informationresearch

VOL. 21 NO. 4, DECEMBER, 2016

Privacy and trust attitudes in the intent to volunteer for data-tracking research

Catherine L. Smith.

Abstract

Introduction. *The analysis of detailed interaction records is fundamental to development of user-centred systems. Researchers seeking such data must recruit volunteers willing to allow tracking of their interactions. This study examines privacy and trust attitudes in the intent to volunteer for research requiring installation of tracking software.*
Method. *A quasi-experimental survey was used to determine how privacy and trust attitudes and the intent to volunteer differ depending on whether tracking software is installed on one's own computer or a university lab computer.*
Analysis. *Data from 110 valid responses were analysed using SPSS. Responses were compared between three levels of intent to volunteer (open, closed, unsure) and installation requirements.*
Results. *Comparing those who decided on installation in the lab to those who decided on installation on their own computers, the acceptability of data tracking differed significantly and differences in the intent to volunteer approached significance. Attitudes on technology, information privacy, trust and research participation differed only with the intent to volunteer.*
Conclusion. *Few people are likely to be open to volunteering when required to install data-tracking software on their own computers. Addressing privacy concerns and conditions of trust requires understanding the dependencies between these factors through further research with broader populations.*

Introduction

In daily life, people use search engines, social networking sites, and other electronic resources as a matter of course. Companies that provide these services record their users' activities for purposes of modelling and predicting needs and preferences. In exchange for valuable services, users grant companies permission to access, record (log), and analyse highly personal and detailed information such as the content of email, search engine query terms, and URLs of Websites visited (Kellar, Hawkey, Inkpen and Watters, 2008). Collected data may be anonymised, or users may grant permission for the retention of identifiable data for the construction of individualised profiles. With these data, commercial enterprises such as Google, Facebook, and Microsoft have acquired detailed and powerful views

on many aspects of human information behaviour.

For academic researchers, understanding how current systems are used in the wild is fundamental. One approach to this is ethnographic methods (Rieh, 2004), which are time-consuming to analyse and often focus on small samples that may not generalise. Research participants may be invited to a lab for observation, but the completion of assigned tasks is unlikely to reflect typical user needs and behaviour, even when participants are asked to perform their own tasks (Hearst and Degler, 2013). Beyond the need for records of interaction during authentic problem solving for domains such as health care (Mamlin and Tierney, 2016) and disaster recovery (Spence, Lachlan and Rainear, 2016), long-term longitudinal data are critical to understanding changes in usage over time. Obtaining such data requires access to shared collections (e.g., USEWOD2012, n.d.), collaborative work across industry and academia (Dumais, Jeffries, Russell, Tang and Teevan, 2014; Yang and Soboroff, 2016), or the deployment of data tracking processes developed by and for academic researchers (Feild, Allan and Glatt, 2011).

One solution for academics is a collaborative approach such as a living laboratory (Kelly, Dumais and Pedersen, 2009; Smith, 2011, 2013). Enterprises of this type are shared among researchers and may engage volunteers in the co-design of information systems (Pallot, Trousse, Senach and Scapin, 2010). In this paper, we focus specifically on the concept of a virtual lab, where collaboration occurs online and participants are remote. Here, ideal volunteers would grant permission for the tracking of detailed interaction data across all personal digital devices. From the perspective of privacy and trust, the development of such a facility faces two interdependent challenges. First, the privacy of volunteers must be safeguarded through techniques such as anonymisation and differential privacy (Ohm 2010; Yang and Soboroff, 2016). Second, in a chicken-and-egg problem, testing these privacy techniques requires a sufficient number of volunteers (Feild and Allan, 2013). Researchers in both the academy and industry have found it difficult to recruit volunteers willing to knowingly install tracking software on their computers (Guo, White, Zhang, Anderson and Dumais, 2011; Community Query Log Project, 2010; Russell and Oren, 2009). Challenges in recruiting research volunteers extend to other domains (Close, Smaldone, Fennoy, Reame and Grey, 2013; Koo and Skinner, 2005), but it is also likely that privacy concerns associated with tracking cause specific impediments. This paper addresses these concerns and other factors hypothesised to affect the decision to volunteer.

This paper is organised as follows. First, we briefly review selected literature on privacy and trust. Following this background information, we state four specific research questions and then describe the method of the study and results. We then discuss our findings and implications before concluding. The paper contributes findings on factors affecting a potential research volunteer's decision to participate in research, with specific findings on the requirement to download and install tracking software on one's own computer.

Background

There are many obvious considerations in a decision to volunteer for research where explicit disclosure of private information is required. Two basic aspects are one's views on personal information privacy and trust that one's privacy will not be violated. While these are straightforward concerns, the study of privacy and trust is not, particularly in light of the many issues raised when modern information technologies are involved (Stutzman, Gross and Acquisti, 2013). There are many studies on privacy and trust in various disciplines and social contexts: law, business, marketing, psychology, computer science, information science and so forth (see, Bélanger and Crossler, 2011; Wang, Min and Han, 2016). While privacy and trust have been treated separately, recent work has examined the combined role of each in human affairs. Many useful conceptualisations flow from this work. Discussions from the law (Nissenbaum, 2001, 2004) are written with the goal of developing a theoretical framework for discussion of practical implications. In this background section we introduce central concepts of privacy and trust starting with Nissenbaum's views, and then present work on several major constructs.

Conceptualisation of information privacy

In introducing conceptualisations that underlie our study, we begin with Nissenbaum's paradigm of contextual integrity (2004). In investigating factors in privacy perception, Martin and Nissenbaum (in press) hypothesised that one's sense of privacy is dependent on three aspects of context: the specific *actors* involved (who is sending or receiving information), expectations on the *flow* of information between the actors (when and how the information will be used), and the type or *content* of information within that flow (what is shared). More generally, Nissenbaum's view posits that context forms social and personal norms for privacy, and that privacy violations come about when contextual elements are misaligned.

For example, granting permission to a search engine (actor) for the recording of query terms (content) for the purpose of improving search outcomes (flow) is normative; in this context the searcher perceives some acceptable level of privacy. In contrast, if the query terms are later distributed to a third party for marketing purposes, the flow is altered in violation of the norm, and privacy is diminished.

In the present study, we examine the specific context of a researcher recruiting volunteers for a study that requires the explicit action of downloading, installing and activating tracking software that records search interaction. In this scenario, the actors are the potential volunteer receiving a recruiting communication, the researcher sending it and the researcher's affiliated institution. The content is the verbatim text of search queries and the URLs of Websites visited. The flow of information mirrors that expected with search providers, except that the researcher offers no exchange of services for the right to access (Richards and Hartzog, in press). As suggested by Nissenbaum's view, privacy is a highly complex construct.

Typical factors studied in work on privacy-related decisions include general privacy concerns (e.g., Malhotra, Kim and Agarwal, 2004), context-dependent privacy concerns (e.g., Internet privacy concerns,

Dinev and Hart, 2006) and other situational factors (for a comprehensive review, see Bélanger and Crossler, 2011). Recent work (Dinev, Xu, Smith and Hart, 2013; Kehr, Kowatsch, Wentzel and Fleisch, 2015) has found evidence for the subsuming construct *privacy perception*, which is characterised as '*an individual state, subsuming all privacy-related considerations at a specific point in time*' (Kehr, *et al.*, 2015, para. 1). Kehr *et al.* found privacy perception to be antecedent to privacy-related decisions on information disclosure. Two key findings flow from this work. There is an interdependency of risk and benefit perceptions, whereby the perception of risk to privacy is mitigated by the perception of greater benefit from disclosure (Dinev *et al.*, 2013; Kehr *et al.*, 2015). The same studies found that perceptions of risk and benefit vary with other factors such as general concerns about privacy, the affective valence of communications and trust in technology infrastructure. Dinev *et al.* found that the perception of control over the information involved (i.e., anonymity and secrecy) affected perceptions of privacy.

More generally, trust has been found to be a key factor in decisions on the disclosure of private information. Next, we introduce the general concept of trust, and then briefly review associated factors before concluding with a discussion of models that account for both privacy concerns and trust relationships.

Conceptualisations of trust

In work on privacy, trust has been modelled as an *outcome* on perceptions of risk (e.g., Dinev and Hart, 2006) and as *antecedent* to perceptions of risks and benefits (e.g., Kehr *et al.*, 2015). In a recent meta-analysis of research on trust in decisions on engagement in social media, Wang, Min and Han (2016) examined trust as a causal factor in perceptions of risk to information privacy and security. Given the complexity of interdependencies between trust and privacy, in considering the role of trust in decision making we turn again to Nissenbaum (2001) for views taken from the broader and more practical vantage point of the law. Next, we summarise and paraphrase her characterisations of trust.

Generally, trust is a specific relationship between a trustee (the entity being trusted) and a *trustor* (the person who trusts). Trust forms over time and with experience; however, in order for trust to accrue, there must be sufficient initial trust. Trust is affected by the history and reputation of the trustee. Where the trustor has some basis for personal knowledge of a trustee, perceptions of the trustee's personal characteristics affect trust. Within the *social context* in which trust is sought, a trustee assumes a role. The trustor's knowledge of the trustee's *qualifications* for that role are important to initial trust formation. More generally, the construct of social context includes *norms* for trustworthiness in the relationship, any *penalty* the trustee faces for failing to prove trustworthy, the *likelihood of disclosure* should there be a failure and any *insurance* against the trustor's loss if the trustee proves untrustworthy. Finally, trust is most likely to develop when two parties share a *mutual condition* or risk and there is some expectation of *reciprocity*. In the context of our study, all of these factors may be involved in a potential volunteer's decision on

enabling data tracking.

As with work on privacy, trust has a large literature covering many models and conceptualisations. McKnight, Choudhury and Kacmar (2002) summarise these constructs in a review of work on trust in the context of e-commerce. We apply these concepts to the perceptions involved in volunteering as a research participant, an act which requires some level of trust or willingness to be vulnerable (Mayer, Davis and Schoorman, 1995). Note that vulnerability implies acceptance of risk.

The study presented in this paper involves three trustees: the *individual researcher* seeking volunteers, the researcher's *affiliated institution* and the *information technologies* used to communicate about and conduct the study.

The work of McKnight and others (McKnight, Carter, Thatcher and Clay, 2011; McKnight and Chervany, 2001; McKnight, Choudhury and Kacmar, 2002) suggest the following conditions for trust in e-commerce and technology-enabled contexts. With respect to the trustworthiness of individuals, the researcher must be perceived as *benevolent* and possessing sufficient *competence* and *integrity* to perform as promised (McKnight *et al.*, 2001). The university, as an institution, must be perceived as providing *structural assurance* (mitigation of risk by social constructs such as rules and regulations) and *situational normality* (proper, customary, and understandable roles) (McKnight *et al.*, 2002). Finally, the specific technologies involved must be perceived as having the *functionality* required to perform as promised, sufficient *reliability* to assure predictability and a quality of *helpfulness* (McKnight *et al.*, 2011). In recruiting volunteers through online means, only electronic or digital communication is available for conveying these qualities of trustworthiness.

Privacy and trust in decision making

We conclude our review on privacy and trust by considering elements involved in the recruitment of research volunteers, where participation requires the installation of tracking software. We focus on two papers that have modelled privacy and trust factors in the context of engagement with specific software applications. These papers use the constructs mentioned above while introducing additional factors.

In a synthesis of prior findings on interrelated constructs on trust and risk, Wang, Min and Han (2016) conducted a meta-analysis of forty-three studies drawn from the literature on social media. In reviewing the work, the authors found the perception of risk often measured using privacy constructs. Their analytical framework examined associations between trust and risk, and the associations of each on data sharing behaviour, among other outcomes. Trust was found to have a larger effect than risk. While the perception of risk was associated with diminished sharing, the larger effect of trust was associated with diminished risk perception and more sharing. In examining moderators on sharing behaviour, trust in the technology platform (the site, community, or service provider) was found to have

greater effect on sharing than did trust in members of the community or network.

The specific situation of trust required for agreement to tracking involves a sufficient belief that privacy will be protected in a complex *information relationship* between the three hypothesised trustees and the volunteer. Richards and Hartzog (in press) conceptualised information relationships involving entrustment of private information to a service provider; however, in a research scenario, there is no direct service relationship. For the volunteer, benefits are likely to be short-term rewards such as cash or other credits, and possibly anticipation of long-term value from new knowledge or improved outcomes. Also, an altruistic volunteer may place value on benefits that accrue to the public good (Edwards *et al.*, 2009; Stunkel and Grady, 2011). In modelling the decision to disclose private information through a smartphone app, Kehr *et al.* (2015) found that the perception of greater benefit mitigated the perception of risk. The perceived sensitivity of the information to be disclosed had a compounding effect on perceptions, so that where more sensitive information was involved, the perception of benefit was diminished and the perception of risk was enhanced. The model also included measures of trust in underlying smartphone technology (termed institutional trust), finding greater trust associated with increased perception of benefit. These findings suggest that for research studies involving sensitive information and few direct benefits, participation is likely to hinge on a sufficient level of trust.

Another aspect of trust and privacy concerns for information technologies is communication of privacy protection from the trustee to the trustor. For academic researchers, this involves disclosure and informed consent meeting the legal and ethical standards of institutional review boards (Eynon, Fry and Schroeder, 2008; Kraut, Olson, Banaji, Bruckman, Cohen and Couper, 2004). Communication on risks associated with tracking software may be considered a unique form of *fear appeal* (Maddux and Rogers, 1983), where the goal is to invoke concerns about risk sufficient to result in reasoned consideration of a decision to take protective action. In the case of recruiting research volunteers, the goal of the fear appeal is to delineate the risks and benefits of participation in a manner that conveys the nature of the threat to privacy while informing on promised protections. Ethics require that the message be devoid of a persuasive valence of positive affect or social influence (Kehr *et al.*, 2015; Johnston and Warkentin, 2010), which are likely to diminish the perception of risk.

Johnston and Warkentin (2010) studied the effect of a fear appeal intended to motivate the installation of software that detects tracking threats (anti-spyware). Two forms of efficacy were included in the model: self-efficacy with respect to the ability to utilise the software, and perception of the efficacy of the software. Higher levels of efficacy were associated with a greater intention to install the software, but a greater perception of threat was associated with lower efficacy. These findings suggest that a research volunteer's decision to download and install tracking software requires sufficient perception of the efficacy of the promised privacy protections. Downloading tracking software

is an explicit action to accept a threat to privacy, where the alternative is *no action*. Not accepting the threat is likely to be perceived as highly efficacious. These factors are likely to put an additional burden on the role of trust in the decision to volunteer.

Johnston and Warkentin's (2010) model also included the perception of the susceptibility to threats. No significant association was found between the perception of susceptibility and efficacy; the authors speculated that prior experience moderated perceived susceptibility, such that people with no prior experience do not feel susceptible. Elhai and Hall (2016) examined anxiety about data breaches, self-reported use of nine security precautions and prior personal experience with four types of breaches. No significant association was found between prior experience and anxiety. Only one protective behaviour had a significant association; greater anxiety was associated with a higher probability of the use of a password or fingerprint reader on one's smartphone.

The study presented in this paper draws on the earlier work discussed above to explore conditions of trust and data privacy attitudes in the context of a potential research volunteer's intent to volunteer in a study requiring the installation of tracking software. While our work draws on the concepts mentioned above, it was not designed to test theory or to develop predictive models. Rather, the goal was to explore the salience of privacy and trust attitudes in the practical context of a respondent's intent to volunteer as detailed in the questions listed below.

Research questions

1. Does requiring the download and installation of tracking software on one's own computer affect the intent to volunteer in a research study, as compared with the same study without the requirement for installation on one's own computer? We hypothesise that the rate of volunteering will be lower for a study that requires data tracking on one's own computer.
2. Comparing those who say 'yes' to volunteering, those who are 'unsure', and those who say 'no', what, if any, are significant differences in privacy protection behaviour and prior exposure to privacy violations?
3. Comparing those who say 'yes' for a study requiring installation of tracking software on their own computer and those who say 'yes' to the same study without installation on one's own computer, what, if any, are the significant differences in privacy protection behaviour and prior exposure to privacy violations?
4. What, if any, are the associations between the attitudes on privacy and trust, and the intent to volunteer in a study requiring installation of tracking software on one's own computer?

Method

Overview

Previous work suggests that it is expensive to recruit volunteers for an actual study using downloaded tracking software, thus we used a survey approach to gather responses on a hypothetical research study (Russell and Oren, 2009; Smith, 2011). Because rates of volunteering

for research tend to be low in general (Arfken and Balon, 2011; Galea and Tracy, 2007), we sought to separate general factors from those associated with the need to install software on one's own computer. For these reasons, the study used a quasi-experimental design. In a quasi-experimental survey, respondents are assigned to groups and receive different instruments designed for comparison between the groups. Our study used two versions of a questionnaire, which was administered using a web-based online survey service. Respondents were asked about their willingness to participate in one of two versions of the hypothetical study (h-study). Half of respondents received information about an h-study requiring the installation of tracking software on their own computers (OwnV); the other half received information about an h-study requiring an appointment where the software would be installed on a lab computer (LabV). Except for details describing the assigned h-study, the questionnaires were identical.

Participants

The data were collected from an undergraduate research pool comprising students enrolled in a large introductory undergraduate course at a university in the U.S. Midwest. Students received course credit for participation. Recruiting was done through in-class announcements and an online administration system for student research pools. Pool demographics are 78% white, 58% female, with 96% under age 24. Approximately one third of students are in their first semester of college, with only 3% having obtained a prior bachelors' degree. 8% are international students and 95% report English as their native language.

Survey structure

Table 1 provides an overview of the survey structure. After confirmation of consent to participate, the first block asked about smartphone and computer ownership. For the one student who reported not owning a computer and having no access to a computer, the survey ended and course credit was granted. We assumed that those who did not own their own a computer could face impediments to installing software, thus seven students were assigned automatically to LabV. All other respondents were assigned randomly to one of the two groups. The second block of questions asked about privacy protection practices for smartphones, computers and Websites (see Table 2). The third block presented information about the assigned h-study and collected information about the intent to volunteer. The fourth block investigated attitudes involved in the decision. The fifth block asked about knowledge of and personal experience with privacy violations (see Table 3). All survey questions except those in block four were validated for forced completion with the option to refuse an answer. Within blocks, the item order of related questions was randomised automatically.

- Block 1 Smartphone and computer ownership
- Block 2 Privacy practices: smartphone, computer, Websites
- Block 3 h-study version assigned
    - Block 3a Introduction to mock study

- Block 3b Mock email subject line
- Block 3c LabV or OwnV version of mock recruiting email
- Block 3d LabV or OwnV version of mock disclosure and consent statement
- For those responding 'yes' to 3d: LabV: Would you make an appointment? OwnV: Would you download the software?

- Block 4 Fourteen semantic dichotomies
- Block 5 Previous experience with, and knowledge of, privacy violations and threats

Table 1. Block structure of the online survey

Block 3: Questions on intent to volunteer for the h-study

Communication materials were written to comply with institutional review board requirements, with the h-study described as being about how people search for information on the Internet. The block was exactly the same except for the requirement for an appointment and software in the lab for LabV, and the requirement for downloading and installing the software one's own computer for OwnV. The block started with the same mock e-mail subject line, '*Participate in a research study and earn up to $40*', and asked about the likelihood of opening and reading such an e-mail. The next page displayed a mock recruiting e-mail, which contained information about the assigned h-study (see Appendix A). The email contained the sentence: '*To learn more about the study, or to sign up, visit this website: [url]*'. Questions at the bottom of the page asked, '*Do you have enough information to make a decision about participating?*' and '*Would you click the link to learn more?*'. Respondents were instructed to assume they had clicked the link. A disclosure was then displayed for the assigned h-study (see Appendix B). The next questions asked '*Do you have enough information to make a decision about participating?*' and '*Would you agree to participate?*'. Only those who answered '*yes*' to participating were asked a follow-up question about the likelihood of taking action to participate—for LabV: '*Would you click to schedule an appointment in the lab?*', and for OwnV: '*Would you click to download the CrowdLogger [data tracking] software?*'.

	Privacy protection behaviour	% responses			
		Never	Occasionally	Usually	Always
Computer	Manually clear browsing history	21	54	17	7
	Manually delete cookies	51	40	5	4
	Manually clear search history	31	46	17	6
	Manually clear cache	53	39	6	2

Smartphone	Manually clear browsing history	27	41	18	14
	Manually delete cookies	67	22	4	7
	Manually clear search history	35	39	16	10
	Manually clear cache	64	18	11	7
Apps	Check for privacy tool	38	30	13	19
	Ask friends about it	41	27	21	11
	Read privacy policy	57	35	6	2
	Check ratings	9	23	33	35
	Refuse to share with third party	14	28	30	29
	Check for certification	45	28	15	13
	Use privacy setting	9	21	28	42
Websites	Check for privacy policy	36	28	23	14
	Ask friends about it	26	34	26	16
	Read privacy policy	44	47	8	1
	Check ratings	31	33	23	16
	Refuse to share with third party	18	25	32	25
	Check for certification	41	26	16	17
	Use privacy setting	12	28	27	33

Table 2a: Privacy protection practices: self-reported usage rates

	Set device to:	Per cent report using
Computer	Automatically clear browsing history	7
	Automatically delete cookies	12
	Automatically clear search history	6
	Automatically clear cache	6
	Software preventing tracking	8
	Software blocking ads.	57
	Other protection software	33
	Automatically clear browsing history	9

Smartphone	Automatically delete cookies	6
	Automatically clear search history	16
	Automatically clear cache	4

Table 2b: Privacy protection practices through device settings: self-reported usage rates

Knowledge of privacy violations in past two years		Per cent reporting number of instances				
		Never	1 or 2	3 or 4	5 to 10	Over 10
Personal victim of...	privacy invasion	47	41	9	3	—
	stolen credit card, bank account, information	79	21	1	—	—
Known a victim of...	privacy invasion	18	42	23	13	4
	stolen credit card, bank account, information	9	53	22	12	4
Heard or read about...	identify theft	16	38	23	8	15
	potential invasion of privacy through surveillance	16	39	25	11	10
	potential invasion of privacy by hackers	12	38	27	8	15
	misuse of information collected from the Internet	6	19	26	22	27

Table 3: Personal experience with and knowledge of privacy violations: self-reported rates

Block 4: Questions on attitudes toward volunteering

Two types of questions collected information about attitudes. First, after a reminder of their decision about participating ('*Earlier in the survey, you were asked... You answered saying you would/would not/were undecided...*'), an open-text question asked respondents to explain the reason for their answer. The next page displayed fourteen semantic dichotomies in a bi-polar matrix of radio buttons labelled 1 through 10, with opposing ends of the dichotomies displayed at each end of the scale (see Table 4).

The dichotomies were developed using responses from an earlier version of the instrument, which contained scales adapted from the literature reviewed above. The instrument also contained a similar open-response question on the intent to volunteer. The verbatim responses were used in bottom-up content analysis (Krippendorff, 2012), resulting in the bipolar items. Further detail on the development of the dichotomies is presented elsewhere (Smith, 2016). In preparing the instrument for this study, three of the dichotomies were reversed.

In thinking about my desire to participate, [I feel]...			
Area	**Dichotomy name**	**Left (low)**	**Right (high)**
	Download (R).		

Technology	Downloading software makes me feel...	At ease	Worried
	Software. I understand how the software works...	Completely	Not at all
Trustees	*Institutional review boards*. University review board approval [makes me] feel...	Confident	Sceptical
	Email. Email from [university name] is...	Trustworthy	Not trustworthy
	Researcher. The researcher is....	Trustworthy	Not trustworthy
Privacy	*Protection*. The privacy protections are...	Sufficient	Insufficient
	Tracking (R) Tracking my Internet activities is...	Acceptable to me	Unacceptable to me
Information	*Information*. The information I've been given is...	Enough to decide	Not enough
	Opinion (R) Hearing...opinion[s] before volunteering is...	Unessential to me	Essential to me
General	*Time*. The study is (1) ...	Good use of time	Poor use of time
	Interest The study is (2)....	Interesting to me	Not interesting
	Ease (R) Completing the study would be...	Easy	Difficult
	Money. The money is...	Satisfactory	Unsatisfactory
	Volunteering. Helping by volunteering is...	Important to me	Unimportant to me

Table 4: The semantic dichotomies, R indicates reversed item in the instrument

Data preparation

The survey was completed by 404 students. Five steps of data cleaning were used before analysis (see Appendix C). This resulted in the exclusion of 73 per cent of the records, which is consistent with a 75 per cent data-cleaning exclusion for a similar study (De Santo and Gaspoz, 2015). We were mindful that our data cleaning process could bias the final sample; therefore, we tested the distribution of the excluded responses at each step, finding no bias. Respondents were identified as either closed, unsure, or open to volunteering; we refer to this variable as the intent to volunteer, or Intent. With 294 responses excluded, 110 valid records remained: 51 from LabV and 59 from OwnV. We report on analysis of these records only.

Analytical approach

As the measures of behaviour and prior experience are count data, these are compared using the non-parametric, linear-by-linear, chi-squared test for differences in response frequency across the three levels of Intent. Because the measures of attitude were gathered using a ten-point scale, we used an analysis of variance to examine associations between Intent and h-study. As a result of the exploratory nature of this work and the large number of tests performed, we use an alpha level of 0.05 and two tails in our tests for significance. *Post hoc* analysis was performed using the conservative Scheffe method.

Results

Research question 1 on intent to volunteer

Table 5 lists the rates of hypothetical volunteering for each h-study. In answering research question 1, chi-squared analysis was used to test the hypothesis that participants in the OwnV group were less likely to be open to volunteering than those in the LabV group. Although the rate was lower by half (14% vs. 28%), differences were not significant across all three levels of Intent ($\chi 2$ (2) = 3.629, p = 0.159). A comparison between open and closed rates approached significance. ($\chi 2$ (1) = 3.587, p = 0.058).

h-study version	Per cent in response category (Intent)			Per cent total sample (count)
	Open	Unsure	Closed	
LabV	28	37	35	46 (51)
OwnV	14	39	48	54 (59)
Total	20	38	42	100 (110)

Table 5: Percent respondents with intent to volunteer (Intent), by h-study version

Research questions 2 and 3 on protection behaviour and previous experience

In addressing research question 2, we examined associations between intent to volunteer and reported privacy protection behaviour, previous experience, and knowledge of privacy violations (see Tables 2 and 3). As a result of the small sample sises in many cells, exact tests and SPSS Monte Carlo sampling were used to verify results derived through asymptotic analysis. Using a linear-by-linear model, we tested for a significant relationship between behaviour reported as frequencies and intent to volunteer. Each binary response about behaviour was tested by chi-squared analysis. No significant differences were found for any variable for either group.

We addressed research question 3 by comparing the same measures for open respondents in each group. No significant differences were found between open respondents in LabV and open respondents in OwnV.

Research question 4 on attitudes associated with intent to volunteer

Fourteen ten-point bi-polar scales were used to measure attitude in the context of the intent to volunteer. These data were treated as continuous in a series of separate two-way ANOVAs that examined differences within and between h-study groups. Each ANOVA modelled the main effects and interaction of h-study and Intent for one dichotomy.

Dichotomy name	p	Anchor at 1	Mean for response level 2	3	4	5	6	7	8	9	Anchor at 10
Downloading (R)	***	At ease					Oa (5.8)		Ub (8.2)	Cb (8.8)	Worried
Understand	**	Completely			Oa	U	Cb				Not at all

software					(4.3)	(5.4)	(6.3)				
Review board makes me	***	Confident		Oa (3.0)	Ua (4.0)		Cb (5.6)				Sceptical
Email is	*	Trustworthy	Oa (2.0)	U (2.9)	Cb (3.6)						Not trustworthy
Researcher is	***	Trustworthy	Oa (2.1)		Ub (3.9)	Cc (5.4)					Not trustworthy
Protection	***	Sufficient		Oa (2.9)	Ua (4.2)		Cb (5.9)				Insufficient
Tracking (R) †	***	Acceptable				Oa (4.8)	Ub (6.5)		Cc (8.0)		Unacceptable
Information is	*	Enough	Oa (2.3)	Cb (3.0)	Ub (3.7)						Not enough
Opinions (R)	Not.sig.	Unessential									Essential
Use of my time	***	Good			Oa (3.6)	Ua (4.9)		Cb (7.2)			Poor
Study is (R)	**	Easy			Oa (3.7)	U (4.6)	Cb (5.8)				Difficult
Study is	***	Interesting			Oa (4.0)	Ua (5.0)		Cb (7.4)			Not interesting
Money is	***	Satisfactory	Oa (2.1)	Ua (3.5)		Cb (5.2)					Unsatisfactory
Volunteering	Not sig.	Important									Unimportant

† The interaction of Intent and h-study was significant; see Table 7 and the text.
O = Open; U = Uncertain; C = Closed * p < 0.05 ** p < 0.01 *** p < 0.001

Table 6: Comparison of mean response levels for 14 semantic dichotomies by intent to volunteer, with significant subsets as determined by Scheffe's *post hoc* test

h-study group		Mean (s.d.) for response level								
		2	3	4	5	6	7	8	9	
LabV—tracking	Acceptable				O 5.4 (3.0)	U 5.6 (2.6)		C 8.3 (2.1)		Unacceptable
OwnV—tracking				O 3.8 (2.1)			U 7.2 (2.3)	C 7.9 (2.3)		

Table 7: Comparison of mean response levels for measure of unacceptability of tracking, by intent to volunteer and h-study condition.

No significant main effect was found for h-study for any of the measures; however, the main effect of Intent was significant for twelve of the measures. Table 6 lists the dichotomies, the significance of main effects on Intent, and the mean for each at the three response levels across both h-study groups. Subscripts denote significant subgroups, as indicated by Scheffe's post-hoc test. No significant interactions were found except for the acceptability of tracking on the Internet ($F(2,104) = 3.38$, $p < 0.05$), which is detailed in Table 7 and discussed next.

An analysis of simple main effects, with Bonferroni correction, was used to examine the acceptability of tracking. Among unsure respondents, there was a significant difference in the acceptability of tracking between h-study versions ($F(1,104) = 4.237$, $p < 0.05$). Relative to those in LabV, unsure respondents in the OwnV group were more likely to feel that data tracking is unacceptable. Comparison between h-studies for the other two levels of Intent (open, closed), were not significant. Within both versions of the h-study, simple main effects of Intent were significant (LabV: $F(2,104) = 44.8$, $p < 0.001$; OwnV: $F(2,104) = 52.9$, $p < 0.001$). Within LabV,

the difference between open and unsure respondents was not significant, but the unacceptability of tracking was significantly greater for closed respondents ($p < 0.01$). Within OwnV, the difference between closed and unsure respondents was not significant, but the unacceptability of tracking was significantly lower (data tracking was more acceptable) for open respondents ($p < 0.001$).

Given the significant differences in attitude among those with differing levels of intent to volunteer, we investigated associations between the attitudes. Because of the small sample sise, principal components analysis was not defensible. Using separate ordinal logistic regressions, we tested for a linear relationship between each measure and level of Intent. A significant linear relationship was found for all but two measures. Table 8 lists the individual exponential beta coefficients with chi-squared and p values. This result suggests the dichotomies captured meaningful levels for the constructs underlying each bi-polar scale.

Dichotomy name		X^2	df	sig	exp (beta)
Technology	download	24.1	1	***	1.56
	software	8.8	1	**	1.23
Trustee	institutional review board	24.5	1	***	1.55
	e-mail	6.3	1	*	1.29
	researcher	36.5	1	***	1.73
Privacy	protection	22.4	1	***	1.42
	tracking	24.1	1	***	1.42
Information	information	0.3	1	not sig.	—
	opinion	1.7	1	not sig.	—
General	time	41.0	1	***	1.70
	ease	11.7	1	**	1.29
	interest	35.6	1	***	1.61
	money	47.6	1	***	1.50
	volunteering	4.6	1	*	1.17
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$					

Table 8. Analysis of linear relationship between responses to fourteen semantic dichotomies and intent to volunteer, where open = 1, unsure = 2, and closed = 3, with tests of significance and exponential betas.

In the light of the uniform direction and similarity in size of the beta values, we also examined the multicollinearity of the measures, finding three measures with variance inflation factors (VIFs) above 2.5 on all other bi-polar measures. These were: (1) the level of interest in the study, (2) attitude toward spending time on the study and (3) the trustworthiness of the researcher. Inexamining correlations between the measures, we found that the trustworthiness of the researcher was significantly correlated with twelve of the other thirteen measures, with four highly correlated (Spearman's $\rho > 0.600$) and six moderately correlated ($0.400 <$ Spearman's $\rho < 0.599$). Table 9 details the correlations. These results suggest that the trustworthiness of the researcher may summarise or subsume other factors associated with the intent to volunteer. The high correlation (r

= 0.730) between interest in the study and the use of time also suggests that these measures express the same underlying attitude. The small sample size and the high intercorrelation preclude further meaningful statistical analysis of associations between the measures.

Dichotomy	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Researcher	1.000												
2 Software	**0.680****	1.000											
3 Inst. rev. board	*0.596***	0.294**	1.000										
4 E-mail	*0.553***	0.329**	*0.434***	1.000									
5 Protection	**0.632****	*0.451***	*0.410***	*0.520***	1.000								
6 Tracking	*0.447***	0.289**	0.342**	0.225*	0.317**	1.000							
7 Ease	0.361**	*0.429***	0.277**	0.298**	0.396**	0.299**	1.000						
8 Download	*0.482***	0.293**	0.376**	0.234*	*0.413***	*0.474***	0.236**	1.000					
9 Time	**0.665****	*0.479***	*0.458***	0.376**	*0.562***	*0.443***	0.319**	*0.573***	1.000				
10 Money	**0.618****	*0.522***	*0.504***	0.381**	*0.587***	0.289**	*0.407***	0.314**	*0.515***	1.000			
11 Interest	*0.583***	*0.469***	0.371**	0.276**	*0.457***	0.342**	0.343**	0.526**	**0.730****	0.437**	1.000		
12 Volunteering	0.399**	0.314**	0.370**	0.359**	0.333**	0.209*	0.277**	0.163	*0.438***	0.319**	0.339**	1.000	
13 Opinion	0.084	-0.049	0.048	-0.067	0.051	0.200*	-0.022	0.223*	0.128	0.005	-0.022	0.050	1.000
14 Information	*0.414***	*0.418***	0.196*	0.250**	0.199*	0.155	0.255**	0.146	0.158	0.249*	0.157	0.226*	0.056
Bold values = strong correlation (p > 0.600); italic values = moderate correlation (p < 0.600 and > 0.400).													

Table 9: Analysis of correlations among fourteen semantic dichotomies using Spearman's ρ

Discussion

Using two versions of a hypothetical research study, we have examined privacy protection behaviour, prior experience with and knowledge of privacy violations, and attitudes associated with the intent to volunteer. Except for the acceptability of data tracking, no statistically significant results were found for any tests comparing the responses from the two h-study groups. We find no evidence that the protection practices or prior experience are factors in the intent to volunteer. This finding is consistent with that of Elhai and Hall (2016) who found no association on similar measures.

The interaction found for the acceptability of tracking suggests that the requirement to download and install tracking software on one's own computer affected the decision to participate. This finding is consistent with the theory of contextual integrity (Nissenbaum, 2004), which predicts that sensitivity to the implications of data tracking will have different effects in different contexts. Here we find not just different levels for the decision outcomes, but the data also suggest different sensitivities in the effect of the risk perception (see Figure 1). For the OwnV group, openness to volunteering may require a threshold minimum level for the acceptability for data tracking (the minimum being a function of the direction of the dichotomy). For the LabV group, closedness may occur at a minimum threshold level of unacceptability. In viewing acceptability as the perception of risk, this is consistent with prior findings that the perception of greater risk is associated with less willingness to disclose private information and lower trust (Dinev and Hart, 2006; Dinev *et al.*, 2013).

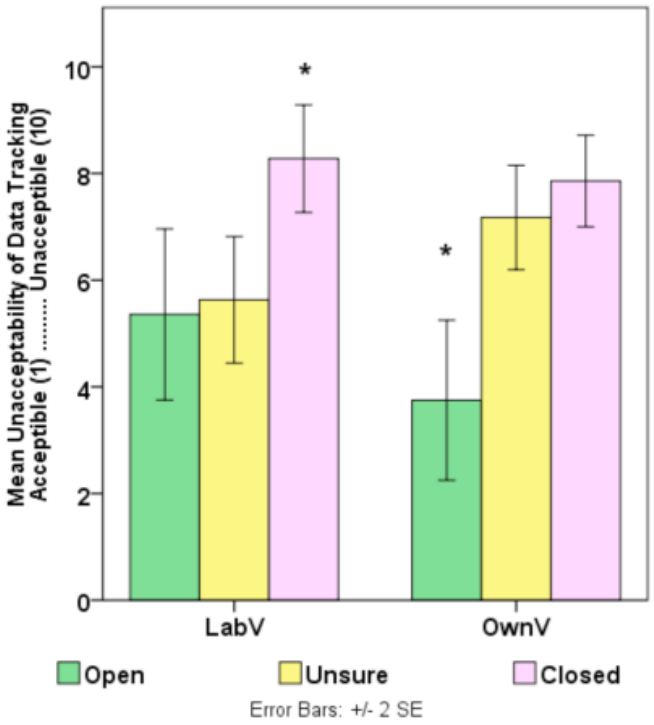

Figure 1: Interaction of h-study version and intent to volunteer for the unacceptability of data tracking, * indicates significantly different group.

While the rate of volunteering for the OwnV study was half that of the LabV study, the large difference was not statistically significant. We speculate that for OwnV, the open response rate is overstated relative to actual recruiting situations. For the OwnV group, only 12% of respondents indicated a positive intent to volunteer (yes) at every step in the decision process (see Figure C.1 in Appendix C). This rate is similar to the only previously reported rate known to us for studies involving the installation of tracking software. Microsoft researchers Guo *et al.* (2011) reported 10% participation among Microsoft employees asked to install tracking software on their computers.

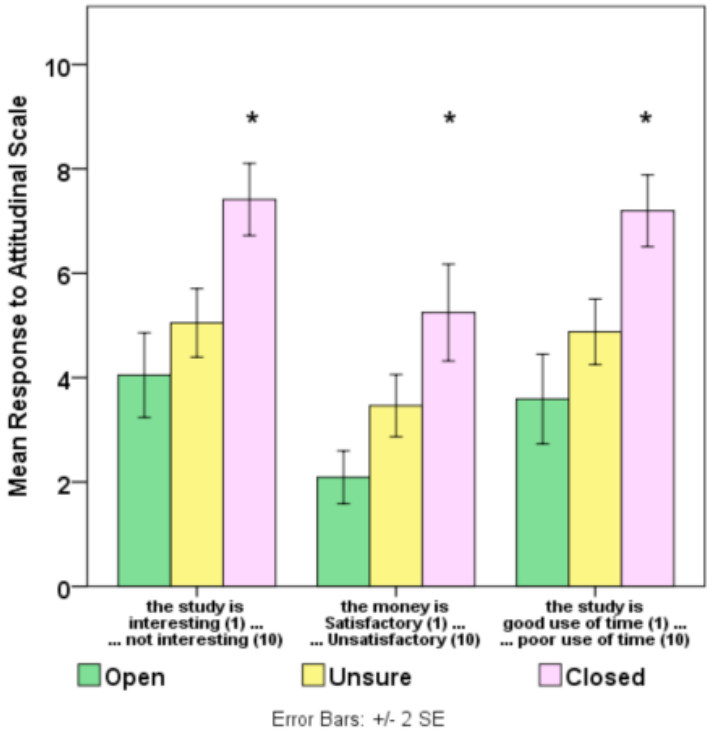

Figure 2: Response to scales on general attitudes toward volunteering (interest, money and time) by intent to volunteer, across all respondents, * indicates significantly different group.

The linear associations between Intent and attitudes for twelve of the bi-polar measures suggest that those factors were meaningful in the contexts presented in the h-study. Among the five general measures on attitude toward participation, three have characteristics suggesting thresholds that separate closed respondents from others (see Figure 2). Closed respondents did not find the money satisfactory, they were not interested in the study and they saw participation as a poor use of their time. With respect to self-efficacy with technology, trust and privacy protection, two measures suggest thresholds for openness and two thresholds for closedness. Open respondents expressed less worry about downloading the software and greater trust in the *researcher* (see Figure 3). Closed respondents derived less confidence from *research board approval* and were less likely to believe that privacy *protection* was sufficient (see Figure 4).

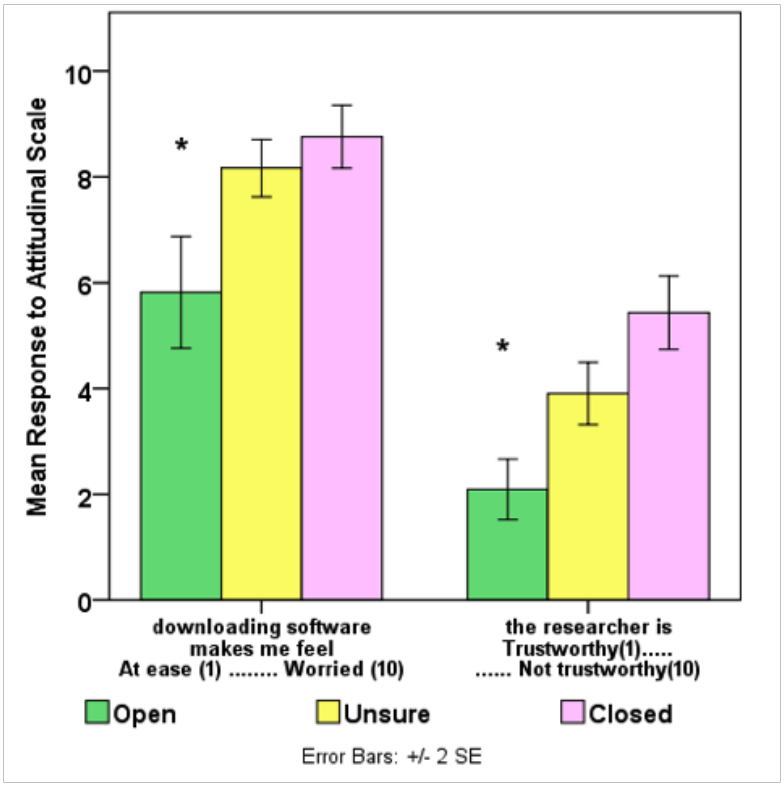

Figure 3: Response to attitudinal scales (downloading, researcher) by intent to volunteer, across all respondents, * indicates significantly different group.

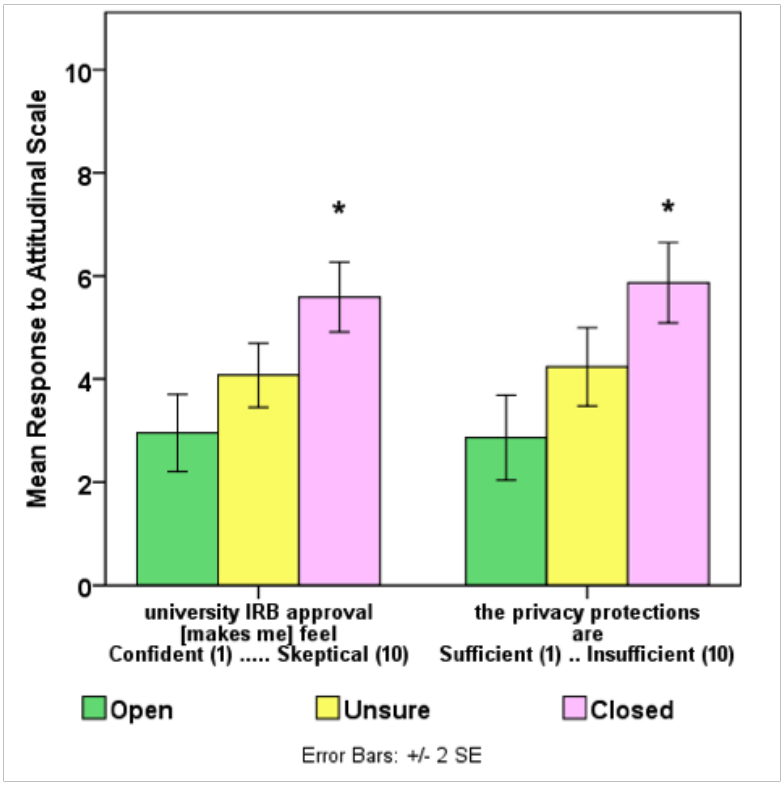

Figure 4: Response to attitudinal scales (IRB, privacy protection) by intent to volunteer, across all respondents, * indicates significantly different group.

Our findings support common sense with respect to the intent to volunteer. Those who are open have the greatest self-efficacy with the

technology, as measured by perceptions on the ease of downloading and understanding of the software. Open respondents also had the greatest sense of the efficaciousness of the software technology, as measured by perception of the sufficiency of the privacy protections. These findings are consistent with those of Johnston and Warkentin (2010), who found that greater technology efficacy was associated with the decision to install protective software; however, the software in our study was the opposite of protective. This suggests that the effect of efficaciousness is independent of the role of the software in mitigating the perception of risk; however, Johnston and Warkentin (2010) also found that the efficacy was diminished by the perception of higher threat severity. This suggests that for the open respondents, other factors mitigated the perception of risk. Trust is a mitigating factor for the perception of risk (Dinev and Hart, 2006; Johnston and Warkentin, 2010; Kehr *et al.*, 2015), hence, we associate open respondents' greater trust in the researcher with a lower perception of risk, and hypothesise that this enhanced the perception of acceptability for data tracking.

With respect to ambivalence, we observe that unsure respondents were closer to open respondents on many measures, including three of the four on general attitudes toward volunteering. This suggests that they were open to volunteering in general, but specific factors affected their intent. Different from those who were open, the unsure were more worried about downloading software although they expressed a similar level of confidence from review board approval and a similar sense that the privacy protections were sufficient. Unsurprisingly, they had the strongest feeling that the information provided was insufficient.

These findings, along with the finding on the acceptability of tracking, suggest complex interdependencies between factors of trust and privacy concern. Understanding the direction of the associations between these factors, and the relative weight of their influence on intent, requires further study using more sensitive survey instruments and larger samples.

Implications, limitations, and future research

Our findings have implications for researchers who wish to use remote methods requiring the installation of logging software on participants' own computers. Participation rates are likely to be low, even when the researcher has an affiliation with the potential participant. Prior research and our results suggest that those likely to participate perceive little risk from participation and are trusting of the recruiting context.

Attempts to enhance participation may be most effective when focused on increasing comfort with downloading and installing the software through additional information and instruction.

Also, because a sizeable share of those receiving initial communications may be unsure and lacking trust in the researcher, follow-up with personal messages sent directly from the researcher may enhance participation.

We have investigated our research question on attitude using scales that reduced complex constructs to fourteen dichotomies. Our instrument presented the measures in a single matrix after the decision to participate was made and respondents had considered their reasoning in a written open response. This approach was likely to have elicited responses with rational coherence among the fourteen measures, which would not necessarily be found using separate, multi-factor scales on the same constructs. Also, we have gathered no data on underlying context-free levels for these measures, so we do not know how the context relates to general perceptions. The above issues may be addressed in future research with a more refined instrument.

Finally, our sample is not representative of a diverse population. To the extent that other populations are likely to be more sensitive to privacy threats, participation is likely to be lower than that reported here. These factors may be addressed in future research.

Conclusion

This study contributes to our understanding of factors associated with the decision to participate in a research study. Using a quasi-experimental survey, we have investigated the roles of privacy and trust attitudes when participants are asked to install tracking software on their own computers. In comparing this scenario to a low-risk scenario (installation on a lab computer) we find that the differing levels of risk affect the decision differently.

- For those who considered the risk of installation on their own computers, agreeing to do so was associated with a strong attitude of acceptance toward data tracking.
- For those who considered the low risk of installation in the lab, not agreeing to participate so was associated with the strong attitude that data tracking is unacceptable.

More generally, across both scenarios we found differences in attitude among those who were open or closed to participation. Our findings are consistent with prior research on the interdependent roles of privacy and trust in the use of information technologies. We have made several recommendations for researchers needing to recruit volunteers for studies requiring software installation. A deeper understanding of associations between attitudinal factors requires further research with more sensitive scales and larger, more diverse samples.

Acknowledgements

The author thanks the anonymous reviewers and editors for their constructive suggestions and comments, which were invaluable in improving the paper. Thanks also to Heather Flynn for her diligent work on the surveys.

About the author

Catherine L. Smith (PhD) is an Assistant Professor at the School of Library and Information Science at Kent State University, Kent, OH 44242-0001, USA. She may be contacted at csmit141@kent.edu.

References

Arfken, C. L. & Balon, R. (2011). Declining participation in research studies. *Psychotherapy and Psychosomatics, 80*(6), 325–328. Retrieved from http://www.karger.com/Article/Pdf/324795 (Archived by WebCite® at http://www.webcitation.org/6iepFmEQL)

Bélanger, F. & Crossler, R. E. (2011). Privacy in the digital age: a review of information privacy research in information systems. *MIS Quarterly*, *35*(4), 1017–1042.

Community Query Log Project Results. (2010). Amherst, MA: University of Massachusetts, Center for Intelligent Information Retrieval. Retrieved from http://web.archive.org/web/20130413140953/http://lemurstudy.cs.umass.edu/ (Archived by WebCite® at http://www.webcitation.org/6jAzZiW0Q)

Close, S., Smaldone, A., Fennoy, I., Reame, N. & Grey, M. (2013). Using information technology and social networking for recruitment of research participants: experience from an exploratory study of pediatric Klinefelter syndrome. *Journal of Medical Internet Research*, *15*(3), e48. Retrieved fromhttp://www.jmir.org/2013/3/e48/ (Archived by WebCite® at http://www.webcitation.org/6idkK97TG)

De Santo, A. & Gaspoz, C. (2015). Influence of risks and privacy literacy on coping responses to privacy threats. In Proceedings of the 20th Association Information et Management Conference, paper 54. Retrieved from http://bit.ly/2dGrUfk (Archived by WebCite® at http://www.webcitation.org/6idl2hAgu)

Dinev, T. & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, *17*(1), 61–80. Retrieved from http://bit.ly/2fcZjUa (Archived by WebCite® at http://www.webcitation.org/6iepPCNFW)

Dinev, T., Xu, H., Smith, J. H. & Hart, P. (2013). Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems*, *22*(3), 295–316.

Dumais, S., Jeffries, R., Russell, D. M., Tang, D. & Teevan, J. (2014). Understanding user behavior through log data and analysis. In J.S. Olson & W.A. Kellogg (Eds.), Ways of knowing in HCI (pp. 349–372). New York, NY: Springer.

Edwards, P. J., Roberts, I., Clarke, M. J., DiGuiseppi, C., Wentz, R., Kwan, I., ... Pratap, S. (2009). Methods to increase response to postal and electronic questionnaires (Review). *Cochrane Database of Systematic Reviews*, *3*, 1–12.

Elhai, J. D. & Hall, B. J. (2016). Anxiety about Internet hacking: results from a community sample. *Computers in Human Behavior, 54*, 180–185. Retrieved from http://bit.ly/2dGo8ml (Archived by WebCite® at http://www.webcitation.org/6iepbs7QB)

Eynon, R., Fry, J. & Schroeder, R. (2008). The ethics of Internet research. In N. Fielding, R.M. Lee & G. Blank (Eds.), *The SAGE handbook of online research methods* (pp. 23-41). Los Angeles, CA: SAGE. Retrieved from http://www.sagepub.in/upm-data/49617_Eynon_et_al,_SHB_of_ORM.pdf (Archived by WebCite® at http://www.webcitation.org/6iepgYvjK)

Feild, H. & Allan, J. (2013). Using CrowdLogger for in situ information retrieval system evaluation. In *Proceedings of the 2013 workshop on living labs for information retrieval evaluation* (pp. 13-14). New York, NY: ACM.

Feild, H. A., Allan, J. & Glatt, J. (2011). CrowdLogging: distributed,

private, and anonymous search logging. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 375–384). New York, NY: ACM. Retrieved 8 March, 2016 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.367.267&rep=rep1&type=pdf (Archived by WebCite® at http://www.webcitation.org/6iepngsA7)

Galea, S. & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of Epidemiology*, *17*(9), 643–653.

Guo, Q., White, R. W., Zhang, Y., Anderson, B. & Dumais, S. T. (2011). Why searchers switch: understanding and predicting engine switching rationales. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 335–344). New York, NY: ACM. Retrieved from http://www.cse.cuhk.edu.hk/irwin.king.new/_media/presentations/why_searchers_switch.pdf (Archived by WebCite® at http://www.webcitation.org/6ieprzF8a)

Hearst, M. A. & Degler, D. (2013). Sewing the seams of sensemaking: a practical interface for tagging and organizing saved search results. In *Proceedings of the symposium on human-computer interaction and information retrieval* (p. 4). New York, NY: ACM. Retrieved from http://people.ischool.berkeley.edu/~hearst/papers/hcir_2013.pdf (Archived by WebCite® at http://www.webcitation.org/6iepw1r0h)

Johnston, A. C. & Warkentin, M. (2010). Fear appeals and information security behaviors: an empirical study. *MIS Quarterly*, *34*(3), 549–566. Retrieved from http://bit.ly/2eCau85 (Archived by WebCite® at http://www.webcitation.org/6ieq2mrxJ)

Kehr, F., Kowatsch, T., Wentzel, D. & Fleisch, E. (2015). Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, *25*(6), 607–635. Retrieved from http://bit.ly/2eIIOMP (Archived by WebCite® at http://www.webcitation.org/6ieq7JrZG)

Kellar, M., Hawkey, K., Inkpen, K. M. & Watters, C. (2008). Challenges of capturing natural Web-based user behaviors. *International Journal of Human–Computer Interaction*, *24*(4), 385–409.

Koo, M. & Skinner, H. (2005). Challenges of Internet recruitment: a case study with disappointing results. *Journal of Medical Internet Research*, *7*(1), e6. Retrieved from http://www.jmir.org/2005/1/e6/ (Archived by WebCite® at http://www.webcitation.org/6idmaTprl)

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J. & Couper, M. (2004). Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, *59*(2), 105. Retrieved from http://bit.ly/2eC9ota (Archived by WebCite® at http://www.webcitation.org/6ieqDb0td)

Krippendorff, K. (2012). *Content analysis: an introduction to its methodology*. (3rd ed.). Thousand Oaks, CA: Sage Publications.

Kelly, D., Dumais, S. & Pedersen, J.O. (2009). Evaluation challenges and directions for information-seeking support systems. *IEEE Computer Society*, *42*(3), 60-66.

McKnight, D. H., Carter, M., Thatcher, J. B. & Clay, P. F. (2011). Trust in a specific technology: an investigation of its components and measures. *ACM Transactions on Management Information*

Systems, 2(2), 12. Retrieved from http://bit.ly/2eO0uFn (Archived by WebCite® at http://www.webcitation.org/6ieqpQ2cC)

McKnight, D. H. & Chervany, N. L. (2001). What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology. *International Journal of Electronic Commerce, 6*(2), 35–59. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.6110&rep=rep1&type=pdf (Archived by WebCite® at http://www.webcitation.org/6ieqt5FK4)

McKnight, D. H., Choudhury, V. & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: an integrative typology. *Information Systems Research, 13*(3), 334–359. Retrieved from http://www.bus.iastate.edu/mennecke/434/S05/TrustScaleISR.pdf (Archived by WebCite® at http://www.webcitation.org/6ieqyxAHr)

Maddux, J.E. & Rogers, R.W. (1983). Protection motivation and self-efficacy: a revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology, 19*(5), 469–479. Retrieved from http://bit.ly/2f4VoJX (Archived by WebCite® at http://www.webcitation.org/6ieqJPbOV)

Malhotra, N. K., Kim, S. S. & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. *Information Systems Research, 15*(4), 336–355. Retrieved from http://bit.ly/2fcU5aX (Archived by WebCite® at http://www.webcitation.org/6ieqcyzdm)

Mamlin, B. W. & Tierney, W. M. (2016). The promise of information and communication technology in healthcare: extracting value from the chaos. *The American Journal of the Medical Sciences, 351*(1), 59–68.

Martin, K. E. & Nissenbaum, H. (in press). Measuring privacy: an empirical test using context to expose confounding variables. *Columbia Science and Technology Law Review*. Retrieved from http://bit.ly/2eP5TNH (Archived by WebCite® at http://www.webcitation.org/6jAyvHJ44)

Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An integrative model of organisational trust. *Academy of Management Review, 20*(3), 709–734. Retrieved from http://bit.ly/2es4Pko (Archived by WebCite® at http://www.webcitation.org/6ieqkYljv)

Nissenbaum, H. (2001). Securing trust online: wisdom or oxymoron? *Boston University Law Review, 81*(3), 635–664.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1), 101-139. Retrieved from http://bit.ly/2es5R03 (Archived by WebCite® at http://www.webcitation.org/6ier4s94b)

Ohm, P. (2010). Broken promises of privacy: responding to the surprising failure of anonymisation. *UCLA Law Review, 57*(6), 1701-1777. Retrieved from http://paulohm.com/classes/techpriv13/reading/wednesday/OhmBrokenPromisesofPrivacy.pdf (Archived by WebCite® at http://www.webcitation.org/6ierAanR8)

Pallot, M., Trousse, B., Senach, B. & Scapin, D. (2010). Living lab research landscape: from user centred design and user experience towards user cocreation. In First European Summer School "Living Labs". Paris: HAL. Retrieved from http://bit.ly/2fcXg2u (Archived by WebCite® at http://www.webcitation.org/6idnlaVPk)

Richards, N. M. & Hartzog, W. (in press). Taking trust seriously in privacy law. *Stanford Technology Law Review*. Retrieved from

http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2655719_code1107005.pdf (Archived by WebCite® at http://www.webcitation.org/6jAz59rWP)

Rieh, S. Y. (2004). On the Web at home: information seeking and Web searching in the home environment. *Journal of the American Society for Information Science and Technology*, *55*(8), 743–753. Retrieved from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/35293/20018_ftp.pdf?sequence=1 (Archived by WebCite® at http://www.webcitation.org/6ierDxR8Z)

Russell, D. M. & Oren, M. (2009). Retrospective cued recall: a method for accurately recalling previous user behaviors. In Ralph R. Sprague, (Ed.). *42nd Hawaii International Conference on System Sciences, 2009. HICSS '09* (pp. 1-9). Los Alamitos, CA: IEEE. Retrieved from http://bit.ly/2dTDHtV (Archived by WebCite® at http://www.webcitation.org/6idoXbJLx)

Smith, C.L. (2011). *Conditions of trust for completely remote methods: a proposal for collaboration.* Paper presented at the Fifth Workshop on Human-Computer Interaction and Information Retrieval (HCIR '11), Mountain View, California. Retrieved from http://tinyurl.com/hn8kpyl (Archived by WebCite® at http://www.webcitation.org/6jb0AfP2o)

Smith, C.L. (2013). Factors affecting conditions of trust in participant recruiting and retention: a position paper. In Proceedings of the 2013 Workshop on Living Labs for Information Retrieval Evaluation, San Francisco, California (pp. 13-14). New York, NY: ACM.

Smith, C.L. (2016). Technical report: development of 14 dichotomous scale items measuring attitudes in the intent to participate in research. Kent, OH: Kent State University. Retrieved from http://bit.ly/2dGqC47 (Archived by WebCite® at http://www.webcitation.org/6jb0KlJ1r)

Spence, P. R., Lachlan, K. A. & Rainear, A. M. (2016). Social media and crisis research: data collection and directions. *Computers in Human Behavior, 54*, 667–672.

Stunkel, L. & Grady, C. (2011). More than the money: a review of the literature examining healthy volunteer motivations. *Contemporary Clinical Trials, 32*(3), 342–352.

Stutzman, F., Gross, R. & Acquisti, A. (2013). Silent listeners: the evolution of privacy and disclosure on facebook. *Journal of Privacy and Confidentiality*, *4*(2), 7-41. Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1098&context=jpc (Archived by WebCite® at http://www.webcitation.org/6ierJ4u9h)

USEWOD2012 (n.d.). International Workshop on Usage Analysis and the Web of Data. Retrieved from http://usewod.org/usewod2012.html (Archived by WebCite® at http://www.webcitation.org/6idofF7Oy)

Wang, Y., Min, Q. & Han, S. (2016). Understanding the effects of trust and risk on individual behavior toward social media platforms: a meta-analysis of the empirical evidence. *Computers in Human Behavior, 56*, 34–44.

Yang, G. H. & Soboroff, I. (2016). Privacy preserving IR 2015: a SIGIR 2015 workshop. *SIGIR Forum*, *49*(2), 98–101. Retrieved from http://www.sigir.org/files/forum/2015D/p098.pdf (Archived by WebCite® at http://www.webcitation.org/6idowKjT3)

Appendices

Appendix A: Mock recruiting email

Dear [university name] student,

I am a professor in [school name] here at [university name]. You can learn more about me at this link: [url].

I invite you to participate in a study about how people search for information on the Internet. Your participation is completely voluntary.

LabV: The study involves a two-hour appointment in a research lab at [university name]. You can be in the study if you are at least 18 years old.

OwnV: The study will take place on your own computer (excluding iPads). You can be in the study if you can install software on your machine and you are at least 18 years old.

In the study, I will ask you to:
Fill out a brief online survey with some background information about you.
LabV: Download some software that runs in the browser. The software records your searches and the webpages you open. The recordings are kept on the computer and you can control what is saved. You can learn more about the software at this link: [url].
OwnV: Download some software that installs on your browser. The software records your searches and the webpages you open. The recordings are kept on your computer and you can control what is saved. You can learn more about the software at this link: [url].
Install the software and turn it on.
LabV: Complete eight assigned web searches.
OwnV: Complete eight assigned web searches within one week of installing the software.
Send the recordings to me.

LabV: As thanks for being in the study, I will email you an electronic Amazon gift card. I will give you $2 for coming to your appointment,

$1.50 for each of the assigned search you complete, $6 for completing all eight searches, and $20 for releasing the records on the computer. The most you can get is $2 + $12 + $6 + $20 = $40.

OwnV: As thanks for being in the study, I will email you an electronic Amazon gift card. I will give you $2 for the first day that the software is turned on, $1.50 for each of the assigned search you complete, $6 more for completing all eight searches, and $20 for sending the records on your computer. The most you can get is $2 + $12 + $6 + $20 = $40.

LabV: You can stop being in the study at any time and you will be paid for the activities you have completed.
OwnV: It is expected that the study will take about two hours. You can stop being in the study at any time and you will be paid for the activities you have completed. When you stop, or within a week of you sending your records, I will send you an email reminding you to remove the software.

After you are paid, the association between your records and your identity will be permanently broken. I will analyze and report on the records only after combining them with records from others in the study. I will disclose your identity to others only in order to pay you.

To learn more about the study, or to sign up, visit this website: [url].

I would be happy to answer any questions you have. Please email me, Dr. Casey Jones, at cjonesxxx@[university name].edu or 330-999-9999. This study has been approved by [university name] Institutional Review Board (IRB). You can contact the IRB at xxx@[university name].edu or xxx-xxx-xxxx.

Thank you for thinking about being in the study.

Sincerely,
Casey Jones, Ph.D.
[school name], [university name]

Appendix B: Mock disclosure and consent form

Consent Form for Participation in a Research Study about Searching for Information on the Internet [university name]

WHAT IS THE STUDY ABOUT?
The study helps us better understand how people search for information on the Internet.

WHERE WILL THE STUDY TAKE PLACE?
LabV: You will come to a research lab at the [location] on the [university name] campus.
OwnV: You will do the study on your computer where you will install software in your Web browser.

WHO IS ELIGIBLE TO PARTICIPATE?
LabV: To be in the study, you must be at least 18 years old and you must be able to use a standard desktop computer with flat screen

monitor and no audio speakers.
OwnV: To be in the study, you must be at least 18 years old and you must own a computer where you can install the software. You cannot use an iPad for this study. You must have the Firefox or Chrome browser on your computer in order to install the software.

HOW LONG WILL THE STUDY LAST?
LabV: The study will last for about two hours.
OwnV: The study will last no more than 1 week. During that time, I will ask you to complete eight assigned searches during that time. You may quit the study at any time and receive any payment you are owed for tasks completed before you quit.

WHAT WILL I BE ASKED TO DO?
I will ask you to do the following:
Click below to agree to be in the study.
LabV: Schedule an appointment to come to the lab.
Fill out a brief survey with background information about you.

LabV:
Receive an appointment confirmation and a secure code in an email from me, Dr. Casey Jones.
At your appointment, you will:
Download the study software. The software is called CrowdLogger. More information about CrowdLogger is below.
Install CrowdLogger in either the Firefox or Google Chrome browser and activate it using your secure code.
When you are ready, open the browser where CrowdLogger is installed, turn recording on, and start working on the first assigned search. You will see the search topic in a special window in the browser. Use the browser as you would normally. You may turn recording off when you are done with an assigned search.
Complete seven more assigned searches with the recording on.

OwnV:
Find your secure code in an email from me, Dr. Casey Jones.
Use a link in the email and your secure code to download the study software. It is called CrowdLogger. More information about CrowdLogger is below.
You must have either a Firefox or Google Chrome browser on your computer. If you need to, install one of these before installing CrowdLogger.
Install CrowdLogger and activate it using your secure code.
When you are ready, open the browser where CrowdLogger is installed, turn recording on, and start working on the first assigned search. You will see the search topic in a special window in the browser. Use the browser as you would normally. You may turn recording off when you are done with an assigned search.
OwnV: Before the end of one week, complete seven more assigned searches with the recording on. When you are done with all of the assigned searchers, use the software to send the recordings to me. You have the option to review and edit the records before sending them. The records will be deleted from your machine after they are sent. Detailed information about the records is below.
Uninstall and delete the software using the uninstall button in

CrowdLogger. Instructions for doing this will be in the email from me and on the CrowdLogger website.
Within one week of you completing or quitting the study, I will send a record of your completed participation and an Amazon gift card for the amount you are due.

LabV: Follow a link in the email to verify that you received the gift card.
OwnV: Follow a link in the email to verify that you received the gift card and that the software is deleted from your machine.

WHAT DOES THE CROWDLOGGER SOFTWARE DO?
When recording is on, the software records the words you use in search engine queries, webpages opened, cursor movements and scrolling, and opening and closing of tabs. You can turn recording on and off. When recording is off, or when you search in private mode, no recordings are made. The recordings are kept on your computer. You can review, edit, and delete the recordings at any time. The software also sends the search topics you are assigned and makes other records while you are working on them. You can learn more about the software at this link: [url].

WHAT IS IN THE RECORDS?
Every record contains a unique identifier, kept so that we can track your participation. The records have your IP address, the date and time the record was made, information you typed or clicked in a search engine, the urls of webpages you opened, buttons you clicked, the position of your cursor on the screen, and the position of the page on the screen (which changes when you scroll). When you work on an assigned search topic, special records are made to identify the assigned search and the times when you start and complete the search.

WHAT ARE THE BENEFITS TO ME FOR BEING IN THIS STUDY?
There are no direct benefits to you except the satisfaction of knowing your participation will help improve search systems.

WHAT ARE THE RISKS TO ME OF BEING IN THIS STUDY?
I believe there are minimal risks to you as a participant in this research study. You may also feel self-conscious about the recoding of your search activities. You may feel frustrated if an assigned search task is difficult. You should be aware that when recording is on, the system records all of your use of the browser, not just activities related to the study. You are responsible for controlling the recording and release of data on your machine. You are also responsible for uninstalling the software at the end of the study. We will provide instructions for uninstalling the software.

CAN I STOP BEING IN THE STUDY?
Continued participation in the study is voluntary. If you agree to be in the study, but later change your mind, you may drop out at any time. There are no penalties or consequences of any kind if you decide that you do not want to participate. You will be paid for the activities you have completed before quitting the study.

HOW WILL MY PERSONAL INFORMATION BE PROTECTED?

I will use the following procedures to protect the confidentiality of your records:

The software must have your permission to send your stored records to me. This is controlled using your secure code. The software can connect only to my research server, and it will send records only when connected to that specific server.

After you grant permission for sending the records, the software will invoke a privacy-protection procedure. The procedure has two steps:

First, it finds the records made while you were working on the assigned searches. All of those records are sent to the server.

Second, it looks at all the other records on your machine and at the records on the machines of the other participants in the study. It then deletes any records from your machine if words in your records (words include letters, numbers and punctuation) have been used by less than four other participants. Through this step, records containing words used by fewer than four other participants will not be sent. This lowers the risk that someone could use your rare words to identify you personally. You can learn more about the details of the software and this procedure at this link: [url]

Once sent to me, your records will be password protected and kept on encrypted secure computers on the [university name] campus or on secure, encrypted and password protected cloud computers contracted by [university name]. Only certified research team members will have access to the passwords and the data.

All information that identifies you personally will be kept separate from your search activity records.

After I determine your level of participation (see more on level of participation below) and you confirm that you have your payment, the ID that links your search records to your identity will be permanently deleted. I will not begin to review your search records until this ID has been deleted.

After data collection ends, I will combine your records with those from others in the study. After analyzing the records, I may publish the results of the study. Information will be presented in summary form and I will not identify you personally.

WILL I RECEIVE ANY PAYMENT FOR TAKING PART IN THE STUDY?

I will pay you for your level of participation in the study. This includes installing and activating the software, finishing the assigned searches with the recording turned on, and sending your records.

At the end of the study, I will e-mail you an account of your participation and, depending on your level of participation, an electronic Amazon gift card. The amount of your payment will be no more than $40, and will be calculated as follows:

$2: For the first day the software is turned on

$1.50: For each assigned search completed with recording turned on, up to $12 total

$6: For completing all eight assigned searches with recording turned on

$20: For sending the records on your computer
The most you can get is $2 + $12 + $6 + $20 = $40.

WHAT IF I AM INJURED?
[university name] does not have a program for compensating subjects for injury or complications related to human subjects research, but I will assist you in getting treatment.

WHAT IF I HAVE QUESTIONS?
I will be happy to answer any questions or concerns you have about any part of the study or the software. Please e-mail me, Dr. Casey Jones, at cjonesxxx@[university name].edu or call me at (xxx) xxx-9999. If you have any questions about your rights as a research subject, you can contact the [university name] Institutional Review Board (IRB) at (xxx) xxx-xxxx or xxx@[university name].edu.

SUBJECT STATEMENT OF VOLUNTARY CONSENT
By clicking “I consent” I am agreeing to enter this study voluntarily. I have had a chance to read this consent form, and it was explained to me in a language that I use and understand. I have had the chance to ask questions and have received satisfactory answers. I understand that I can stop being in the study at any time. I am aware that I can get a copy of this consent form by going to this website: [url].

Appendix C: Data cleaning.

This appendix describes the data cleaning process for 404 respondent records. The steps used are outlined in Table C.1.

No. of responses	step	condition
404		Total responses received
95	1	Whole survey completed in under 7 minutes
37	2	One or more blocks completed in under 12 seconds
74	3	Straight-line responses on 2 or more matrices
68	4	Very unlikely, unlikely, somewhat unlikely to open or read email
20	5	Inconsistent or incomplete responses in Block 3
110		Valid responses for analysis

Table C.1: Steps in data cleaning before analysis

The first step removed 95 records for questionnaires completed in less than seven minutes. The 7-minute cut-off was derived by trial runs of the survey, examination of completion times within blocks, and inspection of distributions. In step 2, 37 records were removed where any single block was completed in fewer than 12 seconds, with the cut-off similarly derived. The third step removed an additional 74 records where two or more question matrices contained straight-line answers (the same value for every response). Step 4 removed 68 records where the respondent indicated that it was highly unlikely, unlikely, or somewhat unlikely that they would open or read an email with the subject line. Step 5 removed another 20 responses where

only skips or refusal were recorded after Block 2 or where Block 3 was incomplete, as well as records containing incomplete or possibly insincere answers, as detailed next.

In our study, respondents faced no consequences for insincere responses, which may result in responses with little or no internal consistency. We removed insincere responses, as follows. Block 3 contained questions on willingness to engage with the communications and to participate. Figure C.1 depicts the decision tree used for excluding records with inconsistencies in responses to those questions. The rules removed any record where the response indicated opening or reading the email was unlikely (EmailOpen or EmailRead), or where the respondent reversed a prior negative response (no) to positive (yes) or unsure (unsure) in a subsequent response (LearnMore, Participate, TakeAction). With this approach, we assumed that those who demonstrated consistency were the most likely to have provided considered and honest responses to all of the questions.

The remaining records were further prepared by recoding reversed questions and classifying the respondents according to their final intent to volunteer (closed, open, unsure). Before beginning analysis, we completed a final check for bias created by data cleaning by comparing answers to questions administered before the introduction of the h-study. Chi-squared analysis found no significant differences.

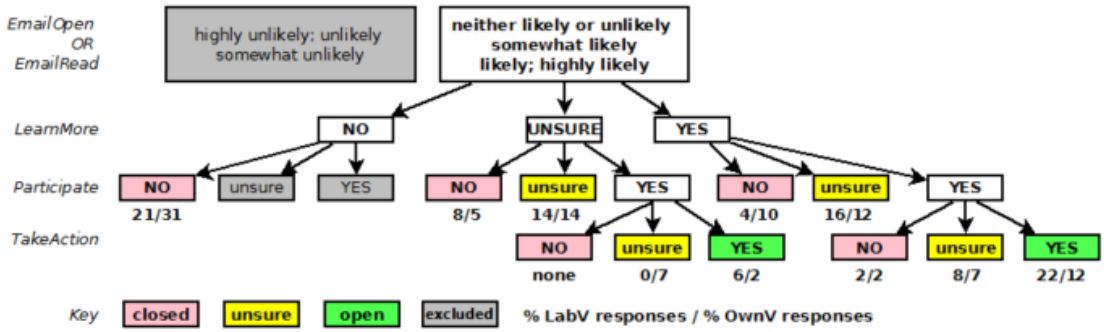

Figure C.1: Data cleaning before analysis, with coding for final intent to volunteer, and for each h-study, percentage retained records at each coding point.ß
