and the availability of means to establish the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

The OECD principle of Individual Participation requires being able to get a confirmation about collected data, or getting data communicated; if either of these is denied then the reasons should be given, and the person should be able to challenge data. To allow compliance questions to be asked unambiguously, this principle was split in this study to two parts, Insight (IN) and Modification (MO). Insight refers to the possibility of inspecting stored data (rather than getting a confirmation or having data communicated). Modification pertains to the possibility to modify or erase collected data.

The scenario describing the healthcare system was a slightly improved version of a scenario used in the pre-study reported in [11]. It was adapted to eliminate some of the comprehension problems that were identified there and to provide a more coherent and natural narrative structure. The scenario describes how a diabetic patient called

John is checking his diabetes condition using his newly installed health-support system. The system informs him that data is being collected to monitor this, and only this. In the menu of his health-support system, John can see that his data can be used for other purposes as well and shows him a list of people and organizations that may access his data. John is not one of these people. He may not inspect, modify or erase data collected about him. In the same menu, he can also see that the system does not protect his data with any security safeguards.

Fig. 1. The text used for both versions of the scenario

John uses a health-support system for monitoring diabetes. The same text (see figure 1) was given in printed form and was read out as a voice-over to a related video. The text describes a system that clearly meets only three of the eight OECD guidelines for privacy and security (namely PS, DQ, OP).

The video shows a character, presumably John, interacting with a machine, presumably the health-support system (see Figure 2). The exact nature of this interaction could not be discerned from the video image since John remained unexpressive and the machine was not shown in detail (to avoid drawing attention to interaction details unimportant for this experiment).



Fig 2. A screenshot of the video-based scenario, showing 'John' taking a physiological measurement; the video does not make more explicit than the text what John is measuring

2.2 Participants

Participants for the face-to-face survey were recruited amongst University employees who were not familiar with the purpose of the experiment. Participants for the online survey were gathered through invitation on email lists and message boards and through asking department secretaries in other departments of the University to forward an email invitation to their colleagues. Because of high level of education and knowledge of the English language, it could be expected that their comprehension should be no worse than participants recruited for the purposes of evaluation studies. (User studies to which these methodological results need to apply, typically do not assume a higher level of education or knowledge of privacy issues; often they concern the general public or large consumer groups, including people of lesser education, or even elderly, with less familiarity with technology than the young adults participating in the present study).

In total 104 participants took part in the study: 25 people participated in the face-to-face survey; 79 participated in the online survey. The used procedure for recruiting ascertained that all participants were familiar with websites/computers and the English language. Participants had various backgrounds.

2.3 Procedure

Participants were first introduced to John and the function of the health-support system in his daily life through a short text before being shown one of the two scenarios. Participants were asked to read or watch each scenario before answering if they agreed with eight statements about the system described to them. These statements described compliance with an individual guideline, so in effect participants were asked whether the system met the guidelines or not and which guidelines it met (see Table 1). Participants were allowed to refer back to the scenario while answering these questions. When finished with answering the questions for one scenario, participants were then presented with the same scenario via the other medium and the experiment was repeated. The experiment was conducted under two conditions: face-to-face and online.

Table 1. Overview of possible system features. Participants were asked to indicate for each feature whether the system provided them or not.

| Provided System Features (Yes/No) |
|---|
| The user is informed about the type of data that will be collected. |
| The user is informed about the main purpose for which the data will be used. |
| The system only collects data that is relevant to the main purpose of the system. |
| The data will be used solely to serve the main purpose of the system. |
| The user is informed about which other parties have access to the collected data. |
| The data is securely stored. |
| The user can inspect the stored personal data. |
| The user has the possibility to make changes in the stored data |

This experimental set-up allowed for a number of experimental controls. First, conducting simultaneous face-to-face and online surveys allowed a double-check for bias. If the results for each of the experimental conditions were the same then it could be concluded that there was no interviewer bias during face-to-face surveys or spoiled results due to the absence of social control during online surveys. Systematic variation was employed during the presentation of the two scenarios. Face-to-face surveys varied presentation of text- or video-based scenarios on a strict every-second-participant basis. Online surveys varied presentation of media randomly per participant.

Importantly, the survey was conducted as a single-blind test. Participants were introduced to the experiment under the belief that they were going to be shown two scenarios and that they were then to be asked questions about these scenarios. They were not informed that they were going to be shown either a text- or video-based scenario or that both scenarios would be 'identical'. This set-up prevented participants from guessing the purpose of the experiment and so, intentionally or not, influencing

its outcome. The face-to-face and online surveys used exactly the same text and formatting and were divided across three identical pages.

Given this set-up the experimental rationale was as follows: If the text- and video-based scenarios are identical, control has been made for the presentation of the media and participants do not suspect that the medium is being tested then it can be concluded that any variance in participants' answers across the media will be due to the effect of that medium.

3 Experimental Results

Since both online and face-to-face research was conducted, it is necessary to examine the differences between these two methods. The face-to-face condition did not involve an interview but it did involve spoken explanation and a continuous presence of the interviewer. However, the tasks, their layout, and the procedure were kept the same in both conditions. This led to the hypothesis that there would be no significant influence upon the test results caused by the method of data gathering. Thus, it was expected that there is no effect caused by potential interview bias or lack of social control.

The first task performed to check this assumption was to conduct t-tests on aggregated results. On the total number of correct responses within each method a t-test showed no significant difference between the two methods ($p=0.878$). After establishing this result for the aggregated data, a t-test was performed on the difference between the two data gathering methods for both the text-based scenarios and the video-based scenarios separately. Within the text based scenario there was no significant difference between the Internet and the face-to-face condition ($p=0.831$). The same was true for the video-based scenario ($p=0.969$).

Finally, before the results were aggregated between these two methods of data collection for further analysis, one extra check of the reliability of the assumption was performed. To check for differences on individual items between the internet and face-to-face surveys, a Chi-square tests for every individual guideline was conducted. The Chi-square test can be used to check for a relation between two variables of nominal measurement level. Table 2 presents the results of these individual Chi-square tests per item by presenting the Chi-squared value, the degrees of freedom and the significance level. Since there is no significant relationship between the data collection method and the number of correct responses the assumption that there is no effect for the method of data collection is supported.

Table 2. Comparison of face-to-face and Internet results per question

| Item | Chi² | Df | Sig. |
|-------------------------------------|------------------------|-----------|-------------|
| Collection Limitation (CL) | 0.42 | 1 | 0.516 |
| Purpose Specification (PS) | 1.08 | 1 | 0.299 |
| Data Quality (DQ) | 0.45 | 1 | 0.501 |
| Use Limitation (UL) | 2.32 | 1 | 0.128 |
| Openness (OP) | 2.31 | 1 | 0.128 |
| Security Safeguards (SS) | 0.40 | 1 | 0.525 |
| Insight (IN) | 1.88 | 1 | 0.171 |
| Modification (MO) | 0.15 | 1 | 0.702 |

Results were obtained by calculating the percentage of correct answers given for each question relating to one of the guidelines in each of the following conditions:

1. The text-based scenario shown first to participants;
2. The video-based scenario shown second (i.e. immediately after the text-based scenario had been shown);
3. The video-based scenario shown first;
4. The text-based scenario shown second (i.e. immediately after the video-based scenario had been shown).

The overall result for the whole scenario in each condition was also calculated.

Initial interpretations showed interesting results. On average, overall scores for each condition were approximately the same, with text performing slightly better. However, large variations in scores were obvious between the guidelines for each condition. Table 3 shows the percentage of good answer per guideline for each condition. Rows in this table show the condition and columns show the percentage of correct results for each guideline. The final column shows the overall percentage of correct answers per condition.

Table 3. The percentage of correct answers and per question by condition

| Condition | CL | PS | DQ | UL | OP | SS | IN | MO | \bar{X} |
|----------------------|------|------|------|------|------|------|------|-------|-----------|
| Text read first | 83.0 | 66.0 | 34.0 | 84.9 | 73.6 | 94.3 | 86.8 | 98.1 | 77.6 |
| Video after the text | 68.0 | 66.0 | 43.4 | 75.5 | 71.7 | 94.3 | 83.0 | 100.0 | 75.2 |
| Video seen first | 47.1 | 78.4 | 58.8 | 72.5 | 70.6 | 82.4 | 82.4 | 96.1 | 73.5 |
| Text after the video | 66.7 | 72.6 | 51.0 | 86.3 | 86.3 | 92.2 | 98.0 | 96.1 | 81.1 |

Of the participants who read the text-based scenario first, 83.0% correctly interpreted the Collection Limitation guideline based on text while only 68.0% correctly interpreted the guideline in the case of video, which the participants were shown second. The other group was exposed to the video-based scenario first. Of these 47.1% of the participants correctly interpreted the guideline in the case of video while 66.7% interpreted it correctly for the text-based version. For the Data Quality guideline, there were large differences in interpretation between the four conditions. This guideline is more often correctly interpreted for the video-based scenarios (51.1%) than for the text-based scenarios (42.5%). It is also clear that for the group who saw text version first, for the Purpose Specification and Secure Storage of data guidelines that there were no differences between text and video version of the scenario. Compared to other systems features, the last three features (Secure Storage, Insight and Modification) were answered most correctly. The least correctly answered question in the test was regarding Data Quality (34.0%). When text was shown first, it slightly improved if video was shown after the text (43.4%). The aggregated percentage of good answers on the text-based questions and aggregated percentages of good answers on the video-based questions was 79.4% and 74.4%. Multiple analysis of variance, using both the order and the medium as independent variables, showed a significant effect for medium (p=0.001).

No significant effect was found for order (p=0.715), showing that results were not affected by learning effects (either because of the purpose of the experiment or the

content of the scenarios). This result confirmed that medium was the main effect. However, an indicative effect was found for interaction between medium and order ($p=0.077$). Taken individually, both text and video performed a little better when shown second – a different measure to checking for an order effect which tested the average score of both media when-shown-first against the average score of both media when-shown-second. This ‘interaction effect’ means, for example, that text performed better when shown second not for the fact of being shown second – it is already known that there was no effect for order – but because it was preceded by the video-based scenario. Scores for video improved also when preceded by text but by a lesser degree. Thus, the best results were obtained when the text scenario was preceded by the video scenario.

The above points are more clearly illustrated in Figure 3, showing the estimated marginal means of the percentage of correct answers for the four different conditions: Text shown first, text shown second, video shown first, and video shown second. Figure 3 shows clearly that there is a higher overall score for the text-based questions, which is the main effect mentioned above. In addition, it shows that the difference between the score of each medium when shown either first or second does not differ that much, explaining why the order effect was not significant.

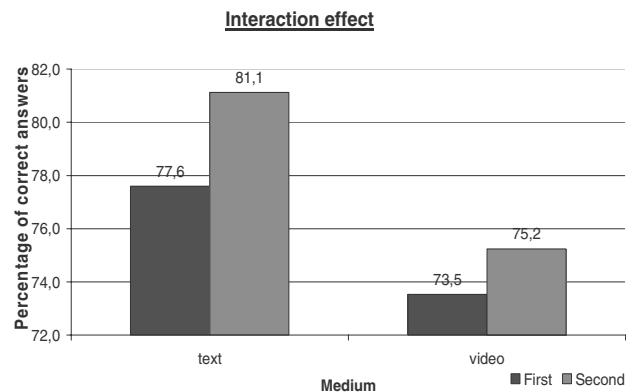


Fig. 3. Graph of the interaction effect

The interaction effect, which was significant at a ten-percent level, is reflected in the difference in increase between the first and second showing of the different media. It clearly shows that the increase in text shown first to being shown second is larger than the increase in video.

4 Discussion

As has been explained, the experiment aimed to investigate the variance in interpretation between video- and text-based scenarios. First of all, both type of scenarios showed poor understanding of the compliance with privacy guidelines. On average the text-based scenario resulted in 79.4% correct interpretation, while for the

video-based scenario this was 74.3%. This means that about one fifth to one quarter of the privacy guidelines in this experiment were interpreted incorrectly. Jensen and Potts [15] explain that text with long words and/or long sentences are more difficult to read. The text used in the scenarios consisted of fairly short sentences and yet proved difficult to read.

Furthermore, the results showed that the interpretation of the both scenarios varied greatly. While on the whole, the text-based scenario performed better than the video-based scenario, it only did so to the order of 6.7%. However, more importantly, underlying these averages were great variations in answers given to individual questions. This variance between individual answers, and thus the interpretation of privacy and security guidelines between text- and video-based scenarios is the most significant finding of this study. It shows that caution must be exercised when surveying users regarding scenarios, whether this is for research purposes or as part of an iterative interaction design process.

The obtained results call into question the reliability of scenarios – both text- and video-based – for surveying users’ attitudes regarding privacy-related issues. As mentioned, it is common practice to use scenarios of this kind as a means to explain a system to users and facilitate discussion regarding privacy. Clearly, results from such studies are contingent upon sufficient user comprehension, so future studies should include appropriate checks of the users’ comprehension of the (sometimes) futuristic scenarios presented to them.

Another striking result is the interaction between media. Video, when shown first, “lifted” the results of the text-based scenario shown afterwards more than the text scenario did for video. This has led to the hypothesis that what may be happening is that video-based scenarios are more suited to act as contextualizing overviews to the more detail-affording qualities of text. Such a hypothesis does not take account for, much less explain, the differences in answers given for the individual guidelines between video and text.

To further analyze the outcomes of the study a hypothetical ‘best-answer’ medium was also looked at. The results of this medium consist of a combination of both text and video. To compute the percentage of correct answers for this hypothetical medium the best scores of text and video (when shown first) for each OECD guideline were combined. When adding all these ‘best-answers’ the following results emerge:

- 73.5% correct overall for video when shown first;
- 75.2% correct overall for video when preceded by text;
- 77.6% correct overall for text when shown first;
- 81.1% correct overall for text when preceded by video;
- 82.2% ‘correct’ overall for a hypothetical ‘best-answer’ medium shown first.

What is striking here is that the combined ‘best-answer’ medium does not score significantly higher than text scenarios when preceded by video. This result suggests that the problems posed by the phenomenon of variation in understanding between media by offering a ‘best of both worlds’ option can be avoided. Rather than make the pitfall of exposing users to one medium or the other, the best qualities of both may be able to be exploited if a text-based scenario is used, and preceded by a video-based version. Use of these scenarios should be framed in such a way that discussion is

limited to the text-based version only, to avoid the shortcomings of video when shown first. In interaction design practice, the implication of this finding is that where a video scenario is made to solicit reactions of users to a particular design concept it can be wise to offer a corresponding text, which they will consult after viewing the video in order to answer detailed questions that require in depth understanding of the concept described.

To help explain the results further, seven of the face-to-face participants were contacted again approximately three months after the experiment to discuss the study's findings with them. These return interviews were undertaken first to double-check for flaws in the experiment design and second to elicit explanations, which might have been overlooked. All seven interviewees could remember the experimental set-up with accuracy and the general function of the device in the scenario. Participants were then presented with the two prototypes once again in the order in which they had been shown in the original experiment. They were then informed that both were the same and provided with the answers, which they had given in the original experiment. These interviews took the form of a discussion of these answers with participants.

A number of participants argued that the video imagery did not provide extra detail regarding interaction with the device. While this point is true, it was necessary in order to maintain equivalence between both media and thus necessary for the experiment as a whole.

On the matter of the experiment results, half of the re-interviewed participants had performed better on the video prototype. However, regardless of their actual results, participants insisted on normative explanations of differences between text and video, contradicting the results data, mainly suggesting that text is vastly easier or that there was a learning effect. One participant did refer to the interaction effect, pointing out that for her, the benefit of video is that it helps frame a subject while text solidifies it. This opinion was reflected in the responses of other participants. They argued that overall they would prefer a mixed-medium, criticizing both video and text on their own. While the benefits of video were that it acted as a guide to a topic and can be digested passively, "once it's gone, it's gone." Another participant linked it to TV news, where text and video appear on screen at once, and others talked of the benefits of being able to "cross-reference" between video and text. These points, raised by participants, support the suggestion for the use of both video and text scenarios in succession.

Further to the methodological implications noted, the flawed understanding of privacy-related consequences of system use is noteworthy. Similar (and usually even more complex) texts to the scenarios are presented as privacy policies in commercial websites; this may lead to wrong assumptions about systems. When users realize the inaccuracy of their assumptions, they experience an invasion of their privacy [3]. It has to be noted that the concepts that were presented to the participants were based on legal documents and perspectives unfamiliar to most people. However, it is precisely such concepts that are translated into privacy policies used on websites. These policies are known to present comprehension problems to users [15, 19] and often are used to inform the interaction design for such systems, e.g., [18, 14].

5 Conclusions

The experiment examined differences in understanding of privacy guidelines between text and video scenarios. This paper has shown that a variation exists, both overall and more importantly in users' understanding of individual issues. It is also shown that the order in which text and video are shown has a significant effect on the level of understanding. Text scenarios resulted in slightly better understanding. However, if a video-based scenario is shown first then the interpretation of a text-based scenario improves. As a result of this study many questions are raised. Why does video appear to support text more than the other way around? What is the underlying process of this phenomenon? And why is there such a great variance in the answers for individual questions between media? This study cannot yet explain the reasons underlying these results.

For the time being, it is recommended that text-based scenarios preceded by a video-based version should be used for the purposes of privacy- and security-related user surveys. Future research into methodologies for investigating privacy should compare how video and text scenarios relate to actual or staged (e.g. Wizard of Oz) experiences of pervasive systems.

A similar advice can be given for the use of text and video scenarios during the interaction design process. Future work should examine whether similar results are obtained outside the specific context of privacy. Extending this inquiry can provide methodological guidance as to when different media are appropriate for presenting scenarios.

References

1. Ackerman, M.S., Cranor, L.F., Reagle, J.: Privacy in e-commerce. Examining user scenarios and privacy preferences. In: ACM Conf. on Electronic Commerce, ACM Press, New York (1999)
2. Acquisti, A., Grossklags, J.: Losses, Gains and Hyperbolic Discounting: An Experimental Approach to Information Security Attitudes and Behaviors. In: 2nd Annual Workshop of on Economics and Information Security (2003)
3. Adams, A., Sasse, M.A.: Privacy in multimedia communications. Protecting users, not just data. In: Blandford, A., Vanderdonck, J., Gray, P. (eds.) People and Computers XV. Interaction without frontiers. Joint Proceedings of HCI2001 and IHM2001, pp. 49–64. Springer, London (2001)
4. Ajzen, I., Fishbein, M.: The Influence of Attitudes on Behavior. In: Albarracín, D., Johnson, B.T., Zanna, M.P. (eds.) The Handbook of Attitudes, pp. 173–221. Erlbaum, Mahwah (2005)
5. Berendt, B., Günther, O., Spiekermann, S.: Privacy in E-commerce: Stated Preferences vs. Actual Behavior. Communications of the ACM 48, 101–106 (2005)
6. Boyle, M., Greenberg, S.: The Language of Privacy: Learning from Video Media Space Analysis and Design. ACM Transactions 12, 2 (2005)
7. Carroll, J.M.: Scenario-Based Design. Wiley, New York (1995)
8. Chung, E., Hong, J., Lin, J., Prabaker, M., Landay, J.A., Liu, A.: Development and Evaluation of Emerging Design Patterns for Ubiquitous Computing. In: Proceedings DIS 2004, pp. 233–242. ACM Press, New York (2004)

9. Culnan, M.J., Armstrong, P.K.: Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science* 10, 104–115 (1999)
10. van de, G-P.E., Markopoulos, P., de Ruyter, B.: On the relative importance of privacy guidelines for ambient health care. In: *The fourth Nordic Conference on Human-Computer Interaction NordiCHI 2006*, pp. 377–380. ACM Press, New York (2006)
11. van de, G-P.E., Markopoulos, P., de Ruyter, B.: Privacy Policies and Text-Based Empirical Research: Methodological Issues. In: *CHI 2006 Workshop on Privacy and HCI. Methodologies for Studying Privacy Issues* (2006)
12. van de, G-P.E., Markopoulos, P., de Ruyter, B.E.R., Eggen, J.H., IJsselsteijn, W.: Investigating privacy attitudes and behavior in relation to personalization. *Social Science Computer Review* 26,1 (in press 2008)
13. Günther, O., Spiekermann, S.: RFID and the Perception of Control. The Consumer's View. *Communications of the ACM* 48, 73–76 (2005)
14. Iachello, G., Abowd, G.D.: Privacy and proportionality: adapting legal evaluation techniques to inform design in ubiquitous computing. In: *Proc. CHI 2005*, pp. 91–100. ACM Press, New York (2005)
15. Jensen, C., Potts, C.: Privacy Policies as Decision-Making Tools. An Evaluation of Online Privacy Notices. *CHI Letters* 6, 471–478 (2004)
16. Jiang, X., Hong, J.I., Landay, J., A.L.: Approximate Information flows. *Socially-Based Modeling of Privacy in Ubiquitous Computing*. In: Miesenberger, K., Klaus, J., Zagler, W. (eds.) *ICCHP 2002. LNCS*, vol. 2398, pp. 176–193. Springer, Heidelberg (2002)
17. Kobsa, A., Teltzrow, M.: Contextualized Communication of Privacy Practices and Personalization Benefits. In: Martin, D., Serjantov, A. (eds.) *PET 2004. LNCS*, vol. 3424, Springer, Heidelberg (2005)
18. Langheinrich, M.: Privacy by Design. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) *UbiComp 2001: Ubiquitous Computing. LNCS*, vol. 2201, pp. 273–291. Springer, Heidelberg (2001)
19. Milne, G.R., Culnan, M.J.: Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing* 18, 15–29 (2004)
20. OECD: *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (1980)
21. Palen, L., Dourish, P.: Unpacking “Privacy” in a Networked World. In: *Proceedings of CHI 2003*, pp. 129–146. ACM Press, New York (2003)
22. Patil, S., Romero, N.A., Karat, J.: Privacy and CHI. Methodologies for studying privacy issues. In: *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pp. 1719–1722. ACM Press, New York (2006)
23. Romero, N.A., Perik, E.M., Patil, S.: Appropriate methodology for empirical studies of privacy. In: Costabile, M.F., Paternó, F. (eds.) *INTERACT 2005. LNCS*, vol. 3585, pp. 87–89. Springer, Heidelberg (2005)
24. Spiekermann, S., Grossklags, J., Berendt, B.: E-privacy in 2nd Generation E-commerce. Privacy Preferences Versus Actual Behavior. In: *Proc. ACM conference on Electronic Commerce*, pp. 38–47. ACM Press, New York (2001)
25. Wright, D. (ed.) *Safeguards in a World of Ambient Intelligence. Final Report, SWAMI Deliverable D4. A report of the SWAMI consortium to the European Commission under contract 006507 August 2006* (2006), <http://swami.jrc.es>