How Much Are “Many People” on Facebook? Interpretations of Vague Quantifiers in Online and Offline Contexts

Ricarda Moll1,2, Anna Jori Lücke1,3, and Rainer Bromme1

Abstract
In an increasingly interconnected world, many people handle large parts of their communication online, often via social networking sites (SNS). In contrast to face-to-face communication, messages on SNS are accessible by potentially unknown and large audiences. However, it is an open question what users actually perceive as a large audience, or else as many people in SNS contexts. Exploring this question from a psycholinguistic perspective, we investigated the meaning of vague quantifiers such as “few” or “many” with regard to audiences in different contexts in two experiments. In Experiment 1, participants assigned numbers to quantifiers describing audiences in online versus offline and private versus public contexts. In Experiment 2, including the same items as Experiment 1, participants rated the appropriateness of specific numbers of people that were described by a quantifier. Our results show, for example, that people assigned larger numbers to quantifiers for online than for offline contexts. This was also true when access to the information was supposed to be restricted which implies a (scalar) change of privacy expectations.

Keywords
natural language quantifiers, social networking sites, privacy, online communication, audiences

From an early age on people extend their mode of communication to digital devices and thus to the online realm, in particular to social media such as social networking sites (SNS; MPFS, 2020). Three quarters of adult U.S. internet users have an account for at least one SNS (Perrin, 2015), which they mostly use on a daily basis. These services change crucial parameters of communication as people know it from offline contexts. As opposed to offline contexts, it is inherently easy to make a message publicly available; in contrast, however, it takes an effort to keep content private when communicating online (boyd, 2010). This is partly caused by the fact that the content is stored on the service provider’s servers from where it is widely accessible by potentially unknown audiences at any point of time in the future (Bernstein et al., 2013; boyd, 2008). While audiences on SNS may on one hand provide social support (Ellison et al., 2007), they may on the other hand cause a user to experience negative consequences such as context collapse (Marwick & boyd, 2010) and adverse events such as gossip, hurtful comments, or social embarrassment. The extent to which users are vulnerable to these risks thereby depends on the size of their online network, with larger audiences typically indicating greater vulnerability (Buglass et al., 2016).

When SNS users decide to disclose information to an online audience, their weighing of risks and benefits (so-called privacy calculus, Trepte et al., 2017) is likely to be biased. For instance, people have limited information about their privacy settings (Moll et al., 2014) and they perceive their own internet privacy risks to be smaller compared with others’ risks (Third Person Effect, Chen & Atkin, 2021). SNS users—as all people—have a preference for immediate gratification and may therefore consider the benefits of self-disclosure rather than the distal risks of that same behavior (Acquisti, 2004; Hallam & Zanella, 2017). In addition, users may have distorted implicit beliefs about the—potentially large—audience’s behavior and seem to think that the actual audience will be small despite the public nature of their shared information (Moll et al., 2017).

Interestingly however, the perception of an audience to be large (or not) may depend on the collectively understood attributes of the particular context, namely the mode of communication. Thus, the fact that online and offline

---

1Westfälische Wilhelms-Universität Münster, Germany
2Hochschule des Bundes für öffentliche Verwaltung, Brühl, Germany
3Ruprecht Karl University of Heidelberg, Germany

Corresponding Author:
Ricarda Moll, Hochschule des Bundes für öffentliche Verwaltung, Willy-Brandt-Str. 1, Brühl D-50321, Germany.
Email: ricarda.moll@hsbund.de

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
communication differ in their defaults regarding information access and distribution is likely to have shaped people's understanding and expectation of what it means, when there are “many” or “few” people in an audience. In the following, we will refer to the psycholinguistic approach of so-called natural language quantifiers or vague quantifiers to elab-
orate this aspect.

Quantiﬁers and Context

Vague quantifiers or Natural Language Quantifiers (NLQs) such as “few,” “some,” “lots of,” or “many” are a fundamental part of our language. Whenever we are not able to speak of specific numbers, we use vague quantifiers to give an approximating impression of our perception. Vague quantifiers are believed to “map onto a second number system that appears to be common to pre-linguistic infants, adults and many animals” (Coventry et al., 2010, p. 222). Since NLQs are vague by definition and purpose, their interpretation inherently depends on the context in which they occur (Newstead & Coventry, 2000). To grasp the meaning of a proposition that includes a vague quantifier, conversation partners need to have some form of shared conceptual understanding of the context (Clark & Brennan, 1991).

For example, to say that “there are many grains of sand” presumes that the receiver of the message has an idea about the quantity in which grains of sand usually occur and that the number related to “many grains of sand” will probably be larger than the one related to the statement about “many houses” (Hörmann, 1983; Pogue et al., 2012). This implies that an object’s cognitive representation includes a scalar concept, namely an understanding of the expected average (expectancy value) of the object’s property as a standard of comparison to the observed value (Syrett et al., 2009). Thus, the expectancy value for an average number of grains of sand will most probably exceed the number of houses, although both are referred to by the same quantifier. As such, the numbers that people map to vague quantifiers mark specific points of the scale inherent in one’s knowledge of the context or, in the above example, knowledge about attributes of the object itself.

Context, however, is not only implied by the attributes of an object but also by the situation in which the object occurs. In that way, set size, functionality, and even unrelated objects in the vicinity can influence the interpretation of vague quantifiers (Coventry et al., 2010; Newstead & Coventry, 2000). Thus, stating that “there are many grains of sand on the beach” will lead to a numerical inference much larger than stating that “there are many grains of sand on the floor.” In other words, mapping numbers to a vague quantifier presumes not only some form of collectively shared knowledge about the object, but also prior expectations about the usual quantity of the named object in a particular situation (Moxey & Sanford, 2000, for a review). Importantly, the present study is based on the idea that when the interpretation of a quantifier depends on context, we can infer people’s prior expectation regarding the context from their interpretation of the quantifier—and thus learn something about their understanding of the context itself.

For instance, we assume that people’s context knowledge entails the (principally valid) assumption that in comparison to public-by-default online communication, offline communication is generally private-by-default and that publicness is only achieved through effort. If this was true, they should associate lower quantities with the expression “many people” for offline than for online contexts. Moreover, if this was also true for situations where information access is restricted as in private groups on SNS, for example, this scalar shift would also imply a tolerance for larger audiences in situations that are considered to be private.

In a typical paradigm in which the meaning of vague quantifiers is empirically investigated, short vignettes are presented to the participants in which a quantifier describes an object within a specific context. Usually, the quantifier describing the object and/or the context in which it occurs is varied experimentally. Participants are asked to indicate, for example, what specific number the quantifier denotes in the situation described in the vignette (Moxey & Sanford, 1993; Newstead & Collis, 1987 see Experiment 1). The changing numerical meaning of one and the same quantifier allows for an inference about different expectations people have across different contexts. With regard to online communication, a most interesting question concerns the interaction between different contextual parameters. For example, would a private context entail different numerical expectations when it takes place within an online as opposed to an offline interaction?

Other experimental designs turn this idea around: Participants are presented with specific numbers or proportions of an object together with a quantifier that describes the specific number in a given context. Participants then indicate how appropriately the quantifier describes the number of objects in the given context (Hörmann, 1983; Newstead & Coventry, 2000; Pogue et al., 2012 see Experiment 2). This method allows for an inference about the extent to which people’s numerical understanding of an object is variable or seems to have an upper or lower limit. With regard to the contextual element of privacy, this method would allow for an inference about the extent to which privacy is a more variable concept online as compared with offline situations.

In the present study, we applied both study designs to investigate the participants’ expectations in different contexts. In the following, we will elaborate and report the rationale, method, results, as well a brief discussion separately for Experiment 1 and then Experiment 2, followed by a general discussion for both studies in reference to the theoretical idea and potential implications.
Experiment 1

Rationale and Hypotheses

The first aim of the present study was to explore people’s expectations regarding the quantity of people that audiences encompass in different situations (i.e., audience expectations). Therefore, we asked participants to map numbers to quantifiers for audiences in different contexts. We chose the quantifiers viel (German for many) and wenig (German for few), which are often used in studies conducted in German to assess people’s numerical understanding of vague quantifiers (e.g., Heim et al., 2020). Importantly, while both mark smaller or larger quantities of an object, they are at the same time not strictly “anchored to the bottom [. . . ] or top of a quantity scale in the way that, say, hardly any, about half and nearly all are anchored” (Moxey & Sanford, 1993, p. 77). The quantifier few as opposed to alternatives such as a few may thereby shift participants’ focus to the complement set of the context. Thus, participants may infer the meaning of few from their expectation of how many are not part of these few (Moxey, 2006; Moxey & Sanford, 1986). While the focus on reference versus complement set may influence the semantics of the broader message (especially in face threatening contexts; see Gulan et al., 2014), it is not ultimately clear in how far it generally changes people’s numerical interpretation of the quantifier in question in substantial ways.

We hypothesized that numbers assigned to the quantifier many would always be larger than numbers assigned to the quantifier few (Hypothesis 1), because the former by definition includes higher ranks of a scale than the latter one (Moxey & Sanford, 1993).

Second, we hypothesized that people’s experiential understanding of online contexts leads them to expect larger numbers of people in online audiences—and thus assign greater numbers to both quantifiers—compared with audiences in offline contexts (Hypothesis 2).

At the same time, audience size is influenced by the extent to which there are physical or technological restrictions to the extent to which a message can be accessed by others, that is, the extent of privacy. Overall, numbers assigned to quantifiers in public contexts should be larger than the ones in private contexts (Hypothesis 3). However, it is an open question, what numerical expectations are triggered when both the mode of communication and the extent of privacy describe the context in which an audience occurs. We therefore explored the question of how both contextual details interact regarding their effect on people’s audience expectations (Research Question 1).

Methods

Participants. We distributed the link to the online experiment via snowball sampling which started on the SNS Facebook. Since the study was conducted in German, the sample was restricted to German-speaking participants. Our final sample consists of \( N = 126 \) participants (dropout via premature termination of the experiment \( = 36.8\% \)). The majority of participants was female (81%) and the mean age was 23.91 years, \( (SD = 3.03; range = 18–30\) years). About three quarters of our sample consisted of students (70.6%). Overall, 95.2% were SNS users and mostly used their SNS at least once a day (79.3%).

Design and procedure. The experiment was implemented as computerized survey via Unipark (QuestBack). As an introduction, participants were informed that we were interested in the numerical meaning of vague words like few or many. They were not given further top-down information which could confound their estimations (Wänke, 2002, see also Walentynowicz et al., 2021). Participants were then presented with the experimental materials. We used a \( 2 \times 2 \times 2 \) design with quantifier (many vs. few), communication mode (online vs. offline), and privacy (public vs. private) as within-subject factors. Participants were presented with two blocks of 22 one-sentence statements (items), respectively. These included a quantifier and a specific context which realized the experimental variations of the factors communication mode and privacy. Each item required an estimate of a number the quantifier would denote. Both blocks included the same items, the only difference being the included quantifier. Block sequence as well as item sequence within each block were counter-balanced to avoid position effects.

Materials. We were particularly interested in expectations regarding audiences in different situations, namely, the number of people who (1) receive a message generally, (2) react to a message, (3) see a photograph, (4) are members of a group, (5) are invited to a party, and (6) the number of friends. Except for item Group 6, every group contained four items, each of which provided different contextual information realizing our experimental variations (Table 1): For the factor communication mode, an online mode included situations that took place online (mostly on an SNS), an offline mode included, for example, face-to-face talks or newspaper ads. For the factor privacy, in a private situation the number of included people was in some way restricted and controllable, whereas such control was not possible in public situations. Each item was presented once for each quantifier.

Results

The descriptive results indicated large variability between the participants’ estimates. For example, mean estimates for the item many Facebook friends was 568.78 with a standard deviation of 320.66 (range: 50–1,583). We therefore report the median for each item’s distribution to facilitate interpretation (Table 1).

Apart from the descriptive results per item (Table 1), we were interested if the numbers assigned to the specific quantifiers differed as a function of our experimental variations of
Table 1. Medians, Lower (Many), and Upper (Few) Quartiles for Items in Experiment 1.

| Privacy | Communication mode | Item groups | Many Q25 | Many Median | Few Q75 | Few Median |
|---------|--------------------|-------------|----------|-------------|---------|------------|
| (1) Number of people who have received a message | | | | | | |
| Private | Online | A person sends a message to ___ people. How many people have received the message? | 18 | 30 | 10 | 3 |
| | Offline | A person talks to ___ people about something that happened to her. To how many people has he/she talked? | 10 | 15 | 6 | 2 |
| Public | Online | ___ people read a blog entry. How many people have read the blog entry? | 71 | 200 | 20 | 10 |
| | Offline | ___ people read an open letter to the editor in a major daily newspaper. How many people have read the letter? | 200 | 1000 | 200 | 70 |
| (2) Number of people who have reacted to a message | | | | | | |
| Private | Online | ___ people leave a comment under a post in a private Facebook group. How many people have commented it? | 10 | 20 | 6 | 2 |
| | Offline | ___ people call someone who has left a note in the hallway of a house. How many people have called her? | 5 | 8.50 | 4 | 2 |
| Public | Online | ___ people leave a comment under a post in a public Facebook group. How many people have commented it? | 50 | 150 | 20 | 7.5 |
| | Offline | ___ people call someone who has published a newspaper ad. How many persons have called her/him? | 20 | 32.50 | 10 | 4 |
| (3) Number of people who have seen a photo | | | | | | |
| Private | Online | ___ people see a photo on a private Facebook timeline. How many people have seen the photo? | 50 | 100 | 24 | 10 |
| | Offline | ___ people have seen a photo in a family album. How many people have seen it? | 15 | 20 | 10 | 3 |
| Public | Online | ___ people see a public Facebook profile photo. How many people have seen the photo? | 200 | 300 | 50 | 20 |
| | Offline | ___ people see a photo on an advertising pillar. How many people have seen the photo? | 375 | 1000 | 100 | 50 |
| (4) Number of people in a group | | | | | | |
| Private | Online | ___ people are part of a group message on Facebook. How many people are part of the message? | 15 | 20 | 8 | 3 |
| | Offline | ___ people stand together and talk to each other. How many people are there? | 8 | 10 | 6 | 3 |
| Public | Online | A public Facebook group has ___ members. How many members are there? | 975 | 3528 | 100 | 30 |
| | Offline | A club has ___ members. How many members are there? | 200 | 500 | 30 | 15 |
| (5) Number of people who have been invited to a party | | | | | | |
| Private | Online | A person has invited ___ people to a party via a private Facebook event. How many people has she/he invited? | 50 | 60 | 12 | 8 |
| | Offline | A person sends ___ invitation cards for a party via the postal service. How many people has she/he invited? | 30 | 50 | 10 | 5 |
| Public | Online | A person has invited ___ people to a party via a public Facebook event. How many people has she/he invited? | 400 | 1000 | 100 | 50 |
| | Offline | ___ people see a poster ad to a public party? How many people have seen the poster? | 273 | 723 | 100 | 50 |
| (6) Number of friends | | | | | | |
| Online | | A person has ___ Facebook friends. How many Facebook friends does she/he have? | 338 | 500 | 100 | 42.5 |
| Offline | | A person has ___ friends (not Facebook friends!). How many friends does she/he have? | 17 | 25 | 10 | 3 |

Note. Blank spaces were filled with the quantifier *many* or *few*. Items were presented in German, Table 1 portrays translations into English.
context. To investigate this, outliers with $z > 3.29$ were adjusted to a back-transformed value of $z = 3.29$ (see Field, 2009, p. 153). We then created means for each experimental condition across item groups (note that we excluded item Group 6 regarding number of friends, because we did not have a full cell realization; see Materials; see Table 1). Based on these means, we ran a $2 \times 2 \times 2$ repeated measure ANOVA with quantifier, communication mode, and privacy as within-subject factors. We also ran a parallel analysis with means for items that were z-standardized based on the mean and standard deviation of the corresponding item group. This produced the same effects as reported in the following with slightly increased effect sizes in the parallel analysis.

Results show significant main effects for quantifier, $F(1, 125) = 29.03, p < .001, \eta^2_p = .19$, communication mode, $F(1, 125) = 12.71, p < .01, \eta^2_p = .09$, and privacy, $F(1, 125) = 29.27, p < .001, \eta^2_p = .19$. As expected (Hypothesis 1), participants assigned higher numbers to the quantifier many ($M = 87449.12, SD = 13976.32, Med = 1140$) than to the quantifier few ($M = 41.71, SD = 53.36, Med = 25$). Confirming Hypotheses 2 and 3, higher numbers were assigned to online contexts ($M = 5450.64, SD = 13212.79, Med = 550$) than offline contexts ($M = 1341.48, SD = 2602.97, Med = 307$), as well as to public contexts ($M = 6757.07, SD = 13960.73, Med = 1107$) compared with private ones ($M = 35.05, SD = 60.44, Med = 22$). These main effects need to be interpreted in the light of several significant interactions.

First, we found a two-way interaction between quantifier and communication mode, $F(1, 125) = 12.83, p < .001, \eta^2_p = .09$. More specifically, many denoted a far higher number in online ($M = 10867.57, SD = 26425.10, Med = 1027$) than in offline modes ($M = 2633.24, SD = 5191.21, Med = 579$), which was not the case for the quantifier few: Few people in online modes ($M = 33.71, SD = 45.37, Med = 18$) denoted a similar number as in offline modes ($M = 49.72, SD = 76.02, Med = 26$).

A two-way interaction between privacy and communication mode, $F(1, 125) = 12.50, p < .01, \eta^2_p = .09$, showed that especially public online contexts ($M = 10867.57, SD = 26397.40, Med = 1062$) yielded larger numbers than public offline contexts ($M = 2669.26, SD = 5205.98, Med = 599$). For private contexts, the absolute difference between online ($M = 56.51, SD = 120.75, Med = 30$) and offline ($M = 13.69, SD = 5.14, Med = 13$) was smaller (Figure 1). Note, however, that the relative difference between online and offline remains similar within the public and the private condition, respectively. An additional three-way interaction between quantifier, communication mode, and privacy, $F(1, 125) = 12.62, p < .01, \eta^2_p = .09$, revealed that for

![Figure 1. Bars indicate the medians of estimated numbers that participants assigned to the quantifiers in the described context conditions (x-axis): privacy (private, public) and mode of communication.](image-url)
public contexts, few people offline \((M = 95.59, SD = 152.05, Med = 49)\) actually meant more than few people online \((M = 61.33, SD = 90.32, Med = 33)\)—turning the predicted pattern around for this particular condition.

Finally, a two-way interaction between quantifier and privacy, \(F(1, 125) = 28.63, p < .001, \eta^2_p = .19\), emphasized that for public communication there was a large difference between the meaning of many \((M = 13435.67, SD = 27922.64, Med = 2165)\) and few \((M = 78.46, SD = 106.52126, Med = 45)\) while for private communication the absolute difference between many \((M = 65.14, SD = 120.93, Med = 38)\) and few \((M = 4.96, SD = 1.80, Med = 4)\) was smaller.

**Discussion**

The results found in Experiment 1 overall support Hypotheses 1 to 3. However, we also found several interesting interactions between the experimental factors.

**Quantifier × privacy.** First, our results showed that many people were always more than few people within the same situation (Hypothesis 1). However, while this result seems straightforward, the extent to which many denotes more than few seems to depend on context conditions. More specifically, when a situation is public, the difference between many and few may be particularly large, while the difference may be smaller for private situations. This implies that private situations are understood to restrict the amount of people in the audience at the upper end of the scale. Thus, in private situations, many people cannot mean that much more than few people, because from a certain quantity on, the situation would be better characterized as public.

**Quantifier × mode of communication.** Similarly, while participants generally expected larger audiences in online than in offline modes of communication (confirming Hypothesis 2), this was particularly true for the quantifier “many”: “Few people” denoted similar numbers in offline as in online contexts. We can conclude that the meaning of “few” is less context dependent than the meaning of “many”: “Many” does not imply a restriction at the upper end of a numerical scale which seems to be different for the meaning of “few.” Thus, when speaking of “many people” in an audience, people’s everyday understanding of online versus offline contexts seems to acknowledge the different core characteristics of these contexts.

**Privacy × mode of communication.** A similarly interesting pattern occurred for the factor privacy. While participants on average expected roughly twice as much people in public than private situations (confirming Hypothesis 3), the difference needs to be interpreted in the light of a significant interaction with mode of communication: For private contexts, the absolute but not the relative difference between online and offline contexts was smaller than for public ones. On one hand, this indicates the context condition “private” restricts the number of expected people at the upper end of the scale, even for online contexts. On the other hand, private online situations still evoke the expectation of larger audiences than in private offline situations.

**Experiment 2**

**Rationale and Hypotheses**

Overall, the descriptive results from Experiment 1 showed a large between-subject variance in the estimations for one and the same item, indicating that people may have very individual expectations toward the meaning of vague quantifiers in certain situations. This can in part be explained by different experiences that people make in one and the same context—experiences that shape their contextual understanding. Importantly, however, regardless of their individual experiences, people should at least roughly agree in their contextual understanding when acting in or talking about an audience, because otherwise they would not properly understand what someone else means by describing many or few people in an online or offline, in a private or public context. We therefore conducted a second experiment, which tested how appropriate people perceived a quantifier to describe a specific number of people in an audience context and to what extent they tolerated deviations from an expectancy value.

In Experiment 2, we used the same vignettes as in Experiment 1, but this time they included specific numbers that were described by the quantifier. The numbers differed in the extent to which they could appropriately be described by the quantifier “few” or “many”. They either met participants’ expectancy value, namely when we filled in the median of the item’s distribution from Experiment 1, or else they deviated more or less from this expectation (see section “Materials”). Participants then rated how appropriately the given quantifier described the specific number in the vignette. We hypothesized that people’s appropriateness ratings would decrease with an increased deviation from the median (Hypothesis 4).

Experiment 2 aimed to validate the results and interpretations from Experiment 1 in two major ways. First, if people mainly agreed with the appropriateness of the median in a particular situation, this would support the validity of the descriptive results from Experiment 1 (also see Hypothesis 4). Second, we aimed to investigate in an explorative manner whether contextual parameters would influence peoples’ acceptance of deviations from an expectancy value. This would imply that context determines in how far people’s numerical expectations are variable.

**Methods**

**Participants.** Participants were school students from Germany who were recruited during their visit of the local university
in Muenster. Ninety-seven participants started the experiment; one person was excluded because of systematic response behavior, seven participants were excluded because they did not explicitly agree that their data could be used for further analysis. Our final convenience sample thus consisted of N = 89 participants (73% female) who were on average 16.96 years old (SD = .81). Overall, 95.5% were SNS users and mostly used their SNS at least once a day (79.8%).

**Design and procedure.** The experiment was implemented as computerized survey via Unipark (QuestBack). We used a 2 × 2 × 2 × 3 design. *Quantifier* (many vs. few) was our between-subject factor and communication mode (online vs. offline) and privacy (public vs. private) were within-subject factors. We added a third within-subject factor, namely the value of the quantifier (target, ambiguous, contrast; see section “Materials”). Participants were introduced to the study and were then randomly assigned to one of two experimental groups. In each group, participants gave appropriateness judgments for 66 items that were presented on a total of three pages. Page sequence and item sequence on each page were counter-balanced. Afterward, participants were debriefed and were offered to leave their email address to receive additional information about the study.

**Materials.** While the context specified within the items remained the same (Table 1), participants in Experiment 2 had to repeatedly judge how appropriately the quantifier described a specific number that we included in each item’s depiction. For example, one item said “Many persons read a blog entry.” Overall, 200 persons read the entry.” Appropriateness of the quantifier (in this case: many; target) was rated on a 7-point-Likert-type scale (1 = not appropriate at all; 7 = completely appropriate).

Each of the 22 items from Experiment 1 was presented in three different versions within each quantifier condition; thus, each participant rated 66 out of a total of 132 items. The three item versions differed in their numeric value describing the quantifier (within-subject factor; target, ambiguous, contrast). The corresponding numbers were taken from the results from Experiment 1. Targets included a number that would likely meet people’s expectancy value for the described context, namely the median of the corresponding item’s distribution in Experiment 1 (the above example realizes a target).

The ambiguous version included a number that would deviate to some extent from the median in the opposite direction. For the quantifier many, we were interested in numbers falling below participants’ expectancy value. Thus, we filled in the lower quartile of each item’s distribution from Experiment 1 (see Table 1, column Many / Q25). The above example would then say: “Many people read a blog entry. Overall, 71 persons read the entry.” For the quantifier few, we were interested in numbers exceeding the expectancy value. Thus, items in the ambiguous condition included the upper quartile for the quantifier few. The above example would then say: “Few people read a blog entry. Overall, 20 persons read the entry” (see Table 1, column Few / Q75).

The contrast version meant to produce a clear inappropriateness of the quantifier. Here, we included the median of the opposite quantifier from Experiment 1. The above example would then entail the following propositions in the condition many: “Many persons read a blog entry. Overall, 10 people read the entry.”

**Results**

The randomized distribution of participants to one of the two quantifier conditions resulted in a slightly uneven cell distribution with 41 participants in the many-condition, and 48 participants in the few-condition. We conducted a a × 2 × 2 × 3 mixed ANOVA as a parametric method, since Levene’s test of equality of variance confirmed the homogeneity of variance across groups. Violations of the assumption of sphericity were corrected using the Greenhouse Geisser estimation when necessary. All statistical analyses were performed using SPSS version 22. *Value* (target, ambiguous, contrast), *mode of communication* (online, offline), and *privacy* (private, public) were the within-subject factors. *Quantifier* (many, few) was the between-subject factor.

As our materials’ design intended, appropriateness ratings for targets lay significantly above the scale’s mean (4), t(88) = 14.25, p < .001. Moreover, there was a significant main effect for the factor value, F(1,537, 133.676) = 779.68, p < .001, η²p = .90. Planned contrasts (method: repeated) showed that targets were judged to be significantly more appropriate than ambiguous distractors, F(1, 87) = 298.98, p < .001, η²p = .77, which again were judged more appropriate than contrast distractors, F(1, 87) = 794.71, p < .001, η²p = .90 (Table 2). The degree to which this was the case, however, depended on the quantifier condition, as indicated by a two-way interaction between value and quantifier, F(2, 174) = 14.50, p < .001, η²p = .14. For targets, ratings in the few- and the many-condition are almost the same (Table 2). However, in the many-condition, the difference between ratings for the target and the ambiguous value is less pronounced than in the few-condition, F(1, 87) = 14.49, p < .001, η²p = .18 (planned contrasts, method: repeated). Thus, for the ambiguous value, appropriateness ratings are higher in the many-condition. This indicates that the meaning of “many” is more gradable than the meaning of “few.” At the same time, the difference between the ambiguous and the contrast value was even more pronounced in the many-condition as compared with the few-condition, F(1,87) = 35.41, p < .001, η²p = .29. Thus, there seems to also be a threshold as to what “many” can numerically mean in a specific context.
| Privacy | Online | Offline |
|---------|--------|---------|
|         | Many (n = 41) | Few (n = 48) | Overall (n = 89) | Many (n = 41) | Few (n = 48) | Overall (n = 89) | Overall |
| Private |        |        |        |        |        |        |        |
| Target  | 4.97 (0.99) | 5.40 (0.96) | 5.20 (0.99) | 5.07 (1.08) | 5.27 (1.13) | 5.42 (1.15) | 5.31 (0.99) |
| Ambiguous | 4.26 (1.02) | 3.93 (1.05) | 4.09 (1.05) | 4.42 (1.03) | 3.83 (0.99) | 4.10 (1.05) | 4.09 (0.93) |
| Contrast | 1.98 (0.64) | 2.55 (0.85) | 2.29 (0.81) | 1.90 (0.64) | 2.48 (0.90) | 2.22 (0.84) | 2.25 (0.76) |
| Overall  | 3.74 (0.77) | 3.96 (0.76) | 3.86 (0.77) | 3.80 (0.75) | 4.01 (0.76) | 3.91 (0.76) | 3.88 (0.71) |

**Public**

| Target  | 5.24 (0.80) | 4.53 (0.94) | 4.86 (0.94) | 5.36 (0.90) | 4.73 (1.00) | 5.02 (1.00) | 4.94 (0.85) |
| Ambiguous | 4.43 (0.90) | 3.68 (0.95) | 4.03 (0.99) | 4.45 (0.93) | 3.80 (0.86) | 4.10 (0.95) | 4.06 (0.82) |
| Contrast | 2.12 (0.66) | 2.08 (0.75) | 2.10 (0.71) | 2.58 (0.70) | 2.28 (0.78) | 2.42 (0.76) | 2.26 (0.64) |
| Overall  | 3.93 (0.65) | 3.43 (0.63) | 3.66 (0.69) | 4.13 (0.71) | 3.60 (0.66) | 3.84 (0.73) | 3.75 (0.63) |
| Overall  | 3.83 (0.59) | 3.70 (0.59) | 3.76 (0.59) | 3.96 (0.63) | 3.81 (0.53) | 3.88 (0.58) | 3.81 (0.55) |

**Note.** Appropriateness was rated on a 7-point Likert-type scale (1 = very inappropriate, 7 = very appropriate).

We found a significant two-way interaction between value and privacy, $F(1.86, 161.38) = 9.75, p < .001, \eta^2_p = .10$. Mauchly’s test indicated a violation of the sphericity assumption, $\chi^2(2) = 7.00, p < .05$; we therefore corrected the results with the Greenhouse Geisser estimation of sphericity; $\varepsilon = .93$. This indicates that participants found numbers assigned to private contexts more appropriate than in public contexts, but, as planned contrasts (method: repeated) showed, this difference is only present for the target as compared with the ambiguous value, $F(1.87) = 12.67, p < .001, \eta^2_p = .13$ (Table 2).

The factor privacy also interacted with the factor quantifier, $F(1.87) = 24.39, p < .001, \eta^2_p = .22$. It shows that in public contexts the numbers assigned to the quantifier many were judged as more appropriate than the numbers assigned to the quantifier few. The opposite is true, however, in private contexts. Here, appropriateness ratings were higher in the few- than in the many-condition. The only exception to this pattern occurs for the ambiguous value, which is revealed by a significant three-way interaction between privacy, quantifier, and value. The ambiguous value is judged as equally appropriate in all privacy and quantifier conditions.

The significant main effect of the factor mode of communication, $F(1.87) = 7.46, p < .01, \eta^2_p = .08$, indicates that participants gave higher appropriateness ratings for offline than for online contexts. A three-way interaction between mode of communication, value, and privacy, $F(2, 174) = 12.67, p < .001, \eta^2_p = .046$, reveals, that in the public condition, this was true regardless of the value condition. In the private condition, however, ratings are only higher for offline than for online contexts when presented with the target condition, that is, a number of people that could realistically be expected in the described situation (Table 2).

**Discussion**

While Experiment 1 showed that people can have a different understanding of vague quantifiers in a specific context, results from Experiment 2 suggest that there is some kind of agreement about scale properties of audiences in different contexts, that is, about the meaning of vague quantifiers. Thus, participants rated the medians extracted from the items’ distributions in Experiment 1 to be appropriately described by the two quantifiers. More specifically, their ratings were not only higher than the scale’s mean, but also higher than their ratings for the two distracters (Hypothesis 4).

**Value × quantifier.** For the quantifier few, smaller deviations from the expectancy value (ambiguous value) were judged less appropriate than smaller deviations from the expectancy value of “many” in a given context. This pattern indicates—and confirms our results from Experiment 1—that people’s understanding of “many” is more gradable than their understanding of “few.”

**Privacy × quantifier.** Moreover, the two quantifiers seem to be conceptually tied to different contexts. Thus, participants in the few-condition gave higher appropriateness ratings for private as opposed to public contexts, whereas the opposite was true for participants in the many-condition—they found numbers in public contexts overall more appropriate than in private contexts.

**Value × mode × privacy.** The results showed a main effect for the factor mode of communication, showing that overall numbers in offline contexts were rated to be more appropriately described by the quantifiers than in online contexts. For public contexts, this was true for all quantities and all value conditions. For private contexts, however, only the target condition produced higher appropriateness ratings offline than online. On one hand, this pattern shows that regardless of the mode of communication, the conceptual understanding of “private” remains narrow. Thus, even smaller deviations from the expectancy value carry a lot of weight as to the perceived appropriateness of the number that is described by the quantifier. On the other hand, since
appropriateness ratings in the public conditions were also higher when combined with the ambiguous or the contrast value, this interaction suggests that people do not necessarily have fixed quantitative expectations once communication is online and public.

**General Discussion**

We argued that the conditions of online communication—in particular on SNS—have changed people’s expectations and numerical understanding of audiences in online contexts and presented two scenario-based experiments to explore this matter. In both experiments, we varied whether the described interaction took place online or offline, and whether access was restricted (private) or unrestricted (public). To the best of our knowledge, this is the first study systematically assessing the question of the numerical meaning of vague quantifiers in the context of online and offline audiences. When using this theoretical and methodological approach as a framework to research people’s expectations of audiences, several interesting patterns occur:

First, both results from Experiments 1 and 2 suggest that regardless of communication mode there seems to be an upper limit restricting the number of people that can be present when a situation is considered as private. Corroborating this finding, there is a conceptual connection between the quantifier “few” and the context attribute of privacy. In that way, the public-by-default nature of online communication seems to not have changed people’s deeper conceptualization of what a private situation means as compared with a public one.

Second, while our findings do not support a literal change in the concept of privacy, they do support a change of the expectancy value, or else, the scalar representation: In online communication, more people are expected to be in the audience as opposed to offline communication, including online situations that are characterized by explicit access restriction. Although the absolute difference between the number of people expected in offline and online audiences is much larger for public situations, the relative difference is not. Thus, online situations in which the audience consists of a median of 29 people—the equivalent of an average school class—are considered to be private (offline private: 13 people). It is an interesting question how much change in scalar representation is possible without the concept itself changing its meaning.

Third, we need to consider the finding from Experiment 2, showing that appropriateness ratings were systematically higher for offline than for online situations. Since this was true for all value conditions (including ambiguous and contrast values), it seems that numeric expectations for offline contexts are more gradable than online ones, especially when they are public. This result seems counterintuitive, because audiences in online situations are typically much more difficult to predict and the audience size of public online information is potentially limitless. This should actually foster a more gradable scalar concept of audiences in online contexts. After all, the limitless of the potential audience leaves room for plausible explanations of why a public message does not have a large audience (in Experiment 2 this would be the case in the contrast conditions). For instance, people may believe that access to public information is restricted by the limited reading capacity of the potential audience (Lundblad, 2004; Moll et al., 2017).

However, our results require a different explanation. The higher flexibility in public offline as opposed to public online contexts may derive from the fact that in online situations people usually have some sort of—inaccurate—metric indicating the size of the audience, for instance, the number of comments or Likes beneath a post. This metric, however, is completely absent in offline contexts. Thus, people will rarely have any evidence relating to the number of people who may have seen a poster on an advertising pillar. Therefore, their assignments of quantifiers to numbers is based purely on probability instead of experience. This would explain higher appropriateness ratings even for ambiguous and contrast values and thus higher gradability. Supporting this argument, the exceptions to this pattern were private situations. Here, ambiguous and contrast values were not rated as systematically more appropriate in offline than in online contexts. This makes perfect sense, because for private offline situations, people do have an experiential metric for the number of people in the audience. After all, in private situations the audience is usually fully controlled by the sender of the message (for instance, when someone sends invitations to a party via the postal service). In that way, although there are less boundaries and thus less actual certainty in online communication, the scalar concept of audiences is less gradable in online as compared with offline communication.

**Limitations**

The interpretation of our results has several limitations. First, our results are based on rather small convenience samples. As such, we cannot rule out the possibility of beta errors; moreover, generalizations to other populations may be limited. At the same time, participants in Experiment 2 found the medians created by participants from Experiment 1 highly appropriate although the two samples differed in their demographic properties, such as their age. This implies that the overall pattern of results is likely to occur in other samples as well.

Second, the short vignettes describing the different contextual parameters in both Experiments 1 and 2 may in some cases differ in more respects than intended by the experimental factors. This problem is inherent in the research question and hypotheses posed in the beginning, because online contexts simply are very different from offline ones. In that way, it might not actually be possible to describe an offline situation that only differs from an online situation in the described respects.
Third, it can be argued that the interpretation of quantifiers is not only based on context, but also on the individual experiences people have made in these contexts. Thus, due to cross-linguistic and cultural differences, individual usage patterns or privacy preferences, people may be used to very different audience sizes in online communication. While this leaves room for future research, usage patterns and privacy preferences do not necessarily confound our findings: First, since people rarely act in accordance with their privacy preferences, these should not systematically lead to different experiences online (so-called privacy paradox; for a critical review, see Solove, 2020). Second, people in a similar age group with roughly similar usage patterns (which are implied by our sample description) should—regardless of their individual experiences—have shared knowledge of the most basic conditions of online communication, and should thus at least roughly agree on the meaning of “many” and “few.” This notion is supported by results from Experiment 2, where distracters are systematically rated as less appropriate than targets.

Conclusion and Future Research

The present study shows that people’s understanding of online audiences differs from offline ones in that the former has increased people’s expectation of larger audiences. This shift in scalar concept can also be observed for situations that are considered to be private. We might argue that this shift is a simple reaction to the reality of online communication, specifically on SNS, which was the context mostly underlying our rationale. However, a shift in scalar concept also implies a less clearly determined concept of a large audience which might then not even be considered as a risk-inducing variable. This, in turn, could lead to a further distortion in users’ risk perception and, ultimately, influence their decision to share information online.

Future research could, on the one hand, explore the influence of users’ individual experiences on their expectancy values (see section “Limitations”). On the other hand, it would be interesting to assess the exact scalar limits that people apply in different situations: Which contextual attributes are necessary for a message to still be validly called “private”? Is it more important how many people can potentially read a message, or rather, how many people have actually done so? This line of research would lead to a clearer picture of how people understand the concept of privacy and whether the public-by-default nature of online environments has changed crucial parameters of this concept. In this regard, future research could assess, in how far quantifiers in online and offline events are actually mapped onto one and the same scale, or whether online contexts represent their own ontological category of communication. This again would imply that the concept of privacy itself is not changing per se, but can simply not be applied to online contexts.

Both interpretations may become particularly relevant with regard to online users’ risk perception when disclosing personal information via SNS: When they judge an actually large audience to be “not that large” in that particular context, their risk perception may be constantly lowered—which may have consequences for the way users regulate their privacy.

In that way, our study also has implications for educational practice in the context of digital literacy: When explaining why it might be harmful to disclose personal information in an SNS context, scholars and practitioners should be aware that their educational attempts may fail, because younger and potentially more experienced SNS users have a different conceptual understanding of audiences in online contexts. Thus, digital literacy education aiming to increase users’ awareness of potential privacy problems when communicating online could—as one part of a training—discuss people’s conceptual understanding of audiences, audience sizes, and in how far consequences of online publicity may easily affect one’s “offline” life. One possibility to enhance users’ risk awareness thereby directly relates to the numerical representation of audiences. Thus, providers should, whenever possible, include displays of actual audience sizes to facilitate individual privacy regulation.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the research training group 1712/01 “Trust and Communication in a Digitized World” of the German Research Foundation (Deutscher Forschungsgemeinschaft, DFG).

ORCID iDs

Ricarda Moll https://orcid.org/0000-0002-0714-9102
Anna Jori Lücke https://orcid.org/0000-0003-0596-0644

References

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In EC ’04: Proceedings of the 5th ACM conference on Electronic commerce (pp. 21–29). https://doi.org/10.1145/988772.988777
Bernstein, M. S., Bakshy, E., Burke, M., Karrer, B., & Park, M. (2013). Quantifying the invisible audience in social networks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 21–30). https://doi.org/10.1145/2470654.2470658
boyd, d. (2008). Taken out of context—American teen sociality in networked publics. University of California, Berkeley. http://www.danah.org/papers/TakenOutOfContext.pdf
