Privacy Risk Perceptions and Privacy Protection Strategies

Isabelle Oomen and Ronald Leenes

TILT – Tilburg Institute for Law, Technology, and Society
Tilburg University, the Netherlands
r.e.leenes@uvt.nl

Abstract. Several opinion polls have reported that many people claim to be concerned about their privacy, yet that most people in fact do very little to protect their privacy. Are privacy concerns indeed insufficient motivators to adopt privacy protection strategies? What then characterizes the users of these strategies? On the basis of a large scale survey amongst Dutch students, this paper explores the relation between privacy risk perception and privacy protection strategies in more detail. It elaborates on factors that constitute privacy risk perception, as well as three kinds of strategies adopted by individuals to protect their privacy: behavioral measures, common privacy enhancing technologies (PETs), and more complex PETs. Next, it explores the relation between the respondents' perception and the strategies they employ in more detail to answer the question what characteristics the users of the various strategies have in terms of perception, gender and age. Gender appears not to influence privacy risk perception, yet men are more familiar with the various privacy protection strategies and use them more of-ten than women. In general, a higher privacy risk perception does not lead to the adoption of stronger or more protection strategies, except for the use of pseudonyms, cookie crunchers, anonymous email, safe email, and providing false personal data. Our analysis deepens the understanding of privacy risk perception and privacy protection strategies, yet leaves the privacy paradox unresolved.

1 Introduction

Many opinion polls report that although the overwhelming majority of people claim to be concerned about their privacy, most people in fact do very little to protect it[1]. This paradox is often explained by stating that people lack information, that they don't know the degree to which their information is being collected, and that they are unaware of the potentially harmful consequences of this practice (e.g., [14]). Most studies, however, don't explore the relationship between the privacy risk perception of individuals and their subsequent actions to protect their privacy in depth or in detail. Opinion polls go no further than to mention: X % claims to be concerned about their privacy and Y % says that they have taken action to protect their privacy. This raises

[1] For a list of opinion polls, see: http://www.epic.org/privacy/survey

the question: How are privacy risk perception and the actions people take to protect their privacy related to each other?

In this study, we explore this question in more detail on the basis of empirical research among Dutch students. We particularly focus on privacy risk perception, the strategies students employed to protect their privacy and the relation between the two. We also explore whether age and gender are related to privacy risk perceptions and adopted strategies.

2 From perception to action

Before cheap printed material was widely available, communication was mainly two-way, face to face. Printed material changed this. One-way communication gained importance with the effect that the identity of the other often is unknown, i.e. the author reveals without being revealed and the reader learns without being learned about. The current shift to electronic communication changed the communication back to two-way [14]. Electronic communication inevitably leaves traces which makes it possible for anybody interested enough to collect, organize, and analyze personal information of others [12]. It is increasingly difficult to reveal without being revealed and to learn without being learned about [14]. As the recognition of this phenomenon grows, privacy has increased in salience.

In the context of online behavior, informational privacy is at stake. Informational privacy relates to an individual's right to determine how, when, and to what extent information about herself will be released to others [1-2, 15]. The acquisition, storage, manipulation, mining, sharing, selling and analysis of personal data represent violations of the individual's informational privacy. It may also lead to practices like identity theft, social sorting, and far-going surveillance [8, 9-11, 13]. Privacy risks are regarded as the consequences of the abuse of misuse of personal information. Possible privacy risks can thus be identity theft, loss of freedom, threat to personal safety, threat to dignity, invasion of the private sphere, unjust treatment, or financial loss. The perception of these risks varies from individual to individual based on that person's own values and general perceptions, and her experiences [1, 6].

When people perceive risks, they want to reduce them. We distinguish three groups of strategies that individuals may employ to do so. The first group involves behavioral strategies. Providing incorrect answers when personal information is requested is a strategy that can be applied easily in online environments. Other examples are the use of anonymous email addresses and the use of pseudonyms. The second group comprises well known security measures and PETs. Spam filters, firewalls, and anti spyware have become standard PETs on almost every computer (that runs on a Microsoft platform). The third group of strategies consists of more advanced PETs. These are more difficult to implement and sometimes require cooperation of service providers which limits their use. Examples of these PETs are: encryption tools, anonymous remailers, trust certificates, anonymisers, cookie crunchers, password managers or vaults, and safe email. Because they are easier to implement, we assume that behavioral privacy protection strategies and common PETs are used more often than advanced PETs.

Hypothesis 1: Behavioral privacy protection strategies and common PETs are used more often than advanced PETs.

People with a weak privacy risk perception will have less incentives to use privacy protection strategies and we expect these individuals therefore not to use them. Individuals with stronger privacy risk perceptions, on the other hand, are more likely to adopt one or more privacy protection strategies to counter or limit the risks. Conversely, we assume that people who use privacy protection strategies to have a stronger privacy risk perception (i.e. they perceive their privacy risk higher) than people who don't use these strategies. Because advanced PETs require extra effort (i.e. they are not pre-installed on computers and one has gain knowledge about their existence, get possession of them, and learn how to use them), it is likely that people who use these advanced PETs have a higher privacy risk perception – their concerns outweighs the required efforts – than people who only use common PETs or behavioral strategies.

Hypothesis 2: Internet users who use privacy protection strategies have a stronger privacy risk perception than those who don't use these strategies.

Hypothesis 3: Internet users who (also) use advanced PETs have a higher privacy risk perception than those who only use common PETs or behavioral strategies

The different strategies are, of course, not mutually exclusive. It is likely that the adoption of one strategy will reinforce the use of the other strategies. Therefore, we assume that there is a positive relationship between the use of strategies.

Hypothesis 4: There are positive relationships between the various strategies.

Figure 1 shows the presumed relationships.

3 Measuring privacy risk perception and protection strategies

The data used for this paper is collected as part of the EU FP6 PRIME project[2]. The PRIME project aims to develop privacy-enhancing identity management tools. In order to gain a better understanding of user needs, large scale surveys are conducted among European internet users.

For the present study a sample of Dutch university and college students was chosen due to time and budget restraints. This approach limits the generalizations of the findings. Twenty six Dutch universities and colleges of further education where approached with a request to forward an email invitation to participate in the survey to their students. Five institutions, two universities and three colleges, forwarded our email invitation containing a link to the online questionnaire. Three universities and four colleges, either placed the message on their website or in their newsletter. As an

[2] See http://www.prime-project.eu for information about the project and its public deliverables.

incentive for the students to participate, four iPod nano's (2GB) were raffled among the respondents.

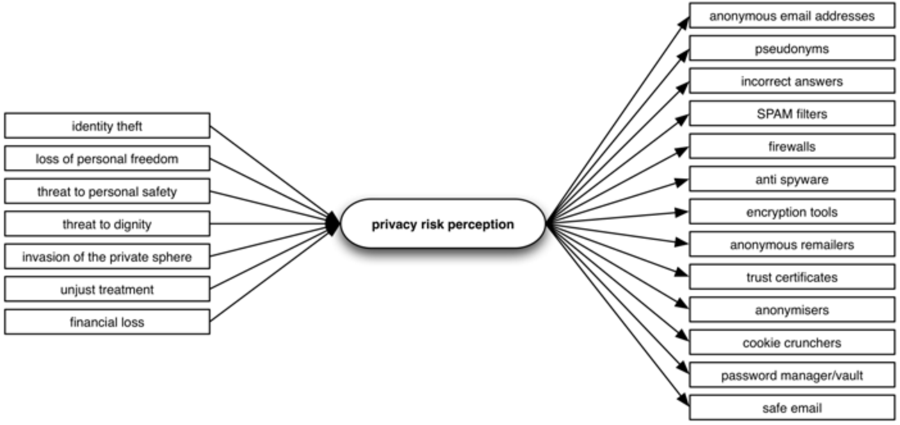

Fig. 1. Privacy risk perception and privacy protection strategies.

The survey ran from Friday December 15th 2006 until Tuesday January 9th 2007. In total, 5541 students participated in the online survey. The response rate for students invited by email was 6.43%, while the response rate for students invited by an announcement on the website or in the newsletter was 0.3%. The overall response rate was 2.31%.

The questionnaire addressed demographics of the respondents, their online activities, their trust with respect to public and private institutions, their general concerns, their concerns with respect to personal data and privacy, and their knowledge and experience with privacy enhancing technologies. The respondents were also randomly assigned to answer questions about one out of four specific online contexts: e-government, online shopping, online banking, and online travel booking. This resulted in a total of 295 variables, of which a few are used in the present study.

Completing the full questionnaire took approximately 30 minutes. Consequently, some 25% (1374) of the respondents did not complete the questionnaire. Respondents with more than 75% missing data were deleted from the sample set. The data of 5139 respondents was used for the analyses and the remaining missing data were handled by using list wise deletion in the analyses.

In the sample 52% of the respondents were male and 48% were female. University students amount for 60.7% of the sample against 39,3% for the colleges of higher education. Due to the participation of the Open University, with students of all ages, the distribution of age of the respondents showed a large variation. Most respondents (85%) are under 28, still 10% of the respondents is older than 32 and 5% of them is even older than 42, one of the students reported 72 years of age. The mean age in the sample was 24.2 (SD = 7.74).

4 Analyses[3] and results

We will first focus on the respondents' perception of specific privacy risks and the theoretical construct privacy risk perception. Next we address the use of privacy protecting strategies by the respondents and whether gender affects the use of different strategies. Finally, the privacy risk perceptions of the groups within the strategies were compared.

4.1 Privacy risk perception

Privacy risk perception of the respondents was measured by seven items on a five point scale as part of the question: ‘How concerned are you in general about the following possible consequences of abuse/misuse of your personal information?’ The concerns listed were: ‘threat to your personal safety’, ‘loss of personal freedom’, ‘invasion of your private sphere’, ‘financial loss’, ‘possibility of identity theft’, ‘unjust treatment’, and ‘threat to you dignity’. The answers were: ‘not at all’ (1), ‘barely’ (2), ‘neutral’ (3), ‘slightly’ (4), and ‘considerably’ (5). The means of these items are displayed in table 1. The respondents are most concerned about invasion in their private sphere, with a mean of 3.42 (on a five point scale). They are least concerned about the possible threat to their dignity and about receiving an unjust treatment as a result of abuse or misuse of their personal data, with means of 2.84 and 2.85 respectively.

Table 1. Privacy risk perception.

	M (*SD*)	Privacy risk perception
Loss of freedom	3.07 (1.198)	.837
Threat to personal safety	2.92 (1.191)	.819
Threat to dignity	2.84 (1.186)	.814
Invasion of the private sphere	3.42 (1.137)	.805
Unjust treatment	2.85 (1.152)	.777
Possibility of identity theft	3.14 (1.223)	.773
Financial loss	3.17 (1.151)	.716
Mean	3.06	
Explained variance	63%	
α	.90	

Notes: N=4424
Values are component loadings
Method: Principal Component Analysis

Moderate to high, positive, and significant (Pearson product-moment) correlations were found between the items, ranging from .463 to .729 (not in the table). There are positive relationships between the specific privacy risk perceptions. This indicated that the items could represent a common concept. A principal component analysis was

[3] Performed with SPSS (v13.0 and v15.0), see [7]. For more information about the mathematics behind the analyses see, for example, [5].

conducted to test this assumption. One component was found, indicating that these seven items together measure the same latent construct: privacy risk perception. The component loadings in the second column of table 1 reflect the contribution of each item to the common component. These range from .716 to .837, so each item contributes sufficient to the component. The component explains 63% of the variance in the seven items, which is high. Cronbach's alpha (α) was calculated to establish the reliability of the scale. An alpha of .90 indicates that items form a very good scale.

For each respondent a score on privacy risk perception was calculated by adding the values on the seven items and dividing the sum by seven. Scores on privacy risk perception thus range from 1 (very low privacy risk perception) to 5 (very high privacy risk perception). The mean of privacy risk perception is 3.06 (see table 1), which is just above the scale mean of 3.00, meaning that on average, the respondents are neutrally concerned about privacy threats.

4.2 The use of privacy protection strategies

The strategies of reducing privacy threats were measured by 13 items in three groups. The first group, the behavioural measures consisted of three items: anonymous email addresses, pseudonyms and providing false information. The use of anonymous email address, was measured by asking: 'How many of these email addresses are anonymous email addresses?'. The figures provided were recoded into 'does not have an anonymous email address' and 'does have an anonymous email address'. The use of pseudonyms and providing incorrect answers were both measured on a five-point scale using the questions: 'Have you ever used a pseudonym, such as another name, alias, or nickname, in online activities when you were asked for your name?' and 'Have you ever intentionally given incorrect answers when you were asked for your personal information?'. The answers 'sometimes', 'frequently', and 'often' were, recoded into 'does use pseudonyms' respectively 'does provide incorrect answers. The answers 'never' and 'rarely' were recoded into 'does not use pseudonyms' and 'does not provide incorrect answers', respectively.

The remaining ten items, measuring the different strategies, were questions about the use of and familiarity with well known ('spam filters', 'firewalls', anti spyware) and lesser known ('encryption tools', 'anonymous remailers', 'trust certificates', 'anonymisers', 'cookie crunchers', 'password managers/vaults', and 'safe email') privacy enhancing technologies. The question was: 'How familiar are you with the following privacy enhancing technologies?'. The possible answers were: 'I use it', 'I'm familiar with it, but I don't use it', and 'I'm not familiar with it'.

Anonymous email addresses are used by 2760 respondents, which amounts for 55.8% of the sample population (see table 2). Within this group 54.5% is male against 45.5% female. The Chi square test (χ^2) reveals that this difference is real. On the basis of a 52% male sample, the expected proportion male anonymous email users is 2760 * 0.52 = 1435, whereas the actual proportion is 2760 * 0.545 = 1504. The same holds for the percentage men in the group who do not use anonymous email addresses, this percentage, 48.6%, is also different from the expected percentage.

Table 2. Use of strategies and gender and mean age in the different groups.

Strategy	Uses it	Does not use it
Anonymous email	2760 (55.8%)	2184 (44.2%)
Males (χ^2 (df=1) = 16.84***)	54.5%	48.6%
Age (t (df=3800) = 7.60***)	23.4	25.1
Pseudonyms	3189 (62.8%)	1885 (37.2%)
Males (χ^2 (df=1) = 240.44***)	60.4%	37.9%
Age (t (df=2937) = 7.77***)	23.5	25.3
Incorrect answers	2157 (45.3%)	2603 (54.7%)
Males (χ^2 (df=1) = 250.98***)	64.2%	41.1%
Age (t (df=4643) = 6.04***)	23.5	24.8

Note: * p < .05; ** p < .01; *** p < .001

The χ^2-test is significant at the p < .001 level, which means that the probability that the difference between the observed and expected frequencies exists due to chance is less than 0.1 percent. So, we may conclude that more men than women use anonymous email addresses and that there are more women than men in the group who do not use anonymous email addresses. The mean age in the group of those who use anonymous email addresses is 23.4 and the mean age of non users is 25.1. A t-test shows that these means differ significantly from each other. So, anonymous email addresses are more often used by younger people.

Pseudonyms are more often used than anonymous email addresses, 62.8% against 55.8%,, and they are more often used by men than women, 60.4% against 39.6%. The differences between men and women in the groups of pseudonym users and non pseudonym users again show a significant difference. Most people provide correct instead of incorrect answers when personal information is asked for, 54.7% against 45.3%, but incorrect answers are more often given by men than women (64.2% against 35.8%). Pseudonyms and incorrect answers are also more often used by younger people, 23.5 and 23.5 against 25.3 and 24.8, respectively.

With respect to the common privacy and security safeguards, it turns out that most respondents use spam filters, firewalls, and anti spyware, 79.0%, 88.5%, and 73.6% respectively (see table 3). Only a small number of people are unfamiliar with these Privacy Enhancing Technologies (ranging from 1.9% for firewalls to 7.9% for anti spyware). More men than women use these technologies, and also more men than women are familiar with spam filters and firewalls (but don't use them). In the groups of those who are unfamiliar with spam filters, firewalls, and anti spyware we see significantly more women than men, 75.7%, 77.9%, and 82.7% respectively. There is a small, but significant, difference between the mean age of those who use spam filters and firewalls and those who are familiar with these technologies. In contrast to using anonymous email addresses and pseudonyms and providing incorrect answers, spam filters and firewalls are more often used by older people, 24.7 and 24.5 against 23.4 and 23.4.

Encryption tools, anonymous remailers, trust certificates, cookie crunchers, password managers and vaults, and safe email are largely unknown to the respondents, percentages range from 42.1% to 69.7%, especially amongst women.

These technologies are more often used by men than women. Of the non-users men are, more often than women, more familiar with these technologies. An exception, however, is the use of password managers or vaults. These are more often used by women than by men. The largest difference between men and women can be seen in the use of encryption tools. No less than 85.3% of the people who use encryption tools are men. The people who use encryption tools or a familiar with them are significantly older than the people who are unfamiliar with this technology, 25.2 and 25.6 against 23.4. No differences in the average age were found between the users of anonymous remailers, anonymisers, cookie crunchers, password managers or vaults, and safe email, and those who are familiar (but don't use them) or are unfamiliar with these technologies.

Table 3. Use of and familiarity with strategies by gender.

Strategy	Uses it	Is familiar with it	Is unfamiliar with it
Spam filters	3169 (79.0%)	738 (18.4%)	103 (2.6%)
Males (χ² (df=2) = 41.72***)	53.5%	58.1%	24.3%
Age (F (df=2;4003) = 7.62***)	24.7	23.4	24.0
Firewalls	3538 (88.5%)	388 (9.7%)	74 (1.9%)
Males (χ² (df=2) = 30.49***)	54.6%	50.8%	23.0%
Age (F (df=2;3993) = 4.57*)	24.5	23.4	23.0
Anti spyware	2932 (73.6%)	740 (18.6%)	313 (7.9%)
Males (χ² (df=2) = 198.68***)	58.4%	49.6%	17.3%
Age (F (df=2;3978) = 1.99)	24.4	24.0	25.0
Encryption tools	667 (16.9%)	118 (30.1%)	2097 (53.1%)
Males (χ² (df=2) = 852.62***)	85.3%	73.8%	32.2%
Age (F (df=2;3945) = 34.29***)		25.2	25.6 23.4
Anonymous remailers	124 (3.1%)	1166 (29.4%)	2677 (67.5%)
Males (χ² (df=2) = 404.68***)	74.2%	76.8%	42.6%
Age (F (df=2;3960) = 4.13)	24.9	24.9	24.1
Trust certificates	1246 (31.3%)	1037 (26.1%)	1697 (42.6%)
Males (χ² (df=2) = 610.85***)	74.0%	65.5%	31.3%
Age (F (df=2;3973) = 2.84)	24.8	24.5	24.1
Anonymisers	113 (2.8%)	1095 (27.5%)	2774 (69.7%)
Males (χ² (df=2) = 431.53***)	75.2%	78.7%	42.7%
Age (F (df=2;3975) = 3.71)	23.4	24.9	24.2
Cookie crunchers	643 (16.1%)	1231 (30.9%)	2108 (52.9%)
Males (χ² (df=2) = 148.64***)	56%	67%	45.3%
Age (F (df=2;3975) = 0.23)	24.5	24.3	24.4
Password managers/vaults	836 (21.1%)	1461 (36.8)	1673 (42.1%)
Males (χ² (df=2) = 308.95***)	47.2%	71.6%	41.1%
Age (F (df=2;3963) = 0.31)	24.2	24.5	24.4
Safe email	380 (9.6%)	1298 (32.8%)	2285 (57.7%)
Males (χ² (df=2) = 222.87***)	59.2%	69.3%	43.7%
Age (F (df=2;3956) = 2.01)	23.6	24.5	24.5

Notes: * p < .05; ** p < .01; *** p < .001

The percentage of people who use behavioral privacy protection strategies or common PETs ranges from 45.3% for providing incorrect answers to 88.5% for firewalls. In contrast, the percentage of people who use advanced PETs ranges from 2.8% for anonymisers to 31.3% for trust certificates. So, hypothesis 1 is confirmed: behavioral privacy protection strategies and common PETs are more often used than advanced PETs.

To check for correlations between pairs of strategies, we used Phi when both strategies had only two answering options (i.e. use or don't use) and Cramer's V when at least one of the strategies had three answering options (i.e. use, familiar, or unfamiliar). The resulting correlation coefficient ranges from 0, no correlation, to -1 or +1, perfect correlation. Table 4 shows the correlation matrix. The three (theoretical) groups outlined previously (behavioral, common measures, more complex PETs) can clearly be distinguished here. The first group consists of using anonymous email addresses and pseudonyms, and providing incorrect answers. Low, positive, and significant correlations were found between these three strategies, ranging from .206 to .355. This indicates that using one of these strategies is slightly enhancing the possibility of the respondent also using the other two strategies as well. The use of anonymous email addresses and pseudonyms, and providing incorrect answers, however, is independent from the use of the more complex PETs, that is, using anonymous email addresses and pseudonyms, and providing incorrect answers do not influence the use of these more complex PETs, or vice versa.

The second group consists of spam filters, firewalls, and anti spyware. Here too, low, positive, and significant correlations between the strategies are found, ranging from .207 to .307. This indicates that using one of these technologies slightly increases the probability of using one of the other two technologies.

The third group consists of the technologies most people are unfamiliar with (i.e. encryption tools, anonymous remailers, trust certificates, anonymisers, cookie crunchers, password managers or vaults, and safe email). Between these technologies, low to moderate, positive and significant correlations were found, indicating that people who are unfamiliar with one of these technologies, are also likely to be unfamiliar with one of the other technologies.

We can, therefore, not clearly accept, nor reject hypothesis 4. There are positive relationships between strategies, but only between strategies within the same group (i.e. behavioral strategies, widely used PETs, and more advanced PETs).

4.3 Does privacy risk perception explain privacy protection strategy

An interesting research question is whether privacy risk perception explains the use of certain privacy protection strategies. On the basis of the current analysis this question can not be answered, we have no insight in the causality or finality of the relations. What we can do is see whether the privacy risk perception of users versus non-users (and within this latter group between those familiar with the technology and those not) of particular strategies differs. For this purpose we have used F-tests. If the F-test is significant, at least two group means are different from each other meaning that the perceptions in the various groups really differ. When the F-test was significant, a

Tukey HSD multiple comparisons Post Hoc test was used to determine which group means differ significantly from each other.

For the spam filters, firewalls, and anti spyware, the F-tests are not significant, which means that the privacy risk perception of the group who uses them is not different from the group who is familiar with them (but not uses them). And the privacy risk perception of those who are unfamiliar with these technologies is not different from those who use or are familiar with them (see table 5). So, hypothesis 2 is rejected for spam filters, firewalls, and anti spyware.

For encryption tools, the F-test is significant and the Tukey test shows that the mean privacy risk perception in the group who uses encryption tools is significantly different from the mean privacy risk perception in the group who is unfamiliar with this technology, 3.12 against 3.01. So, people who use encryption tools, on average, have a stronger privacy risk perception than those who are unfamiliar with encryption tools. The mean privacy risk perception of those who are familiar, but don't use, encryption tools is neither different from those who use this technology, nor from those who are unfamiliar with it. The same holds for trust certificates and password managers or vaults. The privacy risk perception of those who use these technologies is significantly stronger than the privacy perception of those who are unfamiliar with them.

For anonymous remailers, a significant difference in mean privacy risk perception was not only found between those who use them and those who are unfamiliar with them, but also between those who are familiar with them and those who are unfamiliar with them, but not between those who use them and those who are familiar with (but not use) them. Apparently, the privacy risk perception of those who are familiar with anonymous remailers, independent of whether one uses this technology or not, is stronger than the privacy risk perception of those who are unfamiliar with anonymous remailers. The same holds for anonymisers. Independent of whether the technology is used or not, those who are familiar with anonymisers have, on average, a stronger privacy risk perception than those who are unfamiliar with them.

Both for cookie crunchers and safe email, the F-test was significant and the Tukey test shows that, on average, the privacy risk perception of people who use these technologies is stronger than the privacy risk perception of people who do not use these technologies, irrespective whether one is familiar with the technology or not.

The F-test tests only whether the groups differ, but does not show how big the difference is, that is, what the effect size is. The effect size can be calculated, but only for two groups. Because there is no difference in privacy risk perception between those who are familiar with cookie crunchers and those who are not, but both groups differ from the group who use them, we have merged these groups to form a new group: 'does not use cookie crunchers'. The same was done for safe email resulting in the new group: 'does not use safe email'. For both anonymous remailers and anonymisers, the group who uses these technologies and the group who is only familiar with these technologies were merged.

Table 4. Correlations between privacy protection strategies

	anon. email	pseudon.	incorr. answ.	spam filters	Firewall	anti spywar.	encrypt.	anonyn. remailers	trust certs.	anony miser	cookie crunch	pass. man/vlt
Pseudonyms	.253^{***a}											
Incorrect answers	.206^{***a}	.355^{***a}										
Spam filters	.020^{b}	.045^{*b}	.061^{**b}									
Firewalls	.018^{b}	.063^{***b}	.072^{***b}	.307^{***b}								
Anti spyware	.073^{***b}	.157^{***b}	.108^{***b}	.207^{***b}	.266^{***b}							
Encryption	.078^{***b}	.183^{***b}	.188^{***b}	.117^{***b}	.095^{***b}	.188^{***b}						
Anonymous remailers	.097^{***b}	.160^{***b}	.165^{***b}	.071^{***b}	.104^{***b}	.131^{***b}	.387^{***b}					
Trust certificates	.124^{***b}	.241^{***b}	.200^{***b}	.123^{***b}	.103^{***b}	.188^{***b}	.387^{***b}	.325^{***b}				
Anonymisers	.125^{***b}	.187^{***b}	.180^{***b}	.063^{***b}	.053^{***b}	.120^{***b}	.374^{***b}	.613^{***b}	.361^{***b}			
Cookie crunchers	.076^{***b}	.127^{***b}	.103^{***b}	.102^{***b}	.063^{***b}	.136^{***b}	.256^{***b}	.321^{***b}	.235^{***b}	.366***		
Password manager/vault	.089^{***b}	.141^{***b}	.133^{***b}	.079^{***b}	.061^{***b}	.123^{***b}	.294^{***b}	.300^{***b}	.267^{***b}	.315***	.295***	
Safe email	.087^{***b}	.139^{***b}	.105^{***b}	.068^{***b}	.042^{***b}	.100^{***b}	.266^{***b}	.330^{***b}	.235^{***b}	.361***	.323***	.380***

Notes: * $p < .05$; ** $p < .01$; *** $p < .001$

[a] value is Phi

[b] value is Cramer's V

Table 5. Differences in privacy risk perception in PET's strategy groups.

	Mean (*SD)*	F (df)	group means that differ significantly from each other[a]
Spam filters			
Uses it (1)	3.06 (.922)		
Is familiar with it (2)	3.00 (.943)		
Is unfamiliar with it (3)	3.08 (1.056)	F(2;3892) = 1.41	no differences
Firewalls			
Uses it (1)	3.06 (.928)		
Is familiar with it (2)	2.98 (.942)		
Is unfamiliar with it (3)	3.05 (.958)	F(2;3881) = 1.07	no differences
Anti spyware			
Uses it (1)	3.07 (.925)		
Is familiar with it (2)	3.02 (.928)		
Is unfamiliar with it (3)	2.96 (.977)	F(2;3869) = 2.31	no differences
Encryption			
Uses is (1)	3.12 (.914)		
Is familiar with it (2)	3.08 (.926)		
Is unfamiliar with it (3)	3.01 (.939)	F(2;3836) = 3.77*	1 and 3
Anonymous remailers			
Uses it (1)	3.28 (1.029)		
Is familiar with it (2)	3.13 (.919)		
Is unfamiliar with it (3)	3.00 (.927)	F(2;3852) = 11.59***	1 and 2, 2 and 3
Trust certificates			
Uses it (1)	3.09 (.904)		
Is familiar with it (2)	3.07 (.939)		
Is unfamiliar with it (3)	3.00 (.939)	F(2;3868) = 3.74*	1 and 3
Anonymisers			
Uses it (1)	3.30 (.977)		
Is familiar with it (2)	3.12 (.915)		
Is unfamiliar with it (3)	3.01 (.931)	F(2;3867) = 9.28***	1 and 3, 2 and 3
Cookie crunchers			
Uses it (1)	3.23 (.924)		
Is familiar with it (2)	3.06 (.917)		
Is unfamiliar with it (3)	3.00 (.934)	F(2;3867) = 15.09***	1 and 2, 1 and 3
Password manager/vault			
Uses it (1)	3.12 (.954)		
Is familiar with it (2)	3.07 (.921)		
Is unfamiliar with it (3)	3.00 (.925)	F(2;3854) = 4.46*	1 and 3
Safe email			
Uses it (1)	3.21 (.949)		
Is familiar with it (2)	3.07 (.916)		
Is unfamiliar with it (3)	3.06 (.931)	F(2;3850) = 7.13**	1 and 2, 1 and 3

Notes: * p < .05; ** p < .01; *** p < .001

[a] Tukey HSD multiple comparisons Post Hoc test was used

A series of t-tests were conducted to establish whether there was a difference in privacy risk perception between the two groups thus constructed for each PET. First the means on privacy risk perception of men were compared with that of women (see table 6). The t-test is not significant, indicating that men and women do not differ in their privacy risk perception. The (Pearson product-moment) correlation coefficient was calculated between age and privacy risk perception, showing that privacy risk perception is independent from age (r = .013, non significant; not reported in a table).

Table 6. Differences in privacy risk perception between men and women and in strategy groups.

	Mean (*SD*)	*t*(df) independent observations	Cohen's d
Men	3.05 (.914)		
Women	3.06 (.953)	*t*(4422) = -0.47	no difference
Anonymous email			
Uses it	3.13 (.912)		
Does not use it	2.95 (.948)	*t*(4069) = -6.40***	0.19
Pseudonyms			
Uses it	3.15 (.923)		
Does not use it	2.88 (.920)	*t*(3289) = -9.56***	0.29
Incorrect answers			
Uses it	3.15 (.890)		
Does not use it	2.97 (.958)	*t*(4230) = -6.45***	0.19
Encryption			
Uses it	3.11 (.914)		
Is unfamiliar with it	3.01 (.939)	*t*(1095) = 2.50*	0.11
Trust certificates			
Uses it	3.09 (.904)		
Is unfamiliar with it	3.00 (.939)	*t*(2662) = 2.59*	0.10
Password manager/vault			
Uses it	3.12 (.954)		
Is unfamiliar with it	3.00 (.925)	*t*(1566) = 2.83**	0.13
Anonymous remailers			
Uses it/is familiar with it	3.15 (.930)		
Is unfamiliar with it	3.00 (.927)	*t*(2466) = 4.51***	0.16
Anonymisers			
Uses it/is familiar with it	3.14 (.921)		
Is unfamiliar with it	3.01 (.931)	*t*(2263) = 3.90***	0.14
Cookie crunchers			
Uses it	3.23 (.924)		
Does not use it	3.02 (.928)	*t*(881) = 5.19***	0.23
Safe email			
Uses it	3.21 (.949)		
Does not use it	3.03 (.926)	*t*(439) = 3.35***	0.19

Note: * p < .05; ** p < .01; *** p < .001

Next, the privacy risk perception of the group 'anonymous email address users' was compared with that of the non-users. The t-test is significant and comparing the means leads to the conclusion that those who use anonymous email addresses, on average, have a stronger privacy risk perception than the non-users, 3.13 against 2.95. Cohen’s d[4] was calculated to establish the effect size. Effect sizes with absolute values between 0.20 and 0.50 refer to small differences, between 0.50 and 0.80 to moderate ones, and above .80 to large ones [3]. The effect size for anonymous email is .19, so there is a small difference in privacy risk perception between those who use anonymous email addresses and those who don't. The difference in privacy risk perception between those who provide incorrect answers and those who don’t is also significant, but again this is only a small difference. So, The groups who use anonymous email addresses and provide incorrect answers both have, on average, a stronger privacy risk perception than the groups who don’t use anonymous email addresses and don’t provide incorrect answers. The privacy risk perception of those who use pseudonyms is stronger than that of those who don’t use pseudonyms, 3.15 against 2.88, but with a Cohen’s d of .29 this is still a small difference. Hypothesis 2 is accepted for anonymous email addresses, pseudonyms, and incorrect answers, but an adjustment was made to the statement. People who use anonymous email addresses, pseudonyms, and incorrect answers have a somewhat stronger privacy risk perception than the people who don’t use these strategies.

The privacy risk perception of those who use encryption tools, trust certificates, or password managers is significantly stronger than those who don’t use these technologies, but these differences are very small. Also the effect sizes for anonymous remailers and anonymisers are very small, which is also apparent when looking at the means in privacy risk perception of those who are familiar with these technologies and those who are unfamiliar with them, 3,15 and 3.14 against 3.00 and 3.01. Hypothesis 2 is rejected for encryption tools, trust certificates, password managers or vaults, anonymous remailers, and anonymisers. The effect sizes for cookie crunchers and safe email are somewhat larger, but they are still small. So, The privacy risk perception of those who use cookie crunchers or safe email is, on average, stronger than that of those who don’t use these technologies, 3.23 and 3.21 against 3.02 and 3.03. Hypothesis 2 is accepted for cookie crunchers and safe email, but an adjustment was made to the statement. People who use cookie crunchers and safe email have a somewhat stronger privacy risk perception than the people who don’t use these strategies.

To establish whether the privacy risk perception of people who use advanced PETs is stronger than that of people who only use behavioral strategies or common PETs, the mean privacy risk perception of the both groups were compared. The t-test (t (df=4276) = -2.66**) shows that the mean privacy risk perception of people who use advanced PETs (3.10) is significantly different from that of people who only use behavioral strategies or common PETs (3.02), but Cohen’s d indicates that this difference is extremely small and therefore negligible (Cohen’s d = 0.08). So, we can therefore conclude that there is no difference with respect to privacy risk perception

[4] Cohen’s d for independent observations.

between those who use advanced PETs and those who only use behavioral strategies or common PETs and hypothesis 3 is therefore rejected.

5 Discussion and conclusion

The analysis shows that privacy risk perception can be measured by the seven specific privacy threats. The correlation between the concerns about the individual privacy risks were high and the seven items form a good scale. This means that the questions employed can be used as an indicator for online privacy concern. When applied to our sample, the students in our sample appear to show only moderate (3.06 on a scale from 1 to 5) privacy concern.

With respect to the privacy protection strategies, unsurprisingly, spam filters, firewalls, and anti spyware are PETs that are used by almost everyone. Of course, these PETs have become standard tools on almost every computer and internet subscription plan, which means that people have easy access to them. The other strategies, behavioral measures and more complex PETs show a different picture. Behavioral measures: using anonymous email addresses, using pseudonyms, and providing false personal data, which although available to everyone, require more deliberation and decision making in their use. One has to make conscious decisions to use a pseudonym in an interaction or to lie about one's personal data. Furthermore, one has to realize that using these strategies is an option and that using this option may serve privacy protection. Judging from experiences we have gained in focus groups and classroom settings these insights are not completely common. Nevertheless, pseudonyms are used by about 63% of our sample, anonymous email addresses by 56% and providing false data by just 45%. Given the moderate privacy concerns of our sample, and the unclear merits of adopting these behavioral strategies these seem relatively large numbers. A partial explanation for the high numbers for pseudonyms and anonymous email addresses may be found in the fact that many students have Hotmail and Yahoo addresses that are also used to register for social network services such as MSN, Hyves, MySpace, Friendster, Flickr, and YouTube. Using made-up names is common in these services, and these networks are also used by their users to experiment with their identities which often implies the use of pseudonyms [4]. The third privacy protection strategy, the more advanced PETs are largely unknown to our respondents (Dutch students) and (consequently?) hardly employed. This is interesting given the responses to other questions in our survey that suggest that our respondents do feel a need for more computer programmes to protect their privacy, as well as tighter government regulation. How does this relate to our respondents' ignorance of existing applications?

All privacy protection strategies represented in our study are more often used by men than by women, with the exception of password managers or vaults, where women are the principal users. Why are women the principal users of password managers? Do they use more online identities which make identity management more complex calling for identity management solutions? Are they better organized? Men are also more often familiar with PETs than women. Does this reflect the fact that men generally are more interested in technology and do other kinds of things online? Why are women lagging behind? These are questions we will explore in future studies using our current survey data set as a starting point.

Looking more closely at the three groups of protection strategies, the first group, the widely used PETs, shows higher correlations with each other than with the other strategies, indicating that the use of either spam filters, firewalls, or anti spyware, reinforces the use of the other two. This may be due to the fact that spam filters and spyware protection are often provided by the user's internet provider. The group of behavioral strategies also shows higher correlations with each other than with the other strategies. So, using anonymous email addresses or pseudonyms or providing incorrect answers when personal information is requested, reinforces the use of the other two strategies. Here an explanation may be that once one has an anonymous email address, say BigHunk@yahoo.com, BigHunk most likely will also be used as a pseudonym in other interactions. This does, however, not explain the correlation with lying about one's personal data. Also the third group of strategies, the more advanced PETs, shows higher correlations amongst them than with the other strategies, indicating that using one more advanced PET reinforces the use of other more advanced PETs. Is the explanation here that many of the advanced PETs still have a 'geek' connotation?

Finally, our introductory question. Does a high privacy risk perception incur the use of privacy protecting measures? The present study does not answer this question directly. But the analysis does show some interesting findings. Firstly, respondents with high privacy risk perception are likely to be aware of measures they can take themselves. This means they are aware of the risks and the measures to lower these risks. Secondly, the privacy risk perception of users versus non-users of a particular protection strategy does not differ much for most strategies. And also there is no difference in privacy risk perception of those who use advanced PETs and those who only use behavioral strategies or common PETs. These two results suggest that a high privacy perception is an insufficient motivator for people to adopt privacy protecting strategies, while knowing these exist. The only (weak) exceptions to this conclusion are, (from highest to lowest difference in privacy risk perception between users and non-users) pseudonyms, cookie crunchers, anonymous email addresses, incorrect answers, and safe email. The people who use these measures have significantly higher privacy perceptions than those who don't use them. But in general, apparently, the use of privacy protection measures (strategies) is also dependent on other factors like access to the actual measures (such as a cookie cruncher). Another explanation may be that when people have a strong privacy risk perception they adopt protection measures which subsequently (or consequently) lowers their privacy risk perception. If high privacy concerns are not sufficient to motivate people, then what does?

Revising the initial model on the basis of the findings results in figure 2.

Some final remarks. We have presented the data for our sample as a whole and have only looked at age and gender differences. In our data the ethnicity of the respondents, the topic of their study and a detailed account of their online activities are recorded. In upcoming work we will analyze the data for different subgroups. From initial studies it is clear that privacy perceptions of various (ethnic) subgroups differs significantly. Whether this has an effect on their adoption of privacy protecting measures is an interesting question.

A second remark concerns our sample. We have surveyed a relatively large sample of university and colleges of higher education students in the Netherlands. The sample therefore is certainly not representative for the entire Dutch population, let alone for Europe. The privacy risk perception and the adoption of privacy protection strategies within a student population may significantly differ from the population as a whole. Students interact a lot with their peers and they behave in a similar way online. How does this compare to 30-40 year olds, or even 50-60 year olds? In our view people should be aware of the privacy risks they run in the online world. They also have to take measures within their reach to protect their own privacy. This desperately calls for a better understanding of what makes people adopt protection strategies.

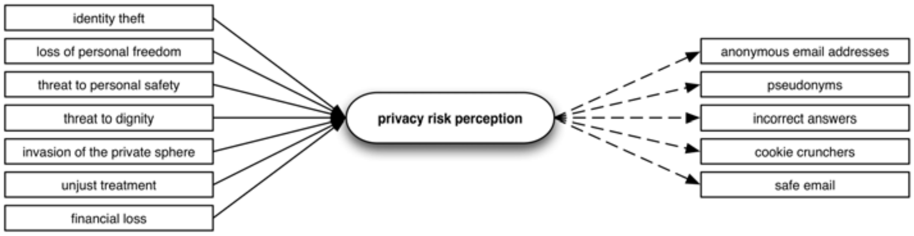

Fig. 2. Privacy risk perception and strategies (dotted lines reflect weak effects).

6 Acknowledgement

The work reported in this paper was supported by the IST PRIME project. The PRIME project receives research funding from the Community's 6th Framework Program (Contract No. 507591) and the Swiss Federal Office for Education and Science.

7 References

1. Buchanan, T., Paine, C., Joinson, A.N., & Reips, U. (2007). Development of Measures of Online Privacy Concern and Protection for Use on the Internet. *Journal of the American Society for Information Sciences and Technology, 58,* 157-165.
2. Burgoon, J.K., Parrott, R., LePoire, B.A., Kelley, D.L., Walther, J.B., & Perry, D. (2007). Maintaining and restoring privacy through communication in different types of relationships. *Journal of Social and Personal Relationships, 6,* 131-158.
3. Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

4. Donath, J. & boyd, d. (2004), Public display of connection, *BT Technology Journal*, *Vol 22, 4,* 71-82.
5. Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1998). *Applied statistics for the behavioural sciences* (4th ed). Houghton Mifflin Company: Boston.
6. Introna, L.D. & Pouloudi, A. (1999). Privacy in the Information Age: Stakeholders, Interests and Values. *Journal of Business Ethics, 22,* 27-38.
7. Pallant, J. (2001). *SPSS Survival Manual. A step by step guide to data analysis using SPSS for Windows (Versions 10 and 11).* Open University Press: Buckingham.
8. Lyon, D. (2004). Globalizing Surveillance. Comparative and Sociological Perspectives *International Sociology, 19,* 135-149.
9. Marx, G.T. (1994). New Telecommunications Technologies And Emergent Norms. In G. Platt and C. Gordon (Eds.). *Self, Collective Behavior and Society. Essays in honour of Ralph Turner.* JAI.
10. Marx, G.T. (2003). A tack in the Shoe: Neutralizing and Resisting the New Surveillance. *Journal of Social Issues, 59,* 369-390.
11. Marx, G.T. (2006). Varieties of personal Information as Influences on Attitudes Toward Surveillance. In K. Haggerty and R. Ericson (Eds.). *The New politics of Surveillance and Visibility.* Toronto: University of Toronto Press.
12. Seničar, V., Jerman-Blažič, B., & Klobučar, T. (2003). Privacy-Enhancing Technologies – approaches and development. *Computer Standards & Interfaces, 25,* 147-158.
13. Solove, D.J. (2003). Identity Theft, Privacy, and the Architecture of Vulnerability. *Hastings Law Journal, 54*, 1227.
14. Stalder, F. (2002). The Failure of Privacy Enhacing Technologies (PETs) and the Voiding of Privacy. *Sociological Research Online, vol. 7, no. 2.* Accessible at http://www. socresonline.org.uk/7/2/stalder.html.
15. Westin, A. (1967). *Privacy and freedom.* New York: Atheneum.

8 Internet sources

http://www.epic.org/privacy/survey