

Anonymity Interacting with Participation on a Q&A site

Malte Paskuda, Myriam Lewkowicz

► **To cite this version:**

Malte Paskuda, Myriam Lewkowicz. Anonymity Interacting with Participation on a Q&A site. *AI & Society: Knowledge, Culture and Communication*, Springer, 2017, 10.1007/s00146-016-0660-9 . hal-02310989

HAL Id: hal-02310989

<https://hal-utt.archives-ouvertes.fr/hal-02310989>

Submitted on 10 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Anonymity Interacting with Participation on a Q&A site

Malte Paskuda · Myriam Lewkowicz

Abstract This article presents a study that investigates how anonymity influences user participation in an online question-and-answer platform (Quora¹). The study is one step in identifying hypotheses that can be used to address a research and design issue concerning the role of anonymity in online participation, particularly in sensitive situations where people are seeking social support. Based on the literature, we present a model that describes the factors that influence participation. These factors were used when analyzing the answers to questions in the health category on Quora. The results of this study were completed by a survey asking Quora users about their use of the anonymity feature. The main result is that the only significant difference between anonymous and non-anonymous answers is that: with anonymous answers, social appreciation correlated with the answer's length.

Keywords Anonymity · Online participation · Online social support · Quora

M. Paskuda (✉) · M. Lewkowicz
ICD, HETIC, Tech-CICO, Troyes University of Technology, UMR 6281, CNRS, 12 Rue Marie Curie - CS 42060, 10004 Troyes Cedex - France
e-mail: malte.paskuda@utt.fr

Myriam Lewkowicz
e-mail: myriam.lewkowicz@utt.fr

¹<https://www.quora.com/>

1 Introduction

Our research is dealing with the role that anonymity could play in the process of (online) community building. In fact, this question intrigued us because there is no easy answer about the identity model that any community platform should allow. On the one hand, user anonymity can diminish credibility (Rains 2007), and anonymity could make it more difficult to organize people who live in the same area to later help each other offline. On the other hand, it could make it easier for community members to talk about the more intimate details of their life. In consequence, we focus on both sides of anonymity: being anonymous and interacting with anonymous people.

We took the opportunity of an ambient assisted living project called TOPIC² (The Online Platform for Informal Caregivers) to tackle this research question because this project aims to build an online platform for social support among elderly informal caregivers (people providing help to their ailing relatives). Informal caregivers would like to share their situations with others, discuss them, and obtain advice and help (Yamashita et al. 2013). However, they do not want to speak ill of their loved ones, and they do not want to fall under the stigma attached to those illnesses. Designing adapted online platforms seems to be a potential solution to this problem, especially for informal caregivers who do not use existing offline support groups and respite care services (Tixier and Lewkowicz 2015). There are different reasons for not using the existing offline services: some caregivers are not aware of their existence, some cannot organize to attend the support groups, or some do not dare to talk openly about their emotions and mixed feelings (which is especially true for older husbands who are caregivers for their wives).

For such a platform, it is crucial that people take part. The more they participate, the higher the chance is that

²<http://www.topic-aal.eu/>

they will receive support. Experience sharing is especially important because it is a key factor in giving and getting social support (Tixier et al. 2009). However, it is difficult to know beforehand the critical factors that influence participation, experience sharing and general user satisfaction with online collaborative systems. Moreover, older caregivers are the target group of the TOPIC platform, but they use the web or even just computers less often than others (Pew Research Center 2014). An online platform for older informal caregivers must not only overcome this but also ensure trust. First, the caregivers need to view the information on the platform as credible. Second, they need to feel comfortable discussing something as private as life with an ailing relative. To sum up, our goal is that such a platform will become more than a place to find information. Members should discuss with each other and help each other by sharing their experiences and giving advice. In short, this place should become a home for a community of informal caregivers (one as described in Wellman and Gulia (1999)).

In this context, our research question is then as follows: what is the influence of anonymity on user participation in general and experience sharing in particular? Specifically, what would be the best option when designing a tool for a community such as the one wanted in the TOPIC project (elderly people who are experiencing stressful situations)? Should users be able to choose their own names, stay completely anonymous, or use their real names?

To address this research and design question, the first goal is to define our expectations for the user identity model of the TOPIC platform. *Anonymous users will participate more in this social support platform* would be an example. To test this type of hypothesis, we defined the following process:

1. Conduct a twofold review of the literature in the field of Computer-Mediated Communication, Human-Computer Interaction and Computer-Supported Cooperative Work, to first define the factors that influence participation and then to search for factors that anonymity influences.
2. By looking at the overlapping factors, select the ones to be tested. It will be necessary to determine how to identify the factors in a text (here: online messages).
3. Collect messages on existing platforms with varying

degrees of anonymity and analyze them for the factors defined in the earlier step.

This process generates hypotheses about how anonymity influences the factors selected in step 2. The hypotheses will then be tested in an experiment that will test anonymity on the TOPIC platform.

In this paper, we extend the work we presented in Paskuda and Lewkowicz (2015), in which we have described how we performed these three steps on the Q&A platform Quora, which has a function for answering anonymously and is thus a mixed environment. This work follows the results of a first study conducted in 2014 on YouTube in which we looked at YouTube's move to Google+; in 2013, Google moved YouTube's comment system from the integrated one that allowed pseudonyms (and thereby a degree of anonymity) to Google+. This move gave us the chance to study the effect of changes in the degree of anonymity.

We present here our analysis of the answers posted on Quora, looking for the differences between anonymous and non-anonymous Quorans. We have categorized the questions into subcategories to look if these potential differences are the same whatever the topic. In order to get a broader understanding of the use of the anonymity function, we completed our observation by an online survey in which we asked Quora users for their impression and usage of this anonymity function.

In the following sections, we describe the related work on the influence of anonymity on participation and we present the model of participation factors that we built to generate hypotheses about our research question and that we used in our studies. We then go on to present and discuss our findings, explain the limitations of this study, and, finally, conclude.

2 Related Work

In this section, we present the related literature about the effect of anonymity. We feature preferentially the effects on online interaction. However, we begin with the two fundamental theories about anonymity in general.

According to the deindividuation theory, a member in a group loses his self-awareness and thus loses his social conscience (Postmes et al. 2002). He can behave in ways he would normally never allow himself to behave because

he is no longer bound by the norms of general society. In the framework of this theory, anonymity could fortify or enable this loss of control. In practice, the tone of the communication becomes less polite. Applying this theory to online interaction was coined *the online disinhibition effect* (Suler 2004).

The reduced social cues approach is another perspective on this theory. In that model, electronic means do not transport the social elements used in face-to-face conversations, which leads to the deindividuation and deregulation of behavior (Sproull and Kiesler 1986).

The social identity model of deindividuation effects (SIDE) regards group behavior more positively, insofar as the actions taken by the members are in line with what is good for the group as a whole because group members identify with the group norms (Cress 2005). Anonymity is then twofold: are the others known, and do the others know me? If the others are anonymous to the member, this increases the likelihood of identifying with the group because individuality—which could hinder that identification—stays hidden. In practice, a user might feel more connected with the group (and this should improve the user experience).

Following these theories, there have been a number of studies on anonymity online. Kilner et al. conducted an important analysis about an online forum for soldiers that gradually changed its account model from anonymity with pseudonyms to asking for the full civil identity (Kilner and Hoadley 2005). Kilner et al. analyzed the comments in the different stages and found that removing anonymity options led to fewer antisocial comments and fewer comments in total. This work heavily influenced our analysis in selecting possible hypotheses.

Another influence comes from a study on the move of the tech site TechCrunch from Disqus to Facebook as a comment system (Omernick and Sood 2013), thereby disabling the option to comment anonymously or under a pseudonym. By comparing comments from before and after the change, Omernick and Sood found evidence for a negative influence of anonymity on comment quality and politeness (what Kilner and Hoadley (2005) would have classified as antisocial, thus underlining this result). However, those changes did not result in a significant decrease in participation; there were fewer comments, but they were longer.

In the area of behavioral science, a number of exper-

iments (such as Lapidot-Lefler and Barak (2012)) have attempted to find effects of anonymity. Wise et al. (2006) measured the effect of having moderation features and changes in the response rate on the intent to participate. This experiment is noteworthy especially because it constructs a connection between interactivity and the reduced social cues model.

In computer science, theoretical models of anonymity and anonymity on a technical level play a more important role (Edman and Yener 2009), and research exists on the link among politeness, civility and anonymity, analyzing that link from a political angle (Santana 2012).

There is also a broad amount of literature describing the factors influence participation. Although anonymity is seldom the focus, it is occasionally mentioned. An example of that is at the same time a main thread in the literature: the common identity and bond theory used by Kraut et al., as described in Ren et al. (2007). The theory assumes two connections between the members of a community—identity and bond—that are influenced by different factors, *social categorization*, *interdependence* and *intergroup comparisons* for identity and *social interaction*, *personal information* and *personal attraction through similarity* for bond. Some of these can be linked to anonymity, and SIDE theory does that explicitly with *personal attraction through similarity*.

Although the research cited here paints a mostly negative picture of the effect of anonymity on online interaction, one should not forget its positive effects. An example is the broad range of different and justified motivations to stay anonymous (Kang et al. 2013). These effects are also deductible by following SIDE theory and noting the positive effects that are possible with easier group identification. The negative results, such as more rude comments in the case of available anonymity, do not unanimously show that this leads to less participation. Additionally, our own study of YouTube's change away from anonymous comments found more rude comments after the change, thus contradicting this specific expectation.

Those contradictions show that we are not able to simply use existing literature to paint a clear picture of the effect of anonymity.

We needed to build a model to show the factors that we found in the literature and extract those factors to use in our studies. The building of the model and how we plan to use it are described in the following section.

3 Building a Model to Design Studies

The current section describes the background of the study that we conducted and explains the choice of measured factors.

In line with Wise et al. (2006), we define online participation as posting contributions into the online environment. On Quora, it means asking or answering questions. Upvoting and downvoting as a peripheral form of participation is invisible to us, and a form of participation sliding into the lurker research Nonnecke and Preece (2001).

To see the effect of anonymity on participation, we needed a model of factors that influence participation. In fact, we assumed that anonymity cannot be the only factor that influences participation. We claim, supported by the literature, that anonymity influences the factors that influence participation. This is the reason the first step in generating our model was to review the existing work to find general factors that influence participation. After having identified these factors, we made a second literature review focusing on the effects of anonymity.

3.1 Factors influencing Participation

We found many factors that might influence participation.

3.1.1 Anonymity itself

In Kilner and Hoadley (2005), a platform moved stepwise from anonymous participation possibilities to making it necessary to disclose one's full civil identity. Kilner et al. observed that although many of the metrics that measured participation did not change, what did change was the number of comments posted.

3.1.2 Common identity

In Ren et al. (2007), it was argued that attachment to the group influences participation. Kraut et al. mentioned two theories to explain that attachment: common identity and common bond.

Common identity theory makes predictions about the causes and consequences of people's attachment to the group as a whole (Ren et al. 2007, p. 377).

The authors also highlighted a number of factors that might help achieve a group identity and therefore foster participation:

Social Categorisation. Simply declaring that people are in a common group based on arbitrary criteria.

Interdependence. Being dependent on others members to achieve a common goal or by a shared fate. An example is described in Ling et al. (2005). In that experiment, users contributed more work when they were told that their work was unique and thus necessary to achieving the group's goal.

Intergroup Comparisons. Comparing group members with other groups.

3.1.3 Common Bond

Its definition:

Common bond theory makes predictions about the causes and consequences of people's attachment to individual group members (Ren et al. 2007, p. 377).

The theory highlights the following factors (Ren et al. 2007):

Social Interaction. Chmiel et al. (2011) described that the expression of negative emotions led to high interactions between users and to a high amount of participation in threads on the BBC forum. Harper et al. (2007) also showed the success of personalized invitations, stressing the social aspect of a forum. However, this success was not observed in Sharma et al. (2011); on the contrary, social aspects in invitations led to fewer registrations with less filled profiles. In van Oostendorp and van Varik (2011), interaction was rather observed as a metric of a successful community. Tausczik et al. (2014) described that depending on the community, direct interaction can be necessary to have an effect, in contrast to only creating social awareness.

The different studies show how difficult it is to distinguish between cause and effect in this area.

Personal Information. Opportunities for self-disclosure. For example, in van Oostendorp and van Varik (2011), the option to have an avatar photo seemed to increase the number of messages and forum threads created.

Personal Attraction through Similarity. Because people like people they share similarities with.

The factors linked to common bond and common identity profit from a design that enables *social presence*, which is defined as making other users visible while using a system. An example would be integrating teams with visible and shared progress (Farzan et al. 2011).

3.1.4 Other Factors

Many other factors exist that are not easy to integrate into common identity and common bond theory.

First to mention are the factors of the *community activity framework* (van Oostendorp and van Varik 2011). Some factors are content based, such as the use of graphical emotions in posts and having rules to guide group discussions. Others are more functional, such as email notifications and posting counts next to posts or related news sections.

A number of authors describe *politeness* as a factor. Burke and Kraut (2008) analyzed a small sample of messages from discussion groups and measured their perceived politeness with an Internet survey. Polite messages received three times more replies in technical groups, but impolite messages received more replies in political groups.

Introductions and *requests* are rhetorical strategies that are analyzed for their impact on responses in Burke et al. (2007). There, they increased the likelihood of replies by 7% and 6%. However, other rhetorical features such as the use of *self-references* also elicited responses:

Posts that included testimonials or requests were more likely to receive a reply. Including self-references (“I”), third-person pronouns, describing cognitive states and processes, and expressing either positive or negative emotions all increased the likelihood that a message received a response. The topical coherence of a message with respect to other recent discussions in the

community also affected the likelihood of getting a reply (Arguello et al. 2006, p. 959).

Timely and positive *feedback* increases the effort put into the task at hand or the general motivation (Zhu et al. 2013). Strong negative feedback decreases motivation to participate (Zhu et al. 2012).

The *behaviour of the founder* of a group can influence its chance of success. For example, groups founded by very controlling people die early (Kraut and Fiore 2014). Similarly, in the context of a learning community, the number of prompts in the course material to answer questions by the organizers led to greater learner participation (Ahn et al. 2013).

How to activate already present members in a community who do not actively participate (lurkers) receives special attention in the literature. Preece and Shneiderman (2009) presented some factors that could activate these members, and these factors were divided into the categories usability and sociability. In its essence, it follows the thought that *easy access to the means to contribute* and *social appreciation of the contribution* will activate lurkers. In contrast, inactive users have a variety of reasons to remain inactive, including privacy concerns (Nonnecke and Preece 2001), and lurking is sometimes simply viewed as a metric that shows that the community does not fit the nonparticipants (Preece et al. 2004). Nonetheless, lurkers are sometimes considered a strong negative factor for the survival of online communities, as in Shiue et al. (2010). Shiue et al. (2010) also proposed that *perceived risks* and *social ties* are sufficient to explain lurking behavior.

3.2 Influence of Anonymity

It is a common thought that anonymity can change situations and that it influences various factors.

A survey of 44 people on the Internet with various backgrounds focused on the self-perceived merits of being anonymous. One such merit is the *emotional benefit*. Additionally, the participants perceived anonymity as something that enables more honest ratings or recommendations (Kang et al. 2013).

Thus, anonymity appears to also influence *Credibility*. Although there are theories in both directions—both more and less credible—Rains (2007) observed that perceived anonymity decreased credibility.

Conformity in a group situation appears to be at least minimally affected by perceived anonymity (Tsikerdekis 2013).

Levmore et al. (2010) attributed *Uncivility* and *Impoliteness* to anonymity.

Shiue et al. (2010) stated that anonymity will result in *stronger social ties*, thus minimizing lurking behavior.

In contrast, Chen et al. (2009) suggested that anonymity leads to more *antisocial behavior* in the context of grieving in online games. This phenomenon had already been mentioned in Kilner and Hoadley (2005), where the removal of anonymity options led to fewer *antisocial comments*.

However, in Kilner and Hoadley (2005) it was also measured whether the change had an *effect on participation*, and the authors found less direct participation but the same number of logins and page views.

An experiment with two groups tested the differences between groups with anonymous and identified members. They found the following:

Depersonalization was associated with greater attitude differentiation than individuation was (Postmes et al. 2002, p. 11).

Participants also *identified more strongly* with their own groups. In a similar vein, as a fitting summary Bodle (2013) states,

The attributes of anonymity, including minimal accountability, disinhibition, and deindividuation, can encourage robust political speech, provide safety from reprisal, permit the freedom to speak freely, and create a strong sense of group identity (Bodle 2013, p. 30).

3.3 The Model

From the factors that influence participation and the factors that are influenced by anonymity, we kept the intersecting factors, i.e., the ones influenced by anonymity that influence participation. Of those, we kept the factors for which we were able to find visible markers in text (see fig. 1).

We observed that a large part of the literature assumes that anonymity influences politeness (see Levmore et al. (2010)). Politeness appears to influence participation, and

Chmiel et al. (2011) showed that impolite comments provoked other comments.

The relationship between anonymity and intergroup comparisons and social interaction is indirect via social presence. Farzan et al. (2011) described that the factors linked to common bond and common identity could both profit from social presence, and Tu (2002) showed that anonymity influences social presence.

Anonymity can change the perception of contributions and can lead to less social appreciation (Rains 2007). However, social appreciation and specific types of feedback foster participation (Preece and Shneiderman 2009; Zhu et al. 2012).

To synthesize the links that we identified in the literature between these factors, we built a model (see fig. 1). It serves as an anchor in our approach to quantifying the effect of anonymity on interaction in an online community. If the model's factors are valid in influencing participation and anonymity truly influences these factors, we can use that model—or, rather, measure markers of the model's factors—to search for differences in anonymous versus non-anonymous contributions. If these differences exist, we can assume that there will be a difference if we allow anonymity in the participation on the TOPIC platform, or in general. If there are no differences, then anonymity should not significantly influence participation in our community.

3.4 Using the Model

One place that appeared to be fit for testing the model was YouTube. This is a place where users interact that hosts videos related to the topic of informal caregiving. It also provides a chance to study a change in anonymity because YouTube moved from an internal comment system with pseudonyms to Google+, where users are required to give their full civil names. To capture the effect of moving away a degree of anonymity, we collected comments from before and after the change from a number of videos related to informal caregiving and Alzheimer's disease (because most of the future users in the city where we are conducting our fieldwork are informal caregivers of people suffering from this disease). We then searched for signs of changes in the model's factors.

As a result, we found a degree of evidence for a link between anonymity and politeness, although it was a pos-

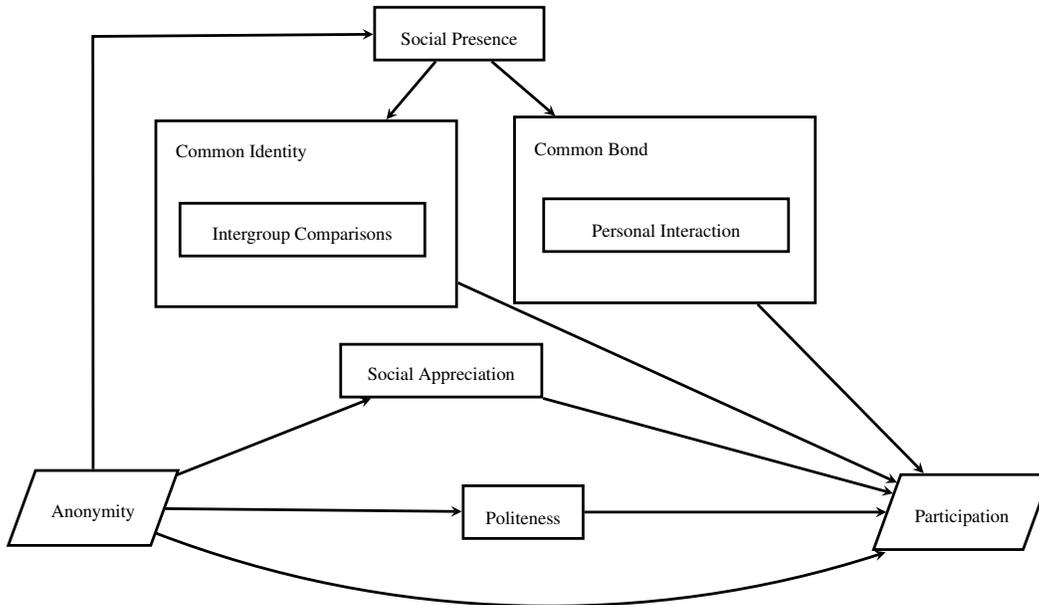


Figure 1: Simplified model showing anonymity interacting with participation

itive one: anonymous comments were more often polite. We also observed a strong link between anonymity and less personal interaction but no significant influence of anonymity on the use of intergroup comparisons.

The YouTube study did not answer the research question adequately. Its conclusion of an increase in rude comments after moving away from anonymity collides with the literature, which led us to expect a positive influence on politeness. The study was of a move from one environment to another, which made it difficult to pinpoint anonymity as the deciding factor in the change. To minimize such other influences, it seemed to be a good idea to repeat a similar study in an environment that by default mixes anonymous and non-anonymous user participation, such as Quora. On Quora, anonymous responses can be given by regular users, who must still be logged in, by toggling the anonymous mode before sending.

Thus, the study on Quora that we describe in the next section had multiple goals:

1. It was another opportunity to apply the model and determine whether the factors' markers are detectable and whether it is possible to identify any dif-

ferences.

2. It was a control for the results of the YouTube study. The generated hypotheses would be stronger if they applied here, and they would be weakened if they did not hold up. Because the final experiment required hypotheses to be possibly disproved, this study helped in choosing them.
3. The fact that Quora already allows anonymity rather than changing its anonymity conditions, provided us with the chance to study the effect of anonymity without having to fear that other changes would influence the measured factors

In a direct comparison with the YouTube study, we were able to see whether the change in the level of politeness would be the same as it was on YouTube, where fewer pseudonymous comments were rude, contradicting the literature. The number of intergroup comparisons between the groups (anonymous and not) could have been the same, as on YouTube, or it could have differed this time. The number of personal interactions with anony-

mous authors—replies to answers or comments—could have been lower as well.

However, there was also the chance to generate new hypotheses given that we were able to include social appreciation through the number of upvotes (which was not possible on YouTube because there were no parsable upvotes). An upvote on Quora usually means that the person who is upvoting an answer believes that the question was properly answered and that the answer contributed in a meaningful way to Quora’s repository of knowledge. However, many user base their upvotes on their own interpretations of what an upvote should represent. Quora’s administration provided guidelines for upvoting, but it appears that many users have their own personal upvoting philosophies³.

We present the findings in the second next section, but social appreciation proved to be an important metric, which hints at a measurable influence of the reduced number of social cues provided by anonymity.

4 Data Collection and Analysis

We collected data in two distinct steps.

4.1 Collection of online data

We selected 375 questions and their 4765 answers, 288 anonymous, by taking the popular questions in the health category (fig. 2). Only questions without answers or those that merged with multiple other question threads were skipped. The obtained HTML was then parsed, and the generated data were saved in a database. No demographic information about users were collected.

The questions were then categorized by two members of the research team as belonging to one of the categories *medical*, *lifestyle*, *joke* and *other*. Those categories were chosen after looking at the types of questions in the dataset.

1. Medical questions were defined as relating to a “real” medical health question, like how to react to a cancer diagnosis.
2. Lifestyle questions were the less serious ones like how to stay fit. They were the majority.

³<http://goo.gl/qf6hBj>

Depression: What is the diagnosis process like for someone to be diagnosed with depression?

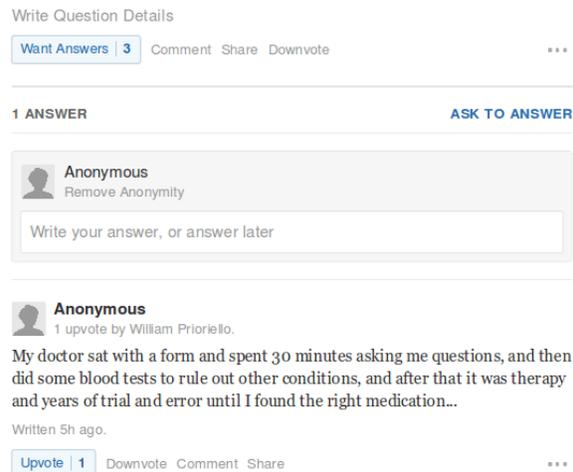


Figure 2: Example of an anonymous question and answer on Quora, <http://goo.gl/md4WJ3>

3. Joke questions are those with either a clear humorous intent or asking about a curiosity, like the idea that one could die from taking a cold shower during the day. However, there were not many questions in this category, those questions were thus not analysed separately.
4. The remaining questions were those with no relation to health.

Only the questions where the two researchers independently agreed upon the category were used in the analysis; 3002 answers remained, of which 148 were anonymous.

The data were analyzed with a number of scripts, in particular, calling a Bayes classifier⁴ and a statistic toolkit⁵.

We searched for a number of factor markers from our participation model:

1. *Politeness*. How polite the message was. We used a Bayes classifier to attempt to categorize the answers

⁴<https://github.com/jeky11/classifier-reborn>

⁵<https://github.com/clbustos/statsample>

into the categories *polite*, *neutral* and *rude*. This was based on reports that algorithmic approaches can work acceptably well for detecting politeness (Wild and Stahl 2007) and our own good experience with the method in the YouTube study.

2. *Intergroup Comparisons*. We searched for the words "we/us/our/them", which show that a group of people is being mentioned (Bramley 2001, p. 86). The idea behind this search is: If people think there are different groups, they have to use pronouns to describe these groups. This will happen when they feel being part of one group, one community, with other groups at the outside, which they are not part of, or which are at least distinct to the current group⁶. In the model, anonymity influences the use of intergroup comparisons through *social awareness*.
3. *Personal Interaction*. To approximate personal interaction, we used the number of comments on an answer. In the model, anonymity influences this through *social awareness*.
4. *Social Appreciation*. The number of upvotes reflected this.

The scripts used and the generated database are available at <https://goo.gl/GRJ6At>.

4.2 Survey

In order to get a broader understanding of the use of the anonymity function, we decided to complete the data collected on the platform by an online survey (see table 5).

Through 12 questions this survey asked Quora users to provide their opinion and impression of the anonymity function and its use. This survey was online for several weeks, but feedback only arrived as long it was linked in a Quora-question on the first two days.

The survey was answered by seven active Quora users that visit Quora every day of the week. Five of them used the anonymity function. These responses are discussed in section Survey and compared to the prior results.

⁶Also see <http://selp.eu/lexique/pronoms-personnels-2>

5 Findings

We analyzed the answers, which means that we searched for significant differences in the selected markers between anonymously and non-anonymously posted answers, overall and again in their respective category. We used multiple t-tests, which we realize increases the chance of having made a Type I error. But we did not see many significant differences, and those we saw hold up against error corrections. The result mainly showed that the two groups did not differ greatly, with one noteworthy exception described in section Social Appreciation.

5.1 Politeness

It was not possible for us to algorithmically analyze the answers for politeness as we had done previously. The algorithm failed to distinguish among the three categories (polite, rude, neutral), categorizing nearly all answers as either all rude or all polite while almost ignoring the much more fitting neutral category. This was a surprise given that the same software and workflow were used in the earlier YouTube study, in which we found 80% accuracy.

A manual examination showed that except for literally one answer, all of them followed a specific tone that appeared to be common on Quora. That is not to say that all answers were equal; there was a great range of quality and length in the sample data. Many responses were factual, and others were filled with pathos, but they all lacked easily distinguishable indicators of politeness. These were present in the comments on YouTube, where it was easy to categorize an insult as rude and many best wishes as polite. We assume that this impression is correct and that there was no difference between the groups. We discuss this further in section Limitations.

5.2 Intergroup Comparisons

There was only a small difference in the number of intergroup comparisons made by anonymous and non-anonymous users (fig. 3, table 1).

The difference was significant with a t-test ($p < 0.05$), which suggests the conclusion that the preference to post anonymously on an otherwise non-anonymous platform does influence identification with the group negatively, at least on Quora. The result of $p = 0.01$ also holds up against a Bonferroni Correction of $\alpha = 0.05/4 = 0.0125$,

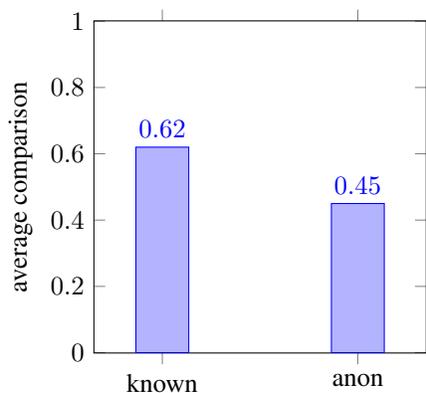


Figure 3: Difference comparisons made

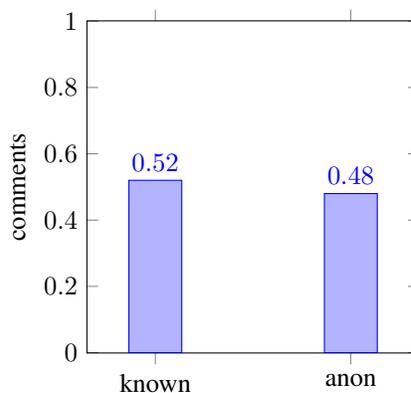


Figure 4: Average amount of received comments

with $m = 4$ for the four initial comparisons without sub-categories. It should however be noted that it would not hold up against a correction with $m = 20$, when including the tests against subcategories.

But the difference was significant only overall, not in category of medical questions, which is an important category for our research context (informal caregivers sharing experience on the care they provide to their sick relative). It is speculation, but if the seriousness of the question counteracts the negative effect, then we could ignore the negative effect when designing a system for informal caregivers.

Table 1: Amount of comparisons

Group	mean	sd	n	category
Known*	0.62	1.97	4477	all
Anonymous*	0.45	0.98	288	
Known**	0.52	1.84	2854	categorized
Anonymous**	0.28	0.70	148	
Known	0.43	1.39	185	medical
Anonymous	0.36	0.91	25	
Known	0.40	1.34	1550	lifestyle
Anonymous	0.24	0.67	78	
Known	0.53	2.22	927	other
Anonymous	0.33	0.66	40	

5.3 Personal Interaction

A t-test showed no significant difference in the number of comments received for the two groups (fig. 4, table 2). In general, comments to answers are not overly common on the platform. In fact, Quora uses elements of the User Interface to not highlight them: they are not visible by default, and they can be made visible only by clicking on a small grey-colored link. As such, an average of roughly one comment for every second answer was already unexpectedly high.

Table 2: Received comments

Group	mean	sd	n	category
Known	0.52	4.45	4477	all
Anonymous	0.48	1.82	288	
Known	0.44	3.55	2854	categorized
Anonymous	0.53	1.71	148	
Known	0.15	0.43	185	medical
Anonymous	0.24	0.52	25	
Known	0.23	1.43	1550	lifestyle
Anonymous	0.46	1.57	78	
Known	0.89	5.91	927	other
Anonymous	0.9	2.39	40	

We should note here that the amount of personal interaction that occurred through direct messages was invisible to us (see also section Limitations).

The non significance of the difference between the

groups was also unexpected. We observed a high increase in personal interaction in the YouTube study after the change to less anonymous comments. The fact that there was no difference here suggests another explanation: that the change on YouTube was not caused by the change in anonymity but by the change in the comment User Interface and the link with the Google+ social network .

5.4 Social Appreciation

The number of received upvotes did not differ significantly between anonymous and non-anonymous questions. However, anonymous answers also received less feedback (fig. 5, table 3).

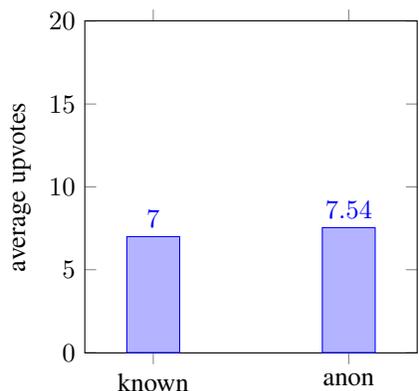


Figure 5: Average upvotes received

There was sufficient literature with different results and theories to expect a stronger difference. In particular, Rains (2007) led us to expect that anonymous answers would be less appreciated. In that study, an experiment measured the assigned persuasiveness of responses that were either linked to a participant or were anonymous. These anonymous responses were viewed as being inferior to the non-anonymous ones, as less trustworthy and less persuasive. We expected that the same would happen here, i.e., that the anonymous responses would receive fewer upvotes.

However, another factor that was measured was the length of the answer, which by itself was not significant (table 4). But the difference between the two groups was significant in a t-test with $p < 0.01$ when looking only at the answers in the lifestyle category. The result of

Table 3: Upvotes

Group	mean	sd	n	category
Known	7	41.36	4477	all
Anonymous	7.54	60.14	288	
Known	6.6	41.76	2854	categorized
Anonymous	6.58	42.63	148	
Known	1.57	3.24	185	medical
Anonymous	1.6	1.44	25	
Known	2.68	26.69	1550	lifestyle
Anonymous	3.83	22.08	78	
Known	14.39	60.53	927	other
Anonymous	15.7	75.94	40	

$p = 0.0013$ also holds up against a Bonferroni Correction of $\alpha = 0.05/20 = 0.0025$, with $m = 20$ for all significance tests made.

Anonymous users wrote significantly shorter answers to this kind of question.

Table 4: Answer length

Group	mean	sd	n	category
Known	712.3	1025.6	4477	all
Anonymous	743.23	1280.83	288	
Known	678.5	987.72	2854	categorized
Anonymous	747.74	1559.52	148	
Known	666.65	685.98	185	medical
Anonymous	943.48	1243.59	25	
Known**	743.28	911.38	1550	lifestyle
Anonymous**	521.06	557.41	78	
Known	574.82	1190.51	927	other
Anonymous	1033.03	2656.04	40	

Using a Pearson correlation, we found a positive correlation $r = 0.384$ with 286 degrees of freedom ($p < 0.01$) between the length of the answer and the number of upvotes, but only for anonymous answers (fig. 6).

This means that for anonymous answers only, the number of upvotes increased with the length of the answer. This is surprising given that that correlation did not exist for the other answers.

It seems plausible to expect that in general, longer answers will receive more upvotes on Quora. They take longer to write, they can contain more relevant informa-

tion, and they show that an effort was made. The fact that there was no correlation between an answer's length and its number of upvotes shows that this is not the case. Perhaps longer answers are more cumbersome to read, or perhaps short answers better convey the needed information to answer a typical Quora question.

However, as soon as answers were anonymous, the correlation became significant: Longer answers by anonymous users received significantly more upvotes. Why is that?

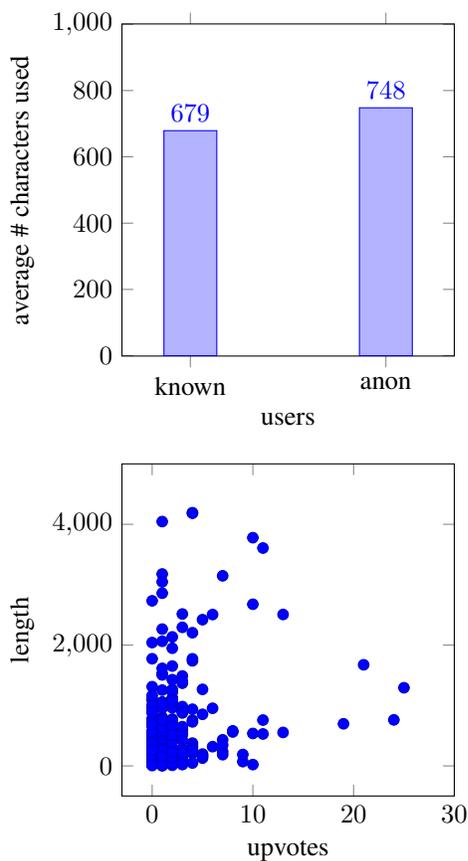


Figure 6: Correlations of answers length and upvotes for anonymous users

A possible explanation is that without the added social cues provided by username and avatar image, readers resorted to comment length as a signal of comment quality.

We base that hypothetical explanation on the assump-

tions of the reduced social cues approach, as exemplified in Wise et al. (2006). A social signal used in non-digital communication can in that model be replaced by a digital signal. Here, it could be the social dimension provided by username, attributability and avatar image that normally influences upvote behavior more than the answer's length. When those signals are missing, a number of things could happen. It could be that the answer's quality becomes more relevant and that quality is linked to length, or it could be that comment length works as a social signal for an answer's credibility, a function that was previously covered by the now hidden peripheral elements.

Other explanations are possible. One could transform the removal of social cues into a positive and argue that, as in Cress (2005), removing the avatar image leads people to identify more strongly with those than with anonymous members. Comment length would again be a more objective factor of comment quality, a factor that was previously covered by negative identification through identity elements such as username and password.

We will not uncover the cause here; the finding is only that the correlation exists.

5.5 Survey

The responses of the online survey give an interesting view of the impression of the anonymity function by its users. One user said:

I go anonymous when I'm revealing something that my family members wouldn't want other people to know about

This was the same sentiment of the other users, except one who said:

Out of all the answers I've written on Quora, I can count on one hand the number of answers I've written anonymously. I stand behind the words I write and do not hide behind the mask of anonymity. On rare occasion, I find that there are exceptions. From my point of view, there is rarely a need for it.

Given those statements, the anonymity function is used exactly for the obvious reason: To distance their civil identity from what is said. No secondary function – like not getting notifications for example – was described.

But how do those users see the quality of anonymous answers?

Table 5: Survey result

Questions	∅ Answer
Quora visits per week	7
Average Participation	3.9 (max: 6)
Followers	4110
Usage of Anonymity	5 (max: 7)
Own anonymous answers are ...	scale: 1 to 6
... polite	5.4
... get comments	3.2
... help community identification	2.2
... get upvotes	3.4
... allow answering some questions	5.2
General answer quality	4.5
Anonymous answer quality	3.7
I appreciate anonymous answers	4.1
I upvote anonymous answers	5.3
Anonymous answers are generally polite	3.1
Happy about getting anonymous answers	3.7

By average, the quality of the answers on Quora on a Likert-Scala (from 1 to 6) is rated as 4.5. The anonymous answers were rated as 3.7, clearly worse. They are also seen as impolite and half of the survey users do not appreciate getting anonymous answers.

This contrasts with the self-judgements of their own anonymous answers, which they rated to be very polite.

The statement *If I could not answer anonymously, I would not answer some questions at all* got a strong approval of 5.2.

Altogether, it can be said that while users see the anonymity function as useful, they see other anonymous answers as critical. Normal answers are preferred.

6 Limitations

Quora did not give us raw sample data, and we did not have access to an API. We collected our data manually and then parsed it with a handwritten parser. Thus, the answers that we collected had already been filtered by Quora’s moderation, with potential consequences for the validity of the question selection. If, for example, anonymous answers were in general less polite, that effect could

have been invisible to us if the rude comments had already been deleted.

Moreover, all of the information that is not available to the public, such as direct messages between members, was also invisible to us. Having this as metadata would have allowed us to measure personal interaction more accurately, and without those data, one should judge the personal interaction metric as an estimate.

Answers that are submitted anonymously are anonymous only to other users. Quora itself could know who made which answer because users must always be logged in to answer. The answers are, as such, not fully anonymous in the strictest sense because that would include anonymity to all possible observers.

There were not enough anonymous answers in the category "joke" to interpret the results in that category.

Finally, the survey was not filled out by enough users to rely on its result heavily.

7 Conclusion and Further Work

Quora provided us with the option to study a mixed environment of anonymous and non-anonymous user-generated texts. In contrast to the YouTube study we conducted earlier, there were fewer differences. On YouTube, the prominent change was in the level of politeness and the increase in social interaction. On Quora, there was only the difference in the correlation between answer length and upvotes for anonymous answers, and a hint for less identification with the community.

Previous work showed the following:

We know that people who have limited motivation to process content are more likely to base evaluations on peripheral cues (Wise et al. 2006, p. 33)

Something similar could have happened here: Quorans could have normally based their upvotes at least partly on the peripheral social cues provided by username and avatar and resorted to answer length as a relevant factor only when those social cues were not present.

Regarding the effect of anonymity on an online community, we understand the result as an argument for the harmlessness of anonymity. Anonymous answers were, in the eyes of the community in general, not worse, and they did not receive significantly fewer upvotes. Overall

they were not significantly shorter, which could be important for community builders. Moreover, in contrast to the expectations generated by Kilner and Hoadley (2005) and deindividuation theory in general, they were not less polite.

Comparing the results by looking at the different topics in the health category uncovered that the seriousness of the question made almost no difference on the perception of anonymous users. But it showed that anonymous users tend to write shorter answers to non-serious questions.

Future work should attempt to test these results in different communities, for example, those with less strict moderation and politeness expectations. Given that SIDE theory predicts an increase in group identity through anonymity, the effect of anonymity could be quite different in communities with different norms. Whether such a group identity for Quora users exists could also be a worthwhile question, and its answer could be useful for interpreting our results.

Acknowledgements This work has been supported by European Union, ANR and national solidarity fund for autonomy through AAL program (project AALI 2012-TOPIC).

We thank our colleagues Nadia Gauducheu and Michel Marcoccia for their help in building the model and identifying the markers, and our colleague Valentin Berthou for the help with categorizing the questions. We also thank the reviewer for his improvements.

References

- Ahn J, Weng C, and Butler B. S (2013). The dynamics of open, peer-to-peer learning: What factors influence participation in the p2p university? In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pages 3098–3107.
- Arguello J, Butler B. S, Joyce E, Kraut R, Ling K. S, Rosé C, and Wang X (2006). Talk to me: Foundations for successful individual-group interactions in online communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06*, pages 959–968, New York, NY, USA. ACM.
- Bodle R (2013). The ethics of online anonymity or zuckerberg vs. "moot". *SIGCAS Comput. Soc.*, 43(1):22–35.
- Bramley N (2001). *Pronouns of Politics: The Use of Pronouns in the Construction of 'self' and 'other' in Political Interviews*. Australian National University.
- Burke M, Joyce E, Kim T, Anand V, and Kraut R (2007). *Introductions and Requests: Rhetorical Strategies That Elicit Response in Online Communities*.
- Burke M and Kraut R (2008). Mind your Ps and Qs: the impact of politeness and rudeness in online communities. In *Conference on Computer Supported Cooperative Work*, pages 281–284.
- Chen V. H.-H, Duh H. B.-L, and Ng C. W (2009). Players who play to make others cry: The influence of anonymity and immersion. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology, ACE '09*, pages 341–344, New York, NY, USA. ACM.
- Chmiel A, Sobkowicz P, Sienkiewicz J, Paltoglou G, Buckley K, Thelwall M, and Hołyst J. A (2011). Negative emotions boost user activity at {BBC} forum. *Physica A: Statistical Mechanics and its Applications*, 390(16):2936 – 2944.
- Cress U (2005). Why member portraits can undermine participation. In *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years!*, CSCL '05, pages 86–90. International Society of the Learning Sciences.
- Edman M and Yener B (2009). On anonymity in an electronic society: A survey of anonymous communication systems. *ACM Comput. Surv.*, 42(1):5:1–5:35.
- Farzan R, Dabbish L. A, Kraut R. E, and Postmes T (2011). Increasing commitment to online communities by designing for social presence. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, CSCW '11*, pages 321–330, New York, NY, USA. ACM.
- Harper F. M, Frankowski D, Drenner S, Ren Y, Kiesler S, Terveen L, Kraut R, and Riedl J (2007). Talk

- amongst yourselves: Inviting users to participate in online conversations. In *Proceedings of the 12th International Conference on Intelligent User Interfaces, IUI '07*, pages 62–71, New York, NY, USA. ACM.
- Kang R, Brown S, and Kiesler S (2013). Why do people seek anonymity on the internet?: Informing policy and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13*, pages 2657–2666, New York, NY, USA. ACM.
- Kilner P. G and Hoadley C. M (2005). Anonymity options and professional participation in an online community of practice. In *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years!*, CSCL '05, pages 272–280. International Society of the Learning Sciences.
- Kraut R. E and Fiore A. T (2014). The role of founders in building online groups. In *CSCW*, pages 722–732.
- Lapidot-Leffler N and Barak A (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, 28(2):434 – 443.
- Levmore S, Levmore S, and Nussbaum M (2010). *THE OFFENSIVE INTERNET*. Harvard University Press.
- Ling K, Beenen G, Ludford P, Wang X, Chang K, Li X, Cosley D, Frankowski D, Terveen L, Rashid A. M, Resnick P, and Kraut R (2005). Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*, 10(4):00–00.
- Nonnecke B and Preece J (2001). Why Lurkers Lurk. In *AMCIS Conference*. Boston.
- Omernick E and Sood S. O (2013). The impact of anonymity in online communities. In *International Conference on Social Computing, SocialCom 2013, SocialCom/PASSAT/BigData/EconCom/BioMedCom 2013, Washington, DC, USA, 8-14 September, 2013*, pages 526–535.
- Paskuda M and Lewkowicz M (2015). Anonymous quorans are still quorans, just anonymous. In *Proceedings of the 7th International Conference on Communities and Technologies, C&T '15*, pages 9–18, New York, NY, USA. ACM.
- Pew Research Center (2014). The Web at 25 in the U.S.
- Postmes T, Spears R, and Lea M (2002). Intergroup differentiation in computer-mediated communication: Effects of depersonalization. *Group Dynamics: Theory, Research, and Practice*, 6(1):3.
- Preece J and Shneiderman B (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 1(1):13–32.
- Preece J. J, Nonnecke B, and Andrews D (2004). The top five reasons for lurking: improving community experiences for everyone. *Computers in Human Behavior*, 20(2):201–223.
- Rains S. A (2007). The impact of anonymity on perceptions of source credibility and influence in computer-mediated group communication: A test of two competing hypotheses. *Communication Research*, 34(1):100–125.
- Ren Y, Kraut R, and Kiesler S (2007). Applying common identity and bond theory to design of online communities. *Organization studies*, 28(3):377–408.
- Santana A (2012). Civility, anonymity and the breakdown of a new public sphere.
- Sharma N, Butler B. S, Irwin J, and Spallek H (2011). Emphasizing social features in information portals: Effects on new member engagement. *Journal of the American Society for Information Science and Technology*, 62(11):2106–2120.
- Shiue Y.-C, Chiu C.-M, and Chang C.-C (2010). Exploring and mitigating social loafing in online communities. *Comput. Hum. Behav.*, 26(4):768–777.
- Sproull L and Kiesler S (1986). Reducing social context cues: Electronic mail in organizational communication. *Management science*, 32(11):1492–1512.
- Suler J (2004). The online disinhibition effect. *Cyberpsychology & behavior*, 7(3):321–326.

- Tausczik Y. R, Dabbish L. A, and Kraut R. E (2014). Building loyalty to online communities through bond and identity-based attachment to sub-groups. In *CSCW*, pages 146–157.
- Tixier M, Gaglio G, and Lewkowicz M (2009). Translating social support practices into online services for family caregivers. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work, GROUP '09*, pages 71–80, New York, NY, USA. ACM.
- Tixier M and Lewkowicz M (2015). Looking for respite and support: Technological opportunities for spousal caregivers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, pages 1155–1158, New York, NY, USA. ACM.
- Tsikerdekis M (2013). The effects of perceived anonymity and anonymity states on conformity and groupthink in online communities: A wikipedia study. *Journal of the American Society for Information Science and Technology*, 64(5):1001–1015.
- Tu C.-H (2002). The relationship between social presence and online privacy. *The Internet and Higher Education*, 5(4):293 – 318.
- van Oostendorp H and van Varik F (2011). Stimulation of activity in online communities. In Dittmar A and Forbrig P, editors, *ECCE*, pages 163–170. Universitätsdruckerei Rostock, Germany / ACM.
- Wellman B and Gulia M (1999). Net surfers don't ride alone: Virtual communities as communities. In Smith M. A and Kollock P, editors, *Communities in Cyberspace*, pages 167–194. Routledge.
- Wild F and Stahl C (2007). Using latent semantic analysis to assess social competence. In Wild, Kalz, van Bruggen, , and Koper, editors, *Mini-Proceedings of the 1st European Workshop on Latent Semantic Analysis in Technology-Enhanced Learning*, pages 3–4. Open University Netherlands.
- Wise K, Hamman B, and Thorson K (2006). Moderation, response rate, and message interactivity: Features of online communities and their effects on intent to participate. *Journal of Computer-Mediated Communication*, 12(1):24–41.
- Yamashita N, Kuzuoka H, Hirata K, and Kudo T (2013). Understanding the conflicting demands of family caregivers caring for depressed family members. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13*, pages 2637–2646, New York, NY, USA. ACM.
- Zhu H, Kraut R, and Kittur A (2012). Effectiveness of shared leadership in online communities. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW '12*, pages 407–416, New York, NY, USA. ACM.
- Zhu H, Zhang A, He J, Kraut R. E, and Kittur A (2013). Effects of peer feedback on contribution: A field experiment in wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13*, pages 2253–2262, New York, NY, USA. ACM.