Likes or Dislikes, Gratifications or Concerns? Analyzing the Effects of Popularity Cues on the Privacy Calculus in Online Communication

Abstract

According to the privacy calculus model, online communication is determined by two factors, privacy concerns and expected benefits. How is the privacy calculus affected by external affordances, for example popularity cues such as like and dislike buttons? And should the privacy calculus model be extended by integrating more socio-psychological measures such as trust, self-efficacy, and privacy deliberation? To this end, a preregistered one-week field experiment with 590 participants was conducted. Participants were randomly distributed to three different websites, on which they discussed a current political topic. The websites included either (a) like buttons, (b) both like and dislike buttons, or (c) no like or dislike buttons. Because of multicollinearity and non-normal distribution of the outcome, the preregistered analyses needed to be updated. The results revealed that implementing both like and dislike buttons decreased online communication by 44%. As predicted by the privacy calculus model, expected benefits and perceived privacy concerns substantially affected communcation. Actively deliberating about privacy by comparing benefits and drawbacks decreased communication, while experiencing self-efficacy and trust strongly increased communication.

Keywords: privacy calculus, communication, popularity cues, field experiment, preregistration

Word count: 4903

Likes or Dislikes, Gratifications or Concerns? Analyzing the Effects of Popularity Cues on the Privacy Calculus in Online Communication

Introduction

According to the privacy paradox, the way users share information online is erratic (Barnes, 2006): People communicate vast amounts of personal information online, although they are concerned about their privacy (Taddicken & Jers, 2011). However, despite its popularity in both academic research (Dienlin, 2019) and the public press (New York Public Radio, 2018), empirical support for the privacy paradox is weak (Baruh, Secinti, & Cemalcilar, 2017; Dienlin & Sun, 2021). Research is therefore increasingly building on the privacy calculus model, which states that communication online can be explained successfully by perceived risks and expected benefits (Bol et al., 2018; Krasnova, Spiekermann, Koroleva, & Hildebrand, 2010; Meier & Krämer, 2022).

Although the privacy calculus has gained momentum in academic research, several important questions remain unanswered. First, we still know little about how the privacy calculus and online communication can be affected by external affordances (Bazarova & Masur, 2020; Bräunlich et al., 2020). For example, do popularity cues such as like and dislike buttons affect the privacy calculus and online communication? Second, the privacy calculus' empirical foundation needs to be improved (Knijnenburg et al., 2017). To date, much research on the privacy calculus builds on (a) self-reports of behavior (Krasnova et al., 2010), (b) vignette approaches (Bol et al., 2018), or (c) one-shot experiments in the lab (Trepte, Scharkow, & Dienlin, 2020). Can the privacy calculus be replicated in more authentic and long-term settings using actual behavioral data? Third, research on the privacy calculus so far did not measure the actual process of weighing pros and cons (Knijnenburg et al., 2017), which I do in this study by elaborating on what I will call the the privacy deliberation process. Fourth, the privacy calculus model is criticized for over-emphasizing rationality (Knijnenburg et al., 2017). Hence, I extend the privacy calculus in analyzing how trust and self-efficacy, two more socio-psychological variables,

affect communication.

The Privacy Calculus

The privacy calculus analyzes why people communicate online. When are we willing to engage in a conversation? It builds on the calculus of behavior (Laufer & Wolfe, 1977) and states that people are weighing risks and benefits before communicating. Communication is closely related to privacy, which is defined as a "voluntary and temporary withdrawal of a person from the general society" (Westin, 1967, p. 7). People regulate their privacy by deciding what and what not to communicate (Dienlin, 2014). Disclosing information can be risky, and if people are more concerned about their privacy they are less likely to communicate online (Baruh et al., 2017; Dienlin & Metzger, 2016; Krasnova et al., 2010; Masur, 2023; Meier & Krämer, 2022). According to a meta-analysis by Baruh et al. (2017), the effect is small but significant (r = -.13).

Outweighing concerns, the most relevant drivers of online communication are expected gratifications (Bol et al., 2018; Dienlin & Metzger, 2016; Kezer, Dienlin, & Baruh, 2022). People are happy to trade in parts of their privacy when they can receive something more valuable in return (Laufer & Wolfe, 1977). In online communication, the most important benefits include social support (Krasnova et al., 2010), social capital (Ellison, Vitak, Steinfield, Gray, & Lampe, 2011), entertainment (Dhir & Tsai, 2017), information-seeking (Whiting & Williams, 2013), and self-presentation (Min & Kim, 2015).

H1: People who are more concerned about their privacy are less likely to communicate actively on a website.

H2: People who obtain more gratifications from using a website are more likely to communicate actively on a website.

Although privacy calculus implies that people explicitly compare benefits and drawbacks before communicating online, research has neglected this aspect (Knijnenburg et al., 2017). Only observing that privacy concerns or expected gratifications and

communication online are related does not prove that an explicit weighing process took place (Knijnenburg et al., 2017). Instead, we need to analyze if and by how much people actively deliberate about their privacy, and whether doing so influences their decision to communicate. Self-disclosure theory (Altman, 1976; Omarzu, 2000) suggests that if the benefits to be attained from communication are attractive, deciding whether or not to communicate is a "conscious and deliberative process" (Omarzu, 2000, p. 183). I hence introduce and investigate a novel concept termed *privacy deliberation*. Privacy deliberation captures the extent to which individuals explicitly compare risks and benefits before communicating with others.

How could deliberating about one's privacy affect communication? On the one hand, it could reduce subsequent communication. Refraining from communication—the primary means of connecting with others (Altman, 1976)—likely requires some active and deliberate restraint (Omarzu, 2000). This is especially true for social media, which are designed to foster communication and participation (Ellison & Vitak, 2015; Masur, DiFranzo, & Bazarova, 2021). Actively thinking about whether communicating is worthwhile might be the first step not to participate. On the other hand, deliberating about privacy might also increase communication. In light of the benefits mentioned above, would it not make sense to conclude that participation is beneficial, thereby fostering communication (Krasnova et al., 2010)?

RQ1: Do people who deliberate more actively about their privacy communicate more or less online?

This weighing process is not flawless or perfect. It is therefore useful to understand the privacy calculus from a perspective of *bounded rationality* (Simon, 1990). Bounded rationality states that "(1) humans are cognitively constrained; (2) these constraints impact decision making; and (3) difficult problems reveal the constraints and highlight their significance." (Bendor, 2015, p. 1303). Although human behavior is partly irrational, bounded rationality does not state that it is completely irrational (Gigerenzer, Selten, &

Workshop, 2002). Rationality is a continuum. In the context of online privacy, for example, rationality is impeded by information asymmetries, presence bias, intangibility, illusory control, or herding (Acquisti, Brandimarte, & Loewenstein, 2020). It follows that to provide a more complete picture, additional factors next to gratifications and concerns should also explain communication. Factors less focused on rationality, but more on socio-psychological aspects.

For example, if users are more familiar, experienced, and knowledgeable in a given online context, they are more likely to navigate that online contexts successfully and to communicate actively (Baruh et al., 2017; Krämer & Schäwel, 2020). People with more privacy self-efficacy also engage more successfully in self-withdrawal (Dienlin & Metzger, 2016). Hence, if users possess more *self-efficacy* to participate, they should also communicate more.

H3: People are more likely to communicate on a website when their self-efficacy to actively use the website is higher.

In situations where people lack experience or competence, a relevant variable to understand behavior is *trust*. Especially in online contexts, users often cannot control the environment or the way their information is handled (Acquisti et al., 2020; Bräunlich et al., 2020). Trust plays a key role in online communication (Metzger, 2004). Trust either captures "*specific* beliefs dealing primarily with the integrity, benevolence, and ability of another party" (Gefen, Karahanna, & Straub, 2003, p. 55, emphasis added) or a "*general* belief that another party can be trusted" (Gefen et al., 2003, p. 55, emphasis added). In online contexts, there are different targets of trust, including (a) the information system, (b) the provider, (c) the Internet, and (d) the community of other users (Söllner, Hoffmann, & Leimeister, 2016). People who put more trust in the providers of networks, for example, disclose more personal information (Li, 2011).

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

The Effects of Popularity Cues

How are the privacy calculus and communication affected by the context, the digital infrastructure? In analyzing this question, it makes sense not to focus on specific features of particular websites, which change and become obsolete (Fox & McEwan, 2017). Instead, it is recommended focusing on the underlying latent structures by analyzing so-called *affordances* (Ellison & Vitak, 2015; Fox & McEwan, 2017). Affordances are mental representations of how objects might be used. They emphasize that it is not the *objective features* that determine behavior, but rather our *subjective perceptions* (Gibson, 2015). There is an ongoing debate on how to best understand and operationalize affordances (Evans, Pearce, Vitak, & Treem, 2017). 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). Popularity cues such as like and dislike buttons, which are of interest here, are categorized as "paralinguistic digital affordances" (Carr, Hayes, & Sumner, 2018, p. 142).

Paralinguistic digital affordances and specifically popularity cues have been shown to affect behavior (Krämer & Schäwel, 2020; Masur et al., 2021; Trepte et al., 2020). Online comments that already have several dislikes are more likely to receive further dislikes (Muchnik, Aral, & Taylor, 2013). When users disagree with a post, they are more likely to click on a button labeled *respect* compared to a button labeled *like* (Stroud, Muddiman, & Scacco, 2017). Popularity cues likely impact the privacy calculus via two underlying theoretical mechanisms (Krämer & Schäwel, 2020): First, the *mere presence* of popularity cues might affect whether people are willing to disclose. Being able to attract likes might motivate users to communicate, while the mere option to receive dislikes might intimidate others. Second, *actually receiving* likes or dislikes might additionally affect subsequent behavior, thereby reinforcing the process (Skinner, 2014).

Likes are affirmative and represent the positivity bias of social media (Schreurs, Meier, & Vandenbosch, 2022). Receiving a like online is similar to receiving a compliment offline (Carr et al., 2018; Sumner, Ruge-Jones, & Alcorn, 2017). Like-buttons afford and emphasize a *gain frame* (Rosoff, Cui, & John, 2013). These gains can be garnered only through active participation. Because like buttons emphasize positive outcomes, it is likely that concerns decrease. In situations where there is more to win, people might also more actively deliberate about whether or not to communicate. Dislikes, instead, represent a punishment, introducing a *loss frame*. Websites featuring both like *and* dislike buttons should therefore be more ambivalent compared to websites without popularity cues. Privacy concerns should not be reduced anymore: People who are more concerned about their privacy are also more shy and risk averse (Dienlin, 2017). Implementing the dislike button might therefore increase privacy concerns, thereby canceling out the positive effects of the like button.

H5. People who use a website with like buttons (a) communicate more, (b) obtain more gratifications, (c) are less concerned about their privacy, and (d) deliberate more about whether they should communicate online.

H6. People who use a website with like *and* dislike buttons (a) communicate 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.

For a simplified overview of the analyzed model, see Figure 1.

Figure 1. Overview of analyzed model.

Methods

Open Science

This manuscript features a companion website, which includes the data, research materials, analyses scripts, and a reproducible version of this manuscript (see https://XMtRa.github.io/like_dislike). The hypotheses, sample size, research materials, analyses, and exclusion criteria were preregistered (see https://osf.io/a6tzc/?view_only=5d0ef9fe5e1745878cd1b19273cdf859). In some cases, the preregistered approach needed to be changed (see companion website). Analyses that were not preregistered are reported as exploratory analyses.

Procedure

The study was designed as an online field experiment with three different groups. The first experimental group used a website that included like buttons; the second

experimental group used an identical website including both like and dislike buttons; and the control group used an identical website without like and dislike buttons. Participants were randomly distributed to one of the three websites in a between-subject design.

The data were collected in Germany. Participants were recruited using the professional panel agency Norstat. As incentive, participants were awarded digital points, which they could use to get special offers from online retailers. Participants had to be above the age of 18 and had to live in Germany. In a first step, the company sent its 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. Members who agreed to participate were sent the first questionnaire (*T1*). The questionnaire asked participants about their sociodemographic background, (b) provided more details about the study, and (c) included a registration link for the website, which was introduced as "participation platform". Afterward, participants were randomly assigned to one of the three websites. After registration was completed, participants were invited (but not obliged) to visit the website and to discuss the topic of the terrorism threat in Germany (*field*). Participants could use the website and write comments over the course of one week. Subsequently, participants received a follow-up questionnaire in which the self-reported measures reported below were collected (*T2*).

The online website was programmed based on the open-source software *discourse* (https://www.discourse.org/). To make sure the website was professional and authentic, several pretests with students from the local university were run. Nine hundred sixty participants created a user account on the website (see below) and used the website actively. Overall, they spent 162 hours online, wrote 1,171 comments, and clicked on 560 popularity cues. All communication was checked, and there were no instances of people providing meaningless text or doubting the experiment. For a screenshot of the landing page and for examples of comments, see Figure 2.

Figure 2. Screenshot of the landing page and of communication taking place

Participants

Sample size was determined using a priori power analyses. The power analyses were based on a smallest effect size of interest approach (SESOI, Lakens, Scheel, & Isager, 2018). When researching aspects of privacy online (e.g., Baruh et al., 2017), effects are often small (i.e., $r = .10$, Cohen, 1992). Hence, the SESOI was set to be $r = .10$. The aim was to be able to detect this SESOI with a probability of at least 95% (i.e., power = 95%). Using the regular alpha level of 5%, power analyses suggested a minimum sample size of N = 1,077. In the end, I was able to include $N = 559$ in the analyses (see below), which was significantly smaller than the original aim. With this sample size, the study had a power of 77% to find an effect at least as large as $r = .10$. Sensitivity analyses showed that the current study could still make reliable results (i.e., with a power of 95%) for effects at least as large as $r = .14$. In conclusion, although not as powerful as planned, the study is still adequately powered to find the small effects reported in the privacy literature (Baruh et al., 2017).

A quota sample that matched the German population in terms of age, gender, and federal state was collected. In sum, 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. The data were connected using tokens and IP addresses. For technical reasons, the data of several participants could not be matched (for example, because they used different devices for the respective steps). In the end, the data of 590 participants could be matched successfully. Considered unreasonably fast, twenty-nine participants were excluded who finished the questionnaire at T2 in less than three minutes. To detect atypical data and response patterns, I calculated Cook's distance. I excluded two participants with clear response patterns (i.e., straight-lining). The final sample included N = 559 participants. The sample characteristics at T1 and T2 were as follows: T1: age = 45 years, gender = 49% male, college degree = 22%. T2: age = 46 years, gender = 49% male, college degree = 29%. One participant did not report their gender.

Measures

Factor validity was assessed using confirmatory factor analyses (CFA). If the CFAs revealed insufficient fit, malfunctioning items were deleted. All items were measured on bipolar 7-point scales. Answer options were visualized as follows: -3 (*strongly disagree*), -2 (*disagree*), -1 (*slightly disagree*), 0 (*neutral*), +1 (*slightly agree*), +2 (*agree*), +3 (*strongly agree*). For the analyses, answers were coded from 1 to 7. All items measuring the same variable were presented in randomized order on the same page.

All measures showed high factorial validity. For an overview of the means, standard deviations, factorial validity, and reliability, see Table 1. For the variables' distributions, see Figure 3. For all items and their distributions, see companion website.

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 was deleted due to poor psychometric properties.

Two types of gratifications were collected. *General gratifications* were measured with five items based on Sun, Wang, Shen, and Zhang (2015). One example item was "Using the

Table 1

Psychometric Properties, Factorial Validity, and Reliability of Measures

	m	sd	chisq	df	p-value	cfi	tli	rmsea	srmr	omega	ave
Privacy concerns	3.21	1.51	11.04	9.00	0.27	1.00	1.00	0.02	0.01	0.96	0.80
General gratifications	4.76	1.22	34.03	5.00	0.00	0.98	0.95	0.10	0.02	0.93	0.74
Specific gratifications	4.71	1.02	269.77	85.00	0.00	0.94	0.93	0.06	0.05	0.95	0.59
Privacy deliberation	3.93	1.29	15.55	5.00	0.01	0.98	0.96	0.06	0.02	0.85	0.53
Self-efficacy	5.21	1.04	3.23	1.00	0.07	0.99	0.96	0.06	0.01	0.83	0.59
General trust	5.08	0.94	2.07	1.00	0.15	1.00	0.99	0.04	0.01	0.87	0.70
Specific trust	5.25	1.12	99.48	26.00	0.00	0.96	0.94	0.07	0.04	0.93	0.62

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

participation platform has paid off for me". Based on Scherer and Schlütz (2002), *specific gratifications* were measured on five dimensions with three items each. Example items were: "Using the participation platform made it possible for me to" ... "learn things I would not have noticed otherwise" (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).

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

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

Two types of trust were measured. *General trust* was operationalized based on

Söllner et al. (2016), addressing three targets (i.e., provider, website, and other users), measured with one item each. One example item 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), 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). Because they were closely related, the two targets "provider" and "website" were combined into one dimension. The updated scale showed adequate fit.

Communication was calculated by counting the number of words each participant wrote in a comment. Communication was zero-inflated and heavily skewed. While 58 percent did not communicate at all, some communicated a lot. On average, participants wrote 77 words.

Data analysis

As preregistered, all hypotheses and research questions were initially tested using structural equation modeling with latent variables. The influence of the three websites was analyzed using contrast coding. Because the assumption of multivariate normality was violated, I estimated the models using robust maximum likelihood (Kline, 2016). As recommended by Kline (2016), to assess global fit I report the model's χ^2, RMSEA (90% CI), CFI, and SRMR. To exclude confounding influences, I controlled all variables for age, gender, and education, which were shown to affect both privacy concerns and online communication (Masur, 2023; Tifferet, 2019). The preregistered hypotheses were tested with a one-sided significance level of 5%; the research questions were tested with a two-sided 5% significance level using family-wise Bonferroni-Holm correction.

As became apparent when analyzing the data, the preregistered analyses had two major problems. First, communication was zero-inflated and gamma distributed. Although

it is possible to analyze non-normal data with structural equation modeling, it is recommended to use analyses that model the variables' distribution, which can be achieved with Bayesian hurdle models (McElreath, 2016). I conclusion, in the exploratory analyses I ran Bayesian hurdle regression models, modeling the outcome as a zero-inflated gamma distribution. Second, in the preregistered analyses several variables were combined that are theoretically and empirically closely related, leading to multicollinearity (Vanhove, 2021). As a remedy, I adopted a causal modeling perspective, controlling only for confounders—in this case, age, gender, and education—but not for mediators (Rohrer, 2018). To assess the effects, I tested whether or not the 95% credibility intervals of the average marginal effects excluded zero. If they excluded zero, effects can be considered "significant" (McElreath, 2016). I also plotted the distribution of the effects. For more information, see online companion website.

The data were analyzed using R (Version 4.2.2; R Core Team, 2018) and the R-packages *brms* (Version 2.18.0; Bürkner, 2017, 2018), *lavaan* (Version 0.6.12; Rosseel, 2012), *marginaleffects* (Version 0.11.0; Arel-Bundock, 2023), *papaja* (Version 0.1.1; Aust & Barth, 2018), *pwr* (Version 1.3.0; Champely, 2018), *quanteda* (Version 3.2.3; Benoit, 2018), *semTools* (Version 0.5.6; Jorgensen et al., 2018), and *tidyverse* (Version 1.3.2; Wickham, 2017).

Results

Descriptive Analyses

I first plotted the bivariate relations of all variables (see Figure 3). All variables referring to the privacy calculus demonstrated the expected bivariate relationships with communication. For example, people who were more concerned about their privacy disclosed less information (r = -.15). Specific gratifications predicted communication better than general gratifications (r = .22 vs. r = .11). The mean of privacy deliberation was m = 3.93. Altogether, 32% of participants reported having actively deliberated about

Figure 3. 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).

their privacy.

The bivariate results showed three large correlations: specific trust and general gratifications ($r = .79$), privacy concerns and privacy deliberation ($r = .61$), and specific gratifications and self-efficacy ($r = .55$). As all six variables were later analyzed within a

single multiple regression, problems of multicollinearity might occur.

Preregistered analyses

First, as preregistered I ran a structural equation model with multiple predictors. The model fit the data okay, $\chi^2(389) = 957.40$, $p < .001$, CFI = .94, RMSEA = .05, 90% CI [.05, .05], SRMR = .05. Regarding H1, results showed that gratifications did not predict communication (β = -.04, b = -0.05, 95% CI [-0.21, 0.10], z = -0.68, p = .247; one-sided). With regard to H2, privacy concerns did not significantly predict communication (β = .05, b = 0.10, 95% CI [-0.24, 0.44], z = 0.60, p = .276; one-sided). RQ1 similarly revealed that privacy deliberation did not predict communication (β = -.10, b = -0.15, 95% CI [-0.34, 0.03], z = -1.66, p = .097; two-sided). Regarding H3, however, I found that experiencing self-efficacy predicted communication substantially (β = .40, b = 0.82, 95% CI [0.52, 1.11], z = 5.40, $p < .001$; one-sided). Concerning H4, results showed that trust was not associated with communication (β = -.11, b = -0.28, 95% CI [-0.80, 0.24], z = -1.05, p = .148; one-sided).

However, these results should be treated with caution. I found several signs of multicollinearity, as evidenced by the large standard errors or "wrong" and reversed signs of predictors (Vanhove, 2021). For example, in the bivariate analysis trust had a *positive* relation with communication, whereas in the multiple regression the effect was *negative*—which should make us sceptical

Next, I analyzed the effects of the popularity cues. It was for example expected that websites with like buttons would lead to more communication, more gratifications, more privacy deliberation, and to less privacy concerns. The results showed that the popularity cues had no effects on communication and on the privacy calculus variables. For an illustration, see Figure 4. For the detailed results of the specific inference tests using contrasts, see companion website.

Figure 4. Overview of the model-predicted values for each variable, separated for the three websites. Control: Website without buttons. Like: Website with like buttons. Like & Dislike: Website with like and dislike buttons.

Exploratory analyses

As explained above, the preregistered results were problematic, because communication was not normally distributed and because the predictors were collinear. I hence ran updated analyses, using Bayesian hurdle models, controlling only for confounders but not mediators. The updated exploratory analyses showed different results.

Hypotheses 1, 2, 3, and 4 were all confirmed: If participants expected more gratifications from participation, they communicated more actively: If their expected gratifications increased by one point, on average they also wrote 27 words more (95% CI:

14, 43). If participants were more concerned about their privacy, they communicated less: With each one-point increase in privacy concerns, on average they wrote 16 words fewer (95% CI: -28, -6). If participants felt more self-efficacious, they also communicated much more: If their self-efficacy increased by one point, on average they wrote 73 words more (95% CI: 56, 95). The relationship was curvilinear, almost exponential: Whereas a change in self-efficacy from 1 to 2 only led to an "increase" of zero words, a change from 6 to 7 led to an increase of 103 words. Next, if participants experienced more trust in the website, provider, and other users, they also communicated much more: If their trust increased by one point, on average they wrote 31 words more (95% CI: 14, 53). The first research question asked how privacy deliberation would affect communication. The results revealed a negative effect. The more people deliberated about their privacy, the less they communicated. If privacy deliberation increased by one point, on average they wrote 34 words less (95% CI: -52, -20).

I then reanalyzed the effects of the popularity cues on communication. Compared to the control condition without popularity cues, implementing like buttons did not significantly affect communication, b = -34 (95% CI: -76, 2). However, implementing both like and dislike buttons affected communication. Contrary to what I expected, implementing both popularity cues decreased communication. If both popularity cues were present, participants on average wrote 47 words less (95% CI: -89, -11). The introduction of both cues hence led to a 45% decline in number of words that were written.

Finally, I tested if the effect of the popularity cues on communication were potentially mediated by the privacy calculus variables. Results showed that this was not the case (see online companion website), suggesting that the effect was either direct or mediated by other variables not included here (Coenen, 2022).

Figure 5. Exploratory analyses. Plotted are the average marginal effects of the Bayesian hurdle models. The difference between the control group and like & disklike is significant.

Discussion

This study analyzed whether like and dislike buttons would affect online communication and the privacy calculus. To this end, a preregistered field experiment was conducted, which lasted one week. Furthermore, the privacy calculus model was extended. The privacy deliberation processes was tested explicitly, and trust and self-efficacy were added as predictors.

The preregistered analyses showed that the popularity cues did not affect communication. Only trust emerged as a significant predictor of online communication. All other variables remained insignificant. However, these analyses need to be treated with caution. The predictors were collinear, which makes their integration in one single model

Table 2

Effects of the privacy calculus variables and the popularity cues on communication. Credibility intervals exluding zero are considered significant.

		CI	
Predictor	Estimate	LL	UL
Privacy calculus			
Privacy concerns	-16	-28	-6
Expected gratifications	27	14	43
Privacy deliberation	-34	-52	-20
Trust	31	14	53
Self-efficacy	73	56	95
Experimental conditions			
Like & dislike vs. control	-47	-89	-11
Like vs. control	-34	-76	2
Like vs. like & dislike	13	-19	44

Note. CI = credibility interval, LL = lower level; UL = upper level. Reported are average marginal effects.

problematic (Vanhove, 2021). In addition, the main variable and outcome of the study, number of communicated words, was zero-inflated and gamma distributed, requiring a different type of analysis. The preregistered analyses using structural equation modeling with multiple predictors were hence problematic

To address these shortcomings, in several exploratory analyses I conducted Bayesian hurdle-gamma models (see Data analysis). This updated approach changed the results. People who were more concerned about their privacy wrote fewer words. To illustrate, people who reported being very much concerned posted only 33 words on average, whereas people who reported being not concerned posted 116 words. Participants who received

more gratifications wrote substantially more words, too. Together, the results provide further support for the privacy calculus and against the privacy paradox (Dienlin, 2019). Communication online does not seem to be overly illogical. To large extents, it is aligned with respondents' concerns and benefits.

Results showed that trust and self-efficacy are important drivers of online communication. Participants who placed more trust into the website, the providers, and the other community members communicated more actively. Self-efficacy was the strongest of all predictors. Participants who felt more self-efficacious disclosed much more than others. To illustrate, if people reported no self-efficacy, they wrote only 1 word on average. However, when they reported high levels of self-efficacy, they wrote 253 words. This finding further supports the underlying premise of bounded rationality (Simon, 1990). Although more rational aspects such as costs and benefits influence behavior, behavior is also determined by more socio-psychological variables such as trust and self-efficacy.

The privacy calculus was criticized for not explicitly analyzing the process of weighing pros and cons before disclosing (Knijnenburg et al., 2017). In this study, I thus elaborated on the privacy deliberation process. The results showed that only one third of all participants agreed to have actively weighed the benefits and drawbacks before communicating on the platform. The figure is comparatively low. Even in new online contexts, the majority of users does not seem to actively deliberate about their online communication, suggesting that online use is to large extents implicit (Acquisti et al., 2020).

Privacy deliberation and privacy concerns were strongly correlated ($r = .61$) and showed comparable correlations with other variables. If we are concerned, we also think and deliberate more about our privacy. And if we are not concerned, we do not deliberate. This finding is aligned with decision theory (Elsbach & Barr, 1999). When in a negative state—such as when being concerned—, we judge more critically. And when in a positive mood, we are more lenient. At this point, it is still unknown if thinking about privacy

increases concerns or, conversely, if growing more concerned about privacy makes us deliberate more carefully.

The updated results showed that implementing both like and dislike buttons decreased communication. This was an unexpected and interesting finding. It suggest that negative feedback, or even only the risks of receiving negative feedback, stifles communication. The effect was strong: Implementing both like and dislikes cues led to a 45% decrease in number of written words. This finding is aligned with studies that reported strong effects of popularity cues on behavior (Muchnik et al., 2013). The negative effects of dislike buttons might explain why almost all existing and successful social network sites have chosen to omit such negative popularity cues. At the time of writing, only a handful of websites have (partially) implemented dislike buttons (e.g., youtube, stackexchange, or reddit). Despite the positivity bias of social media (Schreurs et al., 2022), chances of receiving negative feedback and communication are of course real, as can be seen by moral outrages or "shit-storms". Explicit negative popularity cues are low threshold paralinguistic affordances (Carr et al., 2018). They likely prime or trigger negative experiences or expectations, thereby reducing communication.

Websites only with like buttons had no effect on the number of communicated words. If anything, there was an unexpected (nonsignificant) trend toward reduced communication. Although we would expect that like buttons, being positive feedback cues, increase communication, it also seems plausible that they can decrease communication. Not receiving any likes, although the option exists, can be perceived as ostracizing (Schneider et al., 2017). In the context of this study, participants discussed a political topic. Not receiving likes might be more threatening and intimidating than on regular social media, on which it is more common to discuss every-day topics. Although like buttons are commonplace in social media, the findings suggest that in specific contexts they inhibit communication.

Limitations

The main implications and results discussed above rest on exploratory analyses not registered a priori. Exploratory analyses are part of the research process, compatible with preregistration, and important for scientific progress (Dienlin et al., 2021). The updated analyses represent and document a learning process, which arguably led to an improved analysis. However, the results should still be considered preliminary, to be confirmed in subsequent studies.

Whereas the effects of the popularity cues on all variables can be interpreted from a causal perspective (but see below), more caution is needed regarding the effects of the privacy calculus variables on communication. Although the effects were controlled for age, gender, and education, other variables not included could potentially bias the causal estimates presented here (Coenen, 2022).

In experiments only the treatment variable should be manipulated, while all others should be held constant (assumption of stable unit treatment, Kline, 2016). Being a field experiment, several variables could not be held constant, such as the content of communication by other users, the unfolding communication dynamics, and the characteristics of other users. Future research should repeat the design, preferably using several runs of the same experiment, to further assess generalizability and robustness.

Conclusion

This study provides further support for the privacy calculus model and against the privacy paradox approach. Expected benefits, privacy concerns, comparing benefits and drawbacks, trust, and self-efficacy all affected communication. Like and dislike buttons reduced communication significantly. Users can be considered proactive and reasonable. But, similar to everyday offline contexts, they are also affected by the affordances of their environment, and often act implicitly without pondering the consequences of their actions.

References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and likes: The drive for privacy and the difficulty of achieving It in the Digital Age. *Journal of Consumer Psychology*, *30*(4), 736–758. https://doi.org/10.1002/jcpy.1191

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

Arel-Bundock, V. (2023). *Marginaleffects: Predictions, comparisons, slopes, marginal means, and hypothesis tests*. Retrieved from https://CRAN.R-project.org/package=marginaleffects

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). https://doi.org/10.5210/fm.v11i9.1394

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

Bazarova, N. N., & Masur, P. K. (2020). Towards an integration of individualistic, networked, and institutional approaches to online disclosure and privacy in a networked ecology. *Current Opinion in Psychology*, *36*, 118–123. https://doi.org/10.1016/j.copsyc.2020.05.004

Bendor, J. (2015). Bounded rationality. In *International encyclopedia of the social & behavioral sciences* (pp. 773–776). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.93012-5

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

Bräunlich, K., Dienlin, T., Eichenhofer, J., Helm, P., Trepte, S., Grimm, R., ... Gusy, C. (2020). Linking loose ends: An interdisciplinary privacy and communication model. *New Media & Society*, 1461444820905045. https://doi.org/10.1177/1461444820905045

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

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

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

Coenen, L. (2022). The indirect effect is omitted variable bias. A cautionary note on the theoretical interpretation of products-of-coefficients in mediation analyses. *European Journal of Communication*, 026732312210822. https://doi.org/10.1177/02673231221082244

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

Dhir, A., & Tsai, C.-C. (2017). Understanding the relationship between intensity and

gratifications of Facebook use among adolescents and young adults. *Telematics and Informatics*, *34*(4), 350–364. https://doi.org/10.1016/j.tele.2016.08.017

Dienlin, T. (2014). The privacy process model. In S. Garnett, S. Halft, M. Herz, & J. M. Mönig (Eds.), *Medien und Privatheit* (pp. 105–122). Passau, Germany: Karl Stutz.

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

Dienlin, T. (2019). Das Privacy Paradox aus psychologischer Perspektive [The privacy paradox from a psychological perspective]. In L. Specht, S. Werry, & N. Werry (Eds.), *Handbuch Datenrecht und Digitalisierung* (pp. 305–323). Berlin, Germany: Erich Schmidt Verlag.

Dienlin, T., Johannes, N., Bowman, N. D., Masur, P. K., Engesser, S., Kümpel, A. S., … de Vreese, C. (2021). An Agenda for Open Science in Communication. *Journal of Communication*, *71*(1). https://doi.org/10.1093/joc/jqz052

Dienlin, T., & Metzger, M. J. (2016). An extended privacy calculus model for SNSs—Analyzing 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., & Sun, Y. (2021). Does the privacy paradox exist? Comment on Yu et al.'s (2020) meta-analysis. *Meta-Psychology*, *5*. https://doi.org/10.15626/MP.2020.2711

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* (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 *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

Elsbach, K. D., & Barr, P. S. (1999). The effects of mood on individuals' use of structured

decision protocols. *Organization Science*, *10*(2), 181–198. https://doi.org/10.1287/orsc.10.2.181

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.

Gigerenzer, G., Selten, R., & Workshop, D. (Eds.). (2002). *Bounded rationality: The adaptive toolbox* (First). Cambridge, MA: MIT Press.

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

Kezer, M., Dienlin, T., & Baruh, L. (2022). Getting the privacy calculus right: Analyzing the relations between privacy concerns, expected benefits, and self-disclosure using response surface analysis. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *16*(4). https://doi.org/10.5817/CP2022-4-1

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth). 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

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

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

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.

Masur, P. K. (2023). Understanding the effects of conceptual and analytical choices on "finding" the privacy paradox: A specification curve analysis of large-scale survey data. *Information, Communication & Society*, *26*(3), 584–602. https://doi.org/10.1080/1369118X.2021.1963460

Masur, P. K., DiFranzo, D., & Bazarova, N. N. (2021). Behavioral contagion on social media: Effects of social norms, design interventions, and critical media literacy on self-disclosure. *Plos One*, *16*(7), e0254670.

McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton: CRC Press/Taylor & Francis Group.

Meier, Y., & Krämer, N. C. (2022). The privacy calculus revisited: An empirical investigation of online privacy decisions on between- and within-person levels. *Communication Research*, 00936502221102101.

https://doi.org/10.1177/00936502221102101

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*, *341*(6146), 647–651. https://doi.org/10.1126/science.1240466

New York Public Radio. (2018). *The privacy paradox*. https://project.wnyc.org/privacy-paradox/.

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

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/

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *24*(2), 251524591774562. https://doi.org/10.1177/2515245917745629

Rosoff, H., Cui, J., & John, R. S. (2013). Heuristics and biases in cyber security dilemmas. *Environment Systems and Decisions*, *33*(4), 517–529. https://doi.org/10.1007/s10669-013-9473-2

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.

Schneider, F. M., Zwillich, B., Bindl, M. J., Hopp, F. R., Reich, S., & Vorderer, P. (2017). Social media ostracism: The effects of being excluded online. *Computers in Human Behavior*, *73*, 385–393.

Schreurs, L., Meier, A., & Vandenbosch, L. (2022). Exposure to the positivity bias and adolescents' differential longitudinal links with social comparison, inspiration and envy depending on social media literacy. *Current Psychology*. https://doi.org/10.1007/s12144-022-03893-3

Simon, H. A. (1990). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Utility and probability* (pp. 15–18). London: Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-20568-4_5

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 *Privacy online: Perspectives on privacy and self-disclosure in the social web* (pp. 143–158). Berlin, Germany: Springer.

Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, *93*, 1–12. https://doi.org/10.1016/j.chb.2018.11.046

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

Vanhove, J. (2021). Collinearity isn't a disease that needs curing. *Meta-Psychology*, *5*. https://doi.org/10.15626/MP.2021.2548

Westin, A. F. (1967). *Privacy and freedom*. New York, NY: Atheneum.

Whiting, A., & Williams, D. (2013). Why people use social media: A uses and gratifications approach. *Qualitative Market Research: An International Journal*, *16*(4), 362–369. https://doi.org/10.1108/QMR-06-2013-0041

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