Received 22 July 2015
Revised 28 October 2015
11 November 2015
Accepted 11 November 2015

Sharing personal genetic information: the impact of privacy concern and awareness of benefit

Don Heath
College of Business, University of Wisconsin – Oshkosh, Oshkosh, Wisconsin, USA
Ali Ardestani
Division of General/GI, Department of Surgery, Brigham and Women's Hospital, Harvard University, Boston, Massachusetts, USA, and
Hamid Nemati
Department of Information Systems and Operations Management, The University of North Carolina – Greensboro, Greensboro, North Carolina, USA

Abstract

Purpose – Human genomic research (HGR) demands very large pools of data to generate meaningful inference. Yet, the sharing of one's genetic data for research is a voluntary act. The collection of data sufficient to fuel rapid advancement is contingent on individuals' willingness to share. Privacy risks associated with sharing this unique and intensely personal data are significant. Genetic data are an unambiguous identifier. Public linkage of donor to their genetic data could reveal predisposition to diseases, behaviors, paternity, heredity, intelligence, etc. The purpose of this paper is to understand individuals' willingness to volunteer their private information in this high-risk/high-reward context.

Design/methodology/approach – The authors collect survey data from 273 respondents and use structural equation modeling techniques to analyze responses.

Findings – The authors find statistical support for our theorization. They find that while heightened awareness of the benefits and risks of sharing correlates with increased privacy concerns, the net impact is an increase in intention to share.

Social implications – The findings suggest that prescriptive awareness might be an effective tool with which policy-makers can gain the sufficient voluntary participation from individuals necessary to drive large-scale medical research.

Originality/value – This study contributes a theoretically and empirically informed model which demonstrates the impact of awareness and privacy concern on individuals' willingness to share their genetic data for large-scale HGR. It helps inform a rising class of data sufficiency problems related to large-scale medical research.

Keywords Public policy, TPB, Medicine, Human genomic research, Personal data processing

Paper type Research paper

Journal of Information, Communication and Ethics in Society
Vol. 14 No. 3, 2016
pp. 288-308

1477-996X
DOI 10.1108/JICES-07-2015-0025

1. Introduction

Issues surrounding information privacy are a growing concern to the public. The risks inherent to sharing private information across online information systems are

demonstrable, with spectacular data breaches frequently reported on by the media. For example, the recent data breach at healthcare giant Anthem Inc. is believed to have exposed the names, birthdays, medical IDs/social security numbers, street addresses, email addresses, employment information and income data for 80 million US subscribers – roughly 25 per cent of the population (Riley, 2015). The data breach at Target Corporation data in November of 2013 spilled information on as many as 70 million customers. The stolen information includes names, addresses, email addresses and credit and debit card information including the card verification value (McGrath, 2015). The public is increasingly aware of the potential risks posed by online information systems to their digitized private information.

Human genomic research (HGR) is big data research which is largely reliant on the willingness of individuals to volunteer their genetic information for digitization and sharing with researchers. The information privacy risks associated with this act are significant, as the information being shared is uniquely permanently and personally identifying. Any public linkage of a donor to their genetic sample is unambiguous, potentially revealing that individual's predisposition to certain diseases and behaviors, paternity, heredity, intelligence, etc. Despite the inherent risk to donors, voluntary donation of genetic information is the engine that facilitates large-scale genomic research.

Discovery arising from HGR is elevating scientific understanding and driving innovation in areas such as biotechnology, pharmacogenomics and personalized medicine. HGR holds the potential to improve the human condition. Absent sufficient voluntary donation, HGR risks potential data starvation. The aim of this study is to illuminate the role of awareness and privacy concern as antecedents to willingness to share personal genomic data with researchers, and to inform strategy to encourage willingness to share. According to the most recent UNESCO Broadband Commission report, there will be over 3.2 billion internet users by the end of 2015 (Development, 2015). This represents an 8 per cent increase over the prior year. They estimate over 50 per cent of the world's population will be online by 2018. The steady rise in global interconnectivity suggests the opportunity for large-scale research requiring voluntary participation from individuals will only grow in importance, increasing researcher interest in the delicate balance between information privacy concern, awareness, and individuals' willingness to share their personal data. This research helps inform a rising class of data sufficiency problems.

The aim of this study is to understand individuals' willingness to volunteer their private information for research in the high-risk/high-reward context of HGR. We argue that publicity surrounding promising scientific breakthroughs, discoveries or application of genetic research might heighten individual awareness regarding the benefit of volunteering one's genetic data for research. Similarly, publicity regarding risks to information privacy associated with the use of online information systems might impact an individual's attitude toward volunteering their own private information across these systems. Thus, we consider the impact of *awareness* regarding the benefits of HGR and *individual privacy concern* on individuals' willingness to share their genomic information for research. *Awareness* and *individual privacy concern* each have implication in the shaping an individual attitudes and motivations to share their genetic information with researchers.

To link individual attitudes to behavioral intentions, this study relies upon and extends one of the most robust and well-established theories for understanding behavioral intention, the Theory of Planned Behavior (TPB) (Fishbein and Ajzen, 1975). While many extensions to the TPB have been proposed, none that we are aware of have considered *awareness* and *privacy concern* as antecedents to behavioral intention. We extend the TPB to include these antecedents and examine the predictive power of the new model regarding individuals' intention to voluntarily share personal information.

This balance of the paper is organized as follows. The second section provides background regarding the issues surrounding HGR. The third section outlines the theoretical model and provides a review of literature surrounding the theories and constructs developed and applied throughout the paper. The forth section introduces the research methodology, and the fifth section the results. The final sections are dedicated to discussion of the research findings, conclusion and the implication of this research toward future studies.

2. Review of literature

2.1 Information privacy

Information privacy has received considerable treatment in recent information systems (IS) literature. Bélanger and Crossler (2011) define information privacy as the interest people have in controlling, or at least significantly influencing, the handling of information about themselves. We adopt their definition for this study. In a recent interdisciplinary review of privacy, Smith *et al.* (2011) highlight two perspectives of privacy. The first is value-based and conceptualizes privacy as an innate societal right or a commodity belonging to the individual which might be bargained away in exchange for some perceived benefit. This view might explain why a consumer would offer their demographic data in exchange for merchant coupons or discounts. In the context of healthcare, it might explain why some patients would voluntarily consent to allow their healthcare providers to store their clinical data as an electronic medical record (EMR). The perceived benefits here would be increased accuracy, reducing switching costs and less expensive care.

The second perspective of privacy is based in cognition. In this conceptualization, privacy is concerned with autonomy or control by the individual, of access to self (Smith *et al.*, 2011, 1996). The cognitive perspective has gained significant traction among researchers, as it lends itself more readily to concepts of information privacy and individual agency.

One phenomenon commonly observed in privacy-related research is that of the privacy paradox. Researchers often find that even though individuals profess strong privacy concerns, they readily submit their personal information as a unit of transactional exchange. The privacy paradox is an artifact of the value-based view of privacy as a commodity. Smith *et al.* (2011) argue that the privacy paradox might present a substantive threat to *theory of reasoned action* (TRA) research, as TRA measures individuals' professed intentions rather than their actual behaviors. However, a reasonable defense against the privacy paradox, in this study, can be drawn from economic theory. In cases where the benefit derived from an action is immediate and the potential negative consequence of that action is in the future, bounded rationality motivates the individual to weight more heavily the value derived from the temporally local event, while undervaluing any future risk. In this study, the reward for sharing

one's genetic information for research is neither immediate nor direct, making it less likely the privacy paradox would hold. However, should it hold, the privacy paradox would cause privacy concern to be overstated in relation to the actual sharing behavior of individuals. This would not threaten the ability of this study to inform strategies meant to induce sharing of personal genomic data.

Numerous IS research has operationalized privacy concern as an independent variable antecedent to behavior. A preponderance of this research adopts the view put forth by Smith *et al.* (1996) that individuals' concern for the information privacy practices (CFIP) of organizations defines the privacy concerns of individuals (Palvou, 2011). However, Malhotra *et al.* (2004, p. 337) define privacy concern more broadly as an "individual's subjective views of fairness within the context of information privacy". This definition provides the basis for the "internet user's information privacy concerns construct", whose central dimensions are data control, awareness and collection. Awareness reflects the degree to which the individual is aware of the organizational privacy practices. Collection regards whether private data collection takes place without the individual's permission. Control concerns the degree to which the individual controls the data that have been collected. Palvou (2011, p. 1020) writes, "Findings suggest information privacy concerns influence individuals' attitudes, such as their preferences for regulatory environments and willingness to be profiled". In this research, we are less concerned with individuals' CFIP. We adopt the posture that individuals' perceptions when dealing with multiple organizational or institutional actors, as might arise should they share their genetic data for research, must be conceived more broadly (Brandimarte *et al.*, 2013).

Several studies directly consider the impact of patient privacy concerns on willingness to "opt-in" and share personal health information (Angst and Agarwal, 2009; Anderson and Agarwal, 2011). Anderson suggests the potential benefit to patients in sharing their EMRs is compelling, promising to reduce medical error, lower healthcare costs, improve patient safety and foster more robust clinical research. However, privacy concerns among patients threaten to slow diffusion of EMR. They argue that while the healthcare context might appear similar to other privacy contexts in which individuals evaluate the costs and benefits of disclosing personal information, the healthcare context is unique in at least two respects. First, the nature and variety of risks inherent in the compromise of sensitive health information are unique. Second, the emotion linked to one's medical state is more intense. Malhotra *et al.* (2004, p. 349) write:

> to have a complete understanding of consumer reactions to information privacy-related issues, researchers should examine not only consumers' privacy concerns at a general level, but also consider salient beliefs and contextual differences at a specific level.

Thus, they make the argument that information privacy behaviors cannot be divorced from their context. We argue that in the high/risk high/reward context of HGR this is particularly true.

2.2 Awareness

Rogers (1995) first introduced awareness to information systems literature with innovation diffusion theory, defining awareness as the degree to which a target population is cognizant of an innovation and formulates a general perception of what it entails. This definition is appropriable but incomplete in the context of HGR. Awareness

in the context of HGR must include awareness of extant innovation, but must also encompass awareness of ongoing research its' potential for discovery. Further, awareness of innovation generally conjures a positive attitudinal valence. HGR awareness might, on the other hand, involve negative issues or conjure negative feelings or perceptions. For example, religious doctrine might cause an individual to believe certain research is inappropriate and violates some doctrine of their faith. This duality is not captured in Rogers' definition. Dinev and Hu (2007, p. 401) describe awareness is "one of the key components of consciousness-raising, and brings about an appreciation of the needs, impetus, and specificity of issues, events, and processes".

Social and medical practitioners have long understood the role of awareness, often initiating public awareness campaigns for the purpose of motivating individuals or groups to behave in a certain way (Rutten *et al.*, 2012). Awareness is well-understood as a motivator of behavior, and is implicit in behavioral models such as the TRA and its derivatives, where awareness is involved in the formation of attitudes.

Much research has been done in the realm of information security research to develop awareness as a construct of behavior modeling, particularly with regard to a user's willingness to tolerate additional security measures if they are made aware of credible security threats. A recent study by Spears and Barki (2010) considers individual awareness and its impact on behavior in the context of information systems security control design. They find individuals' heightened awareness regarding the threats and benefits of these controls greatly impact their behavior, and consequently the efficacy of the controls.

Similarly, Dinev and Hu (2007) develop the construct of technology awareness to explain user behavior around technological issues, defining technology awareness as users' raised consciousness of, and interest in knowing about, technological issues and strategies to deal with them. They argue for an individual to form either positive or negative beliefs about a particular technology, they must first be made aware of the issues surrounding that technology. As their study deals with protective technologies, technology awareness includes understanding potential threats, consequence of having no protective technology, availability of various protective technologies and the effectiveness of these technologies in mitigating risk.

In this study, we adopt the purposely broad definition of HGR awareness as *the degree to which an individual is cognizant of HGR and formulates a general perception of what it entails*. We take this to include the issues, implications, discoveries, potential and societal impact of HGR, and to capture the duality of its attitudinal valence.

3. Theory to predict voluntary sharing of information by individuals

The focus of this paper is on the external variables *individual privacy concern* and *awareness* on the voluntary sharing by individuals of their genetic information with researchers. To create a linkage between *information privacy concern*, *awareness* and *behavioral intention*, we rely upon the TPB. As the TPB is commonly used and well known throughout IS literature, we omit extensive discussion of the theory's fundamentals, and refer the reader to the seminal references instead (Fishbein and Ajzen, 1975; Ajzen, 1991, 2002).

3.1 Foundational model – the theory of planned behavior

We first present a foundational model of the TPB adapted to intention to voluntarily share genetic information for research (Figure 1).

Many TPB studies have demonstrated that attitude toward a behavior is often the strongest single predictor of a behavior. For modeling clarity, attitude in the context of this study is a positive attitude toward HGR. Therefore, we hypothesize the following:

H1. Attitude will have a positive impact on behavioral intentions to share genetic information.

Perceived behavioral control (PBC) denotes the subjective degree of control over the performance of a behavior, not the perceived likelihood that the behavior will produce a given outcome (Ajzen, 2002). With regard to voluntarily sharing genetic information, PBC is the individual's perceived control over sharing or not sharing their genetic information. PBC is a co-determinant of behavioral intention, and subsequently behavior. TPB studies have repeatedly demonstrated that increased perceived control has a positive impact on behavioral intentions. This leads us to the following hypothesis:

H2. Perceived control will have a positive impact on behavioral intentions to share genetic information.

Subjective norm, in TPB, reflects one's desire to act based upon their perception of how others might think or act. The greater an individual perception that significant others think he/she should think or act a certain way, the greater an individual's level of motivation to comply with those others. This is also known as the perceived prevalence effect or the "everyone does it concept" (Simpson *et al.*, 1994, p. 432), where individuals act according to the expectations of others. Subjective norm (SN) in this context reflects an individual's willingness to share their genetic data based on how they believe others might expect them to act, how their actions might shape others' opinion of them or how they believe others might act in the same situation:

H3. Subjective norm will have a positive impact on behavioral intentions to share genetic information.

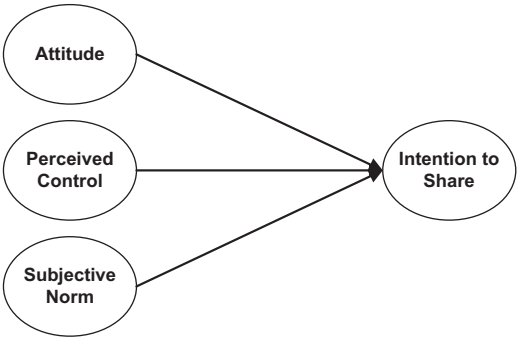


Figure 1.
TPB adapted to intention to share

3.2 Integrating awareness and information privacy into the model

When an individual volunteers their genetic information, they are not sharing tissue, or test results; they are sharing personal information. A great body of work exists that explores the individual's willingness to share personal information; however, due to the unique properties of genetic data, previous research regarding information sharing might not be able to fully explain the complexities associated with the individuals decision to share or withhold their information from researchers in this context. It is unclear the degree to which the public is aware of the risks and benefits associated with sharing genetic information; however, the degree of such awareness might inform and shape an individual's beliefs regarding that sharing.

Publicity surrounding exciting genetic research or any reported misuse of genetic information might greatly influence an individual's attitude toward sharing their genetic information. Similarly, heightened awareness regarding the potential for compromise of one's genetic data might influence intention to share genetic information. Thus, we treat *awareness* and *privacy concern* as antecedents to the TPB, improving its ability to more accurately explain behavioral intention within the context of willingness to share one's genetic information.

While the TPB constitutes a robust framework for understanding and predicting social behaviors, Ajzen has suggested that the TPB should be open to the inclusion of additional variables if it can be shown that they capture a significant proportion of the outcome variance. Numerous researchers have proposed the addition of new predictors to improve the explanatory power of the original TPB (Conner and Armitage, 1998; Parker *et al.*, 1995). We propose expanding the TPB by adding *awareness* and *privacy concern*, conceptualizing them as requisite to the formation of behavioral intention and therefore antecedents to the entire TPB in the context of sharing sensitive personal genetic information (Figure 2). Our primary goal is to explain an individual's willingness to share genetic information by building an extended version of the TPB that is better able to predict human behavior within this context. Ajzen himself provides warrant for this extension by suggesting that an individual's beliefs can be influenced by persuasive communication. As persuasive communication heightens awareness, one can assume that awareness is tacit in his argument and is necessarily an antecedent to the individual's formation/reformation of their beliefs. Absent awareness, one would have no behavioral intention. Further, he writes that attitude is the sum of the individual's beliefs, and that attitude plays a major role in formation of behavioral intention, and consequently behavior. Ajzen's notion that beliefs can be influenced by persuasive communication recognizes the malleability of beliefs.

3.2.1 Awareness. Heightened awareness regarding genomic research increases an individual's likelihood of recognizing the great benefit they might derive in terms of personalized medical care, or pharmacogenomic drug therapies. Awareness necessarily shapes behavioral intention by virtue of its influence on beliefs and attitudes. An individual with some degree of awareness regarding the benefits associated with sharing of their genetic information will be more likely to form a positive behavioral intention regarding sharing their genetic information for research. This leads us to the following hypothesis:

H4. Awareness will have a positive influence on behavioral intentions to share genetic information for research.

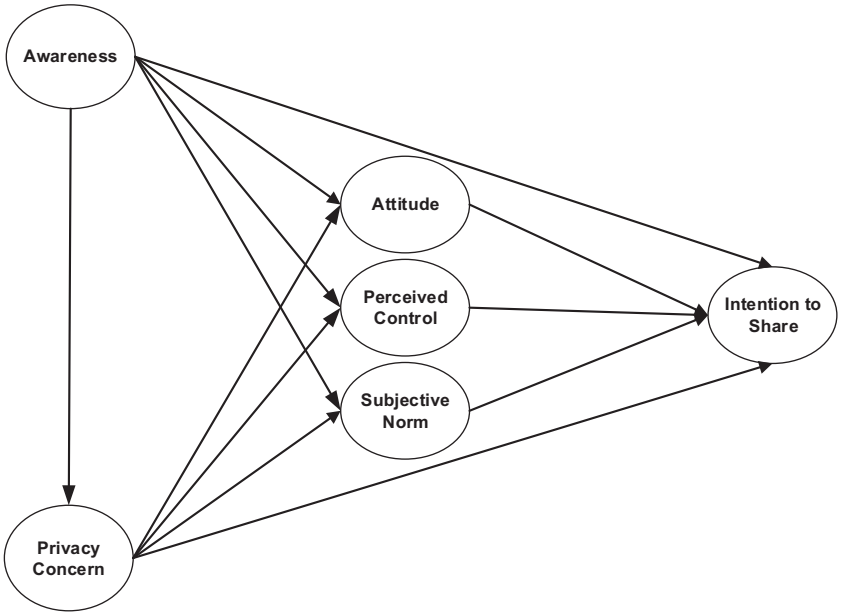


Figure 2.
Fully extended theoretical model

3.2.2 Information privacy. Previous research suggests that individuals with higher levels of concern about information privacy may be more likely in the future to refuse to participate in activities that require the provision of personal information (Stone *et al.*, 1983). An individual's genetic data represent the apex of privacy concern, as it is personally identifying and harbors potential for inescapable discriminatory sanction. This leads us to the following hypothesis:

H5. Privacy concerns will have a negative influence on behavioral intentions to share genetic information.

It has been suggested that media coverage may increase the level of concern about information privacy (Westin, 1967). Individuals who closely follow media reports of breaches of privacy or security might be more sensitive to the possible consequences of a loss of privacy due to accidental, malicious or intentional leakage of personal information. Consequently, one might posit, with respect to privacy concerns, that individuals with heightened awareness will tend to know more about the privacy debate, privacy policies and privacy risks associated with information sharing. These users will have a better understanding of the importance of data privacy. Therefore, individuals with heightened awareness are expected to be more concerned about privacy. This leads us to the following hypothesis:

H6. Awareness will have a positive influence on privacy concerns regarding intentions to share genetic information.

3.3 Full extended model

Attitude toward a particular behavior is the degree to which an individual has a positive or negative evaluation of the behavior. The TPB predicts that the more favorable one's

attitude toward a behavior; the greater one's intention to perform the behavior. Ajzen describes attitude as the sum of the individual's beliefs, with attitude playing a major role in formation of behavioral intention, and consequently behavior. With regard to the focal behavior, an individual's attitude toward sharing genetic information can be expressed as the sum of his or her beliefs. Ajzen posits that an individual's attitudes remain dynamic, subject to revision. Consequently, awareness might shape attitudes by virtue of its influence on beliefs. Therefore, we hypothesize the following:

H7. Awareness will have a positive influence on attitude toward sharing genetic information.

In contrast, information privacy concerns will foster a negative attitude toward sharing of one's individual genetic information for research. The anticipation of potential negative consequences would shape attitude with regard to the behavior of sharing. This leads us to hypothesize the following:

H8. Privacy concern will have a negative influence on attitude toward sharing genetic information.

Heightened awareness with regard to the target behavior will cause an individual to be more attuned to the circumstances in which an individual can express control, and the circumstances in which they cannot. This leads us to the following hypothesis:

H9. Awareness will have a positive influence on perceived control over sharing genetic information.

Privacy represents the control of transactions between person(s) and other(s), the ultimate aim of which is to enhance autonomy and/or minimize vulnerability. Fears associated with unauthorized access to one's genetic information, or unauthorized secondary use of that information, diminish the perception of controllability, not of the data itself, but of the target behavior; sharing one's genetic information with another entity with consent. Controllability is a principal construct of PBC. This leads us to the following hypothesis:

H10. Privacy will have a negative influence on perceived control over sharing genetic information.

The greater an individual perception that significant others think he/she should think or act a certain way, the greater an individual's level of motivation to comply with those others. Subjective norm is therefore concerned with co-orientation of individuals to their environment and represents a powerful survival instinct (Newcomb, 1953). In the case of *individual privacy concern*, individuals would ascribe greater value to opinions of others which were similar to their own. Similarly, in the case of *awareness*, heightened awareness regarding the benefit of sharing one's genetic information with researchers might promote the view by the individual that such behavior is looked upon favorably by their community of peers:

H11. Awareness will have a positive influence on subjective norm regarding sharing genetic information.

H12. Privacy concerns will have a negative influence on subjective norm regarding sharing genetic information.

We operationalize the theoretical model using the methodology that follows.

4. Methodology

A survey instrument was developed to collect quantitative data for modeling and testing hypotheses. Our survey instrument uses probabilistic Likert-style scales of semantic-differential type, which tend to yield highly reliable measures of strength of intention. Scale items are adapted from existing instruments using a five-level interval scale, with anchors from *strongly disagree* to *strongly agree* used to measure each item.

Prior to distribution, the survey instrument used in the study was pretested in a pilot study with 50 individuals. The pilot study revealed no issues with the instrument. Subsequently, a survey was undertaken to empirically test the research model. To identify a diverse pool of participants, we consulted with alumni and industry partners from our department's advisory board to build an email list comprising friends, contacts and associates. This resulted in the collection of 182 potential candidates. In addition, we randomly selected emails from our own university's email directory to strengthen the size and diversity of our sample. Combining these lists, we built a pool of 350 potential participants. We extended each of these an email invitation from which we collected 273 useable responses. Participation was entirely voluntary. All respondents were asked to click on the Web URL link provided in an invitation e-mail, which linked to the online survey instrument. The respondents were offered no incentive for their participation and were assured the results would be reported in aggregate to insure their anonymity. To assess potential non-respondent bias, characteristics of early-respondents were compared to late-responders' using *t*-test statistics. Results showed no statistical differences between late-responders and early-responders; therefore, supporting the assumption there is no response bias. However, some structural bias is inherent in our method as responses were collected entirely over the internet. As each candidate selected for invitation had a valid email address, the potential impact on non-respondent bias is believed to be minimal.

5. Data analysis and results

Our analysis proceeded in a three-step process described below. First, full measurement and structural analyses on the original TBP model were performed to examine whether our data accurately captured the original theory. In Phase 2, the same analyses were performed on our hypothesized model presented in Figure 1. The last phase of our data analysis was to investigate the impact (if any) of moderating variables of *age*,*education* and *gender* on our hypothesized model. A variance-based structural equation modeling using the partial least squares (PLS) method which relies on component-based estimation to maximize the variance explained in the dependent variable was used. PLS allows for assessing psychometric properties of the "measurement" model and estimating the parameters of the "structural" model. PLS does not require multivariate normality of the data and accommodates smaller sample size (Chin, 1998; White-Baker *et al.*, 2007). SmartPLS 2.0, M3 was used for data analysis and model confirmation. This tool accommodates analysis of up to 200 indicator variables and allows examination of interactions among moderator and latent predictor variable indicators. All of the demographic indicators mentioned had a *t*-effect of less than 0.3 (Table I).

Figure 3 shows the structural model results for original TPB constructs measured. The results of this phase of our analysis as shown in Figure 3 confirmed that our data fully supported the original TBP model.

5.1 Measurement model

In the second phase of data analysis, the hypothesized research model was fully analyzed. The measurement model reliability results are provided in Table II. The composite reliabilities of the model measures ranged from 0.70 to 0.91, which exceeded the suggested threshold value of 0.70. The analysis indicates that the measures are quite robust in terms of internal consistency reliability as indexed by composite reliability. In addition, the average variance extracted (AVE) for all measure exceeded 0.50 except those of awareness and perceived control (0.47 and 0.45, respectively).

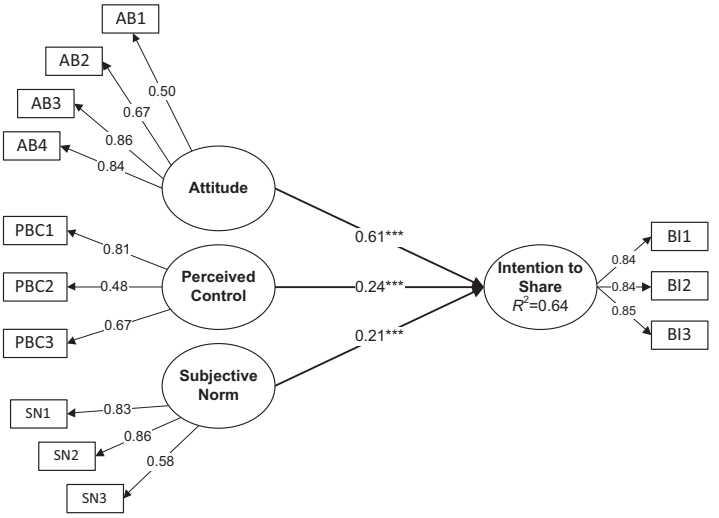
Table III reports the results of testing the latent variable inter-correlation of the measurement constructs. Discriminant validity was assessed by examining whether the squared correlation between a pair of latent variables was less than the AVE of each construct. The diagonal elements of the matrix represent the square roots of the AVE's and are greater than the off-diagonal elements of the corresponding row and column. This supports that discriminate validity of our scales.

We calculated item cross-loadings using Smart PLS by extracting the factor loadings and cross-loadings of all indicator items to their respective latent constructs. These results, presented in Table IV, indicate that each item loaded higher on their principal construct than on other constructs. The factor loading ranged from a lower bound of 0.47 to an upper bound of 0.91. The cross-loading differences were much higher than the suggested threshold of 0.1, with the exception of A1, which was 0.05. Furthermore, each item's factor loading on its respective construct was highly significant ($p < 0.001$) as indicated by the *t*-statistics of the outer model loadings in the Smart PLS output. These values ranged from a low of 3.2 to a high value of 58.1.

The constructs' factor loadings and cross-loadings are presented in Table IV, with the *t*-statistic for each item loading confirming the convergent validity of these

Table I.
Descriptive statistics of survey respondents (n = 273)

Descriptor	Count	% of Total
Gender		
Male	148	55
Female	125	45
Race		
Caucasian	157	58
African-American	60	22
Hispanic	15	5
Asian	21	8
Other	20	7
Age		
Age below 20	113	41
Age 20-40	110	40
Age 40+	50	19
Annual income		
(< 25,000)	217	79
(> 25,000)	56	21
Education		
No college	27	10
Some college	190	70
Bachelors or more	56	20



Notes: *Significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$

Figure 3.
Structural model results of the original TPB model

Construct	AVE	Composite reliability	R^2	Cronbach's alpha
Attitude	0.55	0.82	0.17	0.75
Awareness	0.47	0.86	0.00	0.81
Control	0.45	0.70	0.11	0.46
Behavioral intention	0.73	0.89	0.65	0.82
Privacy	0.59	0.91	0.43	0.88
Subjective norms	0.63	0.84	0.07	0.71

Table II.
Quality assessment of the measurement model

Construct	Attitude	Awareness	Control	Intention	Privacy	Subjective norms
Attitude	*0.74*					
Awareness	0.38	*0.68*				
Control	0.21	0.29	*0.67*			
Behavioral intention	0.73	0.27	0.36	*0.86*		
Privacy	0.12	0.65	0.32	0.03	*0.77*	
Subjective norms	0.34	−0.01	0.10	0.46	−0.21	*0.79*

Table III.
Latent variable correlation – discriminant validity of variable constructs

indicators as representing distinct latent constructs. These results collectively indicate good measurement properties.

5.2 Structural model

The structural model results, including the beta values of all path coefficients, are summarized in Figure 4.

Item	Attitude	Awareness	Control	Intention	Privacy	Subjective norms
AB1	**0.47**	0.03	0.06	0.26	0.04	0.00
AB2	**0.64**	0.09	0.02	0.43	0.04	0.09
AB3	**0.91**	0.47	0.19	0.65	0.19	0.30
AB4	**0.85**	0.28	0.29	0.72	0.09	0.41
A1	0.05	**0.65**	0.15	0.12	0.64	0.19
A2	0.13	**0.72**	0.15	0.03	0.58	0.17
A3	0.44	**0.76**	0.24	0.37	0.43	0.02
A4	0.24	**0.63**	0.18	0.17	0.39	0.12
A5	0.01	**0.46**	0.13	0.04	0.35	0.10
A6	0.43	**0.76**	0.21	0.36	0.36	0.09
A7	0.35	**0.75**	0.29	0.27	0.48	0.07
BI1	0.67	0.29	0.31	**0.85**	0.01	0.37
BI2	0.60	0.20	0.30	**0.84**	0.03	0.39
BI3	0.66	0.21	0.31	**0.87**	0.03	0.42
PC1	0.05	0.51	0.31	0.04	**0.78**	0.18
PC2	0.21	0.55	0.31	0.11	**0.77**	0.02
PC3	0.12	0.50	0.18	0.01	**0.73**	0.07
PC4	0.01	0.53	0.28	0.05	**0.80**	0.30
PC5	0.12	0.53	0.16	0.01	**0.83**	0.26
PC6	0.03	0.50	0.18	0.09	**0.75**	0.25
PC7	0.17	0.38	0.26	0.13	**0.71**	0.03
PBC1	0.14	0.16	**0.56**	0.35	0.00	0.31
PBC2	0.19	0.30	**0.86**	0.22	0.41	0.10
PBC3	0.06	0.02	**0.56**	0.22	0.06	0.19
SN1	0.34	0.08	0.02	0.42	0.09	**0.78**
SN2	0.39	0.00	0.16	0.49	0.17	**0.86**
SN3	0.04	−0.10	0.03	0.13	0.23	**0.74**

Table IV.
Factor loading (Bolded) and cross loading

Note: the bold indicates the relevant loadings of the items from the leftmost column on their respective constructs (factors)

Awareness had a positive, but insignificant, impact on behavioral intention ($\beta = 0.04$), while as hypothesized, privacy concern had a negative influence on behavioral intention ($\beta = -0.10$, $p < 0.05$). All TPB constructs demonstrated a positive influence on behavioral intention; with a beta of 0.64 for attitude, 0.22 for perceived control and 0.19 for subjective norm ($p < 0.001$ for all). Awareness had a positive influence on attitude ($\beta = 0.52, p < 0.001$) and subjective norm ($\beta = 0.22, p < 0.01$), while privacy concern had a negative influence on attitude ($\beta = -0.21, p < 0.01$), a positive influence on perceived control ($\beta = 0.22$, $p < 0.05$) and subjective norm ($\beta = -0.35$, $p < 0.001$). Overall, a significant proportion of the respective variances in behavioral intention ($R^2 = 0.65$) was accounted for in the model, while $R^2 = 0.17$ for attitude, 0.12 for perceived control and 0.07 for subjective norms. Findings and conclusions are summarized in the Table V.

6. Discussion

A review of existing literature suggests many researchers recognize awareness and information privacy concern as necessary and important constructs in explaining

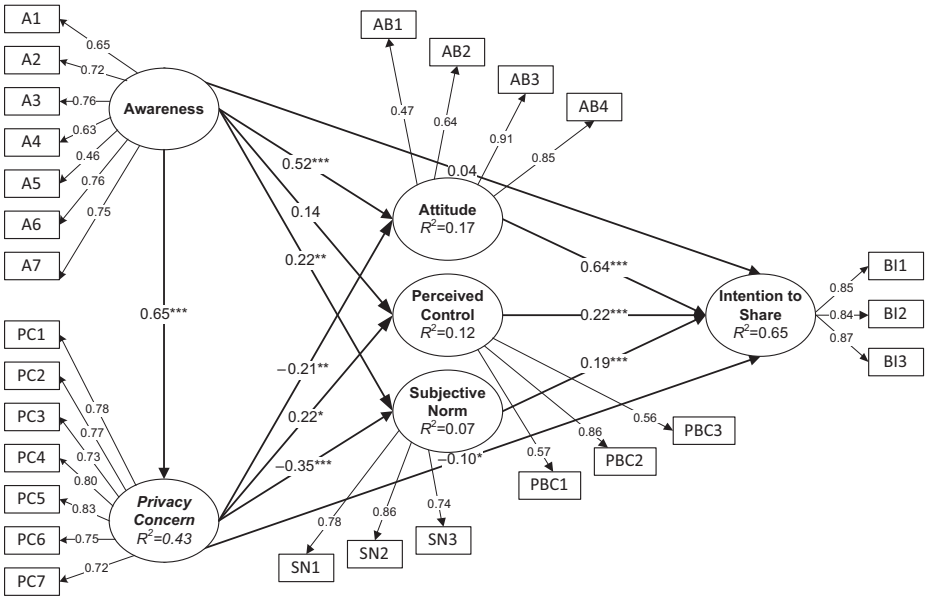


Figure 4.
Structural model results

Notes: *Significant at $p < 0.05$; **significant at $p < 0.01$; ***significant at $p < 0.001$

behavioral intention, although there is considerable divergence regarding their placements within the respective models. We treat awareness and information privacy concern as antecedents to the TPB in its entirety, and additionally to examine their inter-relationship independent of the TPB, believing each alone is important enough to impact the entire model.

Like Jarvenpaa *et al.* (2000), we find a statistically significant negative relationship between information privacy concern and both attitude and behavioral intention. Our results also confirm the findings of Hu and Dinev (2007) with regard to awareness as antecedent to the TPB, demonstrating a positive and statistically significant relationship between awareness and attitude in the moderated model. This supports the role of awareness with regard to voluntary sharing. In total, the results support 9 of the 12 hypotheses. All of the original TPB constructs were well supported, as expected. Contrary to expectation, the model does not support that awareness and information privacy concern significantly impact perceived control, or that awareness in isolation has significant impact on behavioral intention. Numerous papers that have studied the TPB in the context of information sharing have studied only attitude and subjective norm as antecedents of behavioral intention, opting to exclude perceived control (Dinev and Hu, 2007; Lowry *et al.*, 2011). Our findings agree with their assumptions regarding perceived control.

H1, *H2* and *H3* deal specifically with the TPB. The model's results demonstrate the TPB's ability to explain behavioral intention within the context of an individual's intention to share their genetic information. Attitude, perceived control and subjective norm are robust predictors of behavioral intention in this context.

Hypothesis	Finding	Conclusion
H1. Attitude will have a positive influence on behavioral intentions to share genetic information.	Yes: ($\beta = 0.64, p < 0.001$)	Supported
H2. Perceived control will have a positive influence on behavioral intentions to share genetic information.	Yes: ($\beta = 0.22, p < 0.001$)	Supported
H3. Subjective norm will have a positive influence on behavioral intentions to share genetic information.	Yes: ($\beta = 0.19, p < 0.001$)	Supported
H4. Awareness will have a positive influence on behavioral intentions to share genetic information.	Not significant: ($\beta = 0.04$)	Not supported
H5. Privacy concerns will have a negative influence on behavioral intentions to share genetic information.	Yes: ($\beta = -0.10, p < 0.05$)	Supported
H6. Awareness will have a positive influence on privacy concerns regarding intentions to share genetic information.	Yes: ($\beta = 0.65, p < 0.001$)	Supported
H7. Awareness will have a positive influence on attitude toward sharing genetic information.	Yes: ($\beta = 0.52, p < 0.001$)	Supported
H8. Privacy concern will have a negative influence on attitude toward sharing genetic information.	Yes: ($\beta = -0.21, p < 0.01$)	Supported
H9. Awareness will have a positive influence on perceived control over sharing genetic information.	Not significant: ($\beta = 0.14$)	Not supported
H10. Privacy concern will have a negative influence on perceived control over sharing genetic information.	Not significant: ($\beta = 0.22$) non-conforming	Not supported
H11. Awareness will have a positive influence on subjective norm regarding sharing genetic information.	Yes: ($\beta = 0.22, p < 0.01$)	Supported
H12. Privacy concerns will have a negative influence on subjective norm regarding sharing genetic information.	Yes: ($\beta = -0.35, p < 0.001$)	Supported

Table V.
Hypothesis testing results

As noted earlier, our findings did not demonstrate that awareness alone has a direct impact on behavioral intention (*H4*). Awareness does, however, have significant impact on attitude and subjective norm, thus establishing itself as an antecedent of behavioral intention (*H7*, *H11*). Our findings agree with those of Dinev and Hu (2007) regarding the impact of awareness on attitude and subjective norm, however, we could not confirm a statistically significant direct relationship between awareness and behavioral intention. One explanation might be the scope of awareness as a construct in the two works. In their work, awareness was limited to technology awareness, while awareness in our study was more diffuse, dealing with both the social and technological dimensions. We originally hypothesized an individual with a degree of understanding regarding the

potential benefits (to self) of sharing their genetic information for research purposes will be more likely to share that information. However, it might be that increased awareness of HGR might cause individuals to believe instead that the benefit of their donation would flow to others. This warrants additional study. While awareness had a measurable positive impact on perceived control, the impact was not statistically significant (*H9*). Within the model, awareness expresses a significant impact on privacy concern, as hypothesized (*H6*). The impact of awareness on privacy concern conforms to the preponderance of literature, specifically that awareness impacts information privacy concern. The relationship between awareness and information privacy concern accounts for a substantial portion of the variance in information privacy concern, demonstrating the significance of that relationship.

As anticipated, information privacy concern has a negative impact on attitude, subjective norm and behavioral intention (*H8*, *H12* and *H5*). However, its impact on perceived control was statistically insignificant (*H10*). The negative effect of information privacy concern on attitude demonstrated in our findings is similar to that found by Gurung (2006), in his research regarding the impact of information privacy concern on willingness to transact in e-commerce. However, unlike Gurung, who found no statistically significant relationship between information privacy concern and behavioral intention, our findings indicate information privacy concern has a modest but statistically significant impact on behavioral intention. Our findings also indicate a significant negative relationship between information privacy concern and subjective norm, supporting the original hypothesis.

6.1 Limitations and future studies

Although the hypothesized model explains 65 per cent of variability in intention to voluntarily share sensitive genetic data, it can only explain 17, 12 and 7 per cent of variability in attitudes, perceived control and subjective norms, respectively. To improve on these, in Phase 3 of our analysis, we explored the moderating impact (if any) of age, education and gender. Age, education and gender were selected as moderators, as they been shown to moderate the effect of TPB. In addition, a great body of research in the psychology literature has focused on age differences in abilities, traits and performance outcomes (Rhodes, 1983). Morris *et al.* (2000) have established the effect of age on technology adoption and demonstrated the effect of age on different all TPB constructs; they showed that when making the decision to adopt new technology attitudes played a more important role in younger participant, while subjective norms and PBC was more important for older participant. They have also shown a moderating role of age and gender on intention to use technology in their later study, positing that attitude is important to men, while attitude, subjective norm and PBC are all important to women and these differences are even more evident in older individuals (Morris and Venkatesh, 2000; Morris *et al.*, 2005). They mention that the role of moderators have been mostly disregarded, this might be due to different levels of moderation (if any) of these factors on various external variables based on circumstance and the context of the usage of TPB. They further conclude that "it behooves researchers to examine other potential moderators within TPB when applied in other contexts". Education is one of these factors that have been shown to have an impact on behavior. Abu-Shanab (2011) has demonstrated that education level as a moderator of technology adoption.

Age, gender and education were included in most of the previous literature as moderators of original TPB constructs or as antecedents to TPB. The major of focus of these studies were to verify whether these factors have an impact on intention and behavior and therefore should be included in further studies. However, the potential role of these moderators are not fully understood in the context of sharing genetic information, and therefore, we posit that measuring the effect of these factors as moderators of antecedents to TPB constructs (awareness and privacy concerns) would examine the potential role of these factors on the impact of awareness and privacy concerns on TPB. Our hypothesized moderated model is shown in Figure 5.

To model and determine the significance and strength of moderating effect of age, education and gender, we introduced and evaluated the interaction terms in the structural model as discussed in Chin (1998). This approach has been used to understand the moderating effect's direction and the strength of the relation between the predictor variables (in our case, awareness and privacy control) and the dependent variable (s) (in our case, the attitudes, perceived control, subjective norm and intention to voluntarily share genetic information). The interaction terms were calculated by multiplying the corresponding indicators of the predictor and moderator constructs as shown in Table I. Table VI compares the R^2's for the direct effect model with those of the moderated model.

These findings suggest more research is needed around these and other potential moderators to understand the patterns and hypotheses which might arise from their fuller treatment.

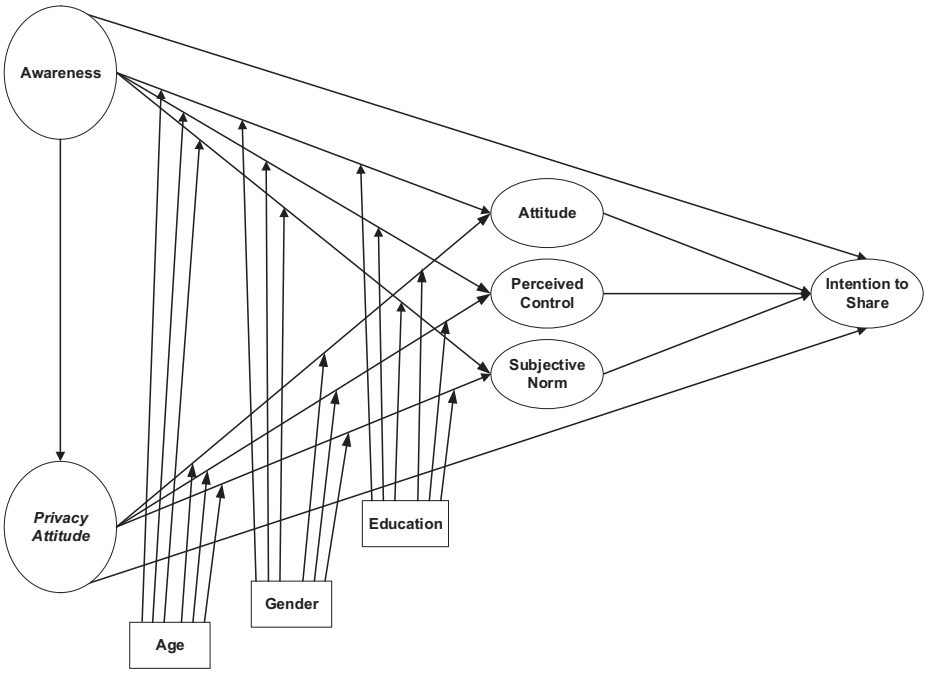


Figure 5.
Hypothesized moderated model

Research on information privacy attitudes is concerned with perceptions of information privacy policies and practices. In this conceptualization, privacy attitudes are seen as influencing behaviors. Many studies around information privacy attitudes are concerned with online information disclosure (Alge *et al.*, 2006; Webster, 1998). A common finding among these studies is that in the presence of privacy attitudes, heightened concern for information privacy no longer influences willingness to disclose personal information. A number of papers confirm that as people's attitudes change in response to being told that fair information practices are used to manage their data, their concern for information privacy is reduced to the degree that they will share their information and not take actions to protect their privacy. Bélanger and Crossler (2011, p. 1021) suggest the majority of research in this area is concerned with *what* leads to a person's reactions to information privacy policies and practices. What is missing, they argue, is *why* individuals react a certain way. They offer a strong call for research to understand why privacy attitudes impact individuals' decisions to disclose their private information.

Despite the robustness of our model, and the empirical support for it, we acknowledge limitations which call for additional research. Although the survey generated 273 responses, and the results were unambiguous, a larger sample size would confirm the validity of our work and offer greater statistical power to the analysis. As all respondents completed the survey online, it is understood each respondent is internet aware, and thus might be more biased toward information technology than those who are not internet savvy. Additional work might be needed to understand how individuals not connected to the internet would respond to our survey. Also, additional work should be done to explore regional, cultural, national and gender specific result sets to affirm the validity of our model with a variety of populations (each has the potential to introduce bias in the study and model interpretation).

7. Conclusions

The goal of this research was to propose and empirically validate a model to explain the relationships between information privacy concern, awareness and the willingness of individuals to share their genetic information for research. While the TPB is an incredibly robust model to explain and predict behavioral intentions, within the context of an individual's willingness to share their genetic information with researchers, the TPB is better able to explain behavioral intention when expanded to include awareness and information privacy concern as antecedents to the model.

The statistical impact of each of these factors as antecedents is far greater than their individual predictive powers. Awareness significantly and positively impacts most components of the TPB. When treated as an antecedent to the model, it is highly predictive of behavioral intention. Heightened awareness increases an individual's intention to share

Construct	R^2 main effect	R^2 Moderated	Effect size (%)	Significant[a]
Attitude	0.17	0.33	24	Moderate
Control	0.12	0.29	24	Moderate
Subjective norms	0.07	0.25	24	Moderate
Intention	0.65	0.63	−5	Weak

Note: [a] significance determined based on Cohen (1988)

Table VI.
Direct effect model vs Moderated model

their genetic information. However, awareness correlates with privacy concerns. Privacy concerns have a negative impact on an individual's willingness to share their genetic information. Despite the correlation between heightened awareness and privacy concern, the net impact of heightened awareness is to increase one's intention to share.

Current and future genomic research demands very large data pools to generate meaningful inference. Therefore, the collection of data sufficient to fuel rapid advancement in genomic research and prevent starvation hinges on individuals' attitudes regarding whether to share their own very personal data. Competing forces shape those attitudes. On the one hand, the information privacy risk in this context is unique in its ability to do lasting harm. On the other, the potential benefit of genomic research is significant. Our findings suggest prescriptive awareness might be an effective mechanism with which to gain the voluntary submissions of genetic information from individuals necessary to populate a sufficiently comprehensive data warehouse of human genetic information for mining by researchers.

According to the most recent UNESCO Broadband Commission report, there will be over 3.2 billion internet users by the end of 2015 (Development, 2015). This represents an 8 per cent increase over the prior year. They estimate over 50 per cent of the world's population will be online by 2018. The steady rise in global interconnectivity suggests the opportunity for large-scale research requiring voluntary participation from individuals will only grow in importance, increasing researcher interest in the delicate balance between information privacy concern, awareness and individuals' willingness to share their personal data. This research helps inform a rising class of data sufficiency problems. It is our best hope that this study influences other researchers to explore additional means of motivating the necessary volunteerism to drive large-scale medical research.

References

Abu-Shanab, E.A. (2011), "Education level as a technology adoption moderator in computer research and development", *International Conference on Computer Research and Development*, *Shanghai*.

Ajzen, I. (1991), "The theory of planned behavior", *Organizational Behavior and Human Decision Processes*, Vol. 50 No. 2, pp. 179-211.

Ajzen, I. (2002), "Perceived behavioral control, self-efficacy, locust of control and the theory of planned behavior", *Journal of Applied Psychology*, Vol. 32 No. 4, pp. 665-683.

Alge, B., Ballinger, G.A., Tangirala, S. and Oakley, J.L. (2006), "Information privacy in organizations: empowering creative and extrarole performance", *Journal of Applied Psychology*, Vol. 91 No. 1, pp. 221-232.

Anderson, C.L. and Agarwal, R. (2011), "The digitization of healthcare : boundary risks, emotion and consumer willingness to disclose personal health information", *Information Systems Research*, Vol. 22 No. 3, pp. 469-490.

Angst, C.M. and Agarwal, R. (2009), "Adoption of electronic health records in the presence of privacy concerns : the elaboration likelihood model and individual persuasion", *MIS Quarterly*, Vol. 33 No. 2, pp. 339-370.

Belanger, F. and Crossler, R. (2011), "Privacy in the digital age: a review of information privacy research in information systems", *MIS Quarterly*, Vol. 35 No. 4, pp. 1017-1041.

Brandimarte, L., Acquisti, A. and Loewenstein, G. (2013), "Misplaced confidences, privacy and the control paradox", *Social Psychology and Personality Science*, Vol. 4 No. 3, pp. 340-347.

Chin, W.W. (1998), "The Partial Least Square Approach to Structural Equation Modeling", in Marcoulides, G.A. (Ed.), *Modern Methods for Business Research*, Lawrence Erlbaum and Associates, Mahwah, NJ.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, (2nd edition), Erlbaum, Hillsdale, NJ.

Conner, M. and Armitage, C.J. (1998), "Extending the theory of planned behavior: a review and avenues for further research", *Journal of Applied Psychology*, Vol. 28 No. 15, pp. 1429-1464.

Development, B.C. for D. (2015), *The State of Broadband*, New York, NY.

Dinev, T. and Hu, Q. (2007), "The centrality of awareness in the formation of user behavioral intention toward protective information technologies", *Journal of the Association for Information Systems*, Vol. 8 No. 7, pp. 386-408.

Fishbein, M. and Ajzen, I. (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Addison-Wesley Publishing Company, Reading, MA.

Gurung, A. (2006), *Empirical Investigation of the Relationship of Privacy, Security and Trust with Behavioral Intention to Transact in E-commerce*, University of Texas at Arlington.

Jarvenpaa, S.L. *et al.* (2000), "Consumer trust in an internet store", *Information Technology and Management*, Vol. 1 No. 1, pp. 45-71.

Lowry, P.B., Cao, J. and Everard, A. (2011), "Privacy concerns versus desire for interpersonal awareness in driving the use of self-disclosure technologies: the case of instant messaging in two cultures", *Journal of Management Information Systems*, Vol. 27 No. 4, pp. 163-200, available at: http://mesharpe.metapress.com/openurl.asp?genre=article&id=doi:10.2753/MIS0742-1222270406 (accessed 2 March 2012).

McGrath, M. (2015), "Target data breach spilled info on as many as 70 million customers", Forbes, available at: www.forbes.com/sites/maggiemcgrath/2014/01/10/target-data-breach-spilled-info-on-as-many-as-70-million-customers/ (accessed 2 September 2015).

Malhotra, N.K., Kim, S.S. and Agarwal, J. (2004), "Internet Users' Information Privacy Concerns (IUIPC): the construct, the scale and a causal model", *Information Systems Research*, Vol. 15 No. 4, pp. 336-355.

Morris, M.G. and Venkatesh, V. (2000), "Age differences in technology adoption decisions: implications for a changing workforce", *Personnel Psychology*, Vol. 53 No. 2, pp. 375-403.

Morris, M.G., Venkatesh, V. and Ackerman, P.L. (2005), "Gender and age differences in employee decisions about new technology: an extension to the theory of planned behavior", *IEEE Tansactions on Management*, Vol. 52 No. 1, pp. 69-84.

Newcomb, T.M. (1953), "An approach to the study of communicative acts", *Pyschology Review*, Vol. 60 No. 6, pp. 393-404.

Palvou, P. (2011), "State of the information privacy literature: where are we now and where should we go", *Management Information Systems Quarterly*, Vol. 35 No. 4, pp. 977-988.

Parker, D., Manstead, A.S.R. and Stradling, S.G. (1995), "Extending the theory of planned behaviour: the role of personal norm", *British Journal of Social Psychology*, Vol. 34 No. 2, pp. 127-138.

Rhodes, S.R. (1983), "Age-related differences in work attitude and behavior: a review of conceptual analysis", *Psychological Bulletin*, Vol. 93 No. 2, pp. 328-367.

Riley, C. (2015), "Insurance giant anthem hit by massive data breach", *CNN Money*, available at: http://money.cnn.com/2015/02/04/technology/anthem-insurance-hack-data-security/ (accessed 2 September 2015).

Rogers, E.M. (1995), *Diffusion of Innovation* 4th edn., Free Press, New York, NY.

Rutten, L., Gollust, S.E., Naveed, S. and Moser, R.P. (2012), "Increasing public awareness of direct-to-consumer genetic tests: health care access, internet use and population density correlates", *Journal of Cancer Epidemiology,* pp. 1-7.

Simpson, P.M., Banerjee, D. and Simpson, C.L. Jr (1994), "Softlifting: a model of motivating factors", *Journal of Business Ethics*, Vol. 13 No. 6, pp. 431-438.

Smith, H.J., Milberg, S.J. and Burke, S.J. (1996), "Information privacy: measuring individuals' concerns about organizational practices", *MIS Quarterly*, Vol. 20 No. 2, pp. 167-196.

Smith, H.J., Dinev, T. and Xu, H. (2011), "Information privacy research: an interdisciplinary review", *MIS Quarterly*, Vol. 35 No. 4, pp. 989-1015.

Spears, J.L. and Barki, H. (2010), "User Participation in information systems security risk management", *MIS Quarterly*, Vol. 34 No. 3, pp. 503-522.

Stone, E., Eugene, F., Gueutal, H.G., Gardner, D.G. and McClure, S. (1983), "A field experiment comparing information prvacy values, beliefs and attitudes across several types of organizations", *Jpournal of Applied Psychology*, Vol. 68 No. 3, pp. 459-468.

Webster, J. (1998), "Desktop videoconferencing: experiences of complete users wary users and non-users", *Management Information Systems Quarterly*, Vol. 22 No. 3, pp. 257-286.

Westin, A.F. (1967), *Privacy and Freedom*, Antheneum, New York, NY.

White-Baker, E., Said, S. and Hubona, G.S. (2007), "The effects of gender and age on new technology implementation in a developing country: testing the theory of planned behavior (TPB)", *Information, Technology and People*, Vol. 20 No. 4, pp. 352-375.

Corresponding author
Don Heath can be contacted at: drheath2@gmail.com