Computers in Human Behavior 75 (2017) 484–493

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Whoever will read it – The overload heuristic in collective privacy expectations

Ricarda Moll[*, 1], Stephanie Pieschl [2], Rainer Bromme

Westfälische Wilhelms-Universität, Institut für Psychologie, Fliednerstr. 21, 48149 Münster, Germany

ARTICLE INFO

Article history:
Received 11 January 2017
Received in revised form
27 April 2017
Accepted 24 May 2017
Available online 26 May 2017

Keywords:
Collective privacy
Audience expectations
Self-disclosure
Social Networking Sites
Information overload

ABSTRACT

Users of Social Networking Sites have the difficulty to regulate their privacy although they have limited knowledge about their audiences. We argue that in the absence of such knowledge, people utilize an *overload heuristic*. Users' own experiences with information overload may lead them to perceive others' messages as redundant *noise*. They might thus expect that their own information is protected because others lack the attentional resources to access them. In four experiments we systematically varied two potentially relevant noise cues, *information density* and *audience size*, utilizing different SNS-contexts and experimental designs. We hypothesized that users should estimate the probability of a single audience member (Experiment 1; $N = 124$) and the proportion of audience members ($N = 120$, 89, and 33 for Experiments 2–4 respectively) to read a specific post to be lower in the presence of higher information density and larger audiences. Results show effects for both cues, thereby confirming our hypothesis that users' expectations towards their audience may be based on an overload heuristic.

1. Privacy on Social Networking Sites

The majority of online adolescents and adults likewise use Social Networking Sites (SNS), many of them on a daily basis (Duggan & Brenner, 2013). Communicating via SNS such as Facebook or Twitter has therefore become an integral part of many people's everyday lives. SNS's popularity can be explained by their associated benefits such as the building and maintaining of social capital (Ellison, Steinfield, & Lampe, 2007). However, in the course of reaping these benefits users often disclose self-related information (self-disclosure) which makes them potentially vulnerable to their (unknown) audiences. The associated vulnerability is amplified in comparison with offline contexts, since users' disclosures are stored in the form of persistent digital data, whose accessibility and distribution – once disclosed – is difficult to control. Therefore, the benefits of online self-disclosures are inherently linked to the risks of that same behavior, namely a loss of privacy.

Researchers have long been interested in the ways people regulate this tension between self-disclosure and privacy (Petronio, 2002). With the rise of digital technologies discussions around this topic seem to have shifted towards the question of in how far privacy-related decisions are *rational*. Acquisti (2004) argued that users' rationality regarding their privacy-related behaviors is likely to be bounded, not only because privacy threats are more distant than social rewards (Hallam & Zanella, 2017), but because they have overall *incomplete information* about all potentially relevant variables to begin with, making it difficult to assess risks related to their online behaviors. More specifically, SNS-users have limited information about their *audiences* in at least two ways: First, while users make little use of the possibility to apply granular privacy settings on SNS, they struggle to monitor what information is potentially accessible to which audience (e.g., Pieschl & Moll, 2016). Second, and posing the main focus of the experiments presented here, SNS-users' have limited information about who and how many people *actually access* their information. More specifically, large parts of the actual audience usually remain silent in leaving no response after having read a user's post. Therefore, visible feedback may provide an invalid cue to the size and nature of the actual audience (Bernstein, Bakshy, Burke, Karrer, & Park, 2013). As a consequence of such limited information, users' disclosure-related decisions may rather be based on mental shortcuts. These may, in the context of audiences, be based on specific *beliefs* about the audience's characteristics and restraints, and a corresponding thumb rule may seemingly reduce the uncertainty of

* Corresponding author. DFG-Research Training Group "Trust and Communication in a Digitized World", Geiststraße 24-26, 48151, Münster, Germany
E-mail addresses: ricarda.moll@uni-muenster.de (R. Moll), pieschl@uni-muenster.de (S. Pieschl), bromme@uni-muenster.de (R. Bromme).
[1] Ricarda Moll is now at Verbraucherzentrale NRW e.V. (National Consumer Centre, North-Rhine Westphalia, registered association), Mintropstr. 27, 402015 Düsseldorf, Germany.
[2] Stephanie Pieschl is now at University of Newcastle, School of Education, University Drive, Callaghan NSW 2308, Australia.

http://dx.doi.org/10.1016/j.chb.2017.05.035
0747-5632/

who and how large their actual audience is (e.g., Marwick & Boyd, 2010; Viégas, 2006). These beliefs about the audience may lead to the probabilistic expectation that unintended others *will not* access one's own information despite having the potential to do so. In the following, we will discuss where beliefs about the audience may derive from and how they might influence perceptions of privacy.

1.1. Beliefs about the audience: information as noise

A plethora of empirical work shows that beliefs about others are often inferred from one's own experiences and behaviors (e.g., Ames, 2004; Nickerson, 1999). In the context of audience-related expectations, users' own experiences with *information overload* may be one of the most important one for their audience-related beliefs and expectations. Information overload denotes situations in which more information is available than people are capable of taking in due to their limited perceptual and processing capacities. Such an overload experience is not only an individual phenomenon in the light of specific tasks (Eppler & Mengis, 2004) but has become a common attribute in modern information societies (Klapp, 1978). When people are confronted with more information than they can actually process, information tends to be perceived as noise, namely redundant or meaningless information that interferes with the goals or expected signals of the receiving person (Klapp, 1978). Naturally, people in the information society learn to cope with such overload situations, which are even more amplified regarding social information exchanged in SNS' microblog streams, for example on Twitter or Facebook (e.g., Bontcheva, Gorrell, & Wessels, 2013). Thus, Hargittai, Neumann, and Curry (2012) conclude that users seem to have developed "skills in engaging a sophisticated mix of attention and inattention" (p.163) when dealing with a stream of incoming information. Empirical results support this idea. For example, results from an eye tracking study in which Twitter users' visual attention was measured while they were reading tweets, showed that users attended to each tweet only for a few seconds and remembered less than 70% of what they saw (Counts & Fisher, 2011).

How is the perception of information as noise relevant for users' audience expectations? When users form beliefs about their audience, it is likely that they take themselves as default model to infer the audience's behavior (see above). It may therefore be that users – on the basis of their own experience – assume that their potential online audience also experiences information overload and perceives incoming information as noise. Thus, users may believe that their potential audience must also be selective in its information consumption. Then, when others in the audience also have to filter out signals from noise, users might perceive a decreased probability that the potential audience's members really retrieve their own self-disclosed information. In other words, users would rely on their potential audience to filter contents appropriately to a given context instead of managing their audiences pro-actively (Litt & Hargittai, 2016). The result of such inferences would consequently be the heuristic that although others *potentially* have access to certain contents, it is unlikely that they would make a time-costly effort to *actually* retrieve and process them in the presence of high information overload (see Lundblad, 2004). Therefore, users might even expect *collective privacy*, namely that their own information is *protected* by the asymmetry between the mass of all online data on the one hand, and the potential audience's limited processing capacities on the other hand (Lundblad, 2004).[3] More specifically, if users infer the risk of their online self-disclosures from the "extent to which [they] are the subject of others' attention" (Gavison, 1980, p. 423), it would explain why people "often act as if [the online audience] is bounded" when it indeed is "potentially limitless" (Marwick & Boyd, 2010). As a consequence, users would regulate their privacy boundaries rather according to their probabilistic audience expectations derived from an overload heuristic than according to their actual needs.

In order to prove our general idea that people rely on an overload heuristic when regulating their privacy concerns we tested the effect of specific cues which might signalize information overload. Since users are unable to experience the actual extent of the audience's attention, they have to use such cues to infer the audience's information overload. In other words, the more information overload the audience is expected to have, the smaller the perceived likelihood that the audience will attend to one's self-disclosed information. This *overload heuristic* might therefore be triggered by specific environmental cues (*noise cues*) indicating the extent of overload of one's audiences.

1.2. Overload heuristic: noise cues

If users' expectations regarding the size of their actual audience indeed exist within an *overload heuristic*, there should be specific effects of cues indicating the audience's overload experience onto their expectation regarding the size of their actual audience. As this question has not been investigated empirically before, we can only speculate about the most relevant noise cues.

One relevant cue from which to infer the potential audience's behavior and thus make inferences about the actual audience might be *information density*, namely the mere amount of content the potential audience is believed to be confronted with. Even basic research about object recognition has shown that people's capability to identify a specific stimulus decreases when presented within a crowd of distractors, because a larger amount of details poses a greater affordance to people's filtering ability (e.g., Pelli, 2008). This may partially be transferred to the context of information overload in SNS-communication, because it is especially the distracting mass of contents that makes a careful differentiation between informational noise and signal necessary (see above). Thus, SNS-users seem to merely roughly screen information when there is too much of it to carefully attend to. Expecting such a behavior from one's potential audience in the light of information density should result in a lowered expectation that the potential audience will actually attend to a specific piece of information.

A second noise cue might be the mere size of the potential audience (*audience size*): While on the one hand larger and more diverse audiences indicate more risks than smaller and less diverse ones (Litt et al., 2014), a users' potential audience size may on the other hand serve as an indirect cue to the audiences' information overload: Being a member of a large audience makes it more likely to have a high number of contacts oneself (often determining the size of the potential audience) – who again tend to produce more information than smaller audiences. Paradoxically then, under conditions of limited attention, a large potential audience may make it less likely that the members actually read a specific post. Thus, Bernstein et al. (2013) found that the *proportion* of the *actual* audience seeing a post *decreased* with an increase of the *potential* audience's size (see also Ugander, Karrer, Backstrom, & Marlow, 2011). In other words, while a larger potential audience may result in higher *absolute* numbers of readers, a reasoning on the basis of the overload heuristic should result in expectations of actual readership that *decrease in proportion* to the potential audience's size (see Bernstein et al., 2013).

Taken together, we proposed that users apply an *overload*

[3] Note that the term *collective privacy* coined by Lundblad (2004) is conceptually different from Petronio's (2002) idea of collective boundary management, in which information is protected through the responsible actions of groups of people who co-own a specific piece of information.