Anonymity, User Engagement, Quality, and Trolling on Q&A Sites

CHENG GUO, Clemson University, USA
KELLY CAINE, Clemson University, USA

In online question and answer (Q&A) communities, people ask questions and share answers at all levels of topic sensitivity. Identity options within these communities range from anonymity to real name. The amount of engagement, and the quality of engagement on Q&A sites may differ depending on the identity options available. In this paper, we investigate the relationship between the amount of engagement, the quality of engagement, and different types of identity by analyzing three Q&A sites with different identity policies. We find that highly sensitive questions are more likely to be asked anonymously. Furthermore, allowing anonymity does not affect answer quality and only has a weak, negative indirect effect on engagement. On the other hand, anonymity leads to more trolling. We suggest online communities provide a way for users to ask highly sensitive questions anonymously and pair this with moderation mechanisms to reduce trolling.

CCS Concepts: • **Security and privacy** → **Usability in security and privacy**; • **Human-centered computing** → **Empirical studies in HCI**.

Additional Key Words and Phrases: Identity; Privacy; Anonymity; Online communities; Q&A sites; Trolling; User Engagement

ACM Reference Format:
Cheng Guo and Kelly Caine. 2021. Anonymity, User Engagement, Quality, and Trolling on Q&A Sites. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 141 (April 2021), 27 pages. https://doi.org/10.1145/3449215

1 INTRODUCTION

Social Q&A sites provide a platform for users from all around the world to exchange personal knowledge and information. Social Q&A sites have been of interest to HCI communities in part because they are examples of computer-mediated knowledge sharing [1, 31, 32]. Compared to asking questions via a search engine, Q&A sites provide unique benefits for users such as social fulfillment [40], expert opinions [88], faster-perceived speed, and better answer quality [58]. However, similar to other ways people may gather information online, the quality of information on social Q&A sites varies from excellent to poor [3]. User behavior on Q&A sites also varies. Some users are more engaged, and they actively interact with other users by asking, answering, upvoting, downvoting, and commenting [89]. Deviant behavior such as trolling is also common in online communities including Q&A sites [15, 44].

141

One factor that may be related to information quality, user engagement, and trolling on Q&A sites is the identity policy, which varies from site to site. Q&A sites have different types of identity options ranging from real name — where a contributor's name (and subsequently the contributor's identity) is associated with their question or answer — to anonymity, where no identifying information

Authors' addresses: Cheng Guo, chengg@clemson.edu, Clemson University, USA; Kelly Caine, caine@clemson.edu, Clemson University, USA.

2573-0142/2021/4-ART141 $15.00
https://doi.org/10.1145/3449215

is associated with specific questions or answers. Choosing which identity policy to implement presents questions for firms seeking to develop their sites, engage users, and maintain the quality of information at a high level. One of the most important questions is about privacy; many Q&A sites choose to provide a privacy feature for users to ask or answer questions anonymously.

Being anonymous online can bring certain benefits such as encouraging users to engage in freedom of thought and expression and helping them to control personal information disclosure [37]. There is a need for users to be able to use the Internet privately, and anonymity is one way to achieve privacy. Common types of anonymous activities online include file sharing and downloading, social networking to exchange help and support, browsing, and searching for information [37]. Anonymity also has drawbacks such as resulting in a lack of consequences for bad behavior [36] and undermining personal contribution in group communications [79]. Being identified online may also have positive effects such as leading users to use less offensive words [27] and avoid discrimination [22]. Given that anonymity has both pros and cons, Q&A sites and other online communities may struggle to achieve balance in the design of their identity policy - especially in regard to the important question of whether to include anonymity as an identity option.

In this paper, we seek to deepen our understanding of the relationship between different types of identity and user engagement, answer quality, and trolling behavior on Q&A sites. Our work can help to assist Q&A sites and other user-centered online communities to choose appropriate identity policies and help users to make decisions about identity options provided by Q&A sites.

We collected a dataset of 3,000 questions and answers from three popular Q&A sites with different identity policies: Yahoo Answers (https://answers.yahoo.com/), Quora (https://quora.com/) and Zhihu (Chinese, https://zhihu.com/). We conducted a post-hoc observational analysis of these questions and answers to uncover the relationship between identity/anonymity and user engagement, answer quality, and trolling behavior.

Our results illustrate: (1) highly sensitive questions are more likely to be asked anonymously; (2) allowing anonymity does not affect answer quality and only has a weak, negative indirect effect on engagement; (3) anonymity leads to more trolling. The primary contributions of this work are threefold.

- First, an empirical study of questions and answers from across three Q&A sites with different identity policies reveals the importance of allowing users to have an anonymity option, especially for highly sensitive topics.
- Second, our analysis reveals that answer quality and user engagement are consistent across different levels of identity on Q&A sites; notably, anonymity does not negatively affect answer quality and user engagement.
- Finally, we provide insights for Q&A sites suggesting that they include an option for users to engage anonymously. Providing an option to participate anonymously does not risk reducing content quality and user engagement, and moderation mechanisms may help reduce trolling.

2 RELATED WORK AND BACKGROUND

2.1 Identity in Questions and Answers

One way that people find it useful to seek and find information online is via Q&A sites [88]. People interact with Q&A sites using various types of identity. Based on the level of disclosure, these types of identity include anonymity [7, 68], pseudonymity [95], and real names [75]. On Q&A sites, a real name is one's real, full name [76, 109]. Being identified by one's real name online can be both negative and positive. For example, revealing real identities can lead to discrimination both in online [22] and offline [28] settings. Real name policies may also lead to a decrease in people's willingness to express disagreement, which can harm fruitful debate [35]. On the other hand, by enforcing real-name policies, sites can reduce offensive behavior. For example, offensive word use was reduced by 11% on one online news commenting platform when it began enforcing a real name policy [27].

On Q&A sites, anonymity means that users can choose to not disclose their user names or real names in questions or answers [75, 106]. Being anonymous online has both advantages and disadvantages. On one hand, being anonymous online helps people feel more relaxed, and feel free to express views, and helps control personal information disclosure [37]. On the other hand, anonymity may undermine perceptions of one's contributions in group communications [79]. On the whole, Internet users are supportive of anonymity. Eighty-six percent of Internet users have taken actions to remove or mask their online activities [78], and 59% of Internet users say that they should be able to use the Internet anonymously. Furthermore, 18% of Internet users already use the Internet anonymously or in a way that their identities are hidden [78].

Besides anonymity and real name, another common identity option on Q&A sites and other online communities is pseudonymity. Unlike anonymity, pseudonymity links a user's online behavior to a single identity (pseudonym), but that pseudonym may not be able to be linked back to a user's identity [80].

To summarize, there is variability in the identity options available to users. There is also variability in the behaviors and comfort of people as they interact online based on identity options. In particular, we are interested in understanding whether the sensitivity of topics people ask and answer questions about depends on the identity policy of a Q&A site. However, we found no systematic study of how the type of identity relates to topic sensitivity and users' asking and answering behavior. Users' asking and answering behavior may result in different degrees of user engagement [89], various distribution of information quality [3] and unequal amount of trolling [15].

2.2 User Engagement in Questions and Answers

In this work, we define user engagement as the interactivity [60] between users and Q&A sites. Basic Q&A features including asking, answering, following, commenting, upvoting, and downvoting can be used as metrics to measure user engagement on Q&A sites [52, 56]. Upvoting and downvoting are two core features that many Q&A sites have [3, 32, 50, 71, 81, 97]. According to Yahoo's community guidelines [100], upvoting an answer is synonymous with a thumbs-up or a vote of approval. This understanding of upvoting is shared by the Yahoo community as evidenced by discussions on Yahoo that reveal users also treat upvoting as liking and agreement [99]. Although Quora and Zhihu do not provide a guideline for upvoting, users' understanding of upvoting is similar to Yahoo [77, 110]. Differences in user engagement may relate to different types of identity. For example, after removing the anonymity option on an online community of practice for U.S. soldiers (i.e., the identity model changed from allowing anonymity to disallowing anonymity), the number of comments decreased significantly (although the overall peripheral participation measured by logins and page views were unaffected) [39]. Similarly, after disabling anonymity and requiring users to log in via Facebook on two online news sites (i.e., the identity model changed from allowing anonymity to disallowing anonymity), the number of comments decreased [27, 64].

These works were focused on comparing engagement with and without anonymity options. However, neither of these works examined engagement with Q&A sites where multiple levels of identity are provided to users (anonymity vs. non-anonymity: pseudonymity, real names, or both). Thus, we are left with the following questions: How is user engagement affected by various identity options? When users are given identity options, will the level of user engagement differ between anonymous and non-anonymous participation? Therefore, we ask:

RQ1: *What is the relationship between different types of identity and user engagement on Q&A sites?*

2.3 Information Quality in Questions and Answers

Prior work evaluates and predicts information quality on Q&A sites [3, 32, 50, 71, 81, 97]. The most relevant predictors of quality include the price required for an expert to post an answer [32], length of answer [3, 71], upvotes [71, 97], downvotes [71], and factual information [50]. Information quality may also be associated with different types of identity. For example, after sites removed the option to engage anonymously (i.e., the identity model changed from allowing anonymity to disallowing anonymity), comment quality improved on online news sites [27] and online communities of practice [39]. However, these studies were focused on comparing information quality with and without anonymity options. This does not help us understand what happens to information quality when sites give users multiple types of identity choices simultaneously. Therefore, we ask:

RQ2: *What is the relationship between different types of identity and answer quality on Q&A sites?*

2.4 Question Topics and Sensitivity

Identity, question topics, and topic sensitivity are related in online communities such as Q&A sites. When people use their real names to ask questions on social networks to their friends, family, and colleagues, they tend to ask less sensitive questions (e.g., they will ask questions about technology and entertainment) [58]. On the other hand, when people ask questions to strangers via anonymous platforms like Facebook confession boards, the topics of questions tend toward more sensitive topics (e.g., health, illegal substances, and sex) [7]. On real-name-enforced Q&A sites like Quora, people use the anonymity feature more frequently to answer highly sensitive questions [68]. Similar to Quora, there are many sensitive questions and answers on Yahoo Answers that users ask and answer anonymously [69]. In this study, we selected ten highly sensitive topics and ten less sensitive topics for analysis. We selected these topics based on prior work on Q&A sites that suggested these topics were highly sensitive or less sensitive, respectively [7, 68, 69]. These topics are listed in Table 1.

Table 1. Highly and less sensitive topics [7, 68, 69]

Highly Sensitive Topics	Less Sensitive Topics
Sex	Sports
Suicide	Travel
Death	Corporations
Orgasms	Food
Masturbation	Music
Female sexuality	Movies
Male sexuality	Television
Rape	Science
Abuse	History
Pornography	Psychology

2.5 Trolling in Questions and Answers

Deviant online behavior such as trolling has been of interest to HCI researchers for decades [10, 11]. Prior studies have examined trolling in a number of cyberspaces, including feminist forums [33], computer-mediated communications [30], online video gaming [94], and crowdsourcing platforms such as Wikipedia [87]. Q&A sites are also subject to trolling [21]. Some Quora users [67] have suggested that Q&A sites that enforce real-name policies are subject to less trolling than Q&A sites that do not enforce real-name policies. However, there are many differences across Q&A sites that could lead to differences in trolling. For example, the typical user population of each site varies. Systematic studies are needed since it remains unclear how different types of identity are associated with trolling within a site. Therefore, we ask:

RQ3: *What is the relationship between different types of identity and trolling behavior on Q&A sites?*

To sum up, we know that people use different types of identity in online communities to ask and answer questions with different topics and different levels of sensitivity. However, we do not know if people's identity choice relate to the topics or sensitivity. We also know that people's behavior in online communities, such as the quality of the information they share and trolling, may relate to the identity they use. However, we do not know how different types of identity relate to these behaviors when multiple identity options are provided to people simultaneously within a site. Thus, we conducted an empirical study across three Q&A sites with different identity policies.

3 METHOD

3.1 Dataset

We selected three popular Q&A sites (Quora, Zhihu, and Yahoo) for two main reasons. First, and of primary importance for our research questions, these three Q&A sites have different identity policies (see Table 2). Quora is the only Q&A site we know of that enforces the use of a real-name identity policy [76]. Zhihu recommends that users use real names as their usernames but does not make this mandatory. Zhihu also allows users to upload supporting documentation to verify their identities [111]. Verified users receive a certification mark attached to their usernames (shown as a blue tick on the interface, similar to Twitter's verified account checkmark), which may incentivize users to use real names. As a result, some Zhihu users register using real names while others register using pseudonyms. Yahoo Answers does not enforce or even recommend a real-name identity policy. Instead, Yahoo asks users to use their Yahoo user ID or any nickname they desire [100]. Thus, we consider Yahoo usernames to be pseudonyms. Despite these differences in identity policies, Yahoo, Quora, and Zhihu all provide privacy features that allow users to ask or answer question anonymously.

Second, these three Q&A sites provide similar (but not identical) interactive features. For example, they all allow users to ask questions, but Yahoo does not allow users to comment on questions. All sites allow users to: ask questions, answer questions, follow questions, upvote answers, downvote answers, and comment on answers. The common site features enable us to capture engagement metrics across sites. All sites also have similar web interfaces (see Fig. 1, Fig. 2, and Fig. 3 for example interfaces for Quora, Zhihu, and Yahoo respectively).

Table 2. Identity policy and anonymity feature on Q&A sites

	At Registration	Offer Anonymity?
Yahoo	Pseudonymity	✓
Quora	Real name (mandatory)	✓
Zhihu	Real name (not mandatory)	✓

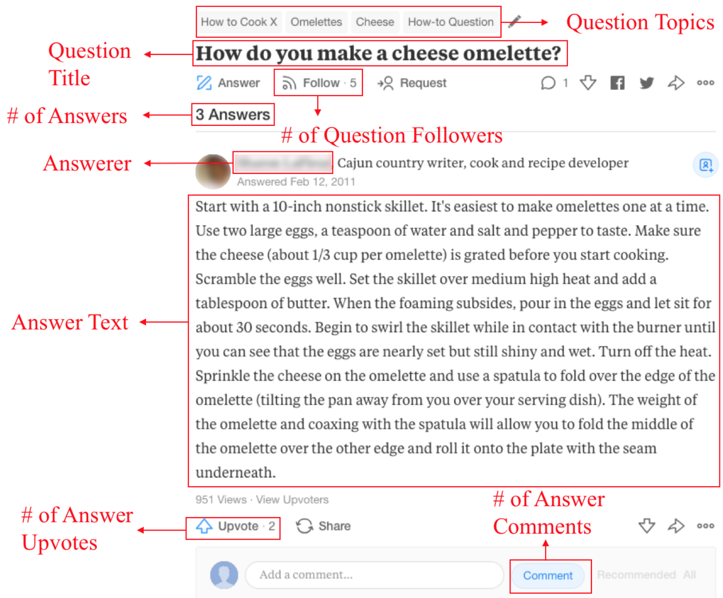


Fig. 1. An example web interface of Quora

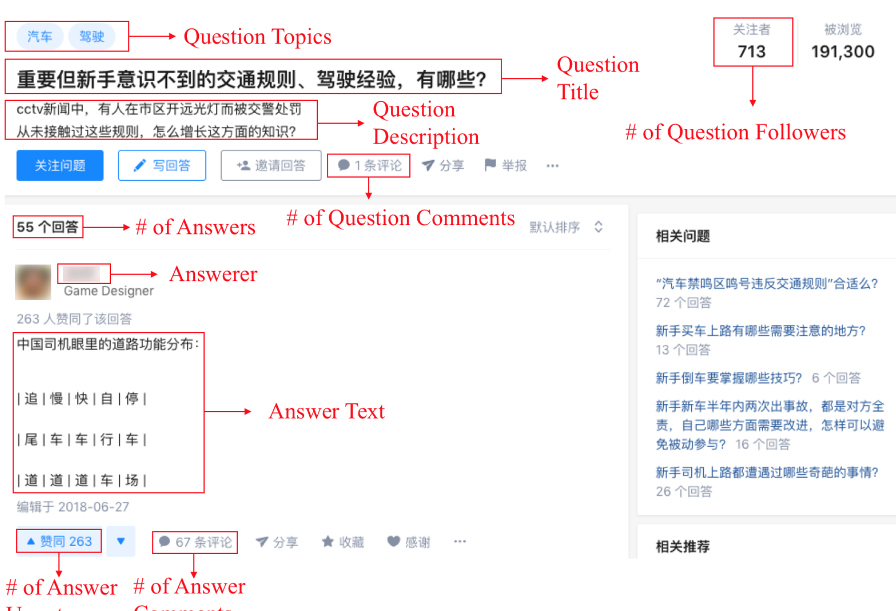


Fig. 2. An example web interface of Zhihu

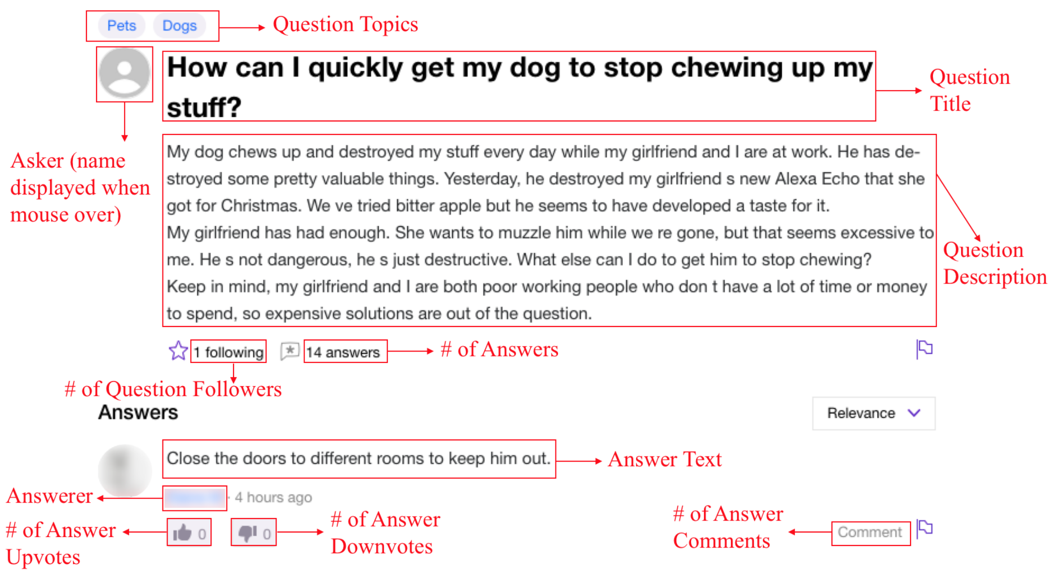


Fig. 3. An example web interface of Yahoo

3.2 Data Collection

We collected data in March 2017. All the questions we collected were asked between January 2014 and March 2017. The data consists of 3,000 questions and all the answers (3,000 of which were analyzed) under these questions on Yahoo, Zhihu, and Quora.

3.2.1 Identifying questions. First, we needed to select questions with similar topics (from the list of highly and less sensitive topics in Table 1) from all sites. To accomplish this, we examined how each site categorizes topics and determined whether the site listed an existing topic for the 20 we identified. The 20 topics we identified were all existing topics on both Quora and Zhihu. To collect questions, we simply navigated to the topic feed page and manually collected the top 50 questions from each based on ranking. However, Yahoo has fewer, and thus broader, topics than Zhihu and Quora [97, 104]. We observed 332 topics listed on Yahoo (https://answers.yahoo.com/dir/index) and found that most of the highly sensitive topics we selected did not exist as high-level topics listed on Yahoo. So, we used the ten highly sensitive topics as keywords and queried Yahoo's web interface using the keywords. Then, we manually selected the top 50 questions from the search results based on the ranking of each query. On the other hand, the ten less sensitive topics were all already existing topics on Yahoo, so we simply collected the top 50 questions from the feed page of each topic based on ranking. Note that it is possible that sites use a different ranking algorithm for the feed page vs. the search results; however, no site specifically discloses their ranking algorithms, so we are not able to measure this. Without access to an API, this was our best option. Thus, using this approach, after a preprocessing (See Section 3.2.2), the 3,000 questions in our dataset are comprised of 20 topics (10 highly sensitive, 10 less sensitive) × 50 questions (one from each topic) × 3 Q&A sites (Yahoo, Quora, and Zhihu).

3.2.2 Filtering questions. We excluded questions that either 1) had a description and/or 2) those with no answers. As shown in Figure 1, 2, and 3, some questions have a description but other questions only have a title. To ensure consistency and avoid the impact of the uneven length and quality of these descriptions, we excluded questions that contained descriptions. Following [66]'s approach, we also excluded questions without any answers because we were interested in analyzing both question-level and answer-level data. As a result, all of the questions we analyzed had at least one answer and had no description.

We understand that excluding questions with a description and without answer limits our ability to extrapolate our results beyond these question types. To clarify the extent to which this sample could plausibly inform policies in the rest of the site, as of April 2020, we collected the same amount of data using the same approach. We found that 81.6% of questions had no question descriptions and 99.9% of questions had at least one answer. Quora has already disabled the question description feature after we collected data [74]. According to Quora, this change provides a number of benefits including: 1) "Prevents answers that appear irrelevant if the writer didn't read the question details"; 2) "Prevents answers that respond specifically to the question details but appear irrelevant to the main question"; 3) "Increases the likelihood that answers to a question will receive upvotes because those answers are more widely relevant" [74]. These arguments support our filtering criteria that exclude questions with a description.

3.2.3 Collecting answers and engagement metrics. After we identified and filtered the questions, we manually collected all the answers associated with these questions as of March 2017. This resulted in more than 385,000 answers in total, providing a large set from which we could sample - 3,000 answers - our goal for analysis (see Section 3.3). We also collected a set of metrics for each question and answer (see Table 3 for the complete list of these metrics which includes upvotes, downvotes, and comments). Note that while answer downvotes on Quora and Zhihu are available features for users, they are not visible on their web interface. Thus, we were not able to collect them.

3.3 Data Processing

We conducted a post-hoc observational analysis on the questions and answers we collected. While the Q&A site metrics listed in Table 3 required no additional processing since we collected them from the Q&A sites' web interface, other metrics including question sensitivity, answer quality and trolling did require processing. We paid seven undergraduate/graduate research assistants to act as human raters for these metrics. Their job was to rate or label question sensitivity, answer quality, and trolling manually. We conducted inter-rater reliability tests for all ratings by human raters to ensure reliability. Because we were limited to human, rather than automatic processing, we needed to limit the number of ratings each human rater needed to complete. So, we randomly selected one answer associated with each question for analysis. This resulted in a final set of 3,000 answers for qualitative and quantitative analysis.

3.3.1 Human raters. Five male and two female human raters volunteered to participate, their ages ranged from 19 to 28. Four raters are Asian, two raters are White, and one rater is African American. Rater #1 is male, speaks both English and Chinese and considers himself to be a native speaker of both languages. Rater #2 is male, speaks English as a native language and speaks Chinese as a second language. Rater #3 and rater #4 are both male and speak English as their native language. Rater #5 and #6 are both native Chinese speakers and speak English as a second language. Rater #7 speaks English as a second language. Both rater #6 and #7 are female. For all measurements, at least two raters participated in each rating. The human raters first went through a 300-question training phase (consisting of a separate set of questions that were not used in later analysis) to rate question sensitivity, answer quality, and trolling until the agreements were over 80%. Then, the human raters used the annotation from the training phase as examples and guidelines to rate individually, at which point inter-rater reliability was calculated again and any disagreements were resolved by a discussion between raters.

Table 3. Q&A sites metrics

	Quora	Yahoo	Zhihu
Question Title	✓	✓	✓
Asker's Name	✓	✓	✓
# of Question Followers	✓	✓	✓
# of Question Comments	✓	×	✓
# of Answers	✓	✓	✓
Answer Text	✓	✓	✓
Answerer's Name	✓	✓	✓
# of Answer Comments	✓	✓	✓
# of Answer Upvotes	✓	✓	✓
# of Answer Downvotes	×	✓	×

We present the measurements resulting from these ratings in the following four sub-sections: Categorizing Identity, Measuring Question Sensitivity, Measuring Answer Quality, and Measuring Trolling on Answers.

3.3.2 Categorizing identity. Identity can be categorized as anonymous, pseudonymous or real name. For Yahoo and Quora, following the approach used by [68], human raters first labeled anonymous askers and answerers (both are shown as "Anonymous" on Yahoo and Quora's web interface). Then, human raters labeled other askers' and answerers' names as pseudonyms on Yahoo or real names on Quora. For Zhihu, human raters first labeled anonymous askers and answerers (shown as "匿名用户" on Zhihu's web interface). Then, we supplemented the approach by [68] by adding human coders to distinguish between real names and pseudonyms. Rater #1 and rater #2 labeled the remaining askers' names individually as pseudonyms or real names. Raters first determined whether the family name of an identity is real by comparing the family name to a reference containing statistics about Chinese names [103]. Then they determined whether the given name is a real name or not. The Fleiss's Kappa indicates that all coders agreed 91% of the time about which names were pseudonyms and which were real names. [68] described the following as a limitation of their approach: "although anecdotal evidence suggests most users use their real name (many users link to their Facebook and Twitter accounts in their Quora profiles), our findings are limited by the extent to which Quora succeeds in enforcing the real names policy". The same limitation applies to our categorization on Quora and Zhihu. On Zhihu, like Quora, many users link to their social media accounts using their real name. However, our work, like [68], is limited in that it is possible that names we categorized as real names are not real names after all.

3.3.3 Measuring question sensitivity. To better understand the relationship between different levels of identity, question topics, and their sensitivity, we measured the sensitivity of questions. Measuring sensitivity has been fraught across disciplines. As Lee and Renzetti describe, "one difficulty with the notion of a 'sensitive topic' is that the term is often used in the literature as if it were self-explanatory" [48]. Perhaps not surprisingly then, we found no guidance about assessing sensitivity or sensitive content in the context of research about Q&A sites. Although prior work on Q&A sites [7, 68, 69] has suggested that some topics are highly sensitive, those work have not formalized a method to assess sensitivity automatically. Emerging work in photo privacy has provided a taxonomy of sensitive content in photos which may guide automatic methods like machine learning in the future [51], but it is not yet clear to what extent that taxonomy applies to Q&A sites. Furthermore, none of the three Q&A sites we analyzed provide a definition or have guidelines about what is considered sensitive, so we could not rely on them to help us assess the sensitivity.

So, for guidance, we turned to the privacy literature. Privacy experts note that the most current definitions are restricted to "data that may give rise to discrimination, for example healthcare data," [59, p. 35]. Similarly, the Organization for Economic Co-operation and Development's (OECD) guidelines of data sensitivity use "the risk of discrimination" as their key criteria in determining whether or not content is sensitive [61, 62].

Since there is no universal definition of sensitivity in the academic literature and Q&A sites do not define sensitivity, following the lead of privacy experts' views [59], the OECD guidelines [61, 62], and experts on sensitive topics [48], we chose to train our human raters with a broad definition of sensitivity that considers anything that could lead to discrimination as sensitive. Lee and Renzetti argue that a broad definition of sensitivity is a "major advantage" that it allows for the "inclusion of topics that ordinarily might not be thought of as 'sensitive' " [48, p. 511].

To assess the sensitivity of questions, we used manual human coding using a using a 5-point Likert scale (1-not at all sensitive, 2-slightly sensitive, 3-moderately sensitive, 4-very sensitive, 5-extremely sensitive). Within each topic, the sensitivity of each question may vary. Thus, it is necessary to manually assess sensitivity at the question-level. Each Yahoo and Quora question was rated by rater #1, #2, and #7. Each Zhihu question was rated by rater #1, #2, and #6. We blinded coders to the asker's identity (i.e., concealed the asker's identity from the rater so as not to influence their rating) during the coding process. We chose the 5-point Likert scale because it has been used in the past to successfully measure the sensitivity of the topics in a variety of survey [83] and interview [38] questions.

To ensure raters were applying this definition consistently, we calculated inter-rater reliability (IRR). Using the guidelines from work on the assessment of IRR that suggest using intraclass correlation (ICC) for Likert scales [29], we performed ICC on the three raters' ratings. The agreement was .815, .834 and .867 for Yahoo, Quora, and Zhihu respectively, indicating good agreement [9, 45].

3.3.4 Measuring answer quality. Previous work measures answer quality using two primary approaches: machine learning [3, 5] and human ratings [14, 16, 23, 90]. Successful machine learning models can "separate high-quality items from the rest with an accuracy close to that of humans" [3]. Thus, since human ratings are the gold standard, we measured answer quality using human raters. As we described in Section 3.3, to limit the number of answers for analysis, we randomly sampled one answer from the entire set of answers we collected along with each question for further analysis. This resulted in 3,000 answers for analysis.

We used the metrics used to measure question quality by [31] to measure answer quality. The items are *writing quality*: "I think this answer is well-written," and *archival value*: "I think this answer provides information of lasting/archival value to others." Three raters rated each answer using a 5-point Likert scale (1-strongly disagree, 2-disagree, 3-neither agree or disagree, 4-agree, 5-strongly agree) with the answerer's identity concealed while referring to the question. Rater #3, #5, and #7 rated 1,000 answers from Yahoo. Rater #4, #5, and #7 rated 1,000 answers from Quora. Rater #4, #5, and #6 rated 1,000 answers from Zhihu. The raters' ratings were consistent using ICC as a measure: For Yahoo, the agreement was .796 and .811 for writing quality and archival value, respectively. For Quora, the agreement was .820 and .793 for writing quality and archival value, respectively. For Zhihu, the agreement was .791 and .807 for writing quality and archival value, respectively.

3.3.5 Measuring trolling on answers. We adopt a definition of trolling on answers that includes flaming, griefing, swearing, personal attacks, or not answering the question, including any behavior that violates the three Q&A sites' community guidelines and terms of use [15]. Pinpointing trolls and trolling posts automatically is challenging [84]. To quantify the amount of trolling behavior on answers, each answer (while referring to the question) was read and labeled by three human raters with answerer's identity concealed. Labeling options were binary: includes trolling vs. does not include trolling. Before coding, raters reviewed three Q&A sites' community guidelines and terms of service to have a sense of what is not allowed on each site. Rater #3, #5, and #7 labeled 1,000 answers from Yahoo. Rater #4, #5, and #7 labeled 1,000 answers from Quora. Rater #4, #5, and #6 labeled 1,000 answers from Zhihu. The Fleiss's Kappa was .802, .842, .814 for Yahoo, Quora, and Zhihu respectively, indicating great agreement.

3.4 Ethical and Privacy Considerations

We followed CSCW's community norms of ethical ways of studying online communities [12, 24, 25, 96]. Our entire research protocol was IRB approved. We also abided by each sites' Terms of Service for the entire study. All the data we collected are stored in a secure location to which only the research team has access. When processing the data, we replaced the identifiable name of users with unidentifiable labels. The names and the avatars of users are blurred in Figure 1, 2, and 3.

4 RESULTS

Table 4. An overview of the question-level metrics analyzed across three Q&A sites. Numbers are N or means (M) with standard errors in parentheses. A = anonymity, P = pseudonym, R = real name.

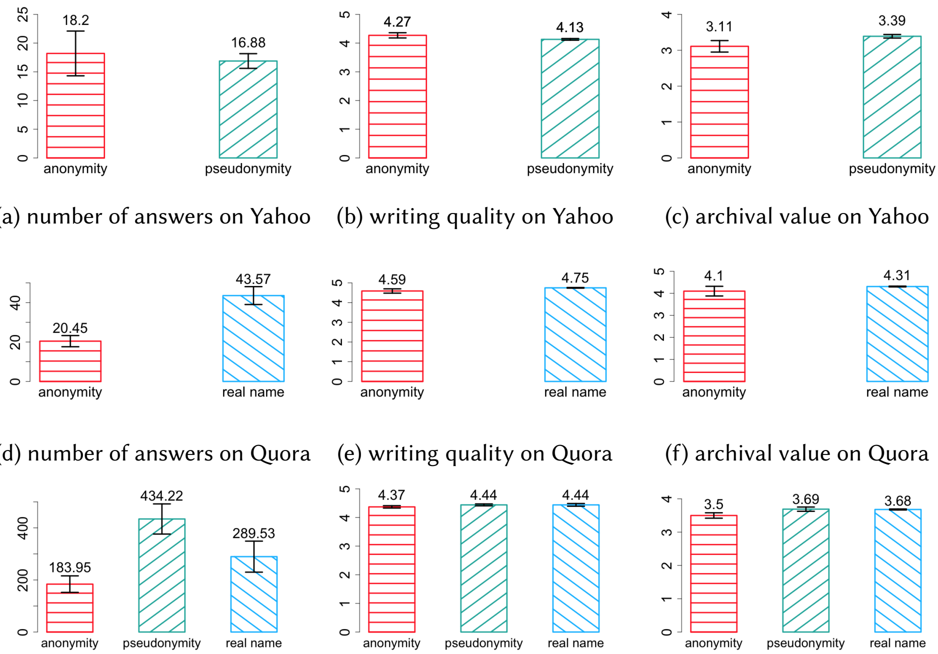
	Yahoo		Zhihu			Quora	
	A	P	A	P	R	A	R
Questions Sampled	334	666	324	483	193	365	635
Answers per Question	18.2 (3.9)	16.9 (1.3)	184.0 (31.8)	434.2 (57.9)	289.5 (59.6)	20.5 (2.8)	43.6 (4.6)
M Question Sensitivity	2.05 (0.05)	1.78 (0.03)	2.49 (0.06)	1.90 (0.05)	1.67 (0.07)	2.12 (0.06)	1.87 (0.04)
M Question Followers	0.3 (0.1)	0.3 (0.04)	2634.2 (380.1)	7520.2 (877.8)	8815.1 (1571.6)	35.4 (10.4)	124.1 (27.4)
M Question Comments	N/A	N/A	8.7 (1.4)	14.1 (1.5)	11.7 (2.3)	0.6 (0.1)	0.9 (0.1)

Table 5. An overview of the answer-level metrics analyzed across three Q&A sites. Numbers are N or means (M) with standard errors in parentheses. A = anonymity, P = pseudonym, R = real name.

	Yahoo		Zhihu			Quora	
	A	P	A	P	R	A	R
Answers Sampled	97	903	208	632	160	41	959
Trolling answers	39	221	54	90	19	3	47
M Answer Comments	0.3 (0.03)	0.3 (0.01)	2.2 (0.5)	3.6 (0.6)	4.1 (1.2)	1.6 (0.6)	2.1 (0.4)
M Answer Upvotes	1.5 (0.1)	1.1 (0.02)	7.2 (2.5)	13.0 (3.6)	23.6 (8.9)	23.9 (11.2)	165.5 (45.2)
M Answer Downvotes	1.1 (0.1)	0.8 (0.02)	N/A	N/A	N/A	N/A	N/A

4.1 Analysis

We considered multiple analysis methods (including both univariate and multivariate methods) to deal with the complex relationships between all the metrics we wanted to analyze. We decided to use path analysis [98] with a robust estimator. Path analysis is a special case of Structural Equal Modeling (SEM) and can be understood as an extension of a regression [92]. A path model lets us determine the structural relations between all variables in a single model. Thus, it will let us see the relationship between the identity of the asker and trolling and archival value. Meanwhile, it will also let us see the relationship of the sensitivity of question topic and the number of answers, for example. Other analysis strategies would not allow us to see and understand the relationship of these variables at the same time. While providing a comprehensive introduction to the analysis method and its interpretation is outside the scope of the paper, please see the following for comprehensive



(a) number of answers on Yahoo (b) writing quality on Yahoo (c) archival value on Yahoo

(d) number of answers on Quora (e) writing quality on Quora (f) archival value on Quora

(g) number of answers on Zhihu (h) writing quality on Zhihu (i) archival value on Zhihu

Fig. 4. Marginal effects of different levels of identity (anonymity is shown in red; pseudonymity is shown in green; real name is shown in blue) on number of answers (left column), answer writing quality (center column) and answer archival value (right column) on Yahoo (first line), Quora (second line) and Zhihu (third line).

guidance: 1) a textbook that gives an excellent description is [42, Chapters 6, 7, 11, 12] and 2) a practical example using path modeling, with an explanation of it in the appendix is [43, p. 489 - 491]).

We chose to create three separate path models, one for each of the sites we analyzed, instead of combining data from all three together into one model. We did this to minimize the effects of confounding variables (e.g., different site features, users, and design of each site). For each site, we started with a saturated path model. A saturated model contains all possible paths between variables. The next step in path model is to trim the path model until the model had a good fit and the effect of each path is statistically significant ($p < .05$). Statistical significance here means that the two variables on each side of the path are significantly associated [92]. To trim each path model, we started with the least significant and least interesting effects (e.g., those that were added for saturation), and removed paths iteratively until all non-significant paths were removed.

Each path of the model is shown using an arrow. Like any regression model, path models make assumptions about the direction of causality for each path. However, since our study is not a randomized experiment, we interpret these as associations, rather than directional causality. In each path model, question-level metrics are shown in green and answer-level metrics are shown in red. The regression weight is predicted by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Significance levels are: ***$p < .001$, **$p < .01$, *$p < .05$. R^2 is the portion of the variance explained by the model. We encoded each categorical variable with the value on the right of "vs." as the baseline.

We conducted multicollinearity tests [65] for all variables for all models. The values of tolerance were all above the threshold (0.2) [65], which means there is no issue of multicollinearity in our models. Note that unlike significance tests (e.g., ANOVAs), in a path analysis there is no need to control familywise error rates [19]. Thus, we did not use any correction for multiple comparisons (e.g, Bonferroni correction).

The *marginal* (unmediated) effects of identity on user engagement (taking number of answers as an example, other metrics following the same trend) and answer quality (writing quality and archival value) are shown in Figure 4, which reflect RQ1 and RQ2, respectively. We first present the results of question sensitivity across the three Q&A sites and then we interpret the trimmed path models with main effects in the next three sub-sections split by site. An overview of the metrics we analyzed, including the number of trolling answers (which reflect RQ3), can be found in Table 4 and Table 5. Note that while only one metric of user engagement is included in the path models, other metrics followed the same trend. Thus, we removed those metrics for simplicity. The full path models with all metrics included for all three Q&A sites can be found in the Appendix.

4.2 Question Sensitivity

Since our sample is not normally distributed, we settled on the use of non-parametric tests (e.g., Mann-Whitney U, Kruskal-Wallis), where appropriate. Across sites, the mean (and its standard error) of question sensitivities varied between highly sensitive topics and less sensitive topics: Yahoo (2.43 ± .035 vs. 1.3 ± .029), Zhihu (3.03 ± .029 vs. 1.03 ± .013) and Quora (2.81 ± .033 vs. 1.06 ± .011). The values of Mann-Whitney U are 219,220 (Yahoo), 246,250 (Zhihu), 244,250 (Quora), respectively. All the values of p are less than 0.001. Questions about highly sensitive topics are more sensitive than questions under less sensitive topics. This variation may be taken as a validation that the topics we identified from previous literature [7, 68, 69] represent highly sensitive and less sensitive topics respectively. An example of a less sensitive question (rated as "not at all sensitive") is: "What is the most clever life hack you've learned?" An example of a highly sensitive question (rated as "extremely sensitive") is: "How painful is death through drinking bleach?"

Table 4 shows the average question sensitivity of three Q&A sites split by different levels of identity. On Yahoo, the mean (and its standard error) of question sensitivity varies between pseudonymous askers (1.78 ± .033) and anonymous askers (2.05 ± .052). The Mann-Whitney U is 129,400, $p < 0.001$. On Zhihu, the mean (and its standard error) of question sensitivity varies between pseudonymous askers (1.90 ± .049), anonymous askers (2.49 ± .058), and real name askers (1.67 ± .071). Kruskal-Wallis is $\chi^2(2) = 86.986$ ($p < 0.001$). However, on Quora, we did not observe the same trend. There is no significant difference between anonymous askers (2.12 ± .055) and real name askers (1.87 ± .039) in terms of question sensitivity. The Mann-Whitney U is 90,462 ($p = 0.127$).

4.3 Yahoo

Figure 5 shows the trimmed path model for Yahoo. The model fit is: $\chi^2(12) = 25.060$, $p = .015$; *RMSEA* = 0.033, 90% CI : [0.014, 0.051], *CFI* = 0.979, *TLI* = 0.963. The model's chi-square should not be statistically significant. However, the chi-square statistic is very sensitive to the sample size [6] and is no longer relied upon [86]. Hu and Bentler [34] propose the cut-off values for other fit indices to be: *RMSEA* < 0.05, with the upper bound of its 90% CI below 0.10, *CFI* > 0.96 and *TLI* > 0.95. The cut-off values indicate that our model has a good model fit.

Using this model we find that on Yahoo, anonymous askers are more likely to ask questions about highly sensitive topics. Notably, questions about highly sensitive topics are more likely to be

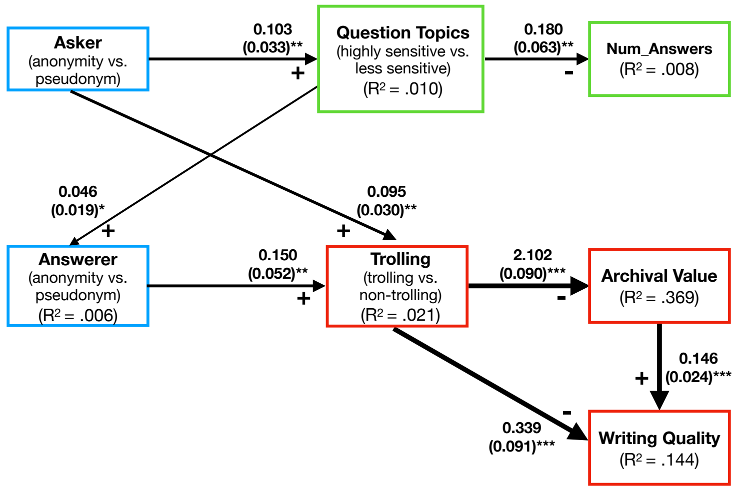


Fig. 5. The trimmed path model for the data of Yahoo. Significance levels: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. R^2 is the portion of the variance explained by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Identities are shown in blue. Question-level metrics are shown in green. Answer-level metrics are shown in red. Anonymity is coded as 1 and pseudonym is coded as 0. Highly sensitive is coded as 1 and less sensitive is coded as 0. Trolling is coded as 1 and non-trolling is coded as 0.

answered anonymously than questions about less sensitive topics. We do not find any direct effect of answerer's identity on answer quality; anonymous and pseudonymous askers ask questions of equivalent quality. The indirect effects are mediated by trolling. The marginal effects of answerer's identity on writing quality and archival value show that the indirect effects are not significant (see Figure 4b and Figure 4c, $p = .051$ and .085 respectively).

We observed 260 (26%) answers are trolling on Yahoo (see Table 5). Anonymous questions are more likely to be trolled. Furthermore, anonymous answers are more likely to be trolling. The trolling answers, in turn, have lower writing quality than non-trolling answers. Answers with higher archival value are more likely to have higher writing quality. Thirty-six point nine percent of the variance in archival value and 14.4% of the variance in writing quality can be explained by this model.

Finally, neither asker's identity nor answerer's identity has a direct effect on user engagement metrics. However, asker's identity has an indirect effect on the number of answers mediated by question topics. Only 0.8% of the variance in the number of answers can be explained by this model. The marginal effect of asker's identity on the number of answers (one of the metrics to measure user engagement) shows that the effect is not significant (see Figure 4a, $p = .748$). Following the same trend, both asker's identity and answerer's identity have a weak, indirect effect on other user engagement metrics (e.g., number of question followers, number of answer upvotes/downvotes/comments, see Table 4 and Table 5). Thus, we removed these metrics (but kept the number of answers to represent user engagement) from the simplified trimmed model for simplicity. The full trimmed path model of Yahoo can be found in Figure 8 in the Appendix.

4.4 Quora

Figure 6 shows the trimmed path model for Quora. The model has a good [34] model fit: $\chi^2(10) = 7.929$, $p = .636$; $RMSEA = 0.000$, 90% CI : [0.000, 0.027], $CFI = 1.000$, $TLI = 1.008$.

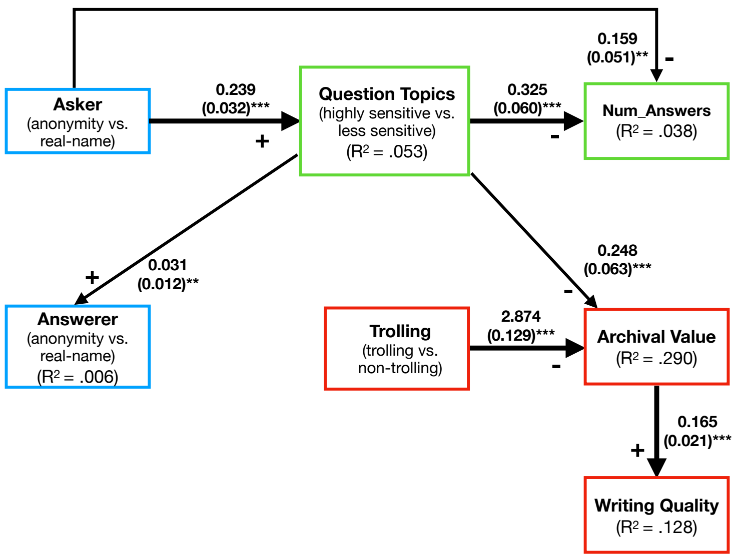


Fig. 6. The trimmed path model for the data of Quora. Significance levels: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. R^2 is the portion of the variance explained by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Identities are shown in blue. Question-level metrics are shown in green. Answer-level metrics are shown in red. Anonymity is coded as 1 and real name is coded as 0. Highly sensitive is coded as 1 and less sensitive is coded as 0. Trolling is coded as 1 and non-trolling is coded as 0.

Using this model we find that on Quora, anonymous askers are more likely to ask questions about highly sensitive topics. Notably, highly sensitive questions are more likely to be answered anonymously than questions about less sensitive topics. We do not find any significant effect of answerer's identity on answer quality (See Figure 4e and Figure 4f, $p = .081$ and .291 respectively).

Unlike Yahoo, on Quora we do not find any significant effect of asker's identity/answerer's identity on trolling. Only 5% of the answers we analyzed on Quora contain trolling compared to the 26% of answers on Yahoo (see Table 5). Trolling answers have a significantly lower archival value than non-trolling answers. Answers with higher archival value are more likely to have higher writing quality. Twenty-nine percent of the variance in archival value and 12.8% of the variance in writing quality can be explained by this model.

Asker's identity and the topic of questions have a direct but weak effect on the number of answers (one of the metrics to measure question-level user engagement). Although the marginal effect is significant (See Figure 4d, $p < .001$), only 3.8% of the variance of the number of answers can be explained by this model. Following the same trend, asker's identity has a weak but indirect effect on other question-level user engagement metrics (e.g., number of question followers/comments, see Table 4). Answerer's identity has no significant effect on answer-level user engagement metrics (e.g., number of answer upvotes/comments, see Table 5). Thus, we removed these metrics (but kept the number of answers to represent user engagement) from the simplified trimmed model for simplicity. The full trimmed path model of Quora can be found in Figure 9 in the Appendix.

4.5 Zhihu

Figure 7 shows the trimmed path model of Zhihu. The model has a good [34] model fit: $\chi^2(13)$ = 30.171, p = .004; $RMSEA$ = 0.036, 90% CI : [0.019, 0.054], CFI = 0.986, TLI = 0.977.

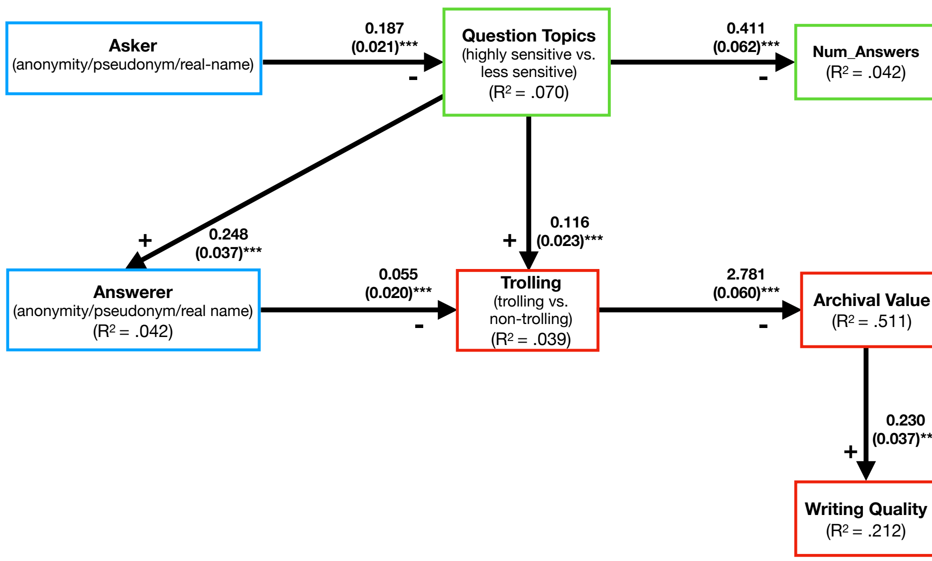


Fig. 7. The trimmed path model for the data of Zhihu. Significance levels: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. R^2 is the portion of the variance explained by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Identities are shown in blue. Question-level metrics are shown in green. Answer-level metrics are shown in red. Anonymity is coded as 1, pseudonym is coded as 2, and real name is coded as 3. Highly sensitive is coded as 1 and less sensitive is coded as 0. Trolling is coded as 1 and non-trolling is coded as 0.

On Zhihu, both asker's identity and answerer's identity have three levels (anonymity, pseudonym, and real name) ordered by the level of identity self-disclosure. Thus, we treat asker's identity and answerer's identity as two linear contrasts. Using this model, we find that on Zhihu, anonymous askers are more likely to ask questions about highly sensitive topics. Meanwhile, questions about highly sensitive topics are more likely to be answered anonymously than questions about less sensitive topics.

We do not find any direct effect of answerer's identity on answer quality (see Figure 4h and Figure 4i, p = .055 and .160 respectively). The indirect effect is mediated by trolling. We observed 163 (16.3%) trolling answers on Zhihu (see Table 5). Similar to Yahoo but unlike Quora, anonymous answers are more likely to be trolling. The trolling answers, in turn, have lower archival value than non-trolling answers. Similar to Yahoo and Quora, answers with higher archival value are more likely to have higher writing quality. Fifty-one point one percent of the variance of archival value and 21.2% of the variance of writing quality can be explained by this model.

Finally, both asker's identity and answerer's identity have no direct effect on user engagement metrics. For example, asker's identity has an indirect effect on the number of answers mediated by the topics of questions. Although the marginal effect is significant (see Figure 4g, $p < .001$), only 4.2% of the variance in the number of answers can be explained by this model. Following the same trend, asker's identity has a weak, indirect effect on other question-level user engagement metrics (e.g., number of question followers/comments, see Table 4). Answerer's identity has no significant effect on answer-level user engagement metrics (e.g., number of answer upvotes/comments, see Table 5). Thus, we removed these metrics (but kept the number of answers to represent user engagement) from the simplified trimmed model for simplicity. The full trimmed path model of Zhihu can be found in Figure 10 in the Appendix.

5 DISCUSSION

The data analysis illustrates that there are many common findings of the identity and Q&A metrics across sites and there are also quite a few different findings across sites. In this section, we explain these findings in terms of identity, user engagement and trolling.

5.1 Anonymity

Across the three Q&A sites we analyzed, anonymous askers are more likely to ask questions about highly sensitive topics. Anonymous questions are then more likely to be answered anonymously. We were not surprised to see people ask and answer questions anonymously about highly sensitive topics. However, before our investigation, we did not know whether it was the sensitivity of the question or the topic of the question that was related to the use and benefits of anonymity. It is possible that questions about highly sensitive topics can be less sensitive and vice versa, questions about less sensitive topics can end up being highly sensitive in certain cases. By analyzing the sensitivity of the questions, we show that the question topic is related to the sensitivity of the question (see Section 4.2). We also show that the use of anonymity is associated with both the topic of the question and the sensitivity of the question on Yahoo and Zhihu (See Sections 4.2, 4.3, and 4.5). On Quora, the use of anonymity is only associated with question topics but not the sensitivity. The identity policy varies across sites. Quora is the only real name enforced Q&A site we analyzed. Thus, users on Quora may be more sensitive to the use of anonymity. Users on Quora may be more likely to use anonymity to ask even less sensitive questions.

There are many varieties of Q&A sites and other forms of online communities and the context of the community itself can vary from highly sensitive to less sensitive. For example, Stack Overflow (https://stackoverflow.com/) is a popular technical Q&A site for software engineers and programmers. As a successful Q&A site [56], Stack Overflow does not provide any anonymity options for its users. Questions and answers on Stack Overflow are technical and therefore may be less likely to be perceived as sensitive. In an online community like Stack Overflow, the use of anonymity may not therefore provide benefits for users. On the other hand, in a more sensitive context (e.g., health), the situation is different. On an Q&A site dedicated to health issues, such as PatientsLikeMe, has more overall sensitive context. In this case, experiential information such as personal health conditions and personal health experience are often revealed and discussed in detail [49]. This information is often considered highly sensitive, personal, and useful. People's trust in these experiential health information is as high as their trust in factual health information, such as published medical papers [18]. A one-level anonymous or non-anonymous identity model may not fit for all. Previous studies on Whisper (an anonymous app) reveals that not all content needs similar levels of anonymity protection and guarantees [17]. By analyzing different topics and sensitivity levels within different Q&A sites, our results further reveal the importance of context-specific anonymity.

Our results suggest that online communities, especially those dealing in highly sensitive information, should provide a way for their users to engage anonymously. Moreover, the need of anonymity, especially in highly sensitive context is probably universal. The users of these sites may be different. For instance, they may have a different cultural background. People with different cultural backgrounds may behave very differently on Q&A sites [2, 63, 102]. The perceived risks may also vary between users. For example, Chinese users and U.S. users are under different threats from surveillance and loss of employment/opportunity [26]. These perceived risks may contribute to different use of anonymity. However, we still find that the way they use anonymity is similar on Q&A sites. For example, the ratio of anonymous questions across the three sites is very similar. Our study reveals that although people from different cultural backgrounds may behave differently on Q&A sites, at least, the way they use anonymity is similar.

Anonymity enables people to freely express views and can help control personal information disclosure [37]. Additionally, it helps reduce social costs such as the psychological barriers that deter solicitation for assistance [53]. Put another way, anonymity enables people to separate the content of their ideas and discussion from their identity, thus reducing many privacy consequences. People are looking for more and more health information (which is usually highly sensitive) online [91]. People may be less willing to use Q&A sites or other online health forums without having anonymity options to protect their privacy. People want to have the options of different levels of anonymity for each post [54]. The anonymity option allows for more negative valence [55], and writing negative aspects of someone's life online provides potential benefits [70]. People have a desire to be simultaneously involved in online communities, and anonymous. However, only a small number of online communities allow users to participate anonymously. When people want to be anonymous in those communities without an anonymity feature, they may engage in workarounds. For example, on a popular social news forum - Reddit, people use "throwaway" accounts as their temporary identities as a workaround to achieve anonymity [47]. On Reddit, users using throwaway accounts are significantly more likely to engage in seeking support for stigmatized context [4].

5.2 Engagement

In December 2013, the Huffington Post changed its identity policy to disallow anonymous comments. Commenters were forced to authenticate their accounts via Facebook, in turn revealing their real names. This change in identity policy resulted in a dramatic decrease in the number of comments [27]. Similarly, when TechCrunch and a site for U.S. soldiers changed their identity policies in the same way, the number of comments also decreased [39, 64]. However, a serious limitation to our interpretation of those findings is that they only focused on the difference of with and without anonymity options. Our results reveal that when users have options of identity, anonymity does not affect engagement directly. Although identity has a weak and indirect effect on user engagement, the effect is mediated by factors like question topic, answer quality, or trolling. Our findings shed new light on the relationship between different levels of identity and engagement, and suggest new research questions such as how different engagement from the community affects users' future identity decision making. Moreover, similar to the findings of anonymity, the findings of user engagement are consistent among the sites, even though the users may be different and have different cultural backgrounds.

Taken together, these results call into question the previously held belief that offering anonymity will result in loss of user engagement for Q&A sites and other online communities. Ideally, we suggest sites should offer anonymity thus enabling highly sensitive questions and discussions, while simultaneously reducing trolling through other measures to protect engagement.

5.3 Answer Quality

We find that different types of identity have no significant effect on answer writing quality or archival value across three Q&A sites. Previous work shows that changing the identity policy from allowing anonymity to only allowing real names results in an increase in online comment quality [27, 39]. However, our work, which is differentiated from prior work in that on the Q&A sites we analyzed, users have multiple identity options available simultaneously, whereas prior work only focus on comparing having anonymity vs. not having anonymity. Also, prior work mainly only focus on one site. Our results find the same trend across sites, even though the users may be different and from different cultural backgrounds. While services providers value anonymous contributions, they often have perceived threats about anonymity such as low-quality contributions to the community [57]. Our results reveal that the quality of the answers on Q&A sites is not affected by including an anonymity option. Because anonymity enables people to feel free to express views [37], it is possible that it may counter some of the negative things associated with other aspects of anonymity, therefore resulting in an equivalent quality of answers between anonymous and non-anonymous users.

5.4 Trolling

On Quora, we found very few trolling answers overall. When we did observe trolling, there was no difference in trolling between different types of identity. On Yahoo and Zhihu, however, anonymous answers were more likely to be trolling than non-anonymous answers. Deviant behavior such as trolling is a common issue in online communities [15], and we know from prior work that anonymity may provide cover for trolls. For example, the theory of online disinhibition describes a phenomenon where people may have a lack of restraint on the Internet. This kind of disinhibition could sometimes be toxic (e.g., trolling) [93]. In this theory, anonymity is one of the factors that could lead to toxic disinhibition. Toxic disinhibition is when the disinhibition someone experiences takes on a negative, potentially harmful nature. Trolling may be one example of toxic disinhibition. Similarly, in the social identity theory of deindividuation [82], people in groups (e.g., on a Q&A site) could lose selfhood and thus lose self-control over behavior. This control loss leads to behaviors like trolling. In this theory, anonymity is one of the factors that could lead to this control loss.

However, not all behaviors like trolling are negative. For example, these behaviors may be considered pro-social when the target has committed some offense [8]. Therefore, even if providing users of Yahoo and Zhihu with the option to ask or answer questions anonymously does enable some trolling behavior, that may not be an entirely bad thing if some of the trolling is pro-social. In an online community (e.g., a Q&A site), trolling requires a certain degree of commitment because trolling requires successful learning and assimilating of a community [20]. This learning, assimilating, and commitment could be beneficial to online communities as it could raise user engagement and the quality of the information users share. While it is outside the scope of this work, future work could explore how much anonymity contributes to the pro-social effects of trolling.

Yahoo, Zhihu, and Quora all moderate content to guard against trolling. For instance, Yahoo maintains the right to delete any content that violates its community guidelines [101]. Quora and Zhihu collapse answers which violate their policies [72, 108]. Quora and Zhihu also delete questions and answers according to their terms of use [76, 105]. One reason Quora has very few trolling answers may be due to its strict moderation mechanisms. Another possible reason may be the real name policy. By analyzing news comments from both anonymous authors and real-name authors, Santana finds that real name comments are more civil than anonymous comments in online newspaper discussion forums [85]. On Facebook Confession Boards, people ask taboo and stigmatized topics anonymously and receive relevant responses without having too much cyberbullying or negativity as the responder's identity is identifiable [7]. Real-name users may be more civil, and thus, the community norm of Quora may be more civil than other Q&A sites that are not a using real-name policy. The community norm could, in turn, affect people's online anonymity behavior [46].

Trolling is common in all online social systems [41]. Since even ordinary people can behave like a troll under certain circumstances [15], Yahoo, Zhihu, and other sites might address trolling using better moderation mechanisms. For example, sites may be able to use longitudinal data from user history to monitor reactions among suspect threads to better detect trolling posts [84].

6 LIMITATIONS AND FUTURE WORK

One major limitation of this work is that we were restricted by the characteristics of existing Q&A sites. As we described in the Section 3.1, we chose the three sites for their shared and unique characteristics. We also noted that there are multiple variables that differ across sites. For example, only one of our sites was in the Chinese language, whereas the other two were in the English language. Although we find that the results are pretty similar across three sites we analyzed, the primary users of these sites have clear language and cultural differences that may affect their behaviors. Furthermore, sites did not have identical policies: all three sites offered the option of asking and answering questions anonymously, whereas only one (Quora) had an enforced real name policy. Future work could examine this further by conducting experiments to control these variables. Another interesting area of future work would be to compare Q&A metrics before and after should any Q&A site change their identity policies. For example, starting in March 2017 (after we collected data), Quora started to review all anonymous content for spam and harassment before being distributed [73]. Starting in June 2017 (after we collected data), Zhihu started to enforce users to attach their mobile phone numbers (which are also attached to Social ID card) with their accounts [107]. These changes in identity policies might alter users' behavior and need further exploration.

Another limitation of this work is the dataset. We asked permission to crawl each of the three sites, but all sites declined or did not respond. Quora[1]is heavily restricted by robots.txt. We can't get all the information we need (e.g., the number of comments and followers) even with crawling. Zhihu does not provide APIs for data crawling. While Yahoo did previously provide an API to crawl, they took their API offline in 2014 [99]. Thus, we were not able to use APIs for data collection, nor did we crawl the sites, which might violate the sites' Terms of Service. Instead, we manually collected data. For some metrics (e.g, number of downvotes on Quora and Zhihu), Q&A sites have features that are not visible to users so we were not able to collect them without APIs. Without the API, our work is also limited by only being able to analyze the top-ranked questions. The low-ranked questions are almost impossible to collect manually. In order to do so, we need to collect the full set of questions. However, it is very difficult even with a web crawler. For example, using a web crawler, [97] was only able to collect about 58% of the questions during a one-month period on Quora. Since each question could have multiple topics, "getting the full set of all questions is difficult"[97]. Future large-scale studies would be facilitated if these Q&A sites would open their APIs to allow researchers to collect data at the question and answer level, rather than requiring researchers to collect these data manually, which limits the amount of data that can be efficiently collected.

Finally, similar to almost all other studies using public online forum data, our study was limited by sites' content moderation practices. As we discussed earlier, all three Q&A sites we analyzed have content moderation mechanisms [72, 76, 101, 105, 108]. Questions and answers may already have been removed before we collected and analyzed the data. Since we only collected questions that have answers, it is possible that some answers were removed by moderation because they violated sites' guidelines and policies prior to data collection. Thus, moderated portions of the Q&As were not accounted for in our analysis. However, we still found that the trolling on answers in different types of identity is consistent on Yahoo and Zhihu.

[1]Quora began allowing users crawl the site on December, 2017 [76], after we collected our data. However, Quora is still heavily restricted by Quora's robots.txt: https://www.quora.com/robots.txt

We found that on Yahoo and Zhihu, anonymous answers are more likely to be trolling; On Quora, there are very few trolling answers. Multiple factors could play a role in this finding. For example, Quora's strict moderation mechanism may have already moderated many of the trolling answers in real time, before we collected the data. It is also possible that Quora's real name policy leads to less trolling behavior. Other factors such as the demographics of Quora's user base could also be behind this finding. Since we were not able to collect raw data from these Q&A sites, we are not able to answer this question in this paper. Future work could examine this by collecting and analyzing both the raw and moderated data if possible. It is difficult for researchers to collect raw data before content moderation manually and may not even be possible in most cases because sites do not provide access to raw data to researchers. Q&A sites and other online communities should consider providing raw data via their APIs so more accurate data can be collected and analyzed by researchers. For example, Reddit now has a streaming API so researchers can collect all the comments that are posted on Reddit on a continuous basis [13].

7 CONCLUSION

By analyzing questions and answers from three Q&A sites that have different identity policies, we find that anonymity is important, especially for sensitive topics and questions. Based on this finding, we suggest Q&A sites and other online communities consider offering users ways to engage anonymously, especially about highly sensitive topics. Our results also show that offering anonymous identity options does not necessarily mean that Q&A sites and other online communities risk losing user engagement or content quality. However, anonymity leads to more trolling, and thus we recommend sites implement moderation mechanisms to guard against trolling.

ACKNOWLEDGMENTS

We would like to thank our anonymous reviewers for invaluable feedback on earlier versions of this paper. We also would like to thank Tory Garland, Brian Justice, Xiang Li, Yifang Li, Jordan Minoda, and Siva Likitha Valluru for their assistance on this work. Finally, we thank Dr. Bart Knijnenburg and Dr. Emily Sidnam-Mauch for their feedback on this work.

REFERENCES

[1] Mark S Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. 2013. Sharing knowledge and expertise: The CSCW view of knowledge management. *Computer Supported Cooperative Work (CSCW)* 22, 4-6 (2013), 531–573.

[2] Ifeoma Adaji and Julita Vassileva. 2017. Towards Understanding Users' Motivation in a Q&A Social Network Using Social Influence and the Moderation by Culture. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization.* ACM, 349–350.

[3] Eugene Agichtein, Carlos Castillo, Debora Donato, Aristides Gionis, and Gilad Mishne. 2008. Finding high-quality content in social media. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*. ACM, 183–194.

[4] Nazanin Andalibi, Oliver L Haimson, Munmun De Choudhury, and Andrea Forte. 2016. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, 3906–3918.

[5] Kohei Arai and Anik Nur Handayani. 2013. Predicting quality of answer in collaborative Q/A community. *Society and culture* 2, 3 (2013), 21–25.

[6] Peter M Bentler and Douglas G Bonett. 1980. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin* 88, 3 (1980), 588.

[7] Jeremy Birnholtz, Nicholas Aaron Ross Merola, and Arindam Paul. 2015. Is it weird to still be a virgin: Anonymous, locally targeted questions on facebook confession boards. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, 2613–2622.

[8] Lindsay Blackwell, Tianying Chen, Sarita Schoenebeck, and Cliff Lampe. 2018. When online harassment is perceived as justified. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 12.

[9] J Martin Bland and Douglas G Altman. 1997. Statistics notes: Cronbach's alpha. *Bmj* 314, 7080 (1997), 572.

[10] Amy Bruckman, Pavel Curtis, Cliff Figallo, and Brenda Laurel. 1994. Approaches to managing deviant behavior in virtual communities. In *Conference companion on Human Factors in Computing Systems*. 183–184.
[11] Amy Bruckman, Catalina Danis, Cliff Lampe, Janet Sternberg, and Chris Waldron. 2006. Managing deviant behavior in online communities. In *CHI'06 extended abstracts on Human Factors in Computing Systems*. ACM, 21–24.
[12] Amy S Bruckman, Casey Fiesler, Jeff Hancock, and Cosmin Munteanu. 2017. CSCW research ethics town hall: Working towards community norms. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 113–115.
[13] Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. 2018. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 32.
[14] Yan Chen, Teck-Hua Ho, and Yong-Mi Kim. 2010. Knowledge market design: A field experiment at Google Answers. *Journal of Public Economic Theory* 12, 4 (2010), 641–664.
[15] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2017. Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 1217–1230.
[16] Alton YK Chua and Snehasish Banerjee. 2013. So fast so good: An analysis of answer quality and answer speed in community Question-answering sites. *Journal of the Association for Information Science and Technology* 64, 10 (2013), 2058–2068.
[17] Denzil Correa, Leandro Araújo Silva, Mainack Mondal, Fabrício Benevenuto, and Krishna P Gummadi. 2015. The many shades of anonymity: Characterizing anonymous social media content. In *Ninth International AAAI Conference on Web and Social Media*.
[18] Jamie L Crawford, Cheng Guo, Jessica Schroeder, Rosa I Arriaga, and Jennifer Mankoff. 2014. Is it a question of trust? How search preferences influence forum use. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. 118–125.
[19] Robert A Cribbie. 2007. Multiplicity control in structural equation modeling. *Structural Equation Modeling* 14, 1 (2007), 98–112.
[20] Angela Gracia B Cruz, Yuri Seo, and Mathew Rex. 2018. Trolling in online communities: A practice-based theoretical perspective. *The Information Society* 34, 1 (2018), 15–26.
[21] Sally Jo Cunningham and Annika Hinze. 2014. Social, religious information behavior: An analysis of Yahoo! Answers queries about belief. *Advances in the Study of Information and Religion* 4, 1 (2014), 3.
[22] Jennifer L Doleac and Luke CD Stein. 2013. The visible hand: Race and online market outcomes. *The Economic Journal* 123, 572 (2013).
[23] Pnina Fichman. 2011. A comparative assessment of answer quality on four question answering sites. *Journal of Information Science* 37, 5 (2011), 476–486.
[24] Casey Fiesler, Alyson Young, Tamara Peyton, Amy S Bruckman, Mary Gray, Jeff Hancock, and Wayne Lutters. 2015. Ethics for studying online sociotechnical systems in a big data world. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work and Social Computing*. 289–292.
[25] Danyel Fisher, David W Mcdonald, Andrew L Brooks, and Elizabeth F Churchill. 2010. Terms of service, ethics, and bias: Tapping the social web for cscw research. *Computer Supported Cooperative Work (CSCW), Panel discussion* (2010).
[26] Andrea Forte, Nazanin Andalibi, and Rachel Greenstadt. 2017. Privacy, anonymity, and perceived risk in open collaboration: A study of Tor users and Wikipedians. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1800–1811.
[27] Rolf Fredheim, Alfred Moore, and John Naughton. 2015. Anonymity and Online Commenting: The Broken Windows Effect and the End of Drive-by Commenting. In *Proceedings of the ACM Web Science Conference*. ACM, 11.
[28] Yanbo Ge, Christopher R Knittel, Don MacKenzie, and Stephen Zoepf. 2016. *Racial and gender discrimination in transportation network companies*. Technical Report. National Bureau of Economic Research.
[29] Kevin A Hallgren. 2012. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology* 8, 1 (2012), 23.
[30] Claire Hardaker. 2010. Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions.
[31] F Maxwell Harper, Daniel Moy, and Joseph A Konstan. 2009. Facts or friends?: distinguishing informational and conversational questions in social Q&A sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 759–768.
[32] F Maxwell Harper, Daphne Raban, Sheizaf Rafaeli, and Joseph A Konstan. 2008. Predictors of answer quality in online Q&A sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 865–874.

[33] Susan Herring, Kirk Job-Sluder, Rebecca Scheckler, and Sasha Barab. 2002. Searching for safety online: Managing" trolling" in a feminist forum. *The Information Society* 18, 5 (2002), 371–384.

[34] Li-tze Hu and Peter M Bentler. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal* 6, 1 (1999), 1–55.

[35] Ni Huang, Yili Kevin Hong, and Gordon Burtch. 2017. Social Network Integration and User Content Generation: Evidence from Natural Experiments. *Management Information Systems Quarterly* 41, 4 (2017), 1035–1058.

[36] Dan Johnson. 2000. Anonymity and the Internet. *The Futurist* 34, 4 (2000), 12.

[37] Ruogu Kang, Stephanie Brown, and Sara Kiesler. 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*. ACM, 2657–2666.

[38] R Kaplan and E Yu. 2015. Measuring Question Sensitivity. *American Association for Public Opinion Research* (2015), 4107–4121.

[39] Peter G Kilner and Christopher M Hoadley. 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!* International Society of the Learning Sciences, 272–280.

[40] Soojung Kim, Jung Sun Oh, and Sanghee Oh. 2007. Best-answer selection criteria in a social Q&A site from the user-oriented relevance perspective. *Proceedings of the American Society for Information Science and Technology* 44, 1 (2007), 1–15.

[41] Ben Kirman, Conor Lineham, and Shaun Lawson. 2012. Exploring mischief and mayhem in social computing or: how we learned to stop worrying and love the trolls. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. ACM, 121–130.

[42] Rex B Kline. 2015. *Principles and practice of structural equation modeling*. Guilford publications.

[43] Bart P Knijnenburg, Martijn C Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 441–504.

[44] Stefan Krappitz. 2012. Troll culture. *Stuttgart, Germany: Merz Academy College of Design, Art and Media* (2012).

[45] Klaus Krippendorff. 2004. *Content analysis: An introduction to its methodology*. Sage.

[46] Stefan Larsson, Måns Svensson, Marcin de Kaminski, Kari Rönkkö, and Johanna Alkan Olsson. 2012. Law, norms, piracy and online anonymity: Practices of de-identification in the global file sharing community. *Journal of Research in Interactive Marketing* 6, 4 (2012), 260–280.

[47] Alex Leavitt. 2015. This is a throwaway account: Temporary technical identities and perceptions of anonymity in a massive online community. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 317–327.

[48] Raymond M Lee and Claire M Renzetti. 1990. The problems of researching sensitive topics: An overview and introduction.

[49] Elizabeth R Lenz. 1984. Information seeking: a component of client decisions and health behavior. *Advances in Nursing Science* (1984).

[50] Lei Li, Daqing He, Wei Jeng, Spencer Goodwin, and Chengzhi Zhang. 2015. Answer quality characteristics and prediction on an academic Q&A Site: A case study on ResearchGate. In *Proceedings of the 24th International Conference on World Wide Web*. ACM, 1453–1458.

[51] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. 2020. Towards A Taxonomy of Content Sensitivity and Sharing Preferences for Photos. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[52] Zhuolun Li, Ke-Wei Huang, and Huseyin Cavusoglu. 2012. Quantifying the impact of badges on user engagement in online Q&A communities. (2012).

[53] Haiwei Ma, Bowen Yu, Hao Fei Cheng, and Haiyi Zhu. 2019. Understanding Social Costs in Online Question Asking. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, LBW2521.

[54] Xiao Ma, Nazanin Andalibi, Louise Barkhuus, and Mor Naaman. 2017. People are either too fake or too real: opportunities and challenges in tie-based anonymity. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1781–1793.

[55] Xiao Ma, Jeff Hancock, and Mor Naaman. 2016. Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3857–3869.

[56] Lena Mamykina, Bella Manoim, Manas Mittal, George Hripcsak, and Björn Hartmann. 2011. Design lessons from the fastest q&a site in the west. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2857–2866.

[57] Nora McDonald, Benjamin Mako Hill, Rachel Greenstadt, and Andrea Forte. 2019. Privacy, Anonymity, and Perceived Risk in Open Collaboration: A Study of Service Providers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 671.

[58] Meredith Ringel Morris, Jaime Teevan, and Katrina Panovich. 2010. What do people ask their social networks, and why?: a survey study of status message q&a behavior. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1739–1748.

[59] Commission nationale de l'informatique et des libertés. 2013. Privacy Towards 2020. Retrieved January 15, 2021 from https://www.cnil.fr/sites/default/files/typo/document/CAHIER_IP_EN2.pdf.

[60] Heather L O'Brien and Elaine G Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology* 59, 6 (2008), 938–955.

[61] OECD. 1980. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Retrieved January 15, 2021 from https://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.

[62] OECD. 2013. OECD Privacy Guidelines. Retrieved January 15, 2021 from https://www.oecd.org/internet/ieconomy/privacy-guidelines.htm.

[63] Nigini Oliveira, Nazareno Andrade, and Katharina Reinecke. 2016. Participation differences in Q&A sites across countries: opportunities for cultural adaptation. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM, 6.

[64] Eli Omernick and Sara Owsley Sood. 2013. The impact of anonymity in online communities. In *2013 International Conference on Social Computing*. IEEE, 526–535.

[65] Robert M O'brien. 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41, 5 (2007), 673–690.

[66] Malte Paskuda and Myriam Lewkowicz. 2015. Anonymous Quorans are still Quorans, just anonymous. In *Proceedings of the 7th International Conference on Communities and Technologies*. ACM, 9–18.

[67] Sharoda A Paul, Lichan Hong, and Ed H Chi. 2012. Who is authoritative? understanding reputation mechanisms in quora. *arXiv preprint arXiv:1204.3724* (2012).

[68] Sai Teja Peddinti, Aleksandra Korolova, Elie Bursztein, and Geetanjali Sampemane. 2014. Cloak and swagger: Understanding data sensitivity through the lens of user anonymity. In *Security and Privacy (SP), 2014 IEEE Symposium on*. IEEE, 493–508.

[69] Dan Pelleg, Elad Yom-Tov, and Yoelle Maarek. 2012. Can you believe an anonymous contributor? On truthfulness in Yahoo! Answers. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*. IEEE, 411–420.

[70] James W Pennebaker. 1997. *Opening up: The healing power of expressing emotions*. Guilford Press.

[71] Luca Ponzanelli, Andrea Mocci, Alberto Bacchelli, Michele Lanza, and David Fullerton. 2014. Improving low quality stack overflow post detection. In *2014 IEEE International Conference on Software Maintenance and Evolution*. IEEE, 541–544.

[72] Quora. 2015. Why are answers on Quora collapsed? Retrieved January 15, 2021 from https://www.quora.com/Why-are-answers-on-Quora-collapsed.

[73] Quora. 2017. Upcoming Changes to Anonymity on Quora. Retrieved January 15, 2021 from https://productupdates.quora.com/Upcoming-Changes-to-Anonymity-on-Quora.

[74] Quora. 2017. Why were question details removed from Quora on August 3, 2017? Retrieved January 15, 2021 from https://www.quora.com/Why-were-question-details-removed-from-Quora-on-August-3-2017/answer/Quora-Official-Account.

[75] Quora. 2018. Do I have to use my real name on Quora? What is Quora's 'real names' policy? Retrieved January 15, 2021 from https://www.quora.com/Do-I-have-to-use-my-real-name-on-Quora-What-is-Quoras-'real-names'-policy.

[76] Quora. 2018. Terms of Service. Retrieved January 15, 2021 from https://www.quora.com/about/tos.

[77] Quora. 2020. What is an upvote on Quora? Retrieved January 15, 2021 from https://www.quora.com/What-is-an-upvote-on-Quora-1.

[78] Lee Rainie, Sara Kiesler, Ruogu Kang, Mary Madden, Maeve Duggan, Stephanie Brown, and Laura Dabbish. 2013. Anonymity, privacy, and security online. *Pew Research Center* 5 (2013).

[79] Stephen A Rains. 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 (2007), 100–125.

[80] Josyula R Rao, Pankaj Rohatgi, et al. 2000. Can pseudonymity really guarantee privacy?. In *USENIX Security Symposium*. 85–96.

[81] Sujith Ravi, Bo Pang, Vibhor Rastogi, and Ravi Kumar. 2014. Great question! question quality in community q&a. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 8.

[82] Stephen D Reicher, Russell Spears, and Tom Postmes. 1995. A social identity model of deindividuation phenomena. *European Review of Social Psychology* 6, 1 (1995), 161–198.

[83] Catherine A Roster, Lorenzo Lucianetti, and Gerald Albaum. 2015. Exploring slider vs. categorical response formats in web-based surveys. *Journal of Research Practice* 11, 1 (2015), D1–D1.
[84] Mattia Samory and Enoch Peserico. 2017. Sizing up the troll: A quantitative characterization of moderator-identified trolling in an online forum. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6943–6947.
[85] Arthur D Santana. 2014. Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice* 8, 1 (2014), 18–33.
[86] Karin Schermelleh-Engel, Helfried Moosbrugger, and Hans Müller. 2003. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online* 8, 2 (2003), 23–74.
[87] Pnina Shachaf and Noriko Hara. 2010. Beyond vandalism: Wikipedia trolls. *Journal of Information Science* 36, 3 (2010), 357–370.
[88] Chirag Shah, Vanessa Kitzie, and Erik Choi. 2014. Modalities, motivations, and materials–investigating traditional and social online Q&A services. *Journal of Information Science* 40, 5 (2014), 669–687.
[89] Chirag Shah, Jung Sun Oh, and Sanghee Oh. 2008. Exploring characteristics and effects of user participation in online social Q&A sites. *First Monday* 13, 9 (2008).
[90] Chirag Shah and Jefferey Pomerantz. 2010. Evaluating and predicting answer quality in community QA. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 411–418.
[91] Elizabeth Sillence, Pam Briggs, Peter Harris, and Lesley Fishwick. 2006. Going online for health advice: changes in usage and trust practices over the last five years. *Interacting with computers* 19, 3 (2006), 397–406.
[92] David L Streiner. 2005. Finding our way: an introduction to path analysis. *The Canadian Journal of Psychiatry* 50, 2 (2005), 115–122.
[93] John Suler. 2004. The online disinhibition effect. *Cyberpsychology & Behavior* 7, 3 (2004), 321–326.
[94] Scott Thacker and Mark D Griffiths. 2012. An exploratory study of trolling in online video gaming. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)* 2, 4 (2012), 17–33.
[95] Andrew W Vargo and Shigeo Matsubara. 2018. Identity and performance in technical Q&A. *Behaviour & Information Technology* (2018), 1–17.
[96] Jessica Vitak, Katie Shilton, and Zahra Ashktorab. 2016. Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing*. 941–953.
[97] Gang Wang, Konark Gill, Manish Mohanlal, Haitao Zheng, and Ben Y Zhao. 2013. Wisdom in the social crowd: an analysis of quora. In *Proceedings of the 22nd International Conference on World Wide Web*. ACM, 1341–1352.
[98] Sewall Wright. 1921. Correlation and causation. *Journal of Agricultural Research* 20, 7 (1921), 557–585.
[99] Yahoo. 2014. Yahoo Answers APIs Will Be Removed as of June 3, 2014. Retrieved January 15, 2021 from https://yahoodevelopers.tumblr.com/post/86260183503/ yahoo-answers-apis-will-be-removed-as-of-june-3.
[100] Yahoo. 2018. Oath Terms of Service. Retrieved January 15, 2021 from https://policies.oath.com/us/en/oath/privacy/products/ mediaservices/index.html#yahooanswers.
[101] Yahoo. n.d.. Report and appeal abuse in Yahoo Answers. Retrieved January 15, 2021 from https://help.yahoo.com/kb/SLN4217.html.
[102] Jiang Yang, Meredith Morris, Jaime Teevan, Lada Adamic, and Mark Ackerman. 2011. Culture matters: A survey study of social Q&A behavior. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 5.
[103] Shuyan Zhang. 1989. 姓氏人名用字的分析统计. (1989).
[104] Zhihu. 2015. 知乎上有多少话题？. Retrieved January 15, 2021 from https://www.zhihu.com/question/27875446.
[105] Zhihu. 2015. 知乎上的不友善内容会被怎样处理？. Retrieved January 15, 2021 from https://www.zhihu.com/question/24039476.
[106] Zhihu. 2017. 如何正确使用知乎上的「匿名身份」功能？？. Retrieved January 15, 2021 from https://www.zhihu.com/question/19581541.
[107] Zhihu. 2017. 知乎也要进一步强化帐号实名制了. Retrieved January 15, 2021 from https://zhuanlan.zhihu.com/p/27257715.
[108] Zhihu. 2018. 为什么知乎的部分回答会被折叠？. Retrieved January 15, 2021 from https://www.zhihu.com/question/20120168.
[109] Zhihu. 2018. 知乎用户信息管理规范是怎样的？. Retrieved January 15, 2021 from https://www.zhihu.com/question/19551824.
[110] Zhihu. 2020. 知乎的「赞同」和「感谢」功能，各在什么情况下使用才合适？. Retrieved January 15, 2021 from https://www.zhihu.com/question/19557420.
[111] Zhihu. n.d.. 个人认证. Retrieved January 15, 2021 from https://www.zhihu.com/question/48510028.

A FULL PATH MODELS

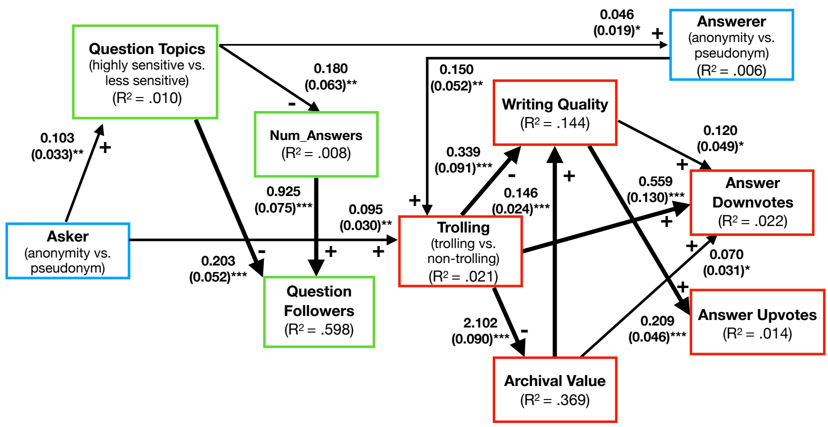


Fig. 8. The full path model for the data of Yahoo. The model fit is: $\chi^2(34) = 67.191$, $p = .001$; $RMSEA = 0.031$, 90% CI : [0.020, 0.042], $CFI = 0.987$, $TLI = 0.980$. Significance levels: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. R^2 is the portion of the variance explained by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Identities are shown in blue. Question-level metrics are shown in green. Answer-level metrics are shown in red. Anonymity is coded as 1 and pseudonym is coded as 0. Highly sensitive is coded as 1 and less sensitive is coded as 0. Trolling is coded as 1 and non-trolling is coded as 0.

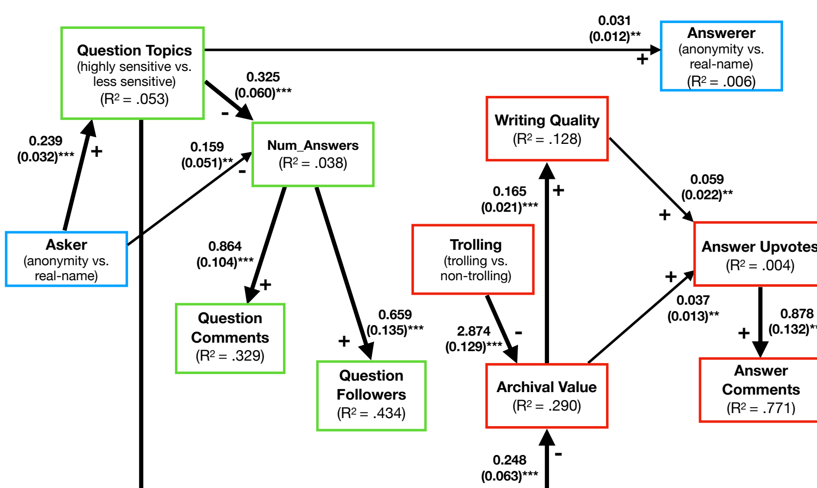


Fig. 9. The full path model for the data of Quora. The model fit is: $\chi^2(39) = 31.422$, $p = .801$; $RMSEA = 0.000$, 90% CI : [0.000, 0.018], $CFI = 1.000$, $TLI = 1.006$. Significance levels: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. R^2 is the portion of the variance explained by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Identities are shown in blue. Question-level metrics are shown in green. Answer-level metrics are shown in red. Anonymity is coded as 1 and real name is coded as 0. Highly sensitive is coded as 1 and less sensitive is coded as 0. Trolling is coded as 1 and non-trolling is coded as 0.

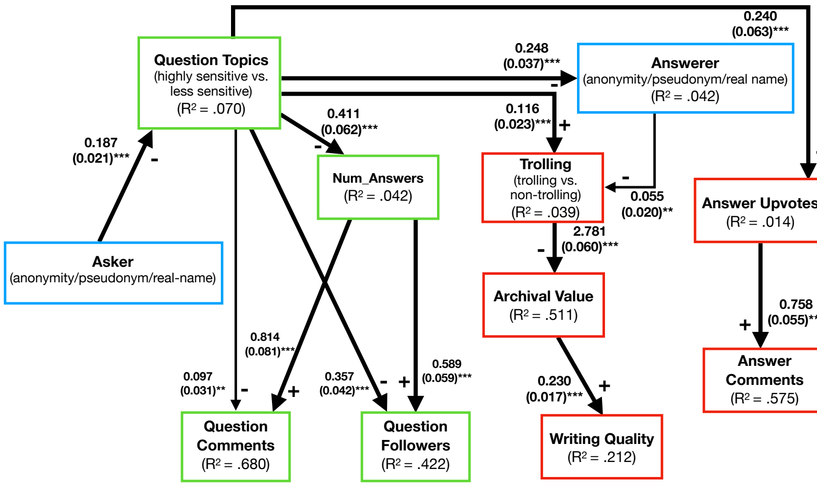


Fig. 10. The full path model for the data of Zhihu. The model fit is: $\chi^2(37) = 43.942$, $p = .201$; $RMSEA = 0.014$, 90% CI : [0.000, 0.026], $CFI = 0.996$, $TLI = 0.994$. Significance levels: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$. R^2 is the portion of the variance explained by the model. Numbers on the arrows (and their thickness) represent the coefficient (and standard error). Factors are scaled to have an SD of 1. Identities are shown in blue. Question-level metrics are shown in green. Answer-level metrics are shown in red. Anonymity is coded as 1, pseudonym is coded as 2, and real name is coded as 3. Highly sensitive is coded as 1 and less sensitive is coded as 0. Trolling is coded as 1 and non-trolling is coded as 0.

Received June 2020; revised October 2020; accepted December 2020