How Do Like and Dislike Buttons Affect Communication? A Privacy Calculus Approach to Understanding Self-Disclosure Online in a One-Week Field Experiment

Abstract

According to the privacy calculus, both privacy concerns and expected gratifications explain self-disclosure online. So far, little is known about whether the privacy calculus can be used to predict observations of actual authentic behavior, and whether the privacy calculus can be influenced by the design of online websites—for example, by implementing popularity cues such as like and dislike buttons. To answer this question, we ran a preregistered one-week field experiment, in which participants were randomly distributed to three different websites where they could discuss a current political topic. The final sample consisted of 590 participants. The results showed that privacy calculus variables predicted a considerable share of actual self-disclosure. The impact of implementing popularity cues was negligible. In conclusion, the results demonstrate that self-disclosure online can be explained by privacy concerns and psychological gratifications. This finding has several implications. For example, it provides further evidence against the privacy paradox.

Keywords: privacy calculus, self-disclosure, popularity cues, structural equation modeling, preregistration

Word count: 5988

How Do Like and Dislike Buttons Affect Communication? A Privacy Calculus Approach to Understanding Self-Disclosure Online in a One-Week Field Experiment

Understanding why people disclose personal information online remains a critical question for both society and academic research. Originally, self-disclosure online was thought to be mostly erratic—for example, it was assumed that self-disclosure cannot be predicted by assessing people's personal beliefs, concerns, or standpoints. Most prominently, the privacy paradox stated that people self-disclose vast amounts of personal information online *despite* having substantial concerns about their privacy (Barnes, 2006; Taddicken & Jers, 2011).

Somewhat surprisingly, despite its popularity in the media (Radio, 2018) the privacy paradox has garnered little empirical support. A recent meta-analysis revealed that the correlation between privacy concerns and self-disclosure on SNS is r = -.13 (Baruh, Secinti, & Cemalcilar, 2017), indicating that privacy concerns are indeed related to self-disclosure online.

Rather than further pursuing the privacy paradox, a large share of current day research posits that self-disclosure online can be explained—at least partly—by means of the so-called *privacy-calculus* (Krasnova, Spiekermann, Koroleva, & Hildebrand, 2010). The privacy calculus builds on the work of Laufer and Wolfe (1977) and claims that both expected risks *and* expected benefits explain self-disclosure. Specifically, by operationalizing expected risks as privacy concerns, several studies have shown that experiencing greater privacy concerns is related to disclosing less information [Heirman, Walrave, and Ponnet (2013); koohikamaliInvestigationDynamicModel2019].

However, although the privacy calculus has gained some momentum several important questions remain unanswered. First, we still know comparatively little about whether the privacy calculus can be replicated with actual behavioral data in an authentic long-term setting (Kokolakis, 2017). Thus far, most research supporting the privacy calculus has used either self-reports of behavior (e.g., Krasnova et al., 2010), vignette approaches (e.g., Bol et al., 2018), or one-shot experiments in the lab (e.g., Trepte, Scharkow, & Dienlin, 2020). However, all three of these approaches significantly hamper external validity.

Second, current research on the privacy calculus is often criticized for not explicitly focusing on the deliberation process of self-disclosure. According to critics (e.g., Knijnenburg et al., 2017), showing that concerns and gratifications both correlate with self-disclosure is not evidence for any substantial or explicit weighing of pros and cons.

We agree and consider it necessary to now explicitly focus on the privacy deliberation process itself. Moreover, and on a more general level, we aim to gauge the usefulness of further extending the privacy calculus model by adding new variables such as privacy deliberation, trust, and self-disclosure self-efficacy.

Finally, we want to determine whether the privacy calculus can be affected by the design of a website. Specifically, we analyze whether *popularity cues* such as like and dislike buttons affect self-disclosure and the privacy calculus.

To test our research questions, we conducted a preregistered online field experiment, drawing from a representative sample of the German population. Participants were randomly distributed to one of three different websites, which either included only a like button, both a like and a dislike button, or no buttons at all. Over the course of one week participants had the chance to discuss a topical issue (i.e., prevention of terrorist attacks in Germany). Afterward, they answered our follow-up questionnaire with items pertaining to the privacy calculus variables.

The Privacy Calculus

Being a primary means of regulating privacy (e.g., Masur, 2018), self-disclosure is our key variable of interest. There are two different understandings of self-disclosure in the literature: The first defines self-disclosure as *deliberate* acts of sharing truthful information about the self with others (Jourard, 1964). The second considers *all* acts of sharing information—whether active or passive, deliberate or unwitting—as self-disclosure, because each piece of information shared allows meaningful inferences to be made about a person (e.g., Watzlawick, Bavelas, Jackson, & O'Hanlon, 2011). In this paper we follow the latter approach, not least because recent years have vividly illustrated how it is possible to derive a plethora of insights about a person simply by analyzing his or her written communication (e.g., Kosinski, Stillwell, & Graepel, 2013). Moreover, independent from which position one chooses to adopt, it is possible to differentiate the content of self-disclosure into three different dimensions: breadth (i.e., number of topics covered), depth (i.e., intimacy of topics covered), and length (i.e., quantity of disclosure) (e.g., Omarzu, 2000). In this study we mainly focus on communication quantity, as we consider communication quantity to be a necessary precondition and hence valid proxy for self-disclosure.

Privacy concerns have been defined as follows: "Concerns about online privacy represent how much an individual is motivated to focus on his or her control over a voluntary withdrawal from other people or societal institutions on the Internet, accompanied by an uneasy feeling that his or her privacy might be threatened" [AUTHOR]. Previous research has found that people who are more concerned about their privacy than others are less inclined to share personal information (Baruh et al., 2017; Dienlin & Trepte, 2015; Heirman et al., 2013; Koohikamali, French, & Kim, 2019).

H1: People are more likely to self-disclose on a website when they are less concerned about their privacy.

Although privacy concerns are related to self-disclosure, one can make the case that since most studies in the literature report only small effects, there should also be additional meaningful factors that contribute to explaining self-disclosure. Most prominently, it has been argued that people trade a loss of privacy for a gain in gratifications such as social capital, entertainment, information, or self-presentation (Ellison, Vitak, Steinfield, Gray, & Lampe, 2011; Taddicken & Jers, 2011). By now, a large body of research has found support for this hypothesis (e.g., Krasnova et al., 2010; Min & Kim, 2015; Trepte et al., 2017).

H2: People are more likely to self-disclose on a website when they obtain more gratifications from using the website.

In the current literature on the privacy calculus there still seems to be a shortage of studies that explicitly analyze the decision process of actively comparing the pros and cons of disclosing information, although this point of criticism has been leveled several times (e.g., Knijnenburg et al., 2017) and although other fields such as behavioral economics have long focused on the underlying problem (e.g., Zhu, Ou, van den Heuvel, & Liu, 2017). This criticism is justified. The observation that both experiencing privacy concerns and expecting gratifications are related to self-disclosure does not bit itself necessitate an explicit weighing process Hence, we argue that the research on the privacy calculus would benefit significantly from analyzing this decision process explicitly. Building on Omarzu (2000) and Altman (1976), we hence address a novel concept that might best be termed *privacy deliberation*, which we define as the extent to which individual people explicitly compare positive and negative potential outcomes before communicating with others.

On the one hand, it seems plausible that deliberating about one's privacy would dampen subsequent self-disclosure, because refraining from regular communication—the primary means of connecting with others—requires at least a minimum of active and hence deliberate restraint. On the other hand, deliberating about one's privacy might also increase self-disclosure, as after having actively deliberated about the potential consequences, a person concerned about his or her privacy might arrive at the conclusion that in this situation self-disclosure is not only appropriate but expedient. In light of the paucity of empirical studies and the plausibility of both effects, we formulate the following research question:

RQ1: Are people more or less likely to self-disclose on a website when they more actively deliberate about whether they should self-disclose?

Several attempts have already been made to expand the privacy calculus (e.g., Dinev & Hart, 2006). Additional variables such as self-efficacy or trust have been introduced. Self-efficacy in the context of the privacy calculus captures whether people believe in their own capability to implement particular privacy behaviors in the future (Dienlin & Metzger, 2016). These privacy behaviors can either refer to self-withdrawal (e.g., deleting inappropriate content) or self-disclosure (e.g., publishing a blog post). Thus far, several studies have found that people who report more privacy self-efficacy also self-withdraw more online than others (e.g., Chen, 2018). In light of our focus on self-disclosure, in this study we investigate the influence of self-disclosure self-efficacy.

Trust can be conceptualized in two different ways (Gefen, Karahanna, & Straub, 2003): It either captures "*specific* beliefs dealing primarily with the integrity, benevolence, and ability of another party" (Gefen et al., 2003, p. 55, emphasis added) or a "*general* belief that another party can be trusted" (Gefen et al., 2003, p. 55, emphasis added). Whereas specific trust beliefs focus on the causes of trust, general trust beliefs focus on the experience of trust. Gefen et al. (2003) prioritize specific trust beliefs (p. 60). In the online context, it is important to differentiate among several targets of trust (Söllner, Hoffmann, & Leimeister, 2016). Potential targets include (a) the information system, (b) the provider, (c) the Internet, and (d) the community of other users (Söllner et al., 2016). Trust plays a key role in online communication (Metzger, 2004). For example, it has been demonstrated that people who put more trust in the providers of networks also disclose more personal information (Li, 2011).

In conclusion, while we expect to find these relations as well, we would also like to determine whether the inclusion of all the other variables mentioned above, including the not yet researched concept of privacy deliberation, might potentially attenuate or even obviate the predictive capacity of self-efficacy and trust.

H3: People are more likely to self-disclose on a website when their self-efficacy about self-disclosing on the website is higher.

H4: People are more likely to self-disclose on a website when they have greater trust in the provider, the website, and the other users.

The Effect of Popularity Cues

What is the effect of the communication context on the privacy calculus and on self-disclosure? First, it has often been noted that researchers should not exclusively focus on specific features of particular websites, for features are prone to change and quickly become obsolete (Fox & McEwan, 2017). Instead, it has been suggested that researchers prioritize underlying latent structures, for example by analyzing what are know as affordances (e.g., Ellison & Vitak, 2015; Fox & McEwan, 2017). The concept of affordances was developed by Gibson (2015), who argued that it is not the objective features of objects that determine behavior but rather subjective perceptions. Affordances are a mental representation of how a given entity might be used; as such, they are by definition subjective. There is much debate in the literature concerning what exactly defines an affordance (Evans, Pearce, Vitak, & Treem, 2017). For example, whereas Evans et al. (2017) propose three affordances for mediated communication (i.e., anonymity, persistence, and visibility), Fox and McEwan (2017) suggest 10 affordances for SNSs alone (i.e., accessibility, bandwidth, social presence, privacy, network association, personalization, persistence, editability, conversation control, and anonymity).

As the privacy calculus states that both benefits and costs determine behavior, we suggest that popularity cues such as like and dislike buttons, which are categorized as "paralinguistic digital affordances" (Carr, Hayes, & Sumner, 2018, p. 142), perfectly epitomize benefits and costs. The like button is positive; it expresses an endorsement, a compliment, a reward (Carr et al., 2018; Sumner, Ruge-Jones, & Alcorn, 2017). However, communication online is also often characterized by negative and critical debates (Ziegele, Weber, Quiring, & Breiner, 2017). As the dislike button is a major means of downgrading content it represents the cost and risk factor of the privacy calculus well. In fact, its stark negative effect might also explain why to date only a handful of major websites have implemented it (e.g., youtube, reddit or stackexchange).

Paralinguistic digital affordances and/or popularity cues have been shown to impact behavior (Krämer & Schäwel, 2020; Trepte et al., 2020). For example, a large-scale field experiment in which 101,281 comments were analyzed found that comments with dislikes were more likely to receive further dislikes (Muchnik, Aral, & Taylor, 2013). Stroud, Muddiman, and Scacco (2017) demonstrated that when users had a different opinion than the one that was communicated in a post, they were more likely to click on a button labelled *respect* compared to a button labelled *like*.

In this vein it seems plausible that popularity cues might also impact the privacy calculus [kramerMasteringChallengeBalancing2020]. First, on a primordial level, popularity cues serve as a means of reward and punishment, affecting behavior via instrumental conditioning (Skinner, 2014). Specifically, being complimented with a like should encourage future self-disclosure, while being punished with a dislike should inhibit future disclosure. Similarly, like buttons should be associated with being able to garner positive feedback, so implementing a like-button—similar to a compliment in the offline world—might leverage gratifications. Implementing a like or a dislike button might also bring people to more actively deliberate about whether or not it is actually worthwhile to disclose information. If both like and dislike buttons are present, privacy deliberation should increase even further. Finally, because people who are more concerned about their privacy are also more shy and risk averse (Dienlin, 2017), implementation of the dislike button should both stir privacy concerns and stifle self-disclosure. For a simplified overview of our theoretical model, see Figure 1.

H5. Compared to people who use a website without like or dislike buttons, people who use a website with like buttons (a) self-disclose more, (b) obtain more gratifications, (c) are less concerned about their privacy, and (d) deliberate more about whether they should communicate online.

H6. Compared to people who use a website without like or dislike buttons, people who use a website with like and dislike buttons (a) self-disclose more, (b) obtain more gratifications, and (c) deliberate more about whether they should communicate online.

H7. Compared to people who use a website with only like buttons, people who use a website with like and dislike buttons (a) are more concerned about their privacy, and (b) deliberate more about whether they should communicate online.

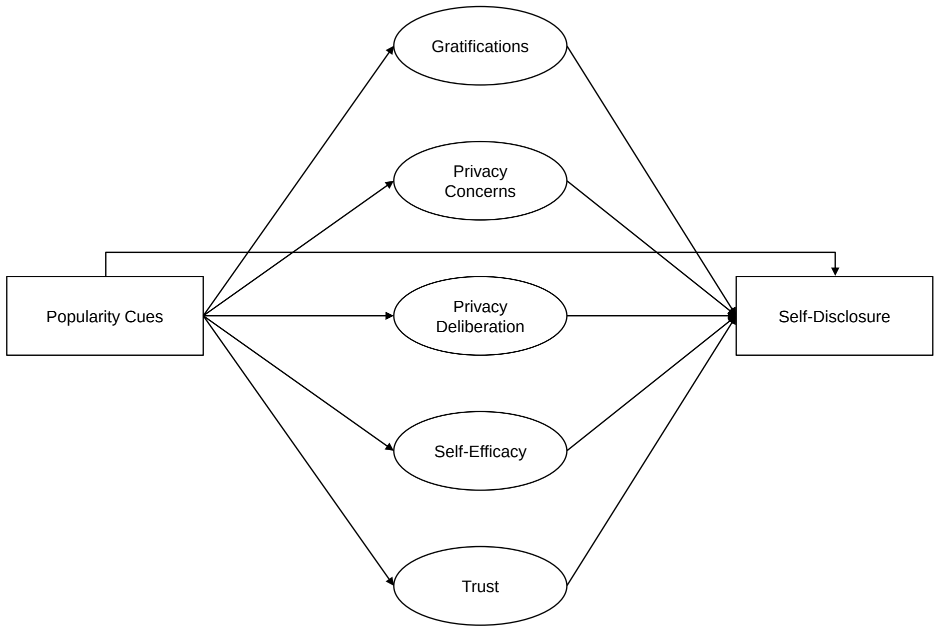

Figure 1. Overview of theoretical model.

Methods

Open Science

The online supplementary material (OSM) of this study include the data, research material, analyses scripts, and a reproducible version of this manuscript (see https://osf.io/hcqat/?view_only=5db35868738d40609b11e58cc343a9b0) We preregistered the study using the registration form *OSF Prereg*, which includes hypotheses, sample size, materials, analyses, and exclusion criteria (see https://osf.io/a6tzc/?view_only=5d0ef9fe5e1745878cd1b19273cdf859). We needed to change our pre-defined plan in some cases. For a full account of all changes, see OSM. New analyses that were not preregistered appear in the section on exploratory analyses. For example, we also measured two additional variables that were not included in the preregistration (e.g., *specific* gratifications and *general* trust; see below), which are included in the exploratory analyses.

Procedure

The study was designed as an online field experiment with three different groups. The first group interacted with a website without like/dislike buttons, the second with a website with only like buttons, and the third with a website with both like and dislike buttons. Participants were randomly distributed to one of the three websites in a between-subject design.

We collaborated with a professional panel agency to recruit participants. As incentive, participants were awarded digital points, which they could use to get special offers from other companies. Participants were above the age of 18 and lived in Germany. In a first step, the agency sent their panel members an invitation to participate in the study (*invitation*). In this invitation, panel members were asked to participate in a study analyzing the current threat posed by terrorist attacks in Germany.[1] Members who decided to take part were subsequently sent the first questionnaire (*T1*), in which we asked about their sociodemographics, provided more details about the study, and included a registration link for the website. Afterward, participants were randomly assigned to one of the three websites. After registration participants had the chance to discuss the topic of the terrorism threat in Germany over the course of one week (*field*). Subsequently, participants received a follow-up questionnaire in which we collected the self-reported measures (*T2*). Measures were collected after and not before the field phase in order not to

[1] Although the terror attack was not of primary interest for this study, the data can and will also be used to analyze perceptions of the terrorism threat. Hence, no deception took place, and in the debriefing participants were informed about our additional research interest in privacy.

prime participants or reveal our primary research interest.

We programmed an online website based on the open-source software discourse (https://www.discourse.org/). We conducted several pretests with students from the local university to make sure the website had an authentic feel (see Figure 2). Participants used the website actively: Overall, they spent 9,694 minutes online, wrote 1,171 comments, and left 560 popularity cues. For an example of communication that took place, see Figure 3.

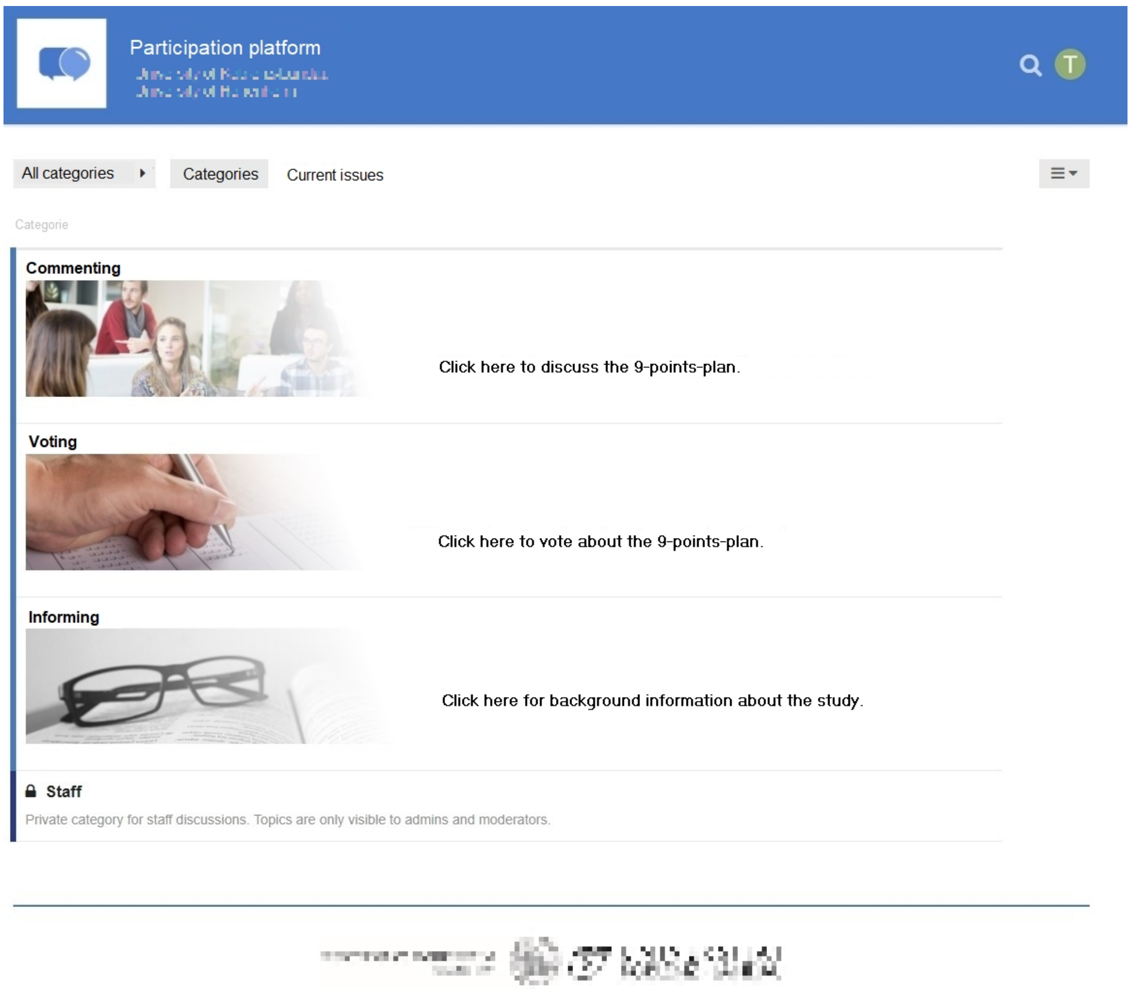

Figure 2. The website's homepage. (Translated to English; university logos pixelated for peer review.)

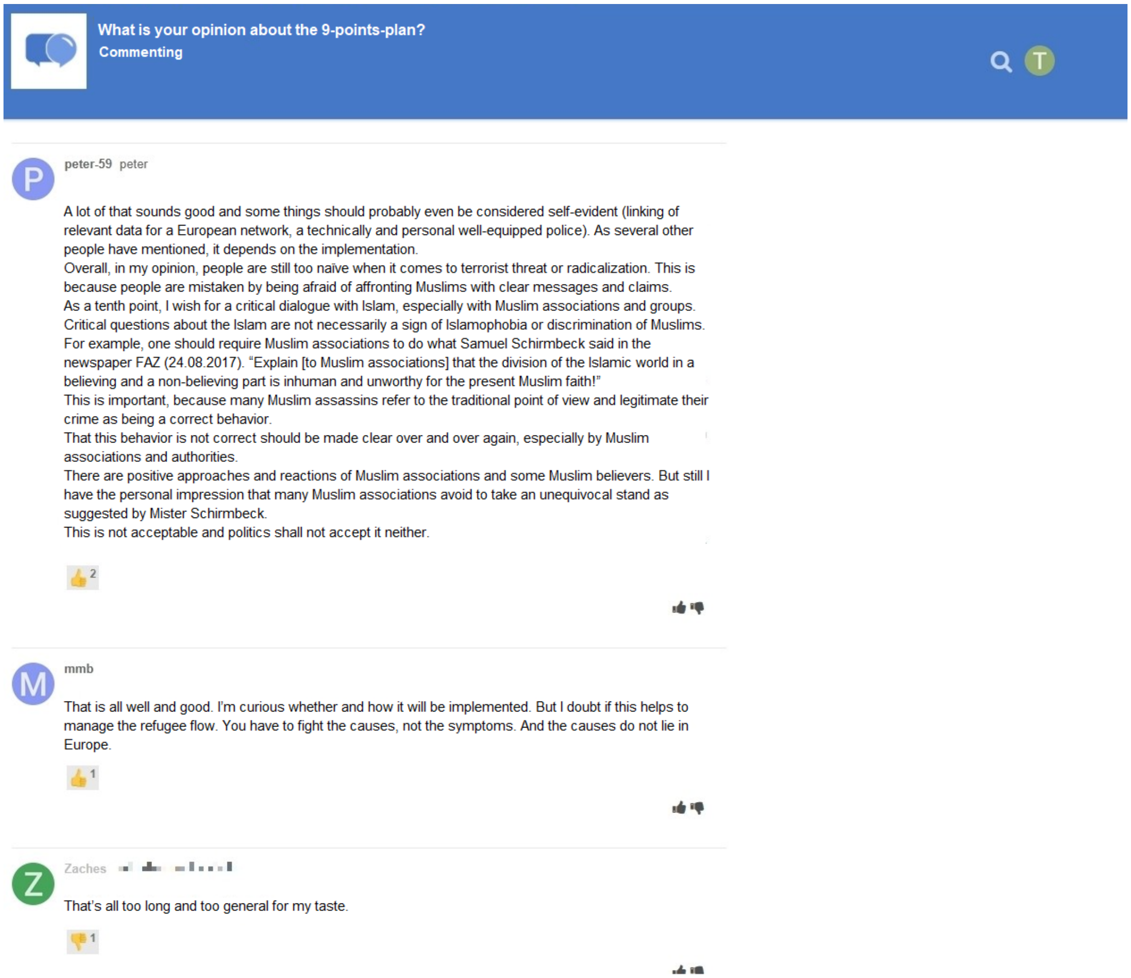

Figure 3. Communication that took place on the website with like and dislike buttons. (Translated to English.)

Participants

We ran a priori power analyses to determine how many participants to recruit. The power analysis was based on the smallest effect size of interest (SESOI; Lakens, Scheel, & Isager, 2018). Thus, we defined an effect size that we would consider enough to support our hypotheses. Because small effects should be expected when researching aspects of privacy online (e.g., Baruh et al., 2017), with small effects beginning at an effect size of r = .10 (Cohen, 1992), we set our SESOI to be r = .10. Our aim was to be able to detect this SESOI with a probability of at least 95%. Using the regular alpha level of 5%, this leads to a minimum sample size of n = 1,077. In the end, we were able to include n = 561 in our analyses (see below). This means that our study had a probability (power) of 77% of finding an effect at least as large as $r = .10$. Put differently, we were able to make reliable inferences about effects at least as big as $r = .14$.

We collected a representative sample of the German population in terms of age, sex, and federal state. 1,619 participants completed the survey at T1, 960 participants created a user account on the website, and 982 participants completed the survey at T2. Using tokens and IP addresses, we connected the data from T1, participants' behavior on the platform, and T2 by means of objective and automated processes. The data for $n = 590$ participants could be matched successfully. We excluded $n = 29$ participants who finished the questionnaire at T2 in less then three minutes, which we considered to be unreasonably fast. The final sample included 561 participants. The sample characteristics at T1 were as follows: Age = 45 years, sex = 49% male, college degree = 22%. The characteristics of the final sample were as follows: Age = 46 years, sex = 49% male, college degree = 0%. Hence, despite dropout, although a bit higher-educated, T2 can also be considered a largely representative sample of the German population.

Measures

In what follows, we present the materials we used to measure our variables. Wherever possible, we operationalized our variables using established measures. Where impossible (for example, to date there exists no scale on privacy deliberation), we self-designed novel items that were pretested in terms of legibility and/or understandability. To gauge the variables' factor validity, we ran confirmatory factor analyses (CFA). If the CFAs revealed insufficient fit, we deleted individual items. All items were formulated as statements to which participants indicated their (dis-)agreement on a bipolar 7-point scale. Answer options were as follows: -3 (*strongly disagree*), -2 (*disagree*), -1 (*slightly disagree*), 0 (*neutral*), +1 (*slightly agree*), +2 (*agree*), +3 (*strongly agree*). In the questionnaire, all

Table 1

Psychometric Properties, Factorial Validity, and Reliability of Measures

	m	sd	chisq	df	pvalue	cfi	tli	rmsea	srmr	omega	ave
Privacy concerns	3.21	1.52	11.04	9.00	0.27	1.00	1.00	0.02	0.01	0.96	0.80
General gratifications	4.76	1.23	34.44	5.00	0.00	0.98	0.95	0.10	0.02	0.94	0.75
Specific gratifications	4.71	1.03	270.68	85.00	0.00	0.94	0.93	0.06	0.05	0.93	0.59
Privacy deliberation	3.93	1.29	14.88	5.00	0.01	0.98	0.96	0.06	0.02	0.85	0.54
Self-efficacy	5.24	1.12	28.53	8.00	0.00	0.97	0.94	0.07	0.03	0.85	0.63
General trust	5.20	1.05	1.64	1.00	0.20	1.00	1.00	0.03	0.01	0.87	0.70
Specific trust	5.07	0.95	71.94	24.00	0.00	0.97	0.95	0.06	0.03	0.92	0.62

Note. omega = Raykov's composite reliability coefficient omega; avevar = average variance extracted.

items measuring a variable were presented on the same page in a randomized order.

For an overview of the means, standard deviations, factorial validity, and reliability, see Table 1. For an overview of the variables' distributions, see Figure 4. For the exact wording of all items and their individual distributions, see OSM.

Privacy concerns. Privacy concerns were measured with seven items based on Buchanan, Paine, Joinson, and Reips (2007). One example item was "When using the participation platform, I had concerns about my privacy". One item had to be deleted due to poor psychometric properties.

Gratifications. Next, we differentiated between two separate types of gratification. *General gratifications* were measured with five items based on Sun, Wang, Shen, and Zhang (2015). One example item was "Using the participation platform has paid off for me". *Specific gratifications* were measured with 15 items on five different subdimensions with three items each. The scaled was loosely based on Scherer and Schlütz (2002). Example items were: "Using the participation platform made it possible for me to" … "learn things

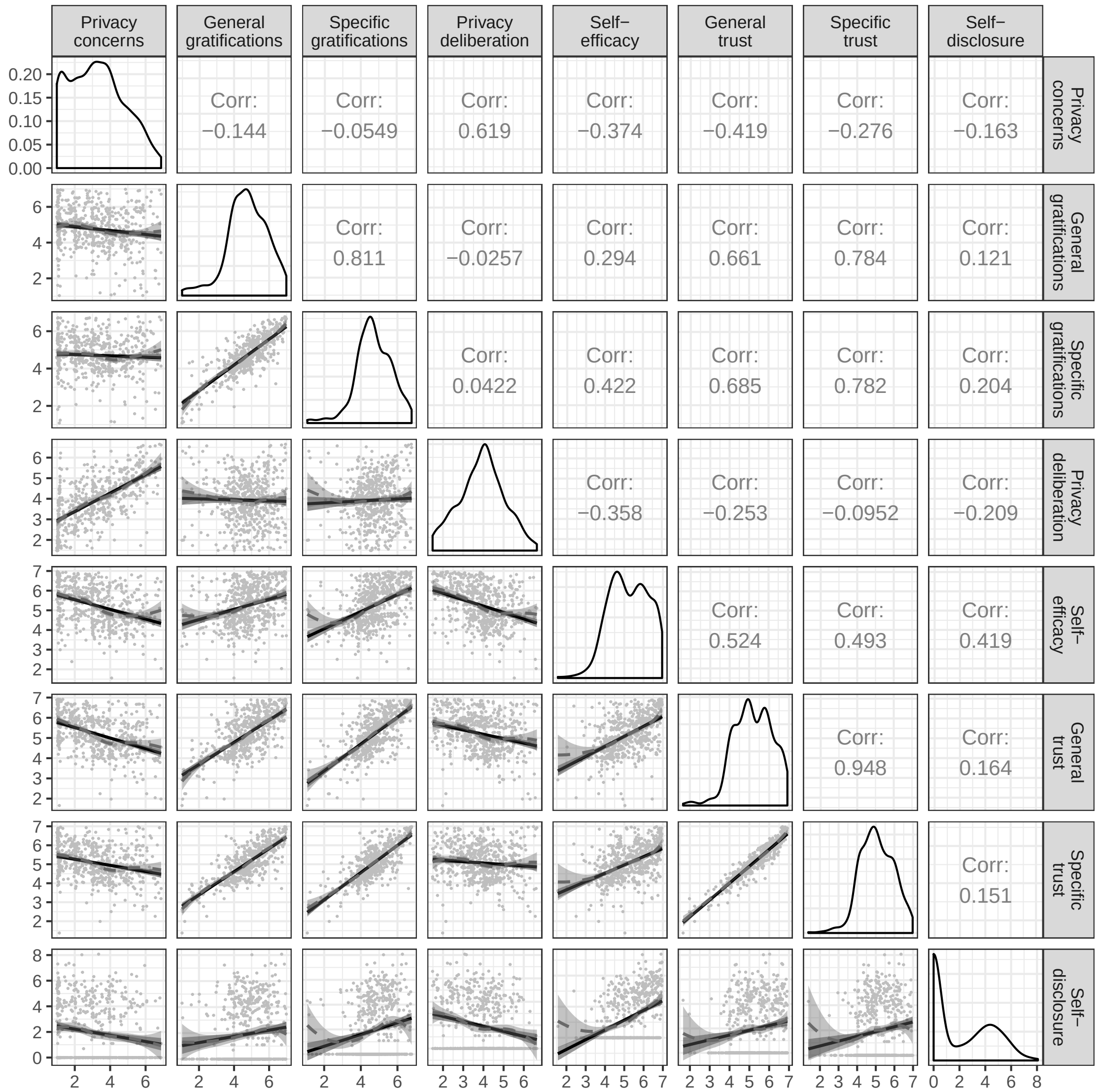

Figure 4. Above diagonal: zero-order correlation matrix; diagonal: density plots for each variable; below diagonal: bivariate scatter plots for zero-order correlations. Solid regression lines represent linear regressions, dotted regression lines represent quadratic regressions. Calculated with the model predicted values for each variable (baseline model).

I would not otherwise have noticed" (information), "react to a subject that is important to me" (relevance), "engage politically" (political participation), "try to improve society" (idealism), and "soothe my guilty consciences" (extrinsic benefits).

Self-disclosure self-efficacy. Self-disclosure self-efficacy was captured with six self-designed items, which captured whether participants felt that they had sufficient self-efficacy to write a comment on the platform. For example, we asked "I felt technically competent enough to write a comment." Two items, which were inverted, had to be deleted due to poor psychometric properties.

Trust. Next, we differentiated between two separate types of trust. *General trust* was operationalized based on Söllner et al. (2016) for three targets (i.e., provider, website, and other users), with one item each. One example items was "The operators of the participation platform seemed trustworthy." *Specific trust* was operationalized for the same three targets with three subdimensions each (i.e., ability, benevolence/integrity, and reliability), which were measured with one item each. Example items were "The operators of the participation platform have done a good job" (ability), "The other users had good intentions" (benevolence/integrity), "The website worked well" (reliability). The results showed that the provider and website targets were not sufficiently distinctive, as was evidenced by the existence of a Heywood case. We hence adapted the scale to combine these two targets. The updated scale exhibited adequate fit.

Privacy deliberation. Privacy deliberation was measured with 5 self-designed items. One example item was "While using the participation platform I have weighed the advantages and disadvantages of writing a comment."

Self-disclosure. Self-disclosure was calculated by taking the log scale of the number of words each participant wrote in a comment plus the number of likes and dislikes, with likes and dislikes being multiplied by two. Like and dislike buttons were multiplied by two because, rudimentarily, like buttons abbreviate the sentence "I like" and dislike buttons the sentence "I dislike". The sum of words and likes/likes was log-scaled because the relative amount of self-disclosure diminishes the more a person has already said.

Data analysis

We tested all hypotheses and research questions using structural equation modeling (SEM). We tested the influence of the three websites using contrast coding, which allows for testing the effects of experimental manipulations within a theoretical framework using latent variables (e.g., Kline, 2016). As the dependent variable (self-disclosure) was not normally distributed, we estimated the model using robust maximum likelihood (Kline, 2016). As recommended by Kline (2016), we report the following global fit indices: χ^2, RMSEA (90% CI), CFI, and SRMR. As sociodemagraphic variables are often related to self-disclosure and other privacy-related variables (e.g., Dindia & Allen, 1992), we controlled all variables for the influence of sex and age. Preregistered hypotheses were tested with a one-sided significance level of 5%. Research questions and exploratory analyses were tested with a two-sided 5% significance level using family-wise Bonferroni-Holm correction. We used R (Version 3.6.1; R Core Team, 2018) and the R-packages *lavaan* (Version 0.6.5; Rosseel, 2012), *papaja* (Version 0.1.0.9942; Aust & Barth, 2018), *pwr* (Version 1.2.2; Champely, 2018), *quanteda* (Version 1.5.2; Benoit, 2018), *semTools* (Version 0.5.2; Jorgensen et al., 2018), and *tidyverse* (Version 1.3.0; Wickham, 2017) for all our analyses.

Results

Descriptive Analyses

First, we measured and plotted all bivariate relations between the study variables (see Figure 4). The results did not reveal any relationships to be particularly curvilinear. Furthermore, all variables making up the privacy calculus demonstrated the expected relationships with self-disclosure. For example, people who were more concerned about their privacy had written fewer posts (r = -.16). Worth noting is that specific gratifications and general trust predicted self-disclosure better than general gratifications and specific trust. The mean of privacy deliberation was m = 3.93. Altogether, 32% of participants reported having actively deliberated about their privacy.

It is important to note that the bivariate results showed three very large correlations: First, between specific trust and general gratifications (r = .78); second, between privacy concerns and privacy deliberation (r = .62); third, between specific gratifications and self-efficacy (r = .42). As all six variables were later analyzed within a single multiple regression, problems of multicollinearity might occur.

Privacy Calculus

Preregistered analyses. First, we ran a model as specified in the preregistration. The model fit our data comparatively well, $\chi^2(389) = 929.46$, $p < .001$, cfi = .94, rmsea = .05, 90% CI [.05, .05], srmr = .05. Regarding H1, we did not find that general gratifications predicted self-disclosure (β = -.05, b = -0.07, 95% CI [-0.22, 0.09], z = -0.84, p = .200). Regarding H2, neither did we find that privacy concerns predicted self-disclosure (β = .09, b = 0.17, 95% CI [-0.14, 0.49], z = 1.09, p = .138). The analyses for RQ1 similarly revealed that privacy deliberation was not correlated with self-disclosure (β = -.10, b = -0.16, 95% CI [-0.34, 0.02], z = -1.74, p = .083). With regard to H3, however, we found that experiencing self-efficacy substantially predicted self-disclosure (β = .38, b = 0.76, 95% CI [0.48, 1.04], z = 5.40, p < .001). Concerning H4, the results showed that trust was not associated with self-disclosure (β = -.14, b = -0.34, 95% CI [-0.84, 0.16], z = -1.33, p = .091).

However, these results should be treated with utmost caution. As mentioned above, we indeed detected problems suggesting multicollinearity. Most prominently, in this multiple regression trust had a negative relation with self-disclosure, even though when analyzed bivariately the relation was positive—which is a sign of multicollinearity (Kline, 2016). As a result, specific trust and general gratifications should not be analyzed within the same model, as the two concepts are empirically too close to one another.

Exploratory analyses. Thus, we slightly adapted our preregistered model on the basis of the insights described above. First, instead of specific trust and general gratifications we now included general trust and specific gratifications (which were also correlated slightly less strongly with one another). The adapted model fit our data comparatively well, $\chi^2(508) = 1517.33$, $p < .001$, cfi = .93, rmsea = .06, 90% CI [.06, .06], srmr = .06.

In the adapted privacy calculus model, we found two additional significant effects. For example, specific gratifications predicted self-disclosure online (β = .18, b = 0.50, 95% CI [0.10, 0.90], z = 2.47, p = .013). Furthermore, in this model people who engaged in more privacy deliberation disclosed less information (β = -.14, b = -0.21, 95% CI [-0.39, -0.02], z = -2.21, p = .027). However, note that the effect was only marginally not significant. Self-efficacy remained substantially correlated with self-disclosure (β = .32, b = 0.64, 95% CI [0.39, 0.90], z = 4.92, $p < .001$). We again found a negative correlation between trust and self-disclosure (β = -.18, b = -0.54, 95% CI [-0.95, -0.12], z = -2.54, p = .011), which implies multicollinearity, and suggests that also general trust and specific gratifications should not be analyzed within the same model either.

Given that we observed several instances of multicollinearity, we also fitted a simple privacy calculus model containing only privacy concerns and specific gratifications. The simple model fit our data well, $\chi^2(202) = 717.70$, $p < .001$, cfi = .95, rmsea = .07, 90% CI [.06, .07], srmr = .05. First, we found that people who experienced more privacy concerns than others also disclosed less information (β = -.15, b = -0.21, 95% CI [-0.32, -0.09], z = -3.46, $p < .001$). Second, people who reported more specific gratifications than others also self-disclosed more information (β = .21, b = 0.61, 95% CI [0.33, 0.88], z = 4.32, $p < .001$). Both effect sizes were above our predefined SESOI of r = .10, implying that the effects were sufficiently large to qualify as support our the hypotheses. All effects labelled as significant were below their individual Bonferroni-Holm corrected significance level. For a visual overview of all results, see Figure 5.

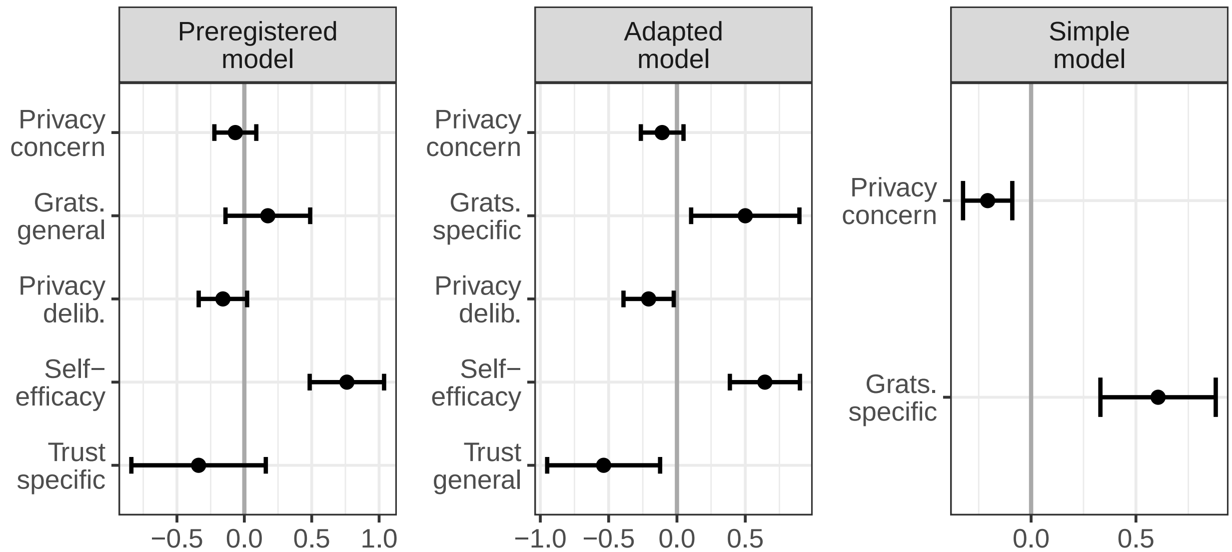

Figure 5. Predictors of self-disclosure. Black lines indicate 90% CIs (for one-sided hypotheses), grey lines 95% CIs (for two-sided hypotheses). Displayed are unstandardized effects.

When comparing the three models with one another, the simple privacy calculus model was the most parsimonious one (BIC = 37292, AIC = 36691), followed by the preregistered model (BIC = 49333, AIC = 48484) and the adapted model (BIC = 57694, AIC = 56729).

Popularity Cues

Preregistered analyses. Somewhat surprisingly, we found no effects of the popularity cues on the privacy calculus variables. For an illustration, see Figure 6, which displays the model-predicted values for each variable (using the baseline model) and shows that the confidence intervals of all preregistered variables overlap. For the results of the specific inference tests using contrasts, see the OSM.

Exploratory analyses. The picture remained mostly the same also when analyzing variables that we did not include in the preregistration. Note that some differences missed statistical significance only marginally (e.g., specific gratifications for the comparison between the website with like buttons and the control website without like and dislike buttons). Nevertheless, we refrain from reading too much into the differences between the three websites and conclude that they were mostly similar with regard to the privacy calculus variables and the amount of self-disclosure.

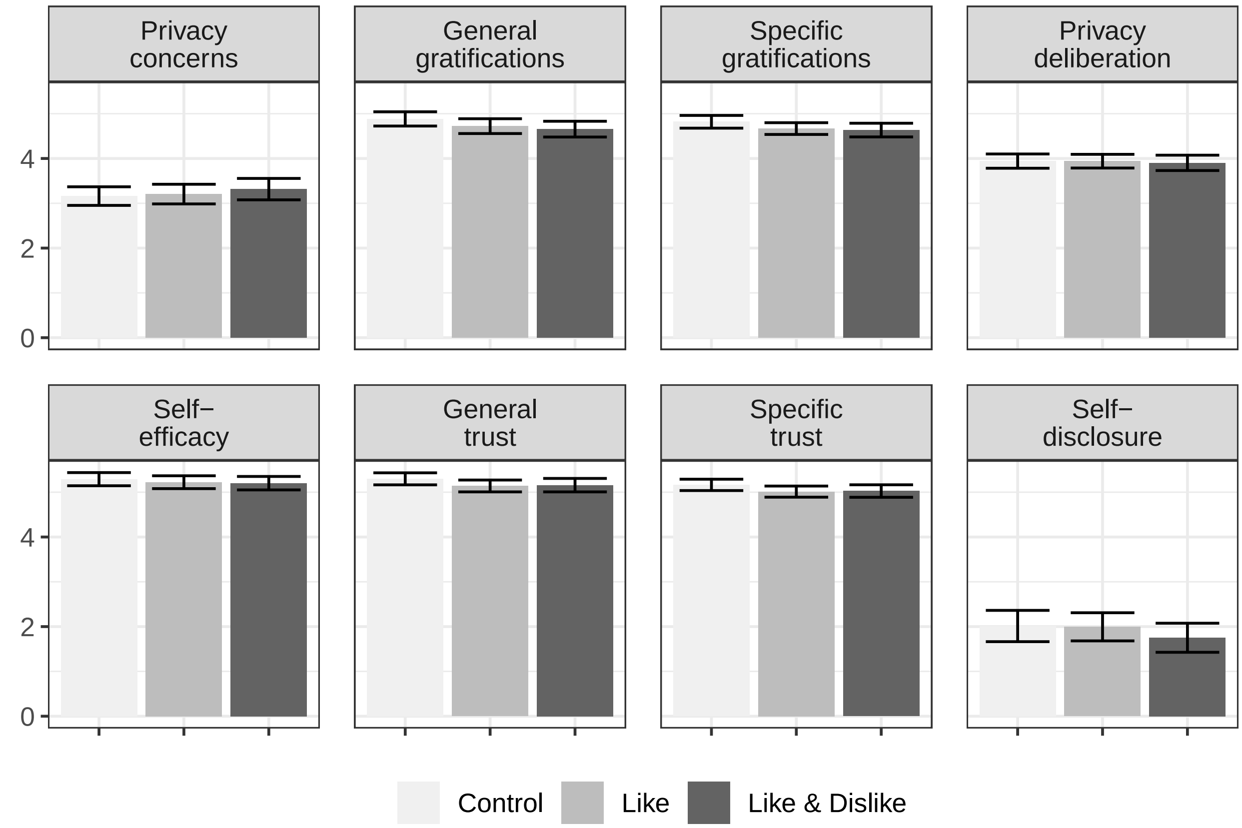

Figure 6. Overview of the variables for the three websites. Control: Website without buttons. Like: Website with like buttons. Like & Dislike: Website with like and dislike buttons.

Discussion

In this study, we analyzed theg e privacy calculus using actual observed behavior in a preregistered field experiment with data from a representative sample of the German population. We additionally sought to determine whether the privacy calculus is dependent on popularity cues such as like and dislike buttons. The data were analyzed using structural equation modeling. In the bivariate analyses, all privacy calculus variables were shown to significantly predict self-disclosure. In the preregistered analyses using multiple regression, in which several variables were analyzed together, self-efficacy turned out to be the strongest predictor of self-disclosure. However, this preregistered model exhibited significant problems with regard to multicollinearity, which is why we also computed a more basal privacy calculus model consisting of only privacy concerns and specific gratifications. In this model, both variables significantly and meaningfully predicted self-disclosure. Taken together, the results add further support to the privacy calculus framework, suggesting that self-disclosure online is not erratic (as for example posited by the privacy paradox, for example) but—at least somewhat—considered.

Our results suggest that in new communication contexts roughly one third of all Internet users actively deliberates about their privacy. Determining whether this figure is large or small is a normative question—for example, one can convincingly argue that this number should be higher and that we as society should still more actively deliberate about our self-disclosure practices online. Interestingly, results showed that privacy deliberation and privacy concerns were remarkably similar—evidenced by their strong correlation with one another and their similar correlations with other variables. This either implies that thinking about one's privacy increases one's concern or, conversely, that being concerned about one's privacy leads one to think about one's options more actively. Future research might tell.

The next major implication is that popularity cues do not seem to have an overly strong influence on the privacy calculus and self-disclosure. Although some studies have found that popularity cues substantially impact behavior (e.g., Muchnik et al., 2013), our results suggest the opposite: Users still disclose the same amount of personal information regardless of whether a website includes like or dislike buttons, potentially highlighting the agency of users.

The results also have several more fine-grained implications. First, we question the tendency to further increase the complexity of the privacy calculus model by adding additional variables (e.g., Dienlin & Metzger, 2016). "Since all models are wrong the scientist cannot obtain a"correct"it one by excessive elaboration. [. . .] Just as the ability to devise simple but evocative models is the signature of the great scientist so overelaboration and overparameterization is often the mark of mediocrity" (Box, 1976, p. 792). Although adding variables can increase the amount of explained variance, it might also introduce spurious results due to multicollinearity—a problem that we think deserves more attention in general. Specifically, we have come to believe that adding self-efficacy to privacy calculus models is of limited value, for self-efficacy is mostly a proxy of behavior and does not offer much epistemic insight. Instead, it might be more interesting to find out *why* some people feel sufficiently efficacious to self-disclose whereas others do not.

Furthermore, we found a remarkably strong correlation between specific trust and expected gratifications (i.e., $r = .79$), which at first glance seemed somewhat peculiar to us. On closer inspection, however, we realized that the way trust is routinely operationalized in the literature is very close to expected gratifications. To illustrate, the trust subdimension *ability* includes items such as "The comments of other users were useful". In fact, the literature often operationalizes trust as a formative construct that directly results from factors such as expected benefits (Söllner et al., 2016). In conclusion, our results suggest that we should not confuse *causes* of trust with *measures* of trust, for this might introduce problems of both homogeneity and/or multicollinearity. Instead, we recommend to measures general and reflective measures of trust, which are less closely related to expected gratifications.

Limitations

The results do not allow for causal interpretation on the within-person level. First, all results are based on analyses of between-person variance. However, between-person relations often do not translate well to within-person effects (e.g. Hamaker, Kuiper, &

Grasman, 2015). While some studies on privacy concerns online have begun to examine both sources of variance (e.g., Dietvorst, Hiemstra, Hillegers, & Keijsers, 2017), finding that intrapersonal changes in privacy concerns are indeed related to intrapersonal changes in self-disclosure, similar analyses are still lacking for the privacy calculus. Second, the self-reported measures were collected *after* the field phase in which the dependent variable was measured. As a result, the coefficients might overestimate the actual relations, because demand effects might have led participants to artificially align their theoretical answers with their practical behavior to reduce dissonance. Nevertheless, we deliberately decided to measure the self-reported variables afterward in order to not bias participants and not prime our specific research interest. Third, experiments should manipulate only the experimental variable while holding all others constant. In this study, we explicitly manipulated the popularity cues. However, as the experiment was conducted in the field, several other variables could not be held constant; for example, the content of communication by other users, the unfolding communication dynamics, or the characteristics of other users. As a result, the assumption of stable unit treatment was violated (Kline, 2016).

It is important to note that our not having found significant effects of like and dislike buttons does not necessarily mean that like and dislike buttons do indeed have no effect on self-disclosure and the privacy calculus. As always, with null-findings one is confronted with the *Duhème-Quinn Problem* (Dienes, 2008), which—put somewhat crudely— states that null findings can either be due to the actual non-existence of effects or, instead, a poor operationalization of the research question. In this case, we were not able send participants notifications when their comments were liked/disliked, significantly decreasing the popularity cues' salience.

This paper analyzes self-disclosure in the context of political participation. Our focus was on understanding self-disclosure, which is why we deliberately excluded variables pertaining to political participation, such as informational self-efficacy (Loy, Masur, Schmitt, & Mothes, 2018). Moreover, operationalizing self-disclosure via communication quantity is, of course, only a proxy.[2] Notably, we did not find any instances of people providing meaningless text and, as mentioned above, in times of big data, every piece of communication allows for increasingly accurate inferences about one's personality.

Finally, there are several interesting research questions that one could address with the data. Most prominently, one could analyze the actual content of the posts to detect whether the three websites might have differed with regard to communication quality. In addition, one can make the case that privacy deliberation rather as a moderator—such that deliberating more actively about one's privacy strengthens the relation between privacy concerns or gratifications and self-disclosure. Upon publication, the data will be made publicly available and we invite researchers from all disciplines to investigate the aforementioned and other interesting research questions.

Conclusion

While some scholars discuss whether we should wish "Death to the privacy calculus?" (Knijnenburg et al., 2017, p. 1), in our opinion the privacy calculus is alive and kicking. This study adds to the growing confirmation of observation that people who are more concerned about their privacy than others disclose less information online, whereas people who receive more gratifications from using a website than others disclose more information online. The results of this study suggest that a substantial share of internet users, approximately 30%, consciously engage in a privacy calculus by actively deliberating about whether or not to disclose information. Popularity cues seem to play a minor role in this process, especially if no means are implemented to guarantee that users are notified about others liking or disliking their communication. In conclusion, our results indicate that internet users are at least somewhat proactive and reasonable—probably no more or less proactive or reasonable than in any other regular everyday situation.

[2] Somewhat fittingly, in the German language there is a saying: "He spoke a lot but didn't say a thing".

References

Altman, I. (1976). Privacy: A conceptual analysis. *Environment and Behavior*, *8*(1), 7–29. https://doi.org/10.1177/001391657600800102

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, *11*(9).

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication*, *67*(1), 26–53. https://doi.org/10.1111/jcom.12276

Benoit, K. (2018). *Quanteda: Quantitative analysis of textual data.* https://doi.org/10.5281/zenodo.1004683

Bol, N., Dienlin, T., Kruikemeier, S., Sax, M., Boerman, S. C., Strycharz, J., … Vreese, C. H. (2018). Understanding the effects of personalization as a privacy calculus: Analyzing self-disclosure across health, news, and commerce contexts. *Journal of Computer-Mediated Communication*, *23*(6), 370–388. https://doi.org/10.1093/jcmc/zmy020

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799. https://doi.org/10.1080/01621459.1976.10480949

Buchanan, T., Paine, C., Joinson, A. N., & Reips, U.-D. (2007). Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American Society for Information Science and Technology*, *58*(2), 157–165. https://doi.org/10.1002/asi.20459

Carr, C. T., Hayes, R. A., & Sumner, E. M. (2018). Predicting a threshold of perceived Facebook post success via likes and reactions: A test of explanatory mechanisms. *Communication Research Reports*, *35*(2), 141–151. https://doi.org/10.1080/08824096.2017.1409618

Champely, S. (2018). *Pwr: Basic functions for power analysis*. Retrieved from https://CRAN.R-project.org/package=pwr

Chen, H.-T. (2018). Revisiting the privacy paradox on social media with an extended privacy calculus model: The effect of privacy concerns, privacy self-efficacy, and social capital on privacy management. *American Behavioral Scientist*, *62*(10), 1392–1412. https://doi.org/10.1177/0002764218792691

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Dienes, Z. (2008). *Understanding psychology as a science: An introduction to scientific and statistical inference*. New York, N.Y.: Palgrave Macmillan.

Dienlin, T. (2017). *The psychology of privacy: Analyzing processes of media use and interpersonal communication*. Hohenheim, Germany: University of Hohenheim.

Dienlin, T., & Metzger, M. J. (2016). An extended privacy calculus model for SNSsAnalyzing self-disclosure and self-withdrawal in a representative U.S. Sample. *Journal of Computer-Mediated Communication*, *21*(5), 368–383. https://doi.org/10.1111/jcc4.12163

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*(3), 285–297. https://doi.org/10.1002/ejsp.2049

Dietvorst, E., Hiemstra, M., Hillegers, M. H. J., & Keijsers, L. (2017). Adolescent perceptions of parental privacy invasion and adolescent secrecy: An illustration of Simpson's paradox. *Child Development*. https://doi.org/10.1111/cdev.13002

Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin*, *112*(1), 106–124.

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, *17*(1), 61–80. https://doi.org/10.1287/isre.1060.0080

Ellison, N. B., & Vitak, J. (2015). Social network site affordances and their relationship to social capital processes. In S. S. Sundar (Ed.), *The handbook of the psychology of communication technology* (Vol. v.33, pp. 205–227). Chichester, MA: Wiley Blackwell.

Ellison, N. B., Vitak, J., Steinfield, C., Gray, R., & Lampe, C. (2011). Negotiating privacy concerns and social capital needs in a social media environment. In S. Trepte & L. Reinecke (Eds.), *Privacy online: Perspectives on privacy and self-disclosure in the social web* (pp. 19–32). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-21521-6_3

Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication*, *22*(1), 35–52. https://doi.org/10.1111/jcc4.12180

Fox, J., & McEwan, B. (2017). Distinguishing technologies for social interaction: The perceived social affordances of communication channels scale. *Communication Monographs*, *9*, 1–21. https://doi.org/10.1080/03637751.2017.1332418

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Q*, *27*(1), 5190.

Gibson, J. J. (2015). *The ecological approach to visual perception*. New York, NY: Psychology Press.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Heirman, W., Walrave, M., & Ponnet, K. (2013). Predicting adolescents' disclosure of personal information in exchange for commercial incentives: An application of an extended theory of planned behavior. *Cyberpsychology, Behavior, and Social Networking*, *16*(2), 81–87. https://doi.org/10.1089/cyber.2012.0041

Jorgensen, D., T., Pornprasertmanit, S., Schoemann, M., A., … Y. (2018). *semTools: Useful tools for structural equation modeling.* Retrieved from https://CRAN.R-project.org/package=semTools

Jourard, S. M. (1964). *The transparent self.* New York, NY: Van Nostrand.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: The Guilford Press.

Knijnenburg, B., Raybourn, E., Cherry, D., Wilkinson, D., Sivakumar, S., & Sloan, H. (2017). Death to the privacy calculus? *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.2923806

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, *64*, 122–134. https://doi.org/10.1016/j.cose.2015.07.002

Koohikamali, M., French, A. M., & Kim, D. J. (2019). An investigation of a dynamic model of privacy trade-off in use of mobile social network applications: A longitudinal perspective. *Decision Support Systems*, *119*, 46–59. https://doi.org/10.1016/j.dss.2019.02.007

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal of Information Technology*, *25*(2), 109–125. https://doi.org/10.1057/jit.2010.6

Krämer, N. C., & Schäwel, J. (2020). Mastering the challenge of balancing self-disclosure and privacy in social media. *Current Opinion in Psychology*, *31*, 67–71. https://doi.org/10.1016/j.copsyc.2019.08.003

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues*, *33*(3), 22–42. https://doi.org/10.1111/j.1540-4560.1977.tb01880.x

Li, Y. (2011). Empirical studies on online information privacy concerns: Literature review and an integrative framework. *Communications of the Association for Information Systems*, *28*, 453–496.

Loy, L. S., Masur, P. K., Schmitt, J. B., & Mothes, C. (2018). Psychological predictors of political Internet use and political knowledge in light of the perceived complexity of political issues. *Information, Communication & Society*, *45*, 1–18. https://doi.org/10.1080/1369118X.2018.1450886

Masur, P. K. (2018). *Situational privacy and self-disclosure: Communication processes in online environments*. Cham, Switzerland: Springer.

Metzger, M. J. (2004). Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *Journal of Computer-Mediated Communication*, *9*(4). https://doi.org/10.1111/j.1083-6101.2004.tb00292.x

Min, J., & Kim, B. (2015). How are people enticed to disclose personal information despite privacy concerns in social network sites? The calculus between benefit and cost. *Journal of the Association for Information Science and Technology*, *66*(4), 839–857. https://doi.org/10.1002/asi.23206

Muchnik, L., Aral, S., & Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science (New York, N.Y.)*, *341*(6146), 647–651. https://doi.org/10.1126/science.1240466

Omarzu, J. (2000). A disclosure decision model: Determining how and when individuals will self-disclose. *Personality and Social Psychology Review*, *4*(2), 174–185. https://doi.org/10.1207/S15327957PSPR0402_5

Radio, N. Y. P. (2018). The privacy paradox. InternetDocument, https://project.wnyc.org/privacy-paradox/.

R Core Team. (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Scherer, H., & Schlütz, D. (2002). Gratifikation à la minute: Die zeitnahe Erfassung von Gratifikationen. In P. Rössler (Ed.), *Empirische Perspektiven der Rezeptionsforschung* (pp. 133–151). Munich, Germany: Reinhard Fischer.

Skinner, B. F. (2014). *Science and human behavior.* Upper Saddle River, NJ: Pearson Education.

Söllner, M., Hoffmann, A., & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. *European Journal of Information Systems*, *25*(3), 274–287. https://doi.org/10.1057/ejis.2015.17

Stroud, N. J., Muddiman, A., & Scacco, J. M. (2017). Like, recommend, or respect?: Altering political behavior in news comment sections. *New Media & Society*, *19*(11), 1727–1743. https://doi.org/10.1177/1461444816642420

Sumner, E. M., Ruge-Jones, L., & Alcorn, D. (2017). A functional approach to the Facebook Like button: An exploration of meaning, interpersonal functionality, and potential alternative response buttons. *New Media & Society*, *20*(4), 1451–1469. https://doi.org/10.1177/1461444817697917

Sun, Y., Wang, N., Shen, X.-L., & Zhang, J. X. (2015). Location information disclosure in location-based social network services: Privacy calculus, benefit structure, and gender differences. *Computers in Human Behavior*, *52*, 278–292. https://doi.org/10.1016/j.chb.2015.06.006

Taddicken, M., & Jers, C. (2011). The uses of privacy online: Trading a loss of privacy for social web gratifications? In S. Trepte & L. Reinecke (Eds.), *Privacy online: Perspectives on privacy and self-disclosure in the social web* (pp. 143–158). Berlin, Germany: Springer.

Trepte, S., Reinecke, L., Ellison, N. B., Quiring, O., Yao, M. Z., & Ziegele, M. (2017). A cross-cultural perspective on the privacy calculus. *Social Media + Society*, *3*(1). https://doi.org/10.1177/2056305116688035

Trepte, S., Scharkow, M., & Dienlin, T. (2020). The privacy calculus contextualized: The influence of affordances. *Computers in Human Behavior*, *104*, 106115. https://doi.org/10.1016/j.chb.2019.08.022

Watzlawick, P., Bavelas, J. B., Jackson, D. D., & O'Hanlon, B. (2011). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. New York, NY: W.W. Norton & Co.

Wickham, H. (2017). *Tidyverse: Easily install and load the 'tidyverse'*. Retrieved from https://CRAN.R-project.org/package=tidyverse

Zhu, H., Ou, C. X. J., van den Heuvel, W. J. A. M., & Liu, H. (2017). Privacy calculus and its utility for personalization services in e-commerce: An analysis of consumer decision-making. *Information & Management*, *54*(4), 427–437. https://doi.org/10.1016/j.im.2016.10.001

Ziegele, M., Weber, M., Quiring, O., & Breiner, T. (2017). The dynamics of online news discussions: Effects of news articles and reader comments on users' involvement, willingness to participate, and the civility of their contributions. *Information, Communication & Society*, *7*, 1–17. https://doi.org/10.1080/1369118X.2017.1324505