

Let's agree to differ: varying interpretations of online privacy policies

Steve McRobb

Centre for Computing and Social Responsibility, De Montfort University, Leicester, UK

Email: smcrobbs@dmu.ac.uk

During the period of growth of e-commerce, e-business and online life in general, trust has been identified by a number of authors as a key factor, the absence of which can act as a powerful disincentive to an individual's engagement in a transaction. This has encouraged a great deal of research into the various facets of trust in an online environment, both theoretical and empirical. One of the many recommendations for business practice that have emerged from this research is the suggestion that online businesses should publish on their website a privacy policy that explains clearly the use that will be made of any personal information collected on the site, the third parties to whom it may be disclosed, and the circumstances under which disclosure may occur. A number of surveys have been conducted that highlight the rather patchy adoption of this recommendation in various countries. We now know, for example, that by no means all online organisations publish an online privacy policy, and that many of those that do exist display a range of serious shortcomings, including poor visibility on the site, incomplete coverage of the main issues of concern and poor readability. However, previous discussion of privacy policies has tended to assume that any particular policy can provide value to its readers by informing them of the privacy practices of its host organisation, and thus also to its publishers through encouraging customers and clients to trust them more than they otherwise would. This assumption is expected to be valid where the policy meets certain criteria, which are either established on the basis of theoretical considerations, or are derived from a kind of best-in-breed comparative exercise. This paper seeks to address the question how far privacy policies can ever achieve the goal of providing clear information to website visitors about the privacy practices of an organisation. It reports on an empirical study that was conducted between November 2005 and April 2006 using two groups of University students as subjects. The subjects were asked to read three privacy policies, selected in advance by the author, and to complete a short questionnaire on what the subject thought each policy had to say about certain key privacy issues. The results reveal that there is surprisingly little agreement about what a policy actually means. This has significant implications both for policy writers and their managers, and also for those who are considering entering into a transaction with the host website. There is a need for further research to investigate this question in more detail, but it is clear from these findings that we know less than we thought we did about the ways in which people interpret the notices that they read on websites.

Keywords: privacy policy, privacy online, personal data

INTRODUCTION

Since the mid to late 1990s, it has been suggested by many that website visitors are concerned about the possible abuse of personal data that they supply online, and that this acts as a disincentive to full engagement in e-commerce and other comparable sorts of online activity. In essence, we need to feel

that we can have some trust in the organisations with which we transact (e.g. Belanger *et al*, 2002; Ranganathan & Ganapathy, 2002; Srinivasan, 2004). Among the various possible indicators of trustworthiness – organisation behaviour, site design, site content, membership of a third-party trust network and so on – some have suggested that publication of a privacy policy is an important element. This is

taken to reassure those who transact with a website that appropriate protection measures will be applied to the transmission, subsequent use and storage of any personal information that they provide during their transactions (e.g. Udo, 2001; Sheehan, 2002; UMIST, 2002; Gauzente, 2004).

Since 2000, the author has conducted an ongoing longitudinal study that has sampled at two-year intervals a set of over 100 policies and analysed their content. Results that have been published to date (McRobb & Rogerson, 2004a; McRobb & Rogerson, 2004b; McRobb & Rogerson, 2005; McRobb & Stahl, 2006) identify what individual policies cover, what they do not cover, how their different characteristics may conflict with each other, how they have changed over time, or not, and also how much some have changed. This work, both in the sense of the survey itself, and also the analysis of the rich dataset collected so far, is still in progress. It has also raised a number of questions that appear to merit closer investigation.

For example, given that there is often an apparent conflict between the meta-characteristics (prominence of display, readability, etc.) of a policy and its informational content, it would be interesting to know the extent to which policies are practised – in other words, do organisations really walk their talk? It would also be interesting to know more about the extent to which those who transact with a website are really worried about the privacy and possible abuse of personal data that they provide. Berendt *et al.* (2005) studied this and observed that their subjects' actual online behaviour was at odds with stated privacy preferences. However, this phenomenon as a whole still remains far from clear.

This paper moves in a different but related direction. This is prompted in part by the work of Milne and Culnan (2002), who noted that there was little understanding of how consumers actually used privacy notices. Another influence is the work of Moores (2005), who found that few consumers either recognised or understood the significance of privacy seals (also commonly known as trustmarks). In addition, McRobb and Rogerson (2004b) found that different readers disagreed about the practical import of a small sample of privacy policies.

The paper reports on a study that seeks to investigate how readers understand the meaning of a small set of privacy policies. It is based on a smaller-scale exercise that was conducted alongside the detailed analysis of the first bi-annual survey reported in McRobb and Rogerson (2004b). That study involved the interpretation by the author of 113 policies, in particular to determine what they

had to say on a number of privacy-related issues. Policies were analysed to determine whether it was clear that personal data would, or would not, be collected, what data was collected, whether there was an opt-out from providing data, how such data would be used, whether it would be revealed to third parties, and so on.

Since all the analysis was initially based on the author's own interpretation, it was apparent that there could be legitimate doubt regarding the validity of the particular views that were reached. To counter this concern, a small validation exercise was designed to test the author's interpretations and thus give some assurance that the study's overall conclusions were valid. Two colleagues who had not been previously involved with the project were asked to replicate the coding for a sub-set of the main sample (one was a Research Fellow at the Centre for Computing and Social Responsibility and the other an undergraduate student who was at that time employed by the Centre on a sandwich placement year). They were instructed in the task by the author and asked to follow the original procedure as closely as possible. Ten policies were chosen as reasonably representative of the range in the sample as a whole, taking into account the geographical spread of the organisations, the types of organisation and the characteristics of the policies themselves – in particular, their readability, length and coverage.

On many issues, there was close agreement; for example, regarding the overall characteristics of a policy, such as language and structure, and on whether personal data was said to be collected. However, on some points there was very little agreement. This applied particularly – and naturally enough – to aspects where the interpretation depended on a subjective interpretation of an imprecise wording. For example, the uses to which personal data is put, the reasons for which it is disclosed, or how the visitor consents to its collection or use, were all occasions for disagreement. Furthermore, the disagreements appeared to be unrelated either to the subjective clarity of the policy or to objective criteria such as word-count or readability index (McRobb & Rogerson, 2004b).

The general conclusion from this part of the earlier study was that little can be taken for granted about how a privacy policy will be interpreted, even when the reader is highly educated and unusually well-informed about privacy issues.

In the next section, we will examine some of the methodological issues that were considered and how these were addressed in the design of the follow-up study. The findings of the empirical phase

of the research will then be presented, analysed and discussed. Conclusions will be drawn from this discussion, and recommendations made to organisations that currently publish, or plan in the future to publish, a privacy policy, and also to those who may carry out online transactions with those organisations. Finally, some suggestions will be made for possible future research directions.

METHODOLOGICAL CONSIDERATIONS

The main aim of the research presented here was to examine the subjective interpretations that different individuals can make of a section of text, specifically of an online privacy policy. The topic is potentially complex, and can be investigated from a number of different perspectives using various research techniques. However, the first, and most important, question that must be answered is essentially factual and requires a simple quantitative answer (or, rather, a quantitative description of a set of Boolean answers). In effect: do our subjects agree on the factual content of a given privacy policy? If they do, and if the sample is large enough for the results to be capable of meaningful generalisation, we would then have arrived at a probably terminal answer for this research direction, and there would be little point to conducting any further investigation. If, on the other hand, they do not agree, then further questions would come into focus. For example, to what extent do the subjects disagree? We could perhaps also ask whether the level or the particular loci of disagreement have any statistical significance. Some of these further questions are more qualitative than quantitative in nature. For

example, we might hypothesise that the disagreement results from some ambiguity in the text. If this is the case, it will be interesting to examine the text itself. Supposing that we can discern what it is in the text that generates ambiguity, then we may be able to take a step or two towards meeting Milne and Culnan's (2002) call for a standard vocabulary (perhaps, we might add, also a standard grammar) of information practices, which we might expect would be principally deployed in an organisation's privacy statement.

If no ambiguity is found to lie in the text, we may then postulate that the origin of the differences in interpretation is not in the text but in the reader. If this is the case, then further investigation should be focused on the characteristics of our subjects. Such a direction has already been suggested by Graeff and Harmon (2002), for example, who concluded that demographic factors affect the response to a privacy statement, while Sheehan (2002) found that individual personality attributes play an important role in trust decisions.

Finally, the sources of the difference may reside neither in the text nor in its reader, but in the interaction between the two.

Since the selection of research method, and, indeed, research tool, depended closely on the explanation under investigation, it appeared sensible to pursue the simplest of the three general categories of explanation first. For this first stage of enquiry, then, a straightforward questionnaire was adopted. This was expected, first of all, to determine whether our subjects interpreted the policies in the same way. A questionnaire would also be able to highlight areas where there was disagreement and measure the extent of any such disagree-

Table 1 Summary of policy characteristics

Organisation	Flesch index	Word count	Analysis of coverage in previous surveys
Federal Express	37.7	1133	Found to have gaps. Lacked coverage of opt-out from collection or disclosure, how data is collected, procedures for consent, checking and amending data, and means of protection. Also acknowledged that all but one of the anticipated reasons for the collection or disclosure of data applied.
Symantec Store	35.4	1527	Found to be exemplary. Coverage of data collection, disclosure, consent, opt-out, protection procedures, amending stored data, etc. However, this policy also acknowledged that personal data was used and disclosed for every single anticipated reason.
Zapspot	36.8	1071	One of the relatively few to adopt an informal tone, but also one of the thinner in coverage. Acknowledged collection and disclosure, but little indication of what data are collected, how, or about opt-outs, consent, amendment and protection.
2004 means:	36.0	1345 ($\sigma = 1158$)	Mean readability and word-count scores are from the 2004 survey, the most recent figures available at the time of writing.

ment. Furthermore, since it could be designed to discriminate between different aspects of the policies, this approach could be combined to a certain extent with textual analysis of the sections that were found to be problematic for our subjects.

The questionnaire was closely based on the form used to code the data collected in the bi-annual surveys. It was edited to remove questions that were not directly relevant to the present research. It was also converted to an electronic form, partly to allow the data recorded by the subjects to be more closely constrained, and partly to make it easier to deploy. The final questionnaire is shown at Appendix A.

The policies were selected from the original validation subset for the 2000 survey. This was partly a matter of convenience, since it was necessary for a baseline interpretation to be available for each policy used. The number used in this study was restricted because the primary purpose was not to compare the policies, but rather to investigate how they were understood. In this context, three seemed to be a sufficient number that was also appropriate in terms of the load on each subject. If the subjects were asked to read too much, this would tend to reduce cooperation, and thus the number of responses that could be collected. Another reason for limiting the number was that it was planned to use the author's students as subjects, since they were thought to be a highly suitable population, and to administer the survey during class time, since this was expected to maximise the number of responses. It was also thought desirable to exclude longer policies, in order to minimise the risk that too many subjects would fail to complete the task in the available time. The policies that were finally chosen were all close to the mean in both length and readability, but had fairly different coverage characteristics, as shown in Table 1. In addition, they were known to have produced some disagreement in the validation exercise, and so could reasonably be expected to provide interesting material for this study.

All subjects were students enrolled on a University computing course. One group was enrolled at undergraduate Higher Diploma/Bachelor's Degree level, the other at postgraduate Masters level. Both groups were studying a module on e-business strategy and systems. The author teaches on both modules, and, in both cases, the survey was included in the class schedule partly to meet a legitimate pedagogic aim of the module. As a result, access to the subjects was good and a high level of cooperation was expected. Given the subjects' educational level and background, they were

expected to have some grasp of the issues covered by a privacy policy, and also a reasonable level of literacy. The readability (or otherwise) of the policies was thus not expected to cause difficulty for these subjects. The two groups were surveyed separately a few months apart.

Each subject was briefed both orally and in writing (a copy of the written briefing can be found at Appendix B) asked to read the three policies in situ on the host website, to answer the questions on the survey form, and then to email the completed form to the author.

The returns were then printed and each subject was assigned a unique code based on the month in which that survey was undertaken and a simple numeric index (while the data collected was unlikely to be sensitive for any of the subjects, it was felt important in such a study to ensure that the promise of anonymity was fulfilled in practice). The results were then entered into a spreadsheet for analysis.

RESULTS

For the November 2005 survey, a total of 18 subjects returned a total of 45 questionnaires. In fact, more students than this were enrolled on the module, but attendance in class was never 100%. Only 12 subjects completed a form for all three policies, but incomplete comparison for a single subject did not appear to be a sufficient reason to exclude the other submissions.

A number of returns showed some inconsistencies between answers, but most of these were very minor and could be attributed to either haste on the part of the subject or slight ambiguities in the instructions given on the survey form. In cases where it appeared possible to infer with reasonable confidence from a subject's other responses their intent with regard to missing or inconsistent data, the author corrected or inserted responses as necessary. For example, in the Symantecstore return for subject 5/11/18 (subject number 18 in the November, 2005 survey), 'Not mentioned' was checked against Q6 ("Is personal data passed to third parties?"). On the same form, 'No' was checked against Q7 ("Can you opt-out from disclosure?"), a response which does not make sense if the policy has said nothing about disclosure in the first place. However, the instructions on the form did not explicitly tell subjects to skip Q7 if the answer to Q6 is not 'Yes'. Since this return also had no checks against any of the possible reasons for disclosure listed for Q9, it seemed safe to assume

Table 2 Summary of results for Federal Express

Question		Nov 05 survey			Apr 06 survey			Author		
		Yes	No	n/m	Yes	No	n/m	Yes	No	n/m
1	Collected?	14	0	0	9	0	0	✓		
2	Opt out?	8	4	2	1	4	4			✓
3	How?	11	3		7	2			✓	
4	What?	13	1		9	0		✓		
5a	Why - contact	9			5			✓		
5b	Why - service	13			8			✓		
5c	Why - evaluation	7			8			✓		
5d	Why - sell adverts	6			3					
5e	Why - targeted marketing	10			9			✓		
5f	Why - other	0			1			✓		
6	Disclose?	10	4	0	7	2	0	✓		
7	Opt out?	4	4	2	1	2	4			✓
8	Why disclosed?	10	2		6	1		✓		
9a	Disclose - service	6			2			✓		
9b	Disclose - tailor	3			1			✓		
9c	Disclose - marketing	8			1			✓		
9d	Disclose - legal	8			5			✓		
9e	Disclose - other	1			1			✓		
10	Consent?	5	9		3	6			✓	
11	Amend?	9	5		1	8			✓	
12	Protection?	8	6		4	5			✓	

that the response to Q7 could be disregarded.

A total of seven other returns showed comparable discrepancies, where the author felt justifying in making a minor correction to the recorded data.

Two forms showed more serious inconsistencies where the author did not feel that it was possible to discern the intent of the subject, and these were discarded from the analysis. The results presented below are therefore based on 14 returns for Federal Express, 16 for Symantecstore and 13 for Zapspot.

A second survey was conducted in April 2006, and this began with a population of 11 students. Two forms could not be analysed since the subjects had failed to indicate which organisation was described. Furthermore, in both cases this was the only submission from that subject. These two returns were discarded. The remaining nine subjects all responded for all three policies. One subject submitted two separate returns for the same policy, but since one was incomplete and the other was fully completed, the former was discarded. There were also a few minor inconsistencies, of the sort described above, which the author corrected in a similar manner.

The overall results are most readily presented in tabular form, and Tables 2 to 4 show a summary of the responses for each organisation in turn. The format of the tables is as follows. Each shows,

reading across the page, the questions and then the number of responses in each category for (i) the November 2005 survey, (ii) the April 2006 survey, and (iii) the author's coding for that policy. In each case, the number of responses that correspond to the author's coding is shown in bold italic.

Note that, for some questions (such as Q1: "Is personal data collected?"), three possible answers could be given: "Yes," "No" and "Not mentioned." For other questions (such as Q4: "What data is collected – is this described?"), only "Yes" and "No" could be given. For the two "Why...?" questions (Q5 and Q9), subjects could tick as many answers as they thought were applicable. Thus the choice for each reason listed was either to tick it or to leave it blank. The greyed-out boxes in Tables 2, 3 and 4 reflect these differences in the possible answers to each question.

For Federal Express, there was complete unanimity regarding Q1 – that personal data was said to be collected – but varying degrees of agreement on the other questions. Only one subject dissented for Q4 (what data is collected) but there were varying levels of disagreement for Q2 (opt out from providing personal data), Q3 (how is the data collected), Q6 (disclosure to third parties), Q7 (opt out from disclosure), Q10 (how consent is given), Q11 (how to amend personal details) and Q12 (methods

Table 3 Summary of results for Symantecstore

Question		Nov 05 survey			Apr 06 survey			Author		
		Yes	No	n/m	Yes	No	n/m	Yes	No	n/m
1	Collected?	16	0	0	9	0	0	✓		
2	Opt out?	12	3	1	5	3	1	✓		
3	How?	16	1		9	0		✓		
4	What?	16	1		9	0		✓		
5a	Why - contact	7			4					
5b	Why - service	13			9			✓		
5c	Why - evaluation	7			2			✓		
5d	Why - sell adverts	9			6					
5e	Why - targeted marketing	8			5			✓		
5f	Why - other	2			0			✓		
6	Disclose?	14	1	1	9	0	0	✓		
7	Opt out?	13	1	0	4	4	1	✓		
8	Why disclosed?	12	2		7	2		✓		
9a	Disclose - service	10			5			✓		
9b	Disclose - tailor	1			0					
9c	Disclose - marketing	8			5			✓		
9d	Disclose - legal	4			5			✓		
9e	Disclose - other	0			0					
10	Consent?	8	8		5	4		✓		
11	Amend?	11	5		7	2		✓		
12	Protection?	13	3		8	1		✓		

* Excludes one subject where the answers were internally too inconsistent to interpret

used to protect data).

For Symantecstore, too, there was unanimity on Q1, and only one dissenter for Q4. Similarly, there were varying levels of disagreement for Q2, Q7, Q8, Q10, Q11 and Q12. But there was noticeably less disagreement than was the case for Federal Express regarding Q4 and Q6.

For Zapspot, again, there was complete unanimity regarding Q1 and a single dissenter for Q4, followed by varying levels of disagreement for the remaining questions.

DISCUSSION

On inspecting the results shown in Tables 2, 3 and 4 above it is immediately clear that, while there is a broad consensus on most points, there are also aspects where there is little agreement at all. This repeats, and thus validates, an earlier finding reported in McRobb and Rogerson (2004b). The policy aspects that produced the most disagreement are shown in Table 5, in each case using the author's own answer as the standard (this may not have any ultimate validity, as we shall see further below, but for the moment it will serve just as well as any other reference point). Note that, in the case of the mul-

ti-ple-answer questions (Qs 5a–5f and 9a–9e), responses have only been counted either if a subject agreed with a tick placed by the author, or if a subject placed a tick where the author did not; in other words, blank responses, which indicate that the subject thought a reason did not apply, are not included in the summary.

This view of the data helps to bring certain factors to greater prominence. The first of these, and perhaps of greatest concern to the authors of those policies, is that, for Federal Express and Zapspot, the vast majority of subjects disagree with the author on what the policy says about their ability to opt out from providing personal data. There is less disagreement regarding the Symantecstore policy on this point, but even here the subjects were far from unanimous. Not surprisingly, this pattern is replicated almost exactly regarding opt out from disclosure of personal data to third parties. There is more agreement on whether the policies warn of disclosure, but even this is still far from unanimous in the case of Federal Express. For no policy was there close agreement on Q10 (“Is there a description of how a customer consents...?”), on Q11 (“Is there a description of how a customer can check and amend...?”) or on Q12 (“Is there a description of how personal data will be protected?”, although

Table 4 Summary of results for Zapspot

Question	Nov 05 survey			Apr 06 survey			Author			
	Yes	No	n/m	Yes	No	n/m	Yes	No	n/m	
1	Collected?	13	0	0	9	0	0	✓		
2	Opt out?	6	3	5	5	2	2			✓
3	How?	11	2		8	1			✓	
4	What?	12	1		9	0			✓	
5a	Why - contact	5			4					
5b	Why - service	11			8			✓		
5c	Why - evaluation	5			1					
5d	Why - sell adverts	6			5					
5e	Why - targeted marketing	8			7			✓		
5f	Why - other	0			1					
6	Disclose?	12	1	1	9	0	0	✓		
7	Opt out?	4	6	2	1	4	4			✓
8	Why disclosed?	11	2		8	1		✓		
9a	Disclose - service	9			4					
9b	Disclose - tailor	5			3					
9c	Disclose - marketing	8			6					
9d	Disclose - legal	4			5			✓		
9e	Disclose - other	1			0			✓		
10	Consent?	4	10		4	5			✓	
11	Amend?	3	10		0	9			✓	
12	Protection?	8	4		4	5			✓	

* Excludes one subject where the answers were internally too inconsistent to interpret

the least disagreement was for Q12 related to the Zapspot policy, with only three dissenters out of 22. However, since in this case most subjects agreed with the author that this policy does not indicate how personal data will be protected, this is unlikely to be reassuring to the policy's authors.

The manifestation of such high levels of disagreement on issues that many researchers (and presumably also most business managers and consumer representatives) see as important – opt outs from provision or disclosure of personal information, the reasons given for its collection or disclosure, and the stated procedures for its management – is significant in itself. Meanwhile, the differing responses to a single question when compared across the three policies suggest interesting lines of investigation that could yield useful recommendations for the authors of privacy policies. For example, on this data the Symantecstore policy appears to be by far the clearest of the three regarding a subject's opt out from providing personal data (see Q2 in Table 5) and also regarding how data is actually collected (Q3 in Table 5). The relevant passages in this policy are as follows:

“We request personal information when you place an order, register a product,

request a service, answer a survey, enter a contest, correspond with us or engage in certain other activities on our website. For example:

“1. If you submit an order on-line with us, we will ask you for information necessary to complete the transaction - such as your name, address, and credit card number.

“2. If you request technical support online, we will ask you for information necessary to complete the transaction – such as your name address and information about your computer hardware, software and the nature of the problem you are experiencing. “You make the decision whether to proceed with any activity that requests personal information. If you do not wish to provide the requested information, however, you may not be able to complete the transaction.”

(Symantecstore, 2006)

On the other hand, the Zapspot policy appears to be the least clear regarding what data is collected (Q4 in Table 5). The following phrase is the only one in the policy that appeared to the author to be relevant to this question. It appears in a section

Table 5 Summary of agreement / disagreement

Question		Federal Express		Symantecstore		Zapspot	
		Agree	Dis-agree	Agree	Dis-agree	Agree	Dis-agree
1	Collected?	23	0	25	0	22	0
2	Opt out?	6	17	17	8	7	16
3	How?	5	18	25	1	3	19
4	What?	22	1	25	1	1	21
5a	Why - contact	14	-	-	11	-	-
5b	Why - service	21	-	22	-	19	-
5c	Why - evaluation	15	-	9	-	-	6
5d	Why - sell adverts	-	9	-	15	-	11
5e	Why - targeted marketing	19	-	13	-	14	-
5f	Why - other	1	-	2	-	-	1
6	Disclose?	17	6	23	2	21	2
7	Opt out?	6	11	17	6	6	15
8	Why disclosed?	16	3	19	4	19	-
9a	Disclose - service	8	-	15	-	-	13
9b	Disclose - tailor	4	-	-	1	-	8
9c	Disclose - marketing	9	-	13	-	-	14
9d	Disclose - legal	13	-	9	-	9	-
9e	Disclose - other	2	-	-	-	1	-
10	Consent?	15	8	13	12	15	8
11	Amend?	13	10	18	7	19	3
12	Protection?	11	12	21	4	9	12

headed “WHAT INFORMATION WE COLLECT AND WHY.”

“...information you provide to us directly, like your name, e-mail, or address.”
(Zapspot, 2006)

The author chose to interpret the construction “information like...” as signifying that this was at best an incomplete description of the data collected, and therefore ticked “No” to this question. However, the great majority of survey subjects disagreed. It appears likely that they were satisfied with this description of the data collected; however, even at the time of writing, the author remains unsatisfied on this point.

If we consider how subjects viewed the policy statements on reasons for collecting personal information, in general the majority agreed with the author (for example, Q5b and Q5e), but there was also some striking disagreement (especially Q5d). Regarding the stated reasons for disclosure, there was rather less agreement (reflected by fewer subjects ticking the same boxes as the author) and also higher levels of disagreement (see especially Q9a, Q9b and Q9c for Zapspot).

All of the foregoing discussion is based to some extent on the implied premise that the author’s interpretation of the policies has some privileged

validity, even in cases where the majority of subjects in this study disagree with that interpretation. There might be grounds for arguing that this premise is legitimate. For example, we could argue that the author has been able because he has more time, or because he is educated to a higher level, or because he has more experience in interpreting privacy policies to take greater care over the interpretation. Or we could argue that he has been motivated to do so (for example, due to his close personal involvement in the research). But even if either of these arguments had force, we should still ask whether it is a reasonable expectation of a privacy policy that it should be read only by highly educated or highly motivated individuals who also have ample time to complete the task with all possible care and diligence. Clearly, such an expectation would be entirely unreasonable. But let us next examine the relative validity of the author’s interpretations in the only sensible way – by examining the texts themselves, in relation to some selected points of interpretation. There is not the space here to be comprehensive in this discussion of policy content, but a few examples should suffice to allow the drawing of some general conclusions.

As noted above (Q2 in Table 2), the author found no mention in the Federal Express policy of an opt out from the collection of personal data; 6 survey

subjects agreed, 9 thought the policy said that there was an opt out, and 8 thought that the policy said there was no opt out. The policy includes the following three statements, which appear to relate to choice about the provision of personal data, but the author was unable to find any others (the reader is invited to visit the policy and to make this judgment for him/herself).

1. "Certain areas of fedex.com... require registration or a password for access. Information obtained in these areas may also be used in accordance with agreements governing access to and use of the specific areas..."
2. "...visitors to fedex.com may choose to register for a fedex.com Login user ID and password to simplify access to some interactive features on fedex.com."
3. "Several areas of fedex.com ask for personal information..."
(Fedex, 2006)

Extract 1 appears to convey with reasonable clarity that certain areas of the website (identified in more detail in the full text) require a visitor to provide personal data through a registration process, that the visitor can expect at the time of registration to view the terms on which each service is offered, and can presumably then decide whether or not to proceed with the registration, although the terms of these separate agreements are not made clear within the policy itself. It is not clear to what extent extract 2 overlaps with extract 1 – i.e. whether they refer to the same areas of the website or to separate areas. Nor is it clear whether the registration referred to in extract 2 is the same as that (or those) mentioned in extract 1. Nor is there any explanation in the policy of what loss of functionality or ease of use a visitor will suffer if they choose not to register. There may be, perhaps, a number of different registrations for distinct services and interactive enhancements of these services. Extract 3 acknowledges that some areas of the site make some sort of request for personal information, but it is not clear how such requests will be presented, nor what alternatives are available to a visitor who chooses not to meet the request. It is also left as an implication, rather than a clear statement, that no other part of the site collects any personal data. Finally, the policy refers at various points to 'visitors,' to 'users,' to 'registered users' and to 'customers' without either defining them or making it explicit whether their use has a semantic significance. The different terms may refer to different sorts of engagement with the company – which would have

an effect on how the policy statements should be interpreted – but we cannot infer this from the policy alone. For all of these reasons, the author felt that, while the policy text may convey some implications on this point, it does not state with any clarity whether or not a customer can opt out from the collection of personal data.

Turning now to the Zapspot policy, the author again found no mention of an opt out from the disclosure of personal data (Q7 in Table 4); 6 subjects agreed, while 5 subjects thought that the policy said there was an opt out and 10 thought that it said there was not an opt out. This policy appears to have the following to say on the subject of disclosure in general, and whether customers have an opt out from it.

1. "ZapSpot will not sell, trade, or rent to others any... Information that identifies you as an individual."
2. "In order for us to provide you with the best service possible, we may... submit certain... Information to... third party agents to whom we outsource certain parts of our business processes, or who perform certain services for us."
3. "ZapSpot reserves the right to release... Information when we believe that such release is necessary to comply with any applicable law or to protect the rights, property, interests, or safety of ZapSpot, our customers or others."
4. "ZapSpot may release Requested, Calculated, and Roaming Information when we, in good faith, believe that such release is reasonably necessary to comply with the law or protect the rights, property or safety of ZapSpot, our customers, or others."
(Zapspot, 2006)

Extract 1, on first glance, might seem to state plainly that no personal data will be disclosed to any third party under any circumstances. However, when taken in conjunction with the other extracts quoted above, it becomes apparent that this really applies only to disclosure for the purposes mentioned, that is for purely commercial reasons. The policy goes on to acknowledge disclosure for a number of other purposes, which are identified more or less clearly in extracts 2, 3 and 4. Only two of 23 subjects disagreed with the author on this (Q6 in Table 4), so it appears that most were not misled. However, regarding a subject's ability to opt out from disclosure, it is hard to construe anything in these extracts as directly stating whether or not a subject can opt out from disclosure. There is,

certainly, a strong implication in extracts 3 and 4 that no opt out can be possible, since the stated reasons (compliance with law and protection of others' rights, safety, etc.) can reasonably be taken as overriding any legitimate interest that a subject might have in protecting their privacy. It is therefore not surprising that a majority of the subjects decided that the policy said no opt out was possible. But, still, the policy does not actually say so; it is left as an implication. What is truly intriguing is that 5 subjects thought the policy said that it was possible to opt out from disclosure – only one less than agreed with the author (that is, they thought that it was not mentioned).

The Symantecstore policy, which the author found to be one of the most informative and also one of the clearest in the larger sample (not just in the subset used presently under discussion), also produced generally higher levels of agreement among our subjects. Yet there were still one or two places where high levels of disagreement were apparent. The most prominent of these was Q10: "Is there a description of how a customer consents...?" Almost half of the subjects who responded on this – 12 in 25 – thought that the policy did not give such a description. The author, together with the narrowest possible majority of subjects, believed that it did. We have already noted above that one section of this policy states:

"You make the decision whether to proceed with any activity that requests personal information. If you do not wish to provide the requested information, however, you may not be able to complete the transaction."

(Symantecstore, 2006)

Another relevant passage appears under the heading of "choice/opt-out." This addresses several issues, including Q2, Q7, Q10 and Q11.

"Symantec provides a number of ways for you to 'opt-out' of receiving additional information from us or having us provide your personal information to our partners. We may offer you these choices at the time you give us your information.

"If you are on a Symantec web page that requires personal information but does not offer communication choices, or would like to change your preferences for information that you have previously provided, please complete the customer profile form at [link] or use one of the other com-

munication channels at the bottom of this statement.

"If you are registering a version of a product that does not provide you the capability to indicate your communication preferences, please complete the customer profile form at [link] or use one of the other communication channels at the bottom of this statement.

"If you do not wish to continue receiving our e-mail newsletters or bulletins you can opt-out of receiving these communications by replying to the unsubscribe [link] in the subject line in the email."

(Symantecstore, 2006)

Taken together, these passages seemed to the author to provide a near-exemplary description of the process by which a visitor gives or withholds consent to the provision of personal data. But not all survey subjects agreed with this.

CONCLUSIONS AND RECOMMENDATIONS

Several conclusions can be drawn from the foregoing discussion. First, a number of privacy policy statements that the author found to be among the clearest and most informative in the larger survey were not necessarily regarded in the same way by a sizeable minority, even a majority, of subjects in this study. Equally, a number of privacy policy statements that the author found to be unclear or uninformative, were apparently seen as adequate by a sizeable minority, and sometimes a majority, of subjects in the study. The broad implication is that neither the organisations that publish privacy policies nor the individuals who read them should be confident that they are being clearly understood in the way that they are meant to be understood. Authors of privacy policies should take note of this finding, since it has serious implications for the effort that is expended on writing and revising privacy policies in the confident expectation that readers will understand what they say. The corollary is that those who are considering surrendering personal data on a transactional website should also take care in how they interpret the site's privacy policy.

It is clear that further work needs to be done in this area. We can speculate on why so many subjects disagreed with the author on the three examples discussed in detail in this paper. The reader may perhaps feel that the explanation is simply that

the author is too cautious. Yet a privacy policy, while not exactly the same thing as a legal statement of terms of use for the website that carries it, nevertheless has something of the same force, not only in terms of how it will be perceived by its readers, but also, very probably, in terms of its import for any subsequent legal action resulting from the use or abuse of personal information. Perhaps the disagreement results mainly from the author's disposition to disbelieve that a privacy policy really means what it merely implies; it may be that many of the subjects have a more trusting attitude. It might also be explained by the relatively limited time in which the subjects were asked to make their judgement. To determine this point with any certainty, it would be necessary to follow up the original investigation, perhaps using interviews to examine in greater detail the attitudes, perceptions and reasoning behind the subjects' responses. That must be left, however, for a possible later study. For the moment, we must be content with raising these as further questions, to which we do not at present know the answers.

ACKNOWLEDGEMENTS

The author is grateful to all the students on HND/BSc (Hons) Business Information Technology, MSc Computing, MSc Information Technology and MSc Information Systems Management who helped with the surveys. He is also grateful to his colleagues Dr. Bernd Carsten Stahl and Prof. Simon Rogerson for their encouragement to complete this study.

REFERENCES

- Belanger, F., Hiller, J.S. and Smith, W.J. (2002) Trustworthiness in electronic commerce: the role of privacy, security and site attributes, *Journal of Strategic Information Systems*, 11: 245–270
- Berendt, B., Günther, O. and Spiekermann, S. (2005) Privacy in E-Commerce: Stated Preferences vs. Actual Behaviour, *Communications of the ACM*, 48(4): 101–106
- Fedex.com (2006) Privacy Policy for fedex.com, online at: <http://www.fedex.com/us/privacypolicy.html>
- Gauzente, C. (2004) Web Merchants' Privacy And Security Statements: How Reassuring Are They For Consumers? A Two-Sided Approach, *Journal of Electronic Commerce Research*, 5(3): 181–198
- Graeff, T.R. and Harmon, S. (2002) Collecting and using personal data: consumers' awareness and concerns, *Journal of Consumer Marketing*, 19(4): 302–318
- McRobb, S.R. and Rogerson, S. (2004) Are They Really Listening? an investigation into published online privacy policies" *Information Technology and People*, 17
- McRobb, S.R. and Rogerson, S. (2004) Privacy Policies Online: reflections on a continuing investigation, *Proc EthiComp 2004*, Syros Greece
- McRobb, S.R. and Rogerson, S. (2005) Privacy Policies Online: further reflections from a continuing investigation, *Proc EthiComp 2005*, Linköping, Sweden
- McRobb, S.R. and Stahl, B.C. (2006) An Essential Problem of the E-Phenomenon: Privacy, *International Journal of IT and Management* (forthcoming special issue)
- Milne, G.R. and Culnan, M.J. (2002) Using the Content of Online Privacy Notices to Inform Public Policy: A Longitudinal Analysis of the 1998–2001 U.S. Web Surveys, *The Information Society*, 18: 345–359
- Moore, T. (2005) Do Consumers Understand the Role of Privacy Seals in E-Commerce? *Communications of the ACM*, 48(3): 86–91.
- Ranganathan, C. and Ganapathy, S. (2002) Key dimensions of B2C websites, *Information and Management*, 39: 457–465
- Sheehan, K.B. (2002) Toward a Typology of Internet Users and Online Privacy Concerns, *Information Society*, 18(1): 21–32
- Srinivasan, S. (2004) Role of trust in e-business success, *Information Management & Computer Security*, 12(1): 66–72
- Symantecstore.com (2006) Privacy Policy, online at: http://www.symantecstore.com/dr/sat4/ec_main.entry25?page=Global-PrivacyPolicy&client=Symantec&sid=49997&CUR=840&DSP=&PGRP=0&ABCODE=&CACHE_ID=0.
- Udo, G.J. (2001) Privacy and security concerns as major barriers for e-commerce: a survey study, *Information Management and Computer Security*, 9(4): 165–174
- UMIST (2002) Study of compliance with the Data Protection Act of 1998 by UK based websites, Office of the Information Commissioner. Available http://www.co.umist.ac.uk/research/tech_reports/trs_2002_008_lam.pdf.
- Zapspot.com (2006) Privacy, online at: <http://www.zapspot.com/about/privacy/>

Appendix A

**User Survey:
Online Privacy
Policies
2006 (1)**



Your *Pnumber* (type or paste into the field):

--

Organisation (copy and paste the URL):

1. Is personal data collected? (Check ONE)

Yes <input type="checkbox"/>	No <input type="checkbox"/>	Not mentioned <input type="checkbox"/>
------------------------------	-----------------------------	--

2. Can you opt-out from providing it? (Check ONE)

Yes <input type="checkbox"/>	No <input type="checkbox"/>	Not mentioned <input type="checkbox"/>
------------------------------	-----------------------------	--

3. How is it collected - is this described? (Check ONE)

Yes <input type="checkbox"/>	No <input type="checkbox"/>	
------------------------------	-----------------------------	--

4. What data is collected - is this described? (Check ONE)

5. IF personal data is collected – what reasons are given? (Check all that apply)

Maintain contact details <input type="checkbox"/>	Provide the service <input type="checkbox"/>	Evaluate / improve site <input type="checkbox"/>
Sell advertising <input type="checkbox"/>	Targeted marketing <input type="checkbox"/>	Other - specify: <input type="checkbox"/>

6. Is personal data passed to 3rd parties?

Yes <input type="checkbox"/>	No <input type="checkbox"/>	Not mentioned <input type="checkbox"/>
------------------------------	-----------------------------	--

7. Can you opt out from disclosure? (Check ONE)

Yes <input type="checkbox"/>	No <input type="checkbox"/>	Not mentioned <input type="checkbox"/>
------------------------------	-----------------------------	--

8. Why is it disclosed - is this explained? (Check ONE).

Yes <input type="checkbox"/>	No <input type="checkbox"/>	If NO, then skip Question 9.
------------------------------	-----------------------------	------------------------------

9. IF personal data is disclosed – – what reasons are given? (Tick all that apply)

To provide service <input type="checkbox"/>	To tailor experience <input type="checkbox"/>	Marketing purposes <input type="checkbox"/>
Legal requirement <input type="checkbox"/>	Other - specify: <input type="checkbox"/>	<input type="checkbox"/>

Miscellaneous (check ONE)

10. Is there a description of how a customer consents for personal data to be collected?

Yes <input type="checkbox"/>	No <input type="checkbox"/>
------------------------------	-----------------------------

11. Is there a description of how a customer can check and amend their personal data?

Yes <input type="checkbox"/>	No <input type="checkbox"/>
------------------------------	-----------------------------

12. Is there a description of how personal data will be protected?

Yes <input type="checkbox"/>	No <input type="checkbox"/>
------------------------------	-----------------------------

Save as: *OrgName-YourPNo.doc* and email to smcrobb@dmu.ac.uk - thanks for your help!

Appendix B

COMP5211 E-Commerce

Laboratory week 9: Investigating Online Privacy Policies

The purpose of this session is:

- To give you an insight into how online organisations handle - or say they handle - personal data collected from customers.
- To give you an opportunity to participate in an ongoing research project with the Centre for Computing and Social Responsibility.

Tasks:

Working individually, visit the websites listed below, and for each site in turn:

1. Open a copy of the questionnaire (follow link on blackboard)
2. Find the site privacy policy (some are easier to find than others).
3. Read the policy and complete the questionnaire by clicking on the button in the appropriate box. All answers should be based entirely on your own judgment. **Do not discuss your answers** until after the class. This research depends on collecting *individual* answers to all questions.
4. Save each questionnaire with a filename in this format: *OrgName-YourPNo.doc*. So your first might be called *Symantec-P04123456.doc*. It doesn't matter where you save them, as long as you can remember where!

Email all your saved questionnaires to smcrobb@dmu.ac.uk. If you do not finish by the end of the class, please complete this task later - your results will be an important contribution to a five year research project.

Organisations to visit:

Symantec Store (www.symantecstore.com)

Zapspot (www.zapspot.com)

Federal Express (www.fedex.com) - you may have to hunt for this one!

After the lab, check the Blackboard site for follow-up work. The findings of the study will also be published here in due course (though it will take some time to analyse and write up the results).

