YES, FIRMS HAVE MY DATA BUT WHAT DOES IT MATTER? MEASURING PRIVACY RISKS

Research paper

Karwatzki, Sabrina, University of Augsburg, Augsburg, Germany, sabrina.karwatzki@wiwi.uni-augsburg.de

Trenz, Manuel, University of Augsburg, Augsburg, Germany, manuel.trenz@wiwi.uni-augsburg.de

Veit, Daniel, University of Augsburg, Augsburg, Germany, daniel.veit@wiwi.uni-augsburg.de

Abstract

In their daily lives, individuals continuously encounter situations where they disclose personal information online. While individuals can largely benefit from personalized, convenient service offerings, many people are at the same time concerned about an invasion of their information privacy based on how organisations access and handle their data. Although we know that specific feared consequences shape our behaviour, little attention has been paid to which noticeable privacy risks can arise for individuals when their privacy is invaded. We differentiate between seven types of negative consequences that individuals perceive if their privacy is invaded, namely physical, social, resource-related, psychological, prosecution-related, career-related, and freedom-related risks. In a comprehensive and rigorous scale development process, we validate scales for our multidimensional privacy risk construct following the approach of MacKenzie et al. (2011). Based on several steps of qualitative and quantitative assessment, we demonstrate the reliability, validity, and usefulness of our measurement instrument.

Keywords: Information Privacy, Privacy Risk, Scale Development, Measurement Instrument.

1 Introduction

Online services have become an integral part of our everyday lives. We shop online, we spend time in social networks, or use search engines to identify relevant information, to name just a few examples. In order to benefit from these convenient ways of buying, communicating, and gathering data, we often share personal data. Companies such as Facebook, Google, and Amazon can use the gained knowledge to improve and personalize their service offerings allowing them to even better address customer interests and needs and also to increase their profits. However, even though people seem to be willing to trade personal information for such benefits, at the same time, surveys continuously find that people are concerned about their privacy in today's digital and data-driven economy (BCG, 2013; TRUSTe, 2013). Thus, the question arises how privacy perceptions influence people's behaviour.

Previous research has addressed this question by investigating how privacy concerns, defined as the worries that individuals have with respect to how their personal information is handled by others (e.g., see Hong and Thong, 2013; Smith et al., 1996), are associated with individuals' behavioural reactions such as their information disclosure behaviour or their engagement in e-commerce (Smith et al., 2011). Another stream of research relies on the construct of privacy risks. This construct is either defined as opportunistic behaviour arising from other parties having access to an individual's infor-

mation (e.g., see Dinev and Hart, 2006; van Slyke et al., 2006; Wu et al., 2009) or as the beliefs of a high potential of loss which is associated with an individual's information disclosure (e.g., see Dinev et al., 2013; Malhotra et al., 2004; Smith et al., 2011; Xu et al., 2009; Yu et al., 2015).

We intend to offer a new perspective on how to investigate the influence of information privacy. In line with Dowling (1986), we argue that individuals only change their behaviour when they think they may be personally impacted by third party behaviour, or, in other words, when individuals believe that third party actions may result in noticeable negative consequences for themselves. However, the well-established and frequently used conceptualisations of privacy concerns (Smith et al., 1996; Malhotra et al., 2004; Hong and Thong, 2013) and privacy risks (Malhotra et al., 2004; Dinev and Hart, 2006) do not explicitly cover this behaviourally relevant component. To give an example, unauthorized secondary use of personal information is one frequently used dimension of privacy concerns (e.g., see Hong and Thong, 2013; Junglas et al., 2008). The secondary use of personal data can steeply increase the value of a personalized information service as it can improve the underlying algorithms of the organisation. Yet, we argue that only if individuals fear to be negatively affected by this secondary use of their information, they actually adapt their information disclosure behaviour accordingly. This is the case if they face specific risks such as a financial loss, a reputational damage, or being manipulated in their behaviour. Thus, the fear of specific risks likely has direct behavioural consequences, while mere concerns may or may not be attached to behavioural consequences.

However, current conceptualisations of privacy risks are not connected to specific consequences. These conceptualisations only mention a general potential for losses when personal information is available to other parties (e.g., see Dinev et al., 2013; Malhotra et al., 2004; Smith et al., 2011; Xu et al., 2009; Yu et al., 2015), leaving unspecified what kind of losses they refer to. In such a case, the associated measurement instruments leave a wide range of interpretation and a uniform meaning cannot be ensured (Converse and Presser, 1986; MacKenzie et al., 2011). For example, some individuals may think of losses in terms of financial losses while others associate them with losses of free choice due to tailored marketing efforts that influence them in their purchase behaviour. However, from *"both practical and research standpoints, what cannot be measured cannot be managed"* (Hille et al., 2015, p.2). We are thus in need of a measurement instrument to empirically assess these different types of risks that individuals may perceive when disclosing information to services or other individuals.

Based on these considerations, we pose the research question: *What are suitable scales for a multi-dimensional conceptualisation of privacy risks?* The objective of this paper is therefore to systematically develop and validate such a scale. We also want to demonstrate the scale's usefulness by assessing it in a nomological network.

The remainder of this paper is structured as follows: The next section gives a short summary of related literature and the state-of-the-art in information privacy measurement instruments. Afterwards, we describe our scale development and validation process, which follows the guidelines and steps of MacKenzie et al. (2011). We therefore present our multi-dimensional privacy risk conceptualisation and the according scales, which we assess, refine, and validate with the help of a quantitative study. Finally, we discuss our results, provide theoretical and practical contributions, and offer avenues for future research.

2 Theoretical Background and Related Literature

Information privacy is a subset of privacy which privacy-related research in the information systems discipline concentrates on (Bélanger and Crossler, 2011). This focus can be explained by the interest of researchers in studying how information technologies and the advent of digital services change the influence individuals have over the gathering and use of their personal information. We rely on the well-established definition of information privacy as "*an individual's self-assessed state in which external [parties] have limited access to information about him or her*" (Dinev et al., 2013, p.299). Fol-

lowing Smith et al. (2011) and Dinev et al. (2013), the term "privacy" always refers to information privacy throughout this study.

As privacy is hardly directly measurable, empirical research relies on privacy-related proxies and has consolidated over time on using privacy concerns and privacy risks as central constructs (Smith et al., 2011). Two established operationalisations for privacy concerns exist. First, the 'concern for information privacy' scale (Smith et al., 1996), which differentiates between four dimensions of privacy concerns, namely the concern that personal data is collected, is internally or externally used in an unauthorized way, is improperly accessed, and is erroneous. Second, the 'internet users' information privacy concerns' scale (Malhotra et al., 2004), which is conceptualised as *"the degree to which an Internet user is concerned about online marketers' collection of personal information, the user's control over the collected information, and the user's awareness of how the collected information is used"* (Malhotra et al., 2004, p.338). These two operationalisations have been consolidated and integrated by Hong and Thong (2013) into one measurement instrument. Overall, privacy concerns are conceptualised to focus on how individuals perceive organisations to handle their data. The conceptualisations do, however, not focus on individuals' perceptions of how these organisational practices may negatively impact individuals, which is in the focus of our study.

The conceptualisations of privacy risks can be broadly divided into two classes. First, privacy risks have been conceptualised as fears about other parties behaving opportunistically when they get access to an individual's information (e.g., see Dinev and Hart, 2006; van Slyke et al., 2006; Wu et al., 2009). This conceptualisation is similar to the conceptualisations of privacy concerns and thus does also not focus on the negative consequences which may arise from such opportunistic behaviour of other parties. Second, privacy risks have been defined as beliefs of a high potential of loss that may occur if an individual's information is disclosed to other parties (e.g., see Dinev et al., 2013; Malhotra et al., 2004; Smith et al., 2011; Xu et al., 2009; Yu et al., 2015). However, the related measurement instruments measure people's worries about a loss of privacy at a rather abstract level and remain unspecific what these losses actually refer to (e.g., see Dinev and Hart, 2006; Xu et al., 2010). Such conceptualisations of a single-dimensional construct leave room for ambiguity and interpretation (Converse and Presser, 1986). They do not incorporate whether and how exactly the loss impacts people. These losses can occur in very different forms. For example, when shopping online, individuals face the risks of a financial loss if their credit card data is abused. In a social networking context, individuals might be more afraid of a reputational damage. Thus, the conceptualisations do not depict the complexity of risks which has been successfully exploited in other areas of research (Dowling, 1986). In marketing and e-commerce, for example, a detailed specification of the negative outcomes is a core element of risk (e.g., see Cunningham, 1967; Dowling, 1986; Glover and Benbasat, 2010; Jacoby and Kaplan, 1972). Risks are mostly defined as being multi-dimensional, comprising performance, financial, social, physical, and psychological risks (Dowling, 1986).

We believe that such a multi-dimensional conceptualisation is also necessary to fully capture the nature of privacy risks. The aforementioned risk dimensions from other research areas are a first indication of possible dimensions, but they need to be adapted and extended to align with the unique context of information privacy. Privacy risks do not refer to product quality or online transactions. Instead, the perceived consequences of information misuse and their likelihood of occurrence are at centre stage. One first step towards a systematic and comprehensive conceptualisation of privacy risks which covers different risk dimensions is provided by Karwatzki et al. (2017). Based on an extensive qualitative study, they identified seven dimensions that describe how privacy-invasive practices such as data collection, improper access, or unauthorized usage might affect individuals physically, socially, resource-related, psychologically, prosecution-related, career-related, and freedom-related. As these dimensions cover the perceived negative consequences that individuals associate with others having access to their information, they can serve as a starting point for our scale development process of multi-dimensional privacy risks.

3 Scale Development

Our research aims at developing scales for privacy risks taking into account the different dimensions of risk. We followed the approach of MacKenzie et al. (2011) to generate, validate, and refine our items. The approach comprises five steps: (1) developing a conceptual definition of the latent variables, (2) generating items that represent the latent variables and qualitatively assessing the content validity of the items, (3) formally specifying the measurement model, (4) evaluating the scales in a pretest and refining them, and (5) validating the final measurement model. The process with the essential activities and outcomes is depicted in Figure 1.

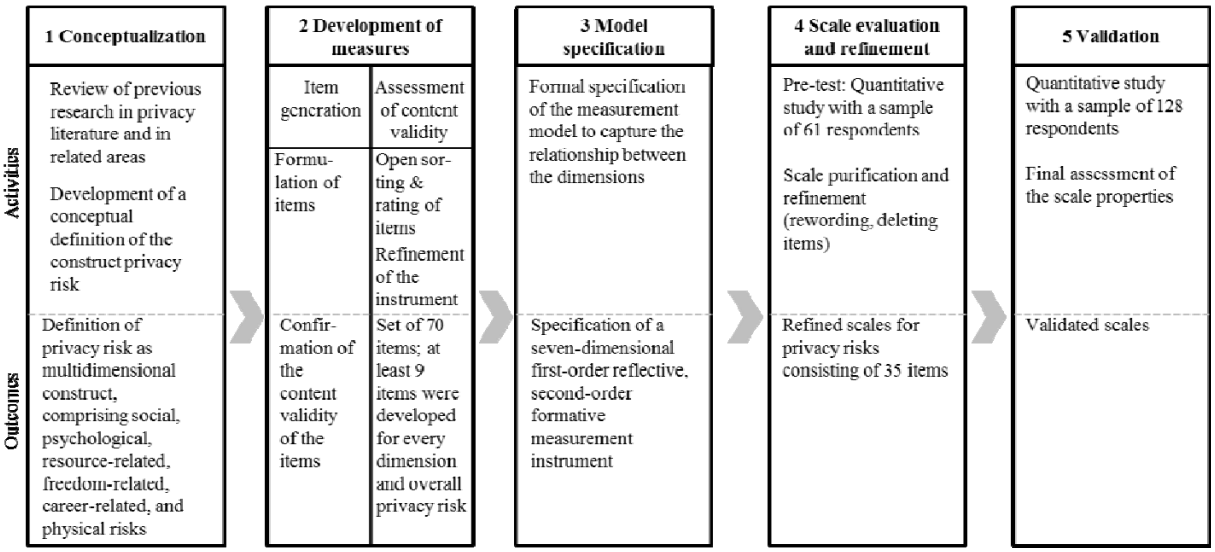

Figure 1. Scale Development Process (adapted from MacKenzie et al. (2011))

3.1 Step 1: Conceptualisation

For conceptualising privacy risks, we draw from existing literature on risks in other contexts (e.g., Featherman and Pavlou, 2003; Glover and Benbasat, 2010; Luo et al., 2010). Risks have been commonly defined as consisting of two components: (1) the severity of adverse consequences of a situation and (2) their probability of occurrence (Cunningham, 1967; Jacoby and Kaplan, 1972; Dowling, 1986; Mitchell, 1999). To contextualize this general risk definition to the privacy area, we extend the conceptualisation of perceived adverse consequences of privacy-invasive practices by Karwatzki et al. (2017) which already captures the severity of negative outcomes that may arise if others have access to individuals' information. We add the probability of occurrence and thus define privacy risks as the *extent to which an individual believes that negative outcomes may arise from others' access to his or her personal information.* This construct refers to a perception as it describes the individually perceived risk in a specific situation. Moreover, we conceptualise privacy risks as a multi-dimensional construct that applies to the entity of individuals. It is not intended to measure the privacy risks of groups or organisations as these parties likely face different risk dimensions than the ones individuals face. Based on the work of Karwatzki et al. (2017), we conceptualise privacy risks to comprise the dimensions physical, social, resource-related, psychological, prosecution-related, freedom-related, and career-related risks. Table 1 depicts the definitions of each dimension, which we extend to cover risk in its entirety. We therefore include the second component of risk, which is the probability of occurrence. These privacy risk dimensions form the basis for our scale development.

Regarding the stability of our construct, we expect it to be varying across individuals and situations. In contrast to constructs such as general privacy dispositions (Li, 2014), privacy risks are an individual's perception of the extent to which negative outcomes may arise out of a specific situation in which oth-

ers may gain access to an individual's information. Thus, we expect privacy risks to naturally differ between individuals and contexts which may depend on, for example, individual characteristics such as privacy dispositions and previous experiences, and situation-specific characteristics such as which information may be shared with whom. Defining privacy risks as a multi-dimensional construct can be useful to analyse these contextual differences in depth.

Dimension	**Definition** Extent to which an individual believes that…
Physical risk	… a loss of physical safety may arise from access to the individual's information.
Social risk	… a change in one's social status may arise from access to the individual's information.
Resource-related risk	… a loss of resources may arise from access to the individual's information.
Psychological risk	… a negative impact on one's peace of mind may arise from access to the individual's information.
Prosecution-related risk	… legal actions that are taken against an individual may arise from access to the individual's information.
Career-related risk	… negative impacts on one's career may arise from access to the individual's information.
Freedom-related risk	… a loss of freedom of opinion and behavior may arise from access to the individual's information.

Table 1. Dimensions of Privacy Risks

3.2 Step 2: Development of Measures

To develop measures, two steps are necessary. First, potential items have to be generated. Second, the content validity of those items must be assessed to ensure their suitability.

3.2.1 Item Generation

For item generation, we relied on existing risk scales (Stone and Grønhaug, 1993; Featherman and Pavlou, 2003; Krasnova et al., 2010; Luo et al., 2010) that were developed for other contexts wherever possible. As the work by Karwatzki et al. (2017) formed the basis for the conceptualisation of our seven risk dimensions, we also relied on their qualitative data set which consists of twenty-two focus groups with 119 participants (see Karwatzki et al. (2017) for more details) to generate suitable items. We particularly drew on the expressions that the focus group participants used in the discussions. Overall, we came up with 70 items that can be allocated to the seven dimensions as displayed in Table 2. We paid special attention to dimensions which were newly developed, which largely differ from existing risk conceptualisations, or for which several subdimensions were identified by Karwatzki et al. (2017). We developed a variety of items to test which of those best capture the nature of the construct.

Dimension	Number of Items	Dimension	Number of Items
Physical risk	9 items	Prosecution-related risk	10 items
Social risk	12 items	Career-related risks	11 items
Resource-related risk	9 items	Freedom-related risk	9 items
Psychological risk	10 items		

Table 2. Initial Set of Items for Each Risk Dimension

3.2.2 Assessment of Content Validity

We conducted two steps to assess the content validity. First of all, we did an open sorting with ten raters based on the guidelines of Moore and Benbasat (1991). The raters got index cards where each item was printed on one card. They were instructed to categorize the items and to label and explain the identified groups. For this task, the raters were not provided with any names or definitions of the underlying constructs. Having sorted the items, we also discussed the identified categories with the raters in detail to gain a deeper understanding of the problems they encountered during the sorting process as well as any difficulties they had with the wording or comprehensibility of the items. Based on their feedback, we dropped some items and slightly adjusted some other items. Next, we provided twenty raters with the definitions of the constructs and asked them to rate the extent to which the refined items belong to each construct domain using a five-point Likert scale. In order to reduce the complexity of this task, we only presented half of our item set to each rater, as they otherwise would be overwhelmed by this huge number of items. Based on these item ratings, we were able to assess the items' content adequacy (MacKenzie et al., 2011). To do so, we conducted a one-way repeated measures analysis of variance for each of the items to investigate whether the item's mean rating on one risk dimension significantly differs from the item's mean ratings on other risk dimensions. Our results show that our raters associated the majority of items with their intended dimensions, yet some items were also associated with more than one dimension. We used the results to slightly adjust the wording of some items and then repeated the rating task to assess the content validity of all adapted and newly added items.

3.3 Step 3: Model Specification

Based on our conceptualisation, we model privacy risk as a seven-dimensional construct: physical, prosecution-related, social, career-related, resource-related, freedom-related, and psychological risk. We measure all risk dimensions reflectively while the dimensions influence overall privacy risk. We used the rules provided by Jarvis et al. (2003) for this decision: The indicators of every first-order construct are manifestations of the construct, they share a common theme and can be used interchangeably, and covariate. Thus, we have a reflective model for our risk dimensions. However, since those constructs are conceptually different and cover separate aspects of the overall privacy risk construct, they cannot be interchanged and do not necessarily covary with each other, so that each of them should have an impact on overall risk. These considerations indicate that we have a formative second-order construct. In sum, we model privacy risks as a reflective first-order, formative-second order construct.

3.4 Step 4: Scale Evaluation and Refinement

To get a first assessment of the proposed scales, we conducted a small pre-study to investigate the reflective first-order constructs. The pre-study was conducted using Amazon Mechanical Turk (MTurk). A total of 61 completed questionnaires were received. The purpose of this pre-study was (1) to test the comprehensibility of the items and of different alternative scenarios that we planned to use in the main study, (2) to do preliminary reliability and validity assessments, and (3) to shorten our instrument so that we had between four to six items per construct. We also included an open text field which allowed the participants to comment on the scenario and on the items as well as on their overall experiences with the questionnaire.

We assessed reliability and validity by doing an exploratory factor analysis (EFA) (settings: principal component analysis, Oblimin rotation) in SPSS (version 24) which showed that all first-order constructs are unidimensional as intended. However, when conducting an EFA with career-related and prosecution-related risk items, a few items had high loadings on both identified factors so that we investigated them in more detail. We also evaluated Cronbach's alpha. For all constructs, the value was above the commonly suggested threshold of 0.7 (even above 0.85). Moreover, we looked at the values for Cronbach's alpha if deleted to identify candidates for elimination and considered the comments in

the open text field to carefully reword our items if necessary. Table 3 gives an overview of the final item set.

Dimension	ID	Item
		If someone has access to the information this app has about me...
Physical risk	PH1	... my physical safety might be impacted.
	PH2	... I might be exposed to physical threats.
	PH3	... the chance of me being physically harmed is increased.
	PH4	... it might endanger my physical safety.
	PH5	... my physical safety might be at risk.
Social risk	SO1	... it might damage my reputation.
	SO2	... it might impact the perception that others have of me.
	SO3	... it might change the way people think about me.
	SO4	... my social status might be influenced.
	SO5	... my peer group might think differently of me.
Resource-related risk	RR1	... it might consume my time or my money.
	RR2	... it might cost me time or money.
	RR3	... it might require efforts or expenditures.
	RR4	... it might cause efforts or financial disadvantages.
	RR5	... it might affect my resources (e.g. time, money) negatively.
Psychological risk	PS1	... I might feel uncomfortable.
	PS2	... it might give me a feeling of anxiety.
	PS3	... it might cause inner restlessness.
	PS4	... I might experience mental tension.
	PS5	... it might burden me mentally.
Prosecution-related risk	PR1	... I might get judicially indictable, either wrongly or rightfully.
	PR2	... I might be prosecuted due to wrongful or rightful suspicions.
	PR3	... I might be sued because of wrongly or rightfully made accusations.
	PR4	... I might be held legally accountable due to wrongful or rightful suspicions.
	PR5	... I might be held responsible due to wrongful or rightful suspicions.
Career-related risks	CR1	... it might reduce my career prospects.
	CR2	... it might affect my career negatively.
	CR3	... it might make it difficult to be successful in my job.
	CR4	... it might result in a negative shift in my career.
	CR5	... it might result in a stagnation of my career development.
Freedom-related risk	FR1	... my opinion or behaviour might get manipulated.
	FR2	... it might influence my decision making.
	FR3	... my thoughts or actions might be influenced externally.
	FR4	... my mindset or my resulting behaviour might get influenced.
	FR5	... my attitude or behaviour might get influenced.

Table 3. Privacy Risk Constructs with their Final Items

3.5 Step 5: Validation

The last step of our scale development process aimed at assessing the developed scales in a larger-scale survey. Moreover, we wanted to investigate privacy risks in a nomological network.

Such a nomological network should include other constructs that are expected to serve as antecedents and consequences of the focal construct which ideally has been shown in previous research (MacKenzie et al., 2011). We followed this recommendation and used a privacy calculus perspective which has been extensively applied in multiple studies (e.g., see Chellappa and Sin, 2005; Dinev et al., 2006; Dinev and Hart, 2006; Hann et al., 2007; Kehr et al., 2015; Krasnova et al., 2010; Sarathy and Li, 2007). The privacy calculus perspective assumes that individuals perform a risk-benefit analysis when having to decide on whether and how much personal information to provide to other parties (Culnan and Bies, 2003; Dinev and Hart, 2006; Smith et al., 2011). We therefore deem it a suitable model for our purposes and apply it to validate our privacy risk measurement instrument. In line with prior privacy research (Bélanger and Crossler, 2011; Smith et al., 2011), we decided to use intention to use an app and willingness to provide information to an app as dependent variables in our nomological network. Regarding predictors of privacy risk, we leverage the construct privacy experiences. It is a widely used antecedent of privacy-related constructs and refers to an individual's prior negative experiences due to being exposed to or having fallen victim to information abuse (Smith et al., 1996; Li, 2014). We expect privacy experiences to have a positive influence on our privacy risk construct. We further expand our nomological network by another commonly used antecedent of privacy constructs, namely individuals' familiarity with the service type (Li, 2014), which we expect to be negatively related to our privacy risk construct. The measures for the additional variables in the nomological network are depicted in Table 6 in the Appendix. The general model is illustrated in Figure 2.

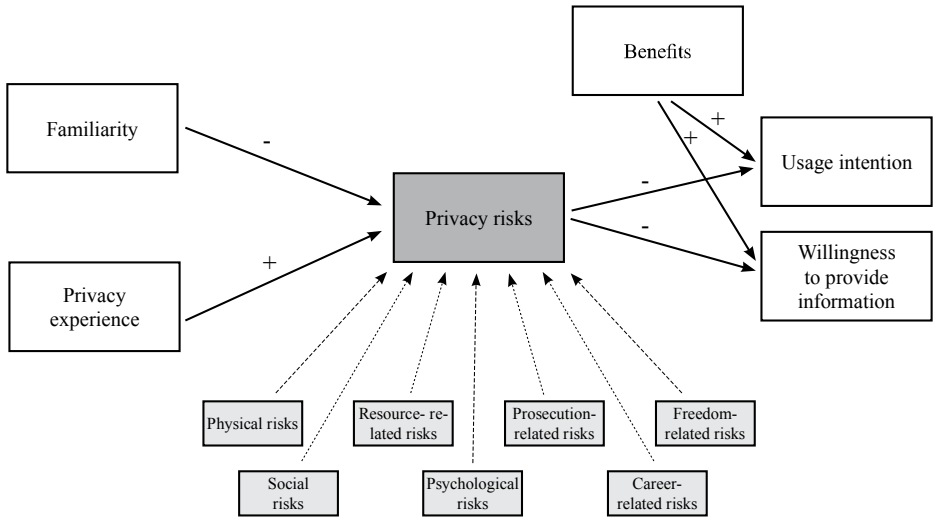

Figure 2. Nomological Network of Validation Study

To test our nomological network, we place our study exemplarily in the context of a health app. We developed a questionnaire that introduced a novel app to participants that requires them to share sensitive information. To make this as realistic as possible, we told all participants that the study was done in cooperation with a start-up that wants to gain market insights before launching their product shortly. The health app would provide interesting insights into one's own well-being and provide suggestions for improvement. In order to do that, it requires sensitive data such as continuous GPS tracking, nutritional information, and sleep tracking. The app also offers the possibility to share health-related data with insurances or social peers. After introducing the service, we then asked for their assessment. First of all, we asked individuals about their general willingness to share data with this service and their intention to use it, once it is released. The participants then had to assess their perceptions of the risks and benefits of the service. The risk dimensions were measured on a seven-point Likert scale ranging from strongly disagree to strongly agree. The items of all risk dimensions were merged in this section and the appearance of the individual items was fully randomized in order to prevent click-throughs or order effects. Establishing reliability and validity using fully randomized items would be a tough test for our conceptualisation and for the developed scales. We then asked the other questions from our

nomological network, a marker variable and finally closed the survey with demographic details and a debriefing.

We recruited 128 participants from MTurk who participated in our survey. MTurk is a well-established platform for behavioural research and experiments (Behrend et al., 2011). A variety of studies have demonstrated high reliability and quality of data and results derived from respondents on MTurk (e.g., Behrend et al., 2011; Buhrmester et al., 2011; Steelman et al., 2014). In addition, MTurk is a suitable platform to reach users familiar with the internet and digital technologies, who are potential adopters of innovative digital services such as the health app in our study. We restricted participation to users with a high reputation (at least 98% approval rate and at least 500 conducted tasks), which is a sufficient measure to ensure high data quality (Peer et al., 2014). Participants were between 20 and 74 years old with a mean age of 39.4 years and a standard deviation of 11.8. 53.9% percent of the participants were female. Around a third of the participants indicated a yearly household income below $35,000 (32.8%). 43% between indicated an income between $35,000 and $75,000 and the remaining 24.2% have more than $75,000 per year at their disposal. Thus, our sample represents a cross-section of the population without being biased towards young people, a particular gender, or an income group.

We first wanted to make sure that our seven dimensions of privacy risks in fact appear in a factor analysis and that there is no major overlap between the dimensions (although of course all dimensions refer to privacy risks, so they should not be fully orthogonal). We thus ran principal component analysis using Oblimin rotation with Kaiser Normalization (see Hair et al., 2014). The pattern matrix reveals the expected seven factors. It shows unique loadings of each item linked to its respective factor. The structure matrix shows correlations of more than 0.7 for the expected relationships between factors and items (most are even above 0.9) and well below 0.7 (most are below 0.6) with factors that should not be linked to the respective risk dimensions. These results indicate that there are no problematic unintended cross loadings. We also investigated Cronbach's alpha to assess the reliability of our constructs. All constructs have values above 0.9, thus they are well above the recommended threshold of 0.7 (Nunnally, 1978). All details on descriptive statistics can be found in Table 4.

We used partial least squares structural equation modelling (PLS-SEM using SmartPLS 3.0 (Ringle et al., 2015)) to examine our research model. Arguments can be made in favour and against the use of PLS-SEM in comparison to covariance-based structural equation modelling (CB-SEM) (Goodhue et al., 2006, 2012; Marcoulides et al., 2012). Both approaches differ in their underlying philosophy and estimation objectives (Gefen et al., 2011). While CB-SEM emphasizes how well the proposed research model accounts for measurement item co-variances, thereby offering various indices how well parameter estimates match sample co-variances (Chin, 1998), PLS-SEM uses the empirical data for estimating relationships with the aim to maximize the explained variance in the endogenous latent variable (Hair et al., 2014). Given the early stage of this investigation, the exploratory character of the study and the primary interest in identifying potential relationships between variables, we decided to use PLS-SEM for evaluating the drivers of actual information disclosure behaviour. However, consistent empirical results are expected when using CB-SEM.

When modelling our nomological network in SmartPLS, we followed the recommendations of Hair et al. (2018) and applied a repeated indicators approach to model our first-order reflective, second-order formative privacy risk construct. This decision has several implications for our analysis, which will be discussed in due course. To assess the measurement model, we first investigated the standardized factor loadings (see Table 4). For reflective constructs to be reliable, composite reliability (CR) and Cronbach's alpha (Alpha) both have to be above 0.7 (Fornell and Larcker, 1981; Nunnally and Bernstein, 1994) which is the case for all of our constructs (see Table 4). To ensure validity at the construct level, average variance extracted (AVE) has to be above 0.5 which means that the latent construct accounts for the majority of the variance of its indicators (MacKenzie et al., 2011). In our sample, AVE even exceeds 0.7 for all risk constructs (see Table 5). We also assessed discriminant validity by using the Fornell-Larcker criterion which says that discriminant validity is sufficient if the square root of

Con-struct	Cr.α / CR	Item ID	Fact. Load.	Mean	STD	Construct	Cr.α / CR	Item ID	Fact. Load.	Mean	STD
Physical risk (PH)	0.98 / 0.98	PH1	0.97	3.70	2.00	Freedom-related risk (FR)	0.92 / 0.94	FR1	0.83	4.21	1.83
		PH2	0.95	3.48	2.01			FR2	0.87	4.56	1.75
		PH3	0.96	3.59	2.00			FR3	0.88	4.39	1.77
		PH4	0.94	3.55	2.00			FR4	0.86	4.41	1.87
		PH5	0.95	3.66	2.05			FR5	0.92	4.30	1.81
Social risk (SO)	0.94 / 0.95	SO1	0.88	3.63	1.87	Benefits (BE)	0.92 / 0.95	BE1	0.97	4.56	1.58
		SO2	0.91	4.05	1.88			BE2	0.96	4.63	1.61
		SO3	0.91	3.90	1.90			BE3	0.87	4.45	1.79
		SO4	0.89	3.63	1.88						
		SO5	0.88	3.51	1.97						
Re-source-related risk (RR)	0.93 / 0.95	RR1	0.88	4.79	1.70	Famili-arity (FA)	0.92 / 0.95	FA1	0.93	4.38	1.64
		RR2	0.93	4.84	1.75			FA2	0.95	4.44	1.64
		RR3	0.81	4.62	1.73			FA3	0.92	4.42	1.69
		RR4	0.89	4.51	1.86						
		RR5	0.90	4.52	1.87						
Psycho-logical risk (PS)	0.94 / 0.95	PS1	0.85	5.81	1.39	Prior negative privacy experi-ence (PP)	0.96 / 0.97	PP1	0.95	3.43	1.93
		PS2	0.92	5.37	1.60			PP2	0.97	3.61	2.10
		PS3	0.91	5.23	1.69			PP3	0.96	3.64	2.05
		PS4	0.91	5.22	1.58			PP4	0.87	3.55	1.98
		PS5	0.90	5.20	1.55						
Prosecu-tion-related risk (PR)	0.96 / 0.97	PR1	0.89	3.24	1.94	Intention to pro-vide in-formation (IP)	0.97 / 0.98	IP1	0.98	4.00	1.97
		PR2	0.95	3.21	2.00			IP2	0.98	4.09	2.08
		PR3	0.93	2.92	2.01			IP3	0.94	3.95	2.07
		PR4	0.95	3.30	2.01						
		PR5	0.92	3.29	1.91						
Career-related risk (CR)	0.97 / 0.98	CR1	0.96	3.02	1.84	Usage intention (UI)	0.95 / 0.97	UI1	0.97	3.81	1.72
		CR2	0.96	3.19	1.87			UI2	0.93	4.71	1.87
		CR3	0.92	2.98	1.84			UI3	0.97	3.88	1.89
		CR4	0.96	2.96	1.77						
		CR5	0.96	2.93	1.79						

Table 4. Measurement Model Results

AVE is larger than the correlations of the construct with any other construct (Fornell and Larcker, 1981). Table 5 depicts that we have an adequate level of discriminant validity.

To assess the measurement model of our second-order privacy risk construct, we needed to assess the weights between the first-order risk dimensions and overall privacy risk, which is depicted as path coefficients in the PLS-SEM analysis (Hair et al., 2018). The analysis shows that all risk dimensions significantly influence the second-order construct and that the effects are of similar size (physical risks: $\beta = 0.21$, $p < 0.001$; social risks: $\beta = 0.18$, $p < 0.001$; resource-related risks: $\beta = 0.17$, $p < 0.001$; psychological risks: $\beta = 0.17$, $p < 0.001$; prosecution-related risks: $\beta = 0.18$, $p < 0.001$; career-related risks: $\beta = 0.21$, $p < 0.001$; freedom-related risks: $\beta = 0.17$, $p < 0.001$). Discriminant validity, however, is not of concern. The modelling of the second-order construct via a repeated indicators approach leads to conceptual and empirical redundancy, so that an assessment of discriminant validity between the

	BE	CR	FA	FR	IP	PH	PP	PR	PS	RR	SO	UI
Benefits (BE)	**0.93**											
Career-r. risk (CR)	-0.19	**0.95**										
Familiarity (FA)	0.12	-0.04	**0.93**									
Freedom-r. risk (FR)	-0.01	0.49	-0.09	**0.87**								
Int. provide info. (IP)	0.60	-0.14	0.26	-0.16	**0.97**							
Physical risk (PH)	-0.22	0.56	-0.09	0.50	-0.26	**0.96**						
Privacy exp. (PP)	0.09	0.21	0.13	0.25	-0.22	0.30	**0.94**					
Prosec. risk (PR)	0.01	0.67	-0.18	0.56	-0.03	0.58	0.18	**0.93**				
Psychol. risk (PS)	-0.05	0.39	-0.09	0.69	-0.17	0.50	0.30	0.38	**0.87**			
Resource-r. risk (RR)	-0.15	0.59	-0.07	0.56	-0.20	0.41	0.37	0.56	0.56	**0.88**		
Social risk (SO)	-0.11	0.74	-0.02	0.55	-0.10	0.49	0.20	0.54	0.57	0.51	**0.89**	
Usage int. (UI)	0.71	-0.22	0.17	-0.17	0.68	-0.21	-0.03	-0.04	-0.15	-0.18	-0.21	**0.95**
Note: The diagonal elements (in bold) represent the square root of AVE												

Table 5. Correlation Matrix and AVE

first-order constructs and the second-order construct is meaningless (Hair et al., 2018). Yet, potential collinearity between the first-order risk constructs needs to be assessed. As the variance inflation factors (VIF) of all risk dimensions are well below the threshold of 5 (physical risks: VIF = 1.94; social risks: VIF = 2.95; resource-related risks: VIF = 2.24; psychological risks: VIF = 2.67; prosecution-related risks: VIF = 2.44; career-related risks: VIF = 3.48; freedom-related risks: VIF = 2.42), collinearity is not an issue (Hair et al., 2014).

The results of the analysis of our nomological model are depicted in Figure 3. We find a negative impact of privacy risks ($\beta = -0.13$, $p<0.05$) on usage intention for the service after controlling for the influence of benefits ($\beta = 0.67$, $p<0.001$). The same holds for the willingness to provide information to the service ($\beta = -0.12$, $p<0.05$), the second variable that we expected to be influenced by privacy risks. Data provision is also strongly depended on the benefits that can be derived from the service ($\beta = 0.58$, $p<0.001$). We also find significant influences of familiarity ($\beta = -0.14$, $p<0.1$) and prior privacy experiences ($\beta = 0.35$, $p<0.001$) on privacy risks as expected. Lastly, we controlled for demographics such as age, gender, and income. The analysis resulted in only one significant relationship: women seem to be more eager to use the service we described ($\beta = 0.15$, $p<0.01$).

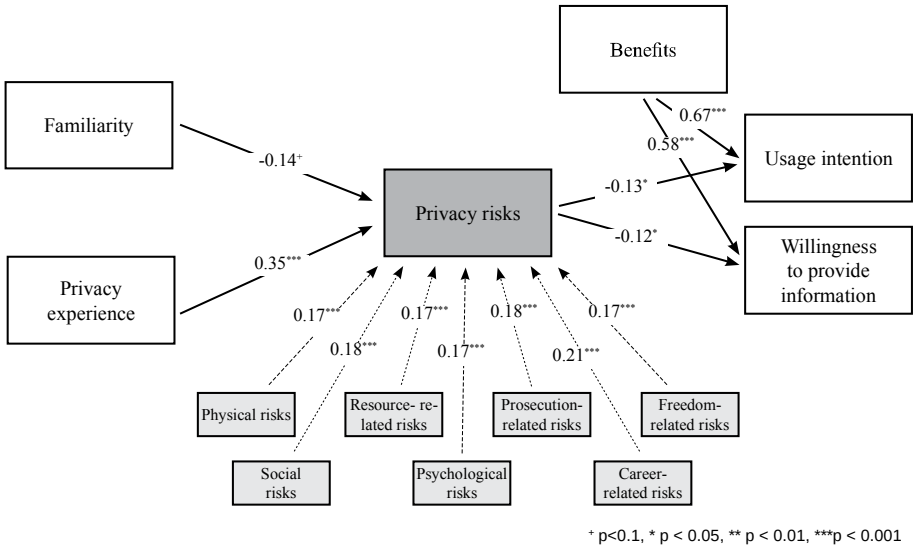

Figure 3. PLS Structural Results

Overall, the research model explains 52.3% of usage intention, and 36.1% of the intention to provide data. As we found support for all hypothesized relationships between our focal construct privacy risks and its antecedents and outcome, we can conclude that nomological validity is given.

4 Discussion

The objective of our study was twofold. First, we aimed at systematically developing and validating scales for privacy risks as multi-dimensional concept. Second, we wanted to do a first empirical assessment of how privacy risks influence individuals' information disclosure and usage intentions. Our results show that we developed suitable scales which offer promising avenues for a further exploration of how privacy perceptions influence individuals' behaviour and thus provide several interesting contributions to theory and practice.

To systematically develop scales for privacy risks, we build our work on the dimensions of privacy risks identified Karwatzki et al. (2017) in a large qualitative study. We conceptualised privacy risks as a multi-dimensional construct comprising the risk dimensions of social, psychological, physical, prosecution-related, freedom-related, career-related, and resource-related risks. We thereby contribute to theory a conceptualisation of privacy risks that describes the extent to which individuals perceive to be affected by a privacy invasion through third parties and specifies the types of impact that may occur. By providing a reliable and valid measurement instrument that captures this multi-dimensional conceptualisation, we offer a novel perspective on how to measure privacy-related perceptions. Previous conceptualisations neglected to capture which negative consequences can arise for individuals when they share information online and how individuals perceive those negative consequences to impact them (e.g., see Dinev et al., 2006; Dinev and Hart, 2006; Hong and Thong, 2013; Malhotra et al., 2004; Smith et al., 1996). Second, we demonstrated the usefulness of our measurement instrument. We used a nomological network to validate our measurement instrument and found support for all hypothesized relationships. The results of our analysis show that our construct exhibits discriminant and nomological validity. Our scale thus captures the complex multi-dimensional nature of privacy risks and thus supplements existing measurement instruments.

Our study has also practical implications. Many business models such as those of innovative apps depend on fast growth rates and on the collection and analysis of user data. Therefore, those service providers are very interested in better understanding the circumstances of individual information disclosure, reasons that might prevent disclosure, and how to mitigate problematic influences. By providing a conceptualisation of privacy risks as a multi-dimensional construct and by showing its influence on usage and information disclosure intention, we provide organisations with opportunities to better understand why consumers might hesitate to share information in certain situations and which risks may have to be mitigated by the service design in order to prevent discouraged users.

We see several promising avenues for future research. While we investigated privacy risks in the context of an innovative health app, in our future research we will explore the influence of privacy risks in other contexts as well, for example in social networking, in e-commerce, personalized newspapers, or in a digital job market with artificial intelligence. Thereby, we could obtain a more fine-grained understanding of how privacy risks are influencing user behaviour in different contexts. With the help of experimental setups, we could explore situation-specific differences in individuals' risk assessments. The multi-dimensional construct of privacy risks also opens up new research opportunities on how third parties can actively mitigate privacy risks. In doing so, we believe it is on the one hand especially interesting to evaluate the effect of existing mitigation mechanisms such as seals, privacy policies, or building up trust and long-term relationships on the single dimensions of privacy risk to understand how and to which extent they work in different contexts. On the other hand, we need to design new mechanisms that effectively mitigate risks. This is also of high practical relevance, as organisations are interested in changing online users' privacy risks to better align their information disclosure behaviours with their aims.

5 Conclusion

In today's digital and data-driven economy, many individuals are concerned about their information privacy. We offer a conceptualisation of privacy risks that incorporates the extent to which individuals perceive negative consequences to occur if third parties invade their privacy. This is a novel and promising perspective to measure the impact of privacy since previous conceptualisations of privacy concerns and privacy risks have not included this component of concrete consequences. We identified seven dimensions of privacy risks, namely physical, social, resource-related, psychological, prosecution-related, career-related, and freedom-related risks and thoroughly developed and empirically validated scales to assess them. Moreover, we demonstrated that privacy risks are a useful predictor of individuals' information disclosure and usage intention. Our research thus provides researchers and practitioners with new avenues for investigating the influence of privacy risks on individuals' online behaviour.

Appendix

Dimension	ID	Item
Benefits	BE1	I would benefit from using [this app].
	BE2	It would be advantageous for me to use [this app] with all its functions.
	BE3	[This app] offers functions I would profit from.
Familiarity	FA1	I know pretty much about apps like this.
	FA2	Compared to most other people, I know less about apps like this. (reversed)
	FA3	I feel familiar with apps like this.
Prior negative privacy experience	PP1	I have had bad experiences with regard to my online privacy before.
	PP2	I was a victim of what I felt was an invasion of my privacy.
	PP3	I believe that my online privacy was invaded by other people or organisations.
	PP4	I experienced my personal information being misused by companies without my authorization.
Intention to provide information	In order to use this service, I would be willing to share the required personal information	
	IP1	Very unlikely/Very likely
	IP2	Not probable/Probable
	IP3	Willingly/Unwillingly (reversed)
Usage intention	UI1	I intend to use [this app] in future.
	UI2	I could imagine using [this app] in future.
	UI3	I predict that I would use [this app] in future.

Table 6. Items for Nomological Network

References

BCG (2013). The Value of Our Digital Identity *www.bcgperspectives.com* URL: https://www.bcgperspectives.com/content/articles/digital_economy_consumer_insight_value_of_our_digital_identity/ (visited on 29/06/2013).

Behrend, T.S., D.J. Sharek, A.W. Meade, and E.N. Wiebe (2011). "The Viability of Crowdsourcing for Survey Research." *Behavior Research Methods* 43 (3), 800.

Bélanger, F. and R.E. Crossler (2011). "Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems." *MIS Quarterly* 35 (4), 1017–1042.

Buhrmester, M., T. Kwang, and S.D. Gosling (2011). "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science* 6 (1), 3–5.

Chellappa, R.K. and R.G. Sin (2005). "Personalization Versus Privacy: An Empirical Examination of the Online Consumer's Dilemma." *Information Technology and Management* 6 (2–3), 181–202.

Chin, W.W. (1998). "Commentary: Issues and Opinion on Structural Equation Modeling." *MIS Quarterly* 22 (1), vii–xvi.

Converse, J.M. and S. Presser (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Newburn Park: Sage.

Culnan, M.J. and R.J. Bies (2003). "Consumer Privacy: Balancing Economic and Justice Considerations." *Journal of Social Issues* 59 (2), 323–342.

Cunningham, S.M. (1967). "The Major Dimensions of Perceived Risk." in Cox, D.F., ed., *Risk Taking and Information Handling in Consumer Behavior* Boston: Harvard University Press, 82–108.

Dinev, T., M. Bellotto, P. Hart, V. Russo, I. Serra, and C. Colautti (2006). "Privacy Calculus Model in E-Commerce - A Study of Italy and the United States." *European Journal of Information Systems* 15 (4), 389–402.

Dinev, T. and P. Hart (2006). "An Extended Privacy Calculus Model for E-Commerce Transactions." *Information Systems Research* 17 (1), 61–80.

Dinev, T., H. Xu, J.H. Smith, and P. Hart (2013). "Information Privacy and Correlates: An Empirical Attempt to Bridge and Distinguish Privacy-related Concepts." *European Journal of Information Systems* 22 (3), 295–316.

Dowling, G.R. (1986). "Perceived Risk: The Concept and Its Measurement." *Psychology & Marketing* 3 (3), 193–210.

Featherman, M.S. and P.A. Pavlou (2003). "Predicting E-Services Adoption: A Perceived Risk Facets Perspective." *International Journal of Human-Computer Studies* 59 (4), 451–474.

Fornell, C. and D.F. Larcker (1981). "Evaluating structural equation models with unobservable variables and measurement error." *Journal of Marketing Research* 18 (1), 39–50.

Gefen, D., E. Rigdon, and D. Straub (2011). "An Update and Extension to SEM Guidelines for Administrative and Social Science Research." *MIS Quarterly* 35 (2), iii-A7.

Glover, S. and I. Benbasat (2010). "A Comprehensive Model of Perceived Risk of E-Commerce Transactions." *International Journal of Electronic Commerce* 15 (2), 47–78.

Goodhue, D.L., W. Lewis, and R. Thompson (2006). "PLS, Small Sample Size, and Statistical Power in MIS Research." in *HICSS 2006 Proceedings* Hawaii, USA.

Goodhue, D.L., W. Lewis, and R. Thompson (2012). "Does PLS have advantages for small sample size or non-normal data?." *MIS Quarterly* 36 (3), 981–1001.

Hair, J.F., W.C. Black, B.J. Babin, and R.E. Anderson (2014). *Multivariate Data Analysis*., 7th edition. ed Harlow: Pearson Education Ltd.

Hair, J.F., G.T. Hult, C.M. Ringle, and C.M. Sarstedt (2014). *A Primer On Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA, USA: SAGE Publications.

Hair, J.F., C. Ringle, M. Sarstedt, and S.P. Gudergan (2018). *Advanced Issues in Partial Least Square Structural Equation Modeling*. Thousand Oaks: Sage Publications.

Hann, I.-H., K.-L. Hui, S.-Y.T. Lee, and I.P.L. Png (2007). "Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach." *Journal of Management Information Systems* 24 (2), 13–42.

Hille, P., G. Walsh, and M. Cleveland (2015). "Consumer Fear of Online Identity Theft: Scale Development and Validation." *Journal of Interactive Marketing* 30, 1–19.

Hong, W. and J.Y.L. Thong (2013). "Internet Privacy Concerns: An Integrated Conceptualization and Four Empirical Studies." *MIS Quarterly* 37 (1), 275–298.

Jacoby, J. and L.B. Kaplan (1972). "The Components of Perceived Risk." in Venkatesan, M., ed., *Proceedings of the Third Annual Conference of the Association for Consumer Research* Presented at the Association for Consumer Research, Chicago, IL, 382–393.

Jarvis, C.B., S.B. MacKenzie, and P.M. Podsakoff (2003). "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research." *Journal of Consumer Research* 30 (2), 199–218.

Junglas, I.A., N.A. Johnson, and C. Spitzmüller (2008). "Personality Traits and Concern for Privacy: An Empirical Study in the Context of Location-Based Services." *European Journal of Information Systems* 17 (4), 387–402.

Karwatzki, S., M. Trenz, V.K. Tuunainen, and D. Veit (2017). "Adverse Consequences of Access to Individuals' Information: An Analysis of Perceptions and the Scope of Organisational Influence." *European Journal of Information Systems* 26 (6), 688–715.

Kehr, F., T. Kowatsch, D. Wentzel, and E. Fleisch (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.

Krasnova, H., S. Spiekermann, K. Koroleva, and T. Hildebrand (2010). "Online Social Networks: Why We Disclose." *Journal of Information Technology* 25 (2), 109–125.

Li, Y. (2014). "The Impact of Disposition to Privacy, Website Reputation and Website Familiarity on Information Privacy Concerns." *Decision Support Systems* 57, 343–354.

Luo, X., H. Li, J. Zhang, and J.P. Shim (2010). "Examining Multi-dimensional Trust and Multi-faceted Risk in Initial Acceptance of Emerging Technologies: An Empirical Study of Mobile Banking Services." *Decision Support Systems* 49 (2), 222–234.

MacKenzie, S.B., P.M. Podsakoff, and N.P. Podsakoff (2011). "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques." *MIS Quarterly* 35 (2), 293–334.

Malhotra, N.K., S.S. Kim, and J. Agarwal (2004). "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model." *Information Systems Research* 15 (4), 336–355.

Marcoulides, G., W.W. Chin, and C. Saunders (2012). "When Imprecise Statistical Statements Become Problematic: A Response to Goodhue, Lewis, and Thompson." *MIS Quarterly* 36 (3), 717–728.

Mitchell, V.-W. (1999). "Consumer Perceived Risk: Conceptualisations and Models." *European Journal of Marketing* 33 (1/2), 163–195.

Moore, G.C. and I. Benbasat (1991). "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation." *Information Systems Research* 2 (3), 192–222.

Nunnally, J. (1978). *Psychometric Theory*. New York: McGraw-Hill.

Nunnally, J.C. and I.H. Bernstein (1994). *Psychometric Theory*. [online], 3rd ed New York: McGraw-Hill URL: http://rds.epi-ucsf.org/ticr/syllabus/courses/46/2005/10/20/Lecture/readings/Psychometric%20Theory.pdf (visited on 09/04/2014).

Ringle, C.M., S. Wende, and J.-M. Becker (2015). *SmartPLS 3*. [online] Boenningstedt: SmartPLS GmbH URL: www.smartpls.com.

Sarathy, R. and H. Li (2007). "Understanding Online Information Disclosure As a Privacy Calculus Adjusted by Exchange Fairness." *ICIS 2007 Proceedings* URL: http://aisel.aisnet.org/icis2007/21.

van Slyke, C., J.T. Shim, R. Johnson, and J. Jiang (2006). "Concern for Information Privacy and Online Consumer Purchasing." *Journal of the Association for Information Systems* 7 (6), 415–444.

Smith, H.J., T. Dinev, and H. Xu (2011). "Information Privacy Research: An Interdisciplinary Review." *MIS Quarterly* 35 (4), 989–1016.

Smith, H.J., S.J. Milberg, and S.J. Burke (1996). "Information Privacy: Measuring Individuals' Concerns about Organizational Practices." *MIS Quarterly* 20 (2), 167–196.

Steelman, Z.R., B.I. Hammer, and M. Limayem (2014). "Data Collection in the Digital Age: Innovative Alterantives to Student Samples." *Mis Quarterly* 38 (2), 355–378.

Stone, R.N. and K. Grønhaug (1993). "Perceived Risk: Further Considerations for the Marketing Discipline." *European Journal of Marketing* 27 (3), 39–50.

TRUSTe (2013). 2013 TRUSTe US Consumer Confidence Index URL: http://www.truste.com/us-consumer-confidence-index-2013/ (visited on 07/08/2013).

Wu, Y., S. Ryan, and J. Windsor (2009). "Influence of Social Context and Affect on Individuals' Implementation of Information Security Safeguards." *ICIS 2009 Proceedings* URL: http://aisel.aisnet.org/icis2009/70.

Xu, H., H.-H. Teo, B.C.Y. Tan, and R. Agarwal (2009). "The Role of Push--Pull Technology in Privacy Calculus: The Case of Location-Based Services." *Journal of Management Information Systems* 26 (3), 135–173.

Yu, J., P.J.-H. Hu, and T.-H. Cheng (2015). "Role of Affect in Self-Disclosure on Social Network Websites: A Test of Two Competing Models." *Journal of Management Information Systems* 32 (2), 239–277.