POSTPRINT VERSION OF:

Büchi, M., Just, N., & Latzer, M. (2016). Caring is not enough: The importance of Internet skills for online privacy protection. *Information, Communication & Society*. Advance online publication. http://dx.doi.org/10.1080/1369118X.2016.1229001

Caring Is Not Enough: The Importance of Internet Skills for Online Privacy Protection

Moritz Büchi[1], Natascha Just[1], and Michael Latzer[1]

Abstract

This article explains Internet users' self-help activities in protecting their privacy online using structural equation modeling. Based on a representative survey of Swiss Internet users, it reveals past experiences with privacy breaches as a strong predictor of current protective behavior. Further, in line with the "privacy paradox" argument, caring about privacy (privacy attitudes) alone does not necessarily result in substantial self-protection. Most strikingly, however, general Internet skills are key in explaining users' privacy behavior. These skills enable users to reduce risks of privacy loss while obtaining the benefits from online activities that increasingly depend on the revelation of personal data. Consequently, Internet skills are an essential starting point for public policies regarding users' self-help in privacy protection.

[1] University of Zurich, IPMZ, Media Change & Innovation Division, mediachange.ch

Introduction

The extent of individual benefits from Internet use highly depends on the disclosure of personal information. This occurs both deliberately, e.g., by self-publishing user profiles, by sharing pictures or commenting, and unknowingly, in the form of digital traces, e.g., left unintentionally during online searches or purchases. Contemporary information societies in which Internet users permanently balance the benefits of disclosure and the risk of privacy incursions (Acquisti, Brandimarte, & Loewenstein, 2015) are marked by several defining features: (1) big (social) data as a new asset class (Boyd & Crawford, 2012; Manovich, 2011; World Economic Forum, 2012), (2) new methods of extracting economic and social value from big data, e.g., variations in algorithmic selection that automatically assigns relevance to selected pieces of information (Latzer, Hollnbuchner, Just, & Saurwein, 2014), (3) a high potential for growth as well as new options and advantages in data-driven management in many sectors (McAfee & Brynjolfsson, 2012; OECD, 2013), and (4) a growing platformization of Internet-based (social media) markets that depend on the revelation of personal data (Geradin & Kuschewsky, 2013; Helmond, 2015).

The unprecedented availability of data on individuals' personal information, behavior, communication, and transactions has prompted much debate and research on the connected risks in terms of loss of privacy, privacy violations, and surveillance. At the same time, information disclosure and sharing also enhance and personalize services, simplify transactions, and strengthen social ties and social capital. It is thus rarely an attractive option for individuals to entirely opt out of services that potentially threaten both their privacy and the control over their own personal data (van Dijck, 2013). From a public-policy perspective, this calls for adequate governance and regulatory provisions for privacy protection (Bennett & Parsons, 2013; Zimmer, 2010). Accordingly, the discussion of privacy governance on the Internet has gained prominence, propelled by widely discussed surveillance scandals (e.g., US National Security Agency, NSA) and sensational law suits (e.g., against Google and Facebook), for example by the resolution of the United Nations General Assembly on the right to privacy in the digital age (A/RES/68/167). The reform of EU data protection rules, which entered into force in May 2016 (Regulation (EU) 2016/679; Directive (EU) 2016/680), and the pending review of the EU ePrivacy Directive (Directive 2002/58/EC) are further responses to privacy challenges in the digital

age. From a user's immediate micro perspective, personally managing and protecting the online self has become an essential part of everyday networked life (Rainie & Wellman, 2012). From an institutional perspective, the governance of privacy comprises a mix of interwoven actors and instruments including (1) market solutions, (2) self-help measures by users and individual companies, (3) collective self-regulation by the industry, (4) private-public co-regulation regimes, and (5) command-and-control regulation by states (Latzer, Just, Saurwein, & Slominski, 2003). In this mix, users' self-help privacy protection is gaining in relative importance, mainly because companies have an economic interest in user data and states are ill-adapted to keep pace with highly dynamic technological developments and corresponding know-how requirements.

Privacy behavior is also related to questions of the digital divide and digital inequality (Park, 2013, 2015). It has been hypothesized and demonstrated that Internet uses and skills are associated with existing social inequalities along socio-economic fault lines (DiMaggio, Hargittai, Celeste, & Shafer, 2004; Helsper, 2012; Robinson et al., 2015; Witte & Mannon, 2010; Zillien & Hargittai, 2009). Such disparities concern various Internet activities and competencies rather than basic questions of Internet access. Inequalities in the way the Internet is used by different social groups, i.e., second-level digital divides, have been analyzed for many countries (Blank, 2013; Brandtzæg, Heim, & Karahasanović, 2011; Büchi, Just, & Latzer, 2015; Hargittai, 2002; Teo, 2001; Van Deursen & Van Dijk, 2014; Wei, 2012). Marginalized or disadvantaged Internet users, meaning those who have not developed advanced digital skills or show constrained usage patterns, are likely to be more vulnerable to privacy threats. Being vigilant about privacy is only the outer layer of a nested digital habitus: social status differences influence access quality and autonomy, which in combination with know-how and Internet skills structure the online experience (Hargittai, 2008; Park, 2013; Robinson, 2009; Smith, Hewitt, & Skrbiš, 2015).

The Pew Research Center outlined future developments of the Internet and predicted that a privacy premium will come into play: only the well-to-do will know how to protect their privacy, while for many the perceived immediate gains of unmanaged information disclosure outweigh concerns (Anderson & Rainie, 2014). Essentially, this could transform privacy and control over personal information into a luxury good (Rainie & Anderson, 2014). But in the modern information society—where online communication is being established as a *fait social* (Schroeder & Ling, 2014)—Internet access, skills, use, and privacy need to

be treated as necessities (Hargittai, 2008; Hoffman, Novak, & Venkatesh, 2004). Not being able to put the Internet to effective and beneficial use may further the digital exclusion of certain social groups. As such, the extent to which Internet users can and do manage their personal online information is a relevant consequence and at the same time a further source of social inequalities. Consequently, users' everyday practical attempts to manage their information online while still benefiting from the use of data-hungry Internet services become the center of public-policy attention.

Focusing on Internet users' self-help

This article focuses on forms of self-help in privacy protection, on variations in users' actions to protect their privacy online. It approaches this phenomenon with the assumption that protective behavior depends on users' attitudes towards privacy, on past experiences with privacy breaches, and on the level of general skills in using the Internet.

In market economies, market solutions, self-help, self-organization and self-regulation are in general preferred over state intervention via command-and-control regulations. State regulation is applied only if private action (e.g., privacy protection by self-help) is not sufficient (subsidiarity). According to a conceptual framework of governance choice that considers contextual factors of governance including incentives, conflicts of interest, or intervention capacity (Latzer et al., 2003; Latzer, Price, Saurwein, & Verhulst, 2007), self-help is an adequate measure against privacy risks. The potential for private solutions is high, ranging from not using problematic services to technical self-help, including the use of privacy-enhancing technologies (PET), cookie-management and do-not-track technologies. In contrast, individual self-organization by companies and collective industry self-regulation lack the necessary incentives, because companies profit from the revelation, collection and trading of personal data (Saurwein, Just, & Latzer, 2015). Accordingly, default privacy settings are generally low and have decreased over time (Acquisti et al., 2015), compelling users to become active. Nonetheless, although users are increasingly concerned about their privacy, research reveals the phenomenon of a "privacy paradox," meaning that despite knowledge and concern about risks and breaches, people readily share information and engage in behavior that could threaten their privacy (Norberg, Horne, & Horne, 2007). Hence it is in the public interest to better understand

active privacy protection by Internet users, including enablers and barriers, as well as options for state interventions concerning privacy protection that might reduce barriers to self-help.

Online privacy behavior may be viewed from two complementary perspectives: information release and information management. Research on the user level has thus focused both on self-disclosure, particularly in the context of social networking sites (e.g., Taddei & Contena, 2013; Taddicken, 2014), and on the active protection of personal information online (e.g., Litt & Hargittai, 2014; Park, Campbell, & Kwak, 2012). Sharing and revealing personal information is key to developing relationships (Utz, 2015), therefore privacy-protective behavior could be seen as socially undesirable. However, seeking privacy even while sharing information online is not necessarily a contradiction (Acquisti et al., 2015). Stutzman, Gross, and Acquisti (2012) showed how Facebook users have over time increased the amount of information shared with friends while increasingly restricting the amounts available to the public, e.g., by changing platform privacy settings. Privacy protection and self-disclosure are not necessarily correlated and these actions are determined by different individual-level factors (Chen & Chen, 2015). Consequently, privacy protection behavior is a highly relevant measure and more useful than assessing self-disclosure in the context of our research interest (see Tufekci, 2007).

Research interest, contribution, and hypotheses

There are numerous studies on privacy actions in SNS (social network sites) (Boyd & Hargittai, 2010; Chen & Chen, 2015; Debatin, Lovejoy, Horn, & Hughes, 2009; Feng & Xie, 2014; Grubbs Hoy & Milne, 2010; Litt, 2013b; Raynes-Goldie, 2010; Tufekci, 2007; Van den Broeck, Poels, & Walrave, 2015; Young & Quan-Haase, 2013), which typically analyze young Facebook users. However, with Internet diffusion rates above 80% in developed countries (International Telecommunication Union, 2015) and with privacy risks becoming apparent not only for applications that entail *explicit* information disclosure, there is a gap in research on privacy-protecting behavior in the general population covering the whole range of Internet activities and not only SNS behavior.

This study aims to close this gap. It empirically addresses the question of how individuals' privacy protection actions are influenced by their overall Internet skills, attitudes towards personal information and past privacy breaches. Its

unique contribution lies in the combination of three crucial features that to our knowledge have not been simultaneously employed in the existing literature. First, we use nationally representative telephone survey data as opposed to many previous studies, which used smaller convenience samples. The advantage is that our data are not biased, e.g., with regard to Internet skills or privacy attitudes, as may be the case with self-selection into online surveys or in specific societal sub-groups such as Facebook users. Implications for public policies that are implemented at the national level need to be based on nationally representative and current data. The study by Park (2013), for example, is one of the few to explicitly address the effect of digital literacy on user privacy protection in light of Internet policy, but is based on 2008 data. Due to the Internet's continued diffusion, the user base is constantly evolving and requires up-to-date research.

Second, the measures of privacy attitudes and breaches as well as self-protective privacy actions relate to Internet activities at large. This reflects the need to treat online privacy as a general requirement in the information society and to expand privacy research beyond voluntary self-disclosure on SNS. Third, we adopt a recently developed Internet skills framework (see Van Deursen, Helsper, & Eynon, 2015) to analyze the role of general skills rather than privacy skills alone, again aiming to increase the generalizability and validity of previous findings. Research on the second-level digital divide has shown that Internet skills are key in explaining different online uses and various forms of participation (e.g., Brake, 2014; De Marco, Robles, & Antino, 2014; Lutz, 2015). Such engagement is impeded if users cannot and do not protect their online information. Systematic research on the link between general Internet skills and privacy protection—both preconditions for effective use and thus potential intensifiers of digital inequalities—has however been scarce. Accordingly, Litt (2013b) explicitly proposes the inclusion of digital skills as a predictor of user privacy protection. Overall, this framework allows us to unravel the relationship between knowing how to perform various online actions and actual usage—two distinct privacy-relevant constructs that have been conflated in previous research (e.g. Van den Broeck et al., 2015).

The structural online privacy model (Figure 1) tests four hypotheses based on previous literature. Hypotheses and corresponding rationales are as follows:

Hypothesis 1 (H1). Internet users with higher levels of general Internet skills will engage in more self-protective privacy behavior online.

Following the digital inequality literature, variations in Internet skills contribute substantially to explaining different types of engagement with and uses of the Internet (e.g. Litt, 2013a; Van Deursen & Van Dijk, 2011). Turning specifically to privacy outcomes, based on a convenience sample of 547 young adults in 2012, Litt and Hargittai (2014) showed that Internet skills reduced the likelihood of having experienced negative consequences of information sharing in SNS. Similarly, using a 2008 probability sample of 419 adult Internet users, Park (2013) found that individuals with higher levels of generic technical familiarity with the Internet were more likely to control their personal information.

Hypothesis 2 (H2). Internet users who place high importance on online personal privacy will engage in more self-protective privacy behavior online.

In addition to knowing how to effectively navigate the web, a precondition for active self-protection of online privacy is high valuation of one's personal information. Research on the privacy paradox has demonstrated inconsistencies between privacy concerns and privacy behavior in the form of revealing personal information online (Norberg, Horne, & Horne, 2007; Taddicken, 2014). However, in the context of SNS use, in a 2011 sample of 515 college students, Chen and Chen (2015) found a strong positive effect of privacy concern on limiting profile visibility. The mixed findings of the extant literature are likely due to different operationalizations of the main concepts *concern* and *behavior* (Kokolakis, 2015). In their convenience sample of 595 Internet users, Dienlin and Trepte (2015) demonstrated that users' privacy behavior was not necessarily paradoxical; informational privacy attitudes had a positive effect on protective behavior.

Hypothesis 3 (H3). Internet users who have experienced privacy breaches will engage in more self-protective privacy behavior online.

Hypothesis 4 (H4). Internet users who have experienced privacy breaches will place higher importance on online privacy.

In addition to skills and attitudes, the research model (Figure 1) includes privacy breaches as a control variable: we suppose that having experienced privacy incursions and perhaps negative emotional consequences in the past greatly increases an individual's awareness and likelihood of self-help privacy protection (Debatin et al., 2009; Dienlin & Trepte, 2015). In their Facebook study using a sample of 119 college undergraduates, Debatin et al. (2009) found that users who had personally experienced privacy invasions were more likely to change privacy settings.

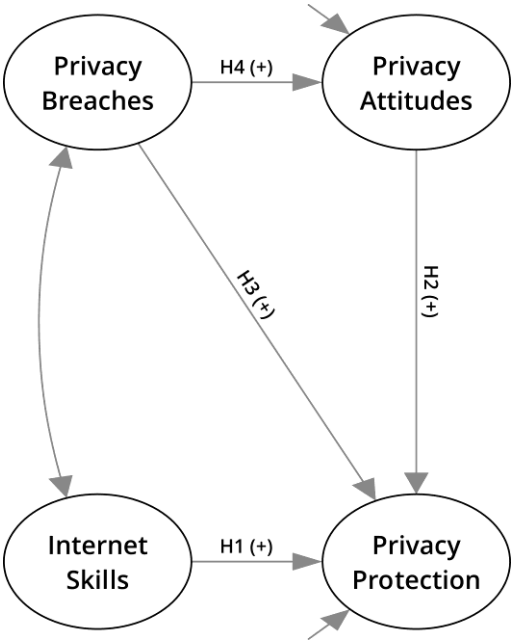

Figure 1. Overview of the structural online privacy model and hypotheses.

Data and methods

The hypotheses and measurement models were tested using structural equation modeling (SEM) and confirmatory factor analysis (CFA) with maximum likelihood estimation, robust Huber–White standard errors and full-information maximum likelihood (FIML) estimation for missing values (Graham, 2009) in the software environment R (version 3.1.0) with the lavaan package (version 0.5-17; Rosseel, 2012). The advantages and hence reason for our choice of SEM lie in the possibility to straightforwardly transfer the theoretical model into a testable statistical model, simultaneously estimate multiple direct and indirect effects, test and include latent variables by means of CFA, and explicitly model measurement errors (see Brown, 2006; Kline, 2011). Furthermore, global goodness of fit measures make it possible to evaluate the match between the conceptual model and the empirical data rather than only calculating the statistical significance of single coefficients. Our hypotheses involve multiple latent variables and mediated relationships, and thus benefit from the greater versatility of SEM as compared to ordinary regression.

Data collection

Data for this study were collected as part of a privacy module that was integrated into a biennial survey of Internet use in Switzerland, investigating the social, political, and economic impact of ICTs. Respondents were contacted on landline and mobile phones between 27 May 2015 and 29 June 2015. Sampling quota were constructed based on age, gender, region, and employment status.

Sample characteristics

A total of 1121 respondents completed the telephone interviews. Table 1 presents several characteristics of the 970 Internet users within this sample. The total sample is representative of the Swiss population between the ages of 14 and 84 who speak German, French, or Italian.

Table 1
Main Characteristics of Survey Respondents Who Use the Internet in Switzerland in 2015

	Percentage (N)
Total	100 (970)
Female	48.14 (467)
Higher education	36.29 (352)
Full-time or part-time employed	70.31 (682)
German-speaking region	65.05 (631)
French-speaking region	22.37 (217)
Italian-speaking region	12.58 (122)
Users of mobile Internet	72.16 (700)
Users of social networking sites	59.01 (573)
	M (SD)
Age	44.39 (17.62)
Years of Internet use	12.75 (6.09)
Minutes of daily Internet use	186.39 (177.16)

Latent variables

The following reports on the measurement models that were tested with confirmatory factor analysis (Brown, 2006) for the four latent variables in Figure 1. Models are evaluated by the minimum function test statistic (χ^2), degrees of freedom (df), χ^2/df, comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Concordant with widely accepted cutoff criteria in CFA and SEM (Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger, & Müller, 2003), values for $\chi^2/df \leq 2$, CFI $\geq .97$, TLI $\geq .97$, RMSEA $\leq .05$, and SRMR $\leq .05$ suggest a good model fit. Given the large sample size, and because it tests the hypothesis of an exact fit between the empirical and the model-implied covariance structure, a significant χ^2 is not sufficient to reject an otherwise fitting model (Byrne, 2010; Kline, 2011).

Table 2
Measurement Items for the Four Latent Variables of the Online Privacy Model

Latent variable	Item	Wording	Scale	M (SD)
Internet Skills	Operational*	*I know how to open downloaded files.*	5-point	4.46 (1.08)
	Information	*I find it easy to decide on the best keywords for web search.*	5-point	3.87 (1.10)
	Social	*I know how to change who I share content with.*	5-point	3.38 (2.24)
	Creative	*I know how to create and upload content.*	5-point	2.96 (2.42)
	Mobile	*I know how to download apps to a mobile device.*	5-point	3.95 (2.37)
Privacy Attitudes	Search*	*How important is it for you that only you or people you authorize know which search queries you perform?*	5-point	3.52 (1.96)
	Location	*…where you are located when using the Internet?*	5-point	3.64 (2.05)
	Websites	*…which websites you visit?*	5-point	3.69 (1.88)
	Correspondence	*…with whom you communicate over the Internet?*	5-point	3.89 (1.90)
	Content**	*…the content of your e-mails or other correspondence?*	5-point	4.13 (1.33)
Privacy Breaches	Violation*	*Has your privacy ever been violated online?*	binary	0.11 (0.22)
	Abuse	*Thinking of the past year, did you feel that your personal data was passed on or abused?*	binary	0.31 (0.10)
Privacy Protection	Settings*	*Do you change settings so that content is only visible to specific people?*	4-point	1.89 (1.42)
	Monitor	*Do you monitor which information is available about you online?*	4-point	1.99 (0.97)
	Fake	*Do you use fake information online such as a fake name?*	4-point	1.51 (0.82)
	Cookies	*Do you block, delete, or deactivate cookies?*	4-point	2.71 (1.55)
	Delete**	*Do you ask other people or service providers to delete personal information about you?*	4-point	1.72 (1.03)

Note. Item wordings are translated to English.
* The latent variable was scaled to this reference item by constraining its unstandardized factor loading to unity.
** Item excluded due to lack of empirical fit with the theoretical construct.

Privacy protection behavior

To measure privacy behavior in the sense of self-protective actions, we adapted four items from the Pew Research Center's Internet & American Life Project (Rainie, Kiesler, Kang, & Madden, 2013) and a Eurobarometer survey on data protection (European Commission, 2011). The reference item of this factor asks whether the respondent changes settings so that content is only visible to specific people on a four-point frequency scale ranging from *never* to *frequently*. Respondents were also asked if they monitor the information available about them online, use fake information such as a fake name, or block, delete or deactivate cookies. A fifth question about requesting the deletion of personal information did not load strongly onto the factor and was excluded from subsequent analyses. The empirical covariance matrix showed a very close fit to the model-implied factor structure (χ^2 (2)=2.06, p=.357, CFI=1.000, RMSEA=.01, SRMR=.01). Protective measures are more frequently employed by men than women and by younger Internet users (Figure 2).

Internet skills

To measure Internet skills, this article adopts a validated survey instrument for general populations (Van Deursen, Helsper, & Eynon, 2014, 2015). The original short scale uses 23 items to measure five different skill factors: operational, information navigation, social, creative and mobile (Van Deursen et al., 2014). To keep the questionnaire shorter, only the reference or highest-loading item of each factor was included in our survey (Table 2). The first measurement model tested a single factor labeled *Internet skills* with the five items as indicators. This showed a poor to acceptable fit (χ^2 (5)=61.22, p<.001, CFI=.936, RMSEA=.11, SRMR=.04). Modification indices revealed that social skill and creative skill are significantly correlated beyond their common factor. Given that both questions ask about content online (Table 2), the covariance between their residual variances was freely estimated. This resulted in an acceptable to good fit of the data to the model (χ^2 (4)=18.55, p=.001, CFI=.983, RMSEA=.06, SRMR=.02) enabling the inclusion of Internet skills as a predictor in the privacy model. Reported Internet skills, similar to privacy protection, clearly fall off with age, and men score slightly higher than women (Figure 2).

Privacy attitudes

The Pew Research Center's Internet & American Life Project covered attitudes towards personal data produced in online communication, surfing, and application use by asking respondents to rate the sensitivity of nine data types (Rainie et al., 2013). We propose that these questions measure individuals' privacy attitudes and adapt five of these items. Respondents were asked how important it was to them that only they or those they authorize know (1) which searches they perform, (2) where they are located when using the Internet, (3) which websites they visit, (4) with whom they communicate over the Internet, and (5) the content of their online correspondence. Modification indices of the initial model revealed a lack of fit due to the content item. After excluding this variable, CFA of the four-item privacy attitudes factor showed a very good model fit (χ^2 (2)=3.30, p=.192, CFI=.998, RMSEA=.03, SRMR=.01). Women tend to be more sensitive about their personal information than men, whereas the effect of age is inconclusive (Figure 2).

Privacy breaches

To assess the degree to which individuals had been subject to privacy breaches, we asked if they had experienced violations of their privacy and if they felt that personal data had been passed on or abused. The correlation coefficient between the two variables is .30 (p<.001). Privacy breaches show very little variation across gender and age groups (Figure 2).

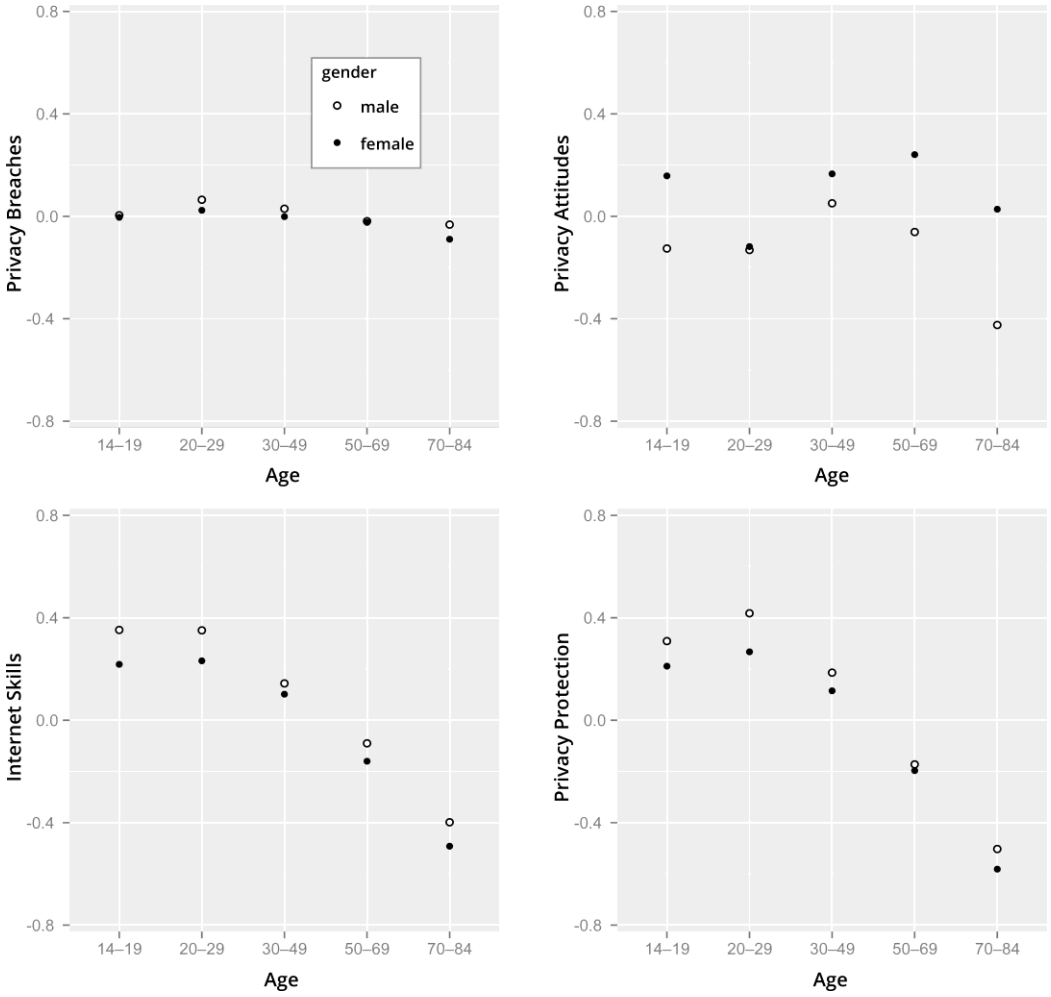

Figure 2. Overview of the predicted factor scores of the latent variables in the model. Means are plotted by gender and age group. The bottom right panel shows the main dependent variable (privacy protection behavior).

Results

The four hypotheses link the latent measures in the structural online privacy model (Figure 3). The implied model fits the observed data very well and all paths are significant at the .001 level.

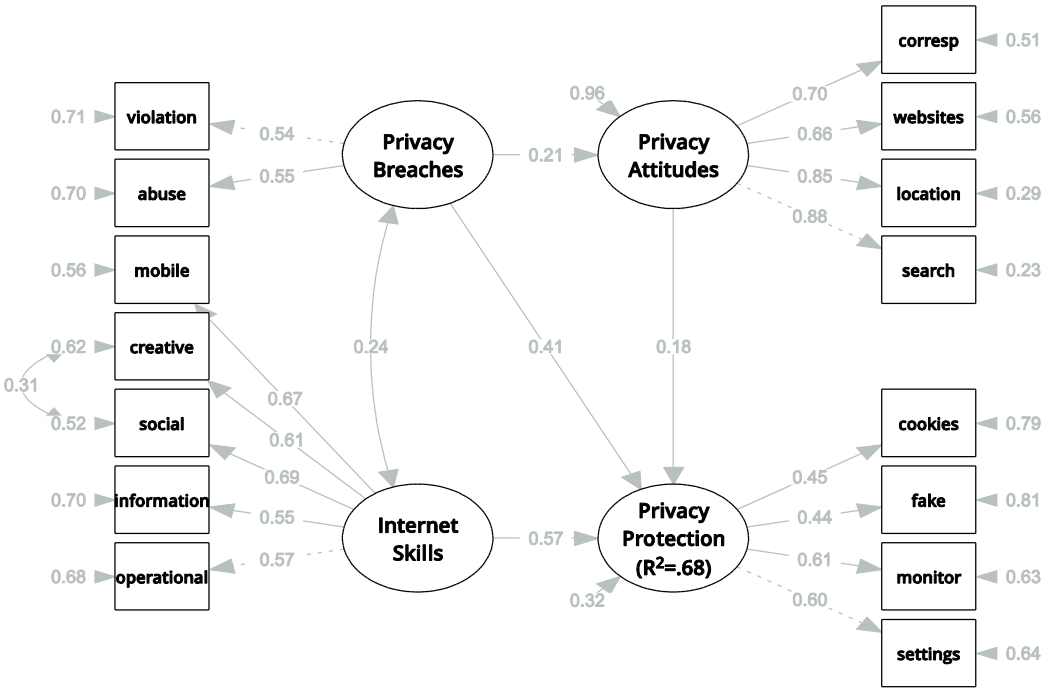

Figure 3. Full online privacy model with standardized parameter estimates. All paths p<.001. The model converged normally after 84 iterations. 51 free parameters, 4 latent intercepts fixed to 0 and 4 reference item factor loadings fixed to one (dashed paths) entered the model. The privacy model shows a very close fit to the data (χ^2 (84)=159.53, p<.001, χ^2/df=1.90, CFI=.980, RMSEA=.03, SRMR=.03; baseline model: χ^2 (105)=3831.11, p<.001, χ^2/df=36.49).

The effect of Internet skills on privacy protection behavior

The primary hypothesis (H1) predicted that the more skilled Internet users are, the more they would engage in privacy protection. The strong positive effect in the model supports this (β=.57). The bivariate correlation between the factor scores for Internet skills and privacy protection is remarkably high (r=.77). This is despite the fact that only one of the skills items is substantially connected to a privacy issue (social skill, "I know how to change who I share content with"). As shown in Figure 4, the relationship is very well represented linearly. The local regression fit line additionally suggests an even greater increase of self-protective behavior for Internet users with above average skills. The inclusion of privacy breaches and attitudes in the multivariate privacy model unsurprisingly weakened this relationship; however, skills remain the strongest predictor of privacy

protection. The model thus demonstrates that caring is not enough—users also need the general skills in navigating the Internet in order to apply self-help measures in their everyday Internet use.

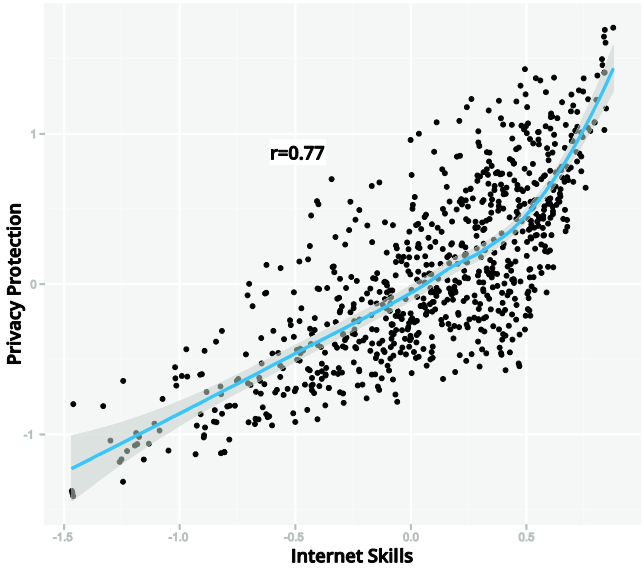

Figure 4. Bivariate relationship between Internet skills and privacy protection (factor scores). The line shows a local regression (LOESS) fit to the data with the surrounding 95% confidence interval.

The roles of privacy breaches and attitudes

H2 predicted that Internet users who place high importance on personal privacy engage in more protective behavior. The significant positive regression coefficient (β=.18) supports this. The comparatively low effect size indicates that attitudes are not the primary explanation for users' varying levels of privacy protection, i.e., valuing the control of personal information does not necessarily induce concrete actions towards this end. H3, which was concerned with the influence of privacy breaches, is supported by a strong positive effect (β=.41). Internet users who report experiences of privacy violation are likely to engage in more privacy protection. Turning to H4, such privacy breaches further predicted stronger pro-privacy attitudes (β=.21). The bivariate correlations between privacy attitudes and protection (r=.38) and privacy breaches and protection (r=.74) are substantially higher than the controlled effects in the structural equation model (Figure 3; β=.18

and β=.41, respectively). This means that Internet skills are an important complementary predictor of self-protective privacy behavior.

Explaining privacy behavior online

The proposed model fits the empirical data very well and can explain a major part of the variance in privacy protection (R^2=.68, Figure 3). The hypotheses regarding the positive direct influence of Internet skills (H1), privacy attitudes (H2), and privacy breaches (H3) were supported, as was the hypothesis that breaches have an indirect effect via attitudes (H4). In determining whether or not individuals perform a certain behavior, three conditions need to be satisfied: the actor must have sufficient motivation and ability, and experience something that triggers the behavior (Fogg, 2009). Applied to the present study, this means that privacy protection behavior is promoted by the attitude that personal information should be protected, by the abilities to perform protective actions, and as a behavioral intensifier by the experience or strong suspicion that one's privacy has been violated. The relative weights of these factors in the privacy model reveal Internet skills as the key driver of self-help privacy protection.

Internet skills and privacy breaches are positively correlated (Figure 3). This may seem paradoxical considering that, by virtue of their digital expertise, more skilled users should be able to better avoid privacy violations and data abuse. However, given that skilled individuals are generally more intensive and extensive Internet users, their exposure to such threats is also much higher (see Litt & Hargittai, 2014). In order to investigate this, we constructed a measure of the amount of Internet use. This is computed by summing the frequency of engaging in 37 diverse online activities, ranging from e-mailing to online banking to instant messaging (also see Blank & Groselj, 2014). The amount of use is positively correlated with privacy breaches (r=.32), suggesting that the mere probability of exposure is an important predictor of negative experience with personal information. Heavy users of the Internet are routinely confronted with privacy issues—accordingly, privacy protection as an adaptive behavior is also strongly correlated with the amount of use (r=.53). General privacy attitudes on the other hand, are largely unrelated to the amount of Internet use (r=.07). This means that highly active users are not more or less sensitive about their personal information than low-use individuals, but since they are more likely to encounter

privacy threats, they have developed strategies to manage their online privacy—contingent upon their skills to do so.

In the interest of parsimonious modeling and focusing on the four hypotheses, sociodemographics are not included in the main model. However, given that previous multi-country research has shown very strong age effects on the way the Internet is used (e.g. Büchi et al., 2015), it makes sense to test the influence of age on privacy protection behavior (see Van den Broeck, Poels, & Walrave, 2015). An additional model introducing age as an exogenous predictor of skills and privacy protection (age is unrelated to breaches and attitudes) resulted in a comparably worse fit (χ^2 (97)=314.78, p=.000, CFI=.948, RMSEA=.05, SRMR=.04), but increased the amount of explained variance in privacy protection to 73%. Older Internet users show lower levels of privacy protection (β=-.33). Additionally, they are less skilled (β=-.50), meaning that age has both a direct effect as well as an indirect effect via Internet skills on the intensity of self-protective privacy behavior. Preliminary analyses of variance suggested a trend for men and highly educated Internet users to engage in more self-protective behavior than women and those with medium or low education—but these differences failed to reach statistical significance in our sample.

Summary, discussion, and policy implications

In Internet governance, users' self-help activities, e.g., personal privacy protection, are gaining in importance. The goal of this article was to explain the variation in Internet users' personal privacy protection in an information society where voluntary and involuntary disclosure of personal data is increasingly important or even a precondition to gain full benefits of Internet use. In order to understand and govern privacy risks, well-founded knowledge is required on factors that influence the degree of individuals' self-help. This study significantly contributes to such knowledge and extends existing research in that it uses representative survey data, covers Internet activities at large, and thus widens the scope beyond SNS, includes a variety of concrete actions to protect privacy, and focuses on the role of general Internet skills as opposed to privacy skills exclusively.

By building upon previous work in the domains of communications policy and governance, psychological social media research, as well as digital inequality and Internet studies, four hypotheses were tested and supported by the estimates

in the privacy model (Figure 3). First, having experienced privacy breaches in the past is a strong predictor of current protective behavior. Perceived data abuse and concrete experiences of privacy violations are associated with higher levels of protection (H3), suggesting a "learning the hard way" mechanism. Second, pro-privacy attitudes as evaluations of the sensitivity of personal information have a significant but relatively weak positive influence on privacy protection (H2). Third, these attitudes are also positively influenced by privacy breaches (H4). Fourth, Internet skills have by far the strongest effect on behavior (H1). In this model, it is thus general Internet skills that best explain the extent to which users actively protect their privacy online. Despite agreement levels well above the scale center for all items measuring privacy attitudes (Table 2), caring about privacy is evidently not sufficient to provoke strong self-protective behavior. This finding is consistent with research on self-disclosure (Taddicken, 2014) and extends the applicability of the "privacy paradox" to self-protective behavior. Future research should also scrutinize the twofold exacerbated vulnerability of older adults, as we found this group to be less skilled in using the Internet and less active in protecting their online privacy (also see Park, 2013). Alongside the investigation of age effects, subsequent research may integrate additional socioeconomic indicators and context factors such as social support into models of privacy behavior.

As Internet access approaches saturation, online privacy threats become ubiquitous. Skills enable users to continue benefiting from their access to and use of the Internet by mitigating the risks of being online. The strong effect of general Internet skills on privacy protection in this study is therefore of great importance for digital inequality. While previous studies have demonstrated the role of privacy knowledge, our results show that those with low levels of general skills will be ill-equipped to selectively reveal and control their online information. These empirical insights are particularly relevant for policy-making and research concerned with digital inclusion: Internet skills are an asset in market economies and are associated with a broad range of beneficial Internet uses, yet they are unequally distributed in the population based on existing stratification by social class and status (Witte & Mannon, 2010). The finding that skills rather than attitudes mainly influence privacy (self-help) behavior leads to the conclusion that policies aimed at empowering users may promise little success if concerned solely with raising *awareness*. Similarly, training less-skilled Internet users primarily for specific tasks such as changing the visibility settings of their Facebook pictures

may not have a long-term empowering effect, due to the fast-changing nature of Internet technologies and services. Furthermore, the mere existence of or the strengthening of *citizens' rights* as recently stipulated, for example by the EU data protection rules (e.g., Regulation (EU) 2016/679), may also have limited effects if the design of the practical implementation and accompanying measures does not pay sufficient attention to skill-based divides. This reform includes provisions for the compulsory notification of data breaches, the encouragement of certification mechanisms and data protection seals for privacy-compliant processes, as well as users' rights to data portability, erasure and rectification of personal data. Most of these provisions require compliance by industry stakeholders, adequate control mechanisms as well as skills on the part of the users to enable them to fully benefit from this potential empowerment. Consequently, there is a need for adaptive public policies that also ensure that *universal* and *transferable* digital skills are constantly developed, maintained, and enhanced in the light of continuous media change.

References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, *347*(6221), 509–514. doi:10.1126/science.aaa1465

Anderson, J., & Rainie, L. (2014). *Digital Life in 2015*. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/2014/02/25/the-web-at-25-in-the-u-s

Bennett, C. J., & Parsons, C. (2013). Privacy and surveillance: The multidisciplinary literature on the capture, use, and disclosure of personal information in cyberspace. In W. H. Dutton (Ed.), *The Oxford handbook of Internet studies* (pp. 486–508). Oxford, UK: Oxford University Press. doi:10.1093/oxfordhb/9780199589074.013.0023

Blank, G. (2013). Who creates content? Stratification and content creation on the Internet. *Information, Communication & Society*, *16*(4), 590–612. doi:10.1080/1369118X.2013.777758

Blank, G., & Groselj, D. (2014). Dimensions of Internet use: amount, variety, and types. *Information, Communication & Society*, *17*(4), 417–435. doi:10.1080/1369118X.2014.889189

Boyd, D., & Crawford, K. (2012). Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication \& Society*, *15*(5), 662–679. doi:10.1080/1369118X.2012.678878

Boyd, D., & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*, *15*(8). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3086/2589

Brake, D. R. (2014). Are we all content creators now? Web 2.0 and digital divides. *Journal of Computer-Mediated Communication, 19*, 591–609. doi: 10.1111/jcc4.12042

Brandtzæg, P. B., Heim, J., & Karahasanović, A. (2011). Understanding the new digital divide—A typology of Internet users in Europe. *International Journal of Human-Computer Studies*, *69*(3), 123–138. doi:10.1016/j.ijhcs.2010.11.004

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

Büchi, M., Just, N., & Latzer, M. (2015). Modeling the second-level digital divide: A five-country study of social differences in Internet use. *New Media & Society*. Advance online publication. doi:10.1177/1461444815604154

Byrne, B. M. (2010). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*. London: Routledge.

Chen, H.-T., & Chen, W. (2015). Couldn't or wouldn't? The influence of privacy concerns and self-efficacy in privacy management on privacy protection. *Cyberpsychology, Behavior, and Social Networking*, *18*(1), 13–19. doi:10.1089/cyber.2014.0456

De Marco, S., Robles, J. M., & Antino, M. (2014). Digital skills as a conditioning factor for digital political participation. *Communications*, *39*(1), 43–65. doi: 10.1515/commun-2014-0004

Debatin, B., Lovejoy, J. P., Horn, A. K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, *15*(1), 83–108. doi:10.1111/j.1083-6101.2009.01494.x

Dienlin, T. & Trepte, S. (2015). Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors. *European Journal of Social Psychology, 45*, 285–297. doi: 10.1002/ejsp.2049

DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). From unequal access to differentiated use. In K. M. Neckerman (Ed.), *Social Inequality* (pp. 355–400). New York: Russell Sage Foundation.

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).OJ L 201, 31.7.2002, p. 37-47.

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA. OJ L 119, 4.5.2016, p. 89–131.

European Commission (2011). Special Eurobarometer 359: Attitudes on data protection and electronic identity in the European Union. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf

Feng, Y., & Xie, W. (2014). Teens' concern for privacy when using social networking sites: An analysis of socialization agents and relationships with privacy-protecting behaviors. *Computers in Human Behavior*, *33*, 153–162. doi:10.1016/j.chb.2014.01.009

Fogg, B. (2009). A behavior model for persuasive design. *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*. doi:10.1145/1541948.1541999

Geradin, D., & Kuschewsky, M. (2013). Competition law and personal data: Preliminary thoughts on a complex issue. Retrieved from http://ssrn.com/abstract=2216088

Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annual Review of Psychology*, *60*, 549–76. doi:10.1146/annurev.psych.58.110405.085530

Grubbs Hoy, M., & Milne, G. (2010). Gender differences in privacy-related measures for young adult Facebook users. *Journal of Interactive Advertising*, *10*(2), 28–45. doi:10.1080/15252019.2010.10722168

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4). Retrieved from http://firstmonday.org/article/view/942/864

Hargittai, E. (2008). The digital reproduction of inequality. In D. Grusky (Ed.), *Social Stratification* (pp. 936–944). Boulder, CO: Westview Press.

Helmond, A. (2015). The Platformization of the web: Making web data platform ready. *Social Media + Society*, *1*(2). doi:10.1177/2056305115603080

Helsper, E. J. (2012). A corresponding fields model for the links between social and digital exclusion. *Communication Theory*, *22*, 403–426. doi:10.1111/j.1468-2885.2012.01416.x

Hoffman, D. L., Novak, T. P., & Venkatesh, A. (2004). Has the Internet become indispensable? *Communications of the ACM*, *47*(7), 37–43.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. doi:10.1080/10705519909540118

International Telecommunication Union. (2015). *ICT facts and figures*. Geneva, Switzerland: International Telecommunications Union. Retrieved from https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Kokolakis, S. (2015). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*. Advance online publication. doi: 10.1016/j.cose.2015.07.002

Latzer, M., Hollnbuchner, K., Just, N., & Saurwein, F. (2014). *The economics of algorithmic selection on the Internet*. Zurich, Switzerland: University of Zurich. Retrieved from http://mediachange.ch/media//pdf/publications/Economics_of_algorithmic_selection_WP.pdf

Latzer, M., Just, N., Saurwein, F., & Slominski, P. (2003). Regulation remixed: institutional change through self and co-regulation in the mediamatics sector. *Communications & Strategies*, *50*(2), 127–157.

Latzer, M., Price, M. E., Saurwein, F., & Verhulst, S. G. (2007). *Comparative analysis of international co- and self-regulation*. Vienna, Austria: Institute of Technology Assessment of the Austrian Academy of Sciences.

Litt, E. (2013a). Measuring users' internet skills: A review of past assessments and a look toward the future. *New Media & Society, 15*(4), 612–630. doi: 10.1177/1461444813475424

Litt, E. (2013b). Understanding social network site users' privacy tool use. *Computers in Human Behavior*, *29*, 1649–1656. doi:10.1016/j.chb.2013.01.049

Litt, E., & Hargittai, E. (2014). A bumpy ride on the information superhighway: Exploring turbulence online. *Computers in Human Behavior*, *36*, 520–529. doi:10.1016/j.chb.2014.04.027

Lutz, C. (2015). *Online but still not taking part? Investigating online participation divides in Germany* (Doctoral dissertation). Retrieved from https://www.alexandria.unisg.ch/246247/

Manovich, L. (2011). Trending: The promises and the challenges of big social data. In M. K. Gold (Ed.), *Debates in the Digital Humanities* (pp. 460–475). Minneapolis, MN: University of Minnesota Press.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The management revolution. *Harvard Buiness Review*, *90*(10), 61–68. doi:10.1007/s12599-013-0249-5

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100–126. doi:10.1111/j.1745-6606.2006.00070.x

OECD. (2013). *Exploring data-driven innovation as a new source of growth. Mapping the policy issues raised by "big data."* Paris, France: Organisation for Economic Co-operation and Development. Retrieved from

http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP(2012)9/FINAL&docLanguage=En

Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research*, *40*(2), 215–236. doi:10.1177/0093650211418338

Park, Y. J. (2015). Do men and women differ in privacy? Gendered privacy and (in)equality in the Internet. *Computers in Human Behavior*, *50*, 252–258. doi:10.1016/j.chb.2015.04.011

Park, Y. J., Campbell, S. W., & Kwak, N. (2012). Affect, cognition and reward: Predictors of privacy protection online. *Computers in Human Behavior*, *28*, 1019–1027. doi:10.1016/j.chb.2012.01.004

Rainie, L., & Anderson, J. (2014). *The future of privacy*. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/2014/12/18/future-of-privacy/

Rainie, L., Kiesler, S., Kang, R., & Madden, M. (2013). *Anonymity, privacy, and security online*. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/

Rainie, L., & Wellman, B. (2012). *Networked: The new social operating system*. Cambridge, MA: MIT Press.

Raynes-Goldie, K. (2010). Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook. *First Monday*, *15*. doi:10.5210/fm.v15i1.2775

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). OJ L 119, 4.5.2016, p. 1–88.

Robinson, L. (2009). A taste for the necessary. *Information, Communication & Society*, *12*(March), 488–507. doi:10.1080/13691180902857678

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., … Stern, M. J. (2015). Digital inequalities and why they matter. *Information, Communication & Society*, *18*(5), 569–582. doi:10.1080/1369118X.2015.1012532

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Saurwein, F., Just, N., & Latzer, M. (2015). Governance of algorithms: options and limitations. *Info*, *17*(6), 35–49. doi:10.1108/info-05-2015-0025

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schroeder, R., & Ling, R. (2014). Durkheim and Weber on the social implications of new information and communication technologies. *New Media & Society*, *16*(5), 789–805. doi:10.1177/1461444813495157

Smith, J., Hewitt, B., & Skrbiš, Z. (2015). Digital socialization: young people's changing value orientations towards internet use between adolescence and early adulthood. *Information, Communication & Society*, *18*(9), 1022–1038. doi:10.1080/1369118X.2015.1007074

Stutzman, F., Gross, R., & Acquisti, A. (2012). Silent listeners: The evolution of privacy and disclosure on facebook. *Journal of Privacy and Confidentiality*, *4*(2), 7–41.

Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, *29*, 821–826. doi:10.1016/j.chb.2012.11.022

Taddicken, M. (2014). The "privacy paradox" in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, *19*(2), 248–273. doi:10.1111/jcc4.12052

Teo, T. S. H. (2001). Demographic and motivation variables associated with Internet usage activities. *Internet Research*, *11*(2), 125–137. doi:10.1108/10662240110695089

Tufekci, Z. (2007). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society*, *28*(1), 20–36. doi:10.1177/0270467607311484

Utz, S. (2015). The function of self-disclosure on social network sites: not only intimate, but also positive and entertaining self-disclosures increases the feeling of connection. *Computers in Human Behavior*, *45*. doi:10.1016/j.chb.2014.11.076

Van den Broeck, E., Poels, K., & Walrave, M. (2015). Older and wiser? Facebook use, privacy concern, and privacy protection in the life stages of emerging, young, and middle adulthood. *Social Media + Society*. Advance online publication. doi: 10.1177/2056305115616149

Van Deursen, A., Helsper, E. J., & Eynon, R. (2014). *Measuring digital skills: From digital skills to tangible outcomes project report*. Retrieved from www.oii.ox.ac.uk/research/projects/?id=112

Van Deursen, A., Helsper, E. J., & Eynon, R. (2015). Development and validation of the Internet Skills Scale (ISS). *Information, Communication & Society*, Advance online publication. doi:10.1080/1369118X.2015.1078834

Van Deursen, A., & Van Dijk, J. (2011). Internet skills and the digital divide. *New Media & Society*, *13*(6), 893–911. doi: 10.1177/1461444810386774

Van Deursen, A., & Van Dijk, J. (2014). The digital divide shifts to differences in usage. *New Media & Society*, *16*(3), 507–526. doi:10.1177/1461444813487959

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford, UK: Oxford University Press.

Wei, L. (2012). Number matters: the multimodality of Internet use as an indicator of the digital inequalities. *Journal of Computer-Mediated Communication*, *17*(3), 303–318. doi:10.1111/j.1083-6101.2012.01578.x

Witte, J. C., & Mannon, S. E. (2010). *The Internet and social inequalities*. New York, NY: Routledge.

World Economic Forum. (2012). *Big Data, big impact: New possibilities for international development*. Geneva, Switzerland: The World Economic Forum. Retrieved from http://www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development

Young, A. L., & Quan-Haase, A. (2013). Privacy protection strategies on Facebook. *Information, Communication & Society*, *16*(4), 479–500. doi:10.1080/1369118X.2013.777757

Zillien, N., & Hargittai, E. (2009). Digital distinction: Status-specific types of internet usage. *Social Science Quarterly*, *90*(2), 274–291. doi:10.1111/j.1540-6237.2009.00617.x

Zimmer, M. (2010). Privacy protection in the next digital decade: "Trading up" or a "race to the bottom"? In B. Szoka & A. Marcus (Eds.), *The next digital decade. Essays on the future of the internet* (pp. 477–482). Washington, DC: TechFreedom.