Effects of User Identity Information On Key Answer Outcomes in Social Q&A

Erik Choi
School of Communication &
Information (SC&I),
Rutgers, The State University of
New Jersey
erikchoi@gmail.com

Craig R. Scott
School of Communication &
Information (SC&I),
Rutgers, The State University of
New Jersey
crscott@rutgers.edu

Chirag Shah
School of Communication &
Information (SC&I),
Rutgers, The State University of
New Jersey
chirags@rutgers.edu

Abstract

Social Q&A (SQA) services have been growing in popularity among health information seekers. Even though research has paid much attention to a variety of characteristics of SQA services to investigate how people interact with each other for seeking and sharing information, the issues of identity and anonymity in these services that might relate to key user outcomes have been understudied. Such issues are especially important when dealing with stigmatized health conditions or sensitive health-related questions where choices are made about the revealing and concealing of identifying information in SQA environments. In the current study, we identified 110 stigmatized health questions from Yahoo! Answers that contained varied amounts and types of identity information corresponding to a framework developed in the study. We found that there are differences for providing personal contact information in one's profile when relating identity information in user profiles to identity information in user questions. Questions with a high amount of demographic information in questions tend to receive slightly higher average number of responses and take shorter time to receive the best answer for stigmatized health questions.

Keywords: social Q&A, health question-answering, identity information, self-disclosure, anonymity

Introduction

The Internet has become an important tool for seeking health information, with a majority of people (74%) finding health information online as opposed to using family and friends (12%) or contacting physicians (4%) to get health information (Hesse et al., 2005). In addition, the Pew Internet Project estimated that 8 million Americans seek health related information online a day and around 75% of U.S. Internet users search for health information online (Fox, 2006, 2008b). As a key part of that Internet search, social Q&A (SQA) services such as Yahoo! Answers (http://answers.yahoo.com[1] or WikiAnswers[2] have significantly increased in popularity over the past decade as an online information seeking methods where an asker's information needs are formed by natural language questions posed to other users who can answer the question or even offer feedback on the given responses (Choi, Kitzie, & Shah, 2012). According to the Hitwise report, the U.S visits to SQA have increased 889 percent from 2006 to 2008 (Tatham, 2008).

Due to popularity of SQA, people have started utilizing SQA to seek health-related information and health has become one of the most popular topics in these services (Oh, 2012). With the popularity of seeking and sharing health information in SQA, previous and current research has paid much attention to various topics related to health information such as information quality (Kim et al., 2008; Oh et al.,

---

[1] http://answers.yahoo.com
[2] http://wiki.answers.com

2011; Stvilia et al., 2009), motivations (Oh, 2012), and systemic approaches (Smedberg, 2007, 2008). Yet, there is a lack of research on how users disclose their own personal information in SQA environments and how that might be related to the responses one receives. This seems especially important in this context because several health topics come with stigma that may make them embarrassing or otherwise threatening to discuss. When dealing with stigmatized health conditions or sensitive health-related questions, choices about the revealing and concealing of identifying information (e.g., name, photo, gender, height, weight, age, location, disease, symptom, etc.) are especially important as users attempt to provide enough information to get answers but also protect their own privacy.

Thus, the purpose of this current study is to focus on the degree to which a SQA user's identity is disclosed and how that relates to various response characteristics in the Q&A process. In particular, the current study aims to address the following research questions:

RQ1: Does user information in a profile influence the amount and/or type of identity information in a question posted to SQA?

RQ2: How does amount and/or type of identity information in a question relate to the number of responses in SQA?

RQ3: How does amount and/or type of identity information in a user's question relate to the time to get the question resolved in SQA?

Related work

Social Q&A

SQA is a web based question-answering service where an asker poses a question and others provide responses to the given question for satisfying an asker's information needs. In other words, SAQ is a form of information retrieval where the users' information need is specified in the form of a natural language questions, and the desired result is self-contained answer (Bian et al., 2008). Shah, Oh, and Oh (2009) argued that research on online SQA can be divided into two major groups—user-based and content-based. For user-based studies, some research has conducted content analysis and divided the roles of users in two ways – seekers and sloths (Gazan, 2007).

The results show that more active users in SQA environments tend to receive more responses than sloths. Other work has attempted to propose a method for automatic identification of authoritative actors based on the number of best answers to given questions that are provided by authoritative actors (Wang, 2008). For content-based studies, Kim, Oh, and Oh (2007) analyzed comments that were left upon the selection of best answers and evaluated them by the best-answer selection criteria – content value, cognitive value, socio-emotional value, information source value, extrinsic value, utility, and general statement. Shah, Oh, and Oh (2008) studied the understanding of various characteristics of user participation; the study identified users in two distinctive roles – consumer and contributor in social Q&A sites. A recent study by Shah, Radford, Connaway, Choi, and Kitzie (2012) investigated why information-seeking questions fail in SQA and developed a typology for explaining why these informational questions failed to get answers.

Health Question-answering in Online Environments

The Pew Internet Center's Internet and American life project announced that approximately 75% of patients attempted to search health-related information online (Fox, 2008a), Additionally, around 40% of respondents in another study (Baker et al., 2003) reported that they use the Internet to look for advice or information about health or health care in 2001. Moreover, to investigate health information related content, a rich body of previous research has paid much attention to online health information quality (e.g., Berland et al., 2001; Bock et al., 2004, Donald, Lindenberg & Humphreys, 1998; Pastore, 2001, Zeng et al., 2004) while other research attempted to focus more on health-related information seekers or answerers.

For example, Gualtieri (2009) argued that some people intend to choose the Internet as their information resource over doctors in order for consultation on health issues, and adolescents sometimes prefer to find information and advice on stigmatized related health questions (e.g., pregnancy, sexuality, etc) from anonymous peers rather than close friends or family members (Suzuki & Calzo, 2004).

Moreover, Oh (2012) focused on how and why health answerers are motivated to answer those questions in order to share their information, knowledge, and experiences in SQA. That study found that the most influential factor for top answerers and health experts is altruism, whereas personal gain is the least influential factor for which top answerers and health experts share their knowledge, information, and experience with others. Unlike those previous studies, the research presented here focuses on identity information in user profile and user question when asking stigmatized health questions and also investigates how that information might be related to responses in SQA.

Identity Information for Self-disclosure

The existing research about identity information in online environments has primarily paid attention to identity sharing and privacy concerns in general (e.g., Gross & Acquisti, 2005; Krishnamurthy & Wills, 2008; Strater & Lipford, 2008; Stutzman, 2006). However, some previous research has examined Internet use and (sharing) identity information with regard to health-related questions. For example, Webber and Wilmot (2012) reported anonymous postings made on the Somazone [3] website where people share and seek information and advice about sexual assault or sexual coercion. The report classified typical comments types – validation, interpretation, condemnation of poster, self-disclosure, personal advice, medical advice, legal advice, resources, and humor; furthermore, young users perceive the site as a safe place where they can disclose themselves to receive professional nonjudgmental answers. Zhang (in press) focused on college students to study how they perceive and use the social networking sites for health and wellness information. The study revealed that 21% of participants responded that they would use the general health-related websites (e.g., WebMD) for getting health and wellness information rather than their existing social ties on Facebook because they do not want to broadcast their personal information to known people. In addition, other research (see Frost & Massagli, 2008, 2009; Brubaker, Lustig, & Hayes, 2010; Wicks et al., 2010) also specifically investigated PatientsLikeMe[4] in an attempt to study how users share their personal information and access to personal health data for acquiring relevant health information.

Anonymity

According to Marx (1999), anonymity is "one polar value of a broad dimension of identifiability versus nonidentifiability" (p.100) and identity knowledge is "an aspect of informational privacy" (p.100) that has multiple components such as legal name, locatability, pseudonyms linked or not linked to name/location, pattern knowledge, social categorization and symbols of eligibility/noneligibility. In computer-mediated communication (CMC), anonymity is often defined in two distinct ways (Qian & Scott, 2007): discursive anonymity concerns the withholding of one's personal information (name, email, gender, location, etc.) and visual anonymity concerns the absence of visual presentation (pictures, video clips, personally-identifying images, etc.) of people.

Moreover, Caspie and Gorsky (2006) found people are sometimes being anonymous to deceive their identity online because of privacy concern (21%) and Azechi (2005) argued that lack of personal specification as a type of anonymity causes communication to become more information-oriented and less social among people.

While previous research argues that being either visually and/or discursively anonymous online tends to create a more information-oriented environment, it is still unclear how different types and amounts of anonymity or identity information might have effects on getting information that satisfies an asker's needs online. Thus, in this research, we focus on investigating how types of identity information and amount of identity information relate to responses from other users in SQA. To do so, we did conduct a content analysis study as described in the next section.

---

[3] http://www.somazone.com.au/

[4] http://www.patientslikeme.com/

Method

Data Collection from Yahoo! Answers

400 resolved questions were collected by the Yahoo! Search Application Programming Interface (API)[5]. We looked at questions and answers in stigmatized health topics from a health category in Yahoo! Answers. Such keywords – abortion, depression, drug addiction, and STD were used to extract 100 resolved questions for each stigmatized health topic from the database.

Tourangeau and Smith (1996) aggregated previous research that focused on sexual behaviors and other sensitive topics in order to compare three methods of collecting survey data for sensitive questions; the study included drug use (Aquilino, 1994; Aquilino & LoSciuto, 1990), HIV risk factors (Locke et al., 1992), and abortion (London & Williams, 1990; Mosher & Duffer 1994; Mott, 1985). Moreover, other research argues that depression is one of significant topics in health informational seeking behaviors among teens (Boldero & Fallon, 1995; Dubow et al., 1990; Puskar, Tusaie-Mumford, Sereika, & Lamb, 1999; Borzekowski & Rickert, 2002).

Identifying Questions within the Framework of Amount and Type of Identity Information

To gain a better understanding of how different amounts and/or type of identity information in questions influences responses, we developed a framework for classifying both amount and type of identity information. Identity information includes profile data that reveals who an asker is and/or language in questions related to one's identity.

"Amount of identity information" can be divided into three levels: questions with no identity information, questions with medium identity information (2 identifiers), and questions with high identity information (4 or more identifiers). "Type of identity information" takes three forms for those questions containing identifiers: demographic only (e.g., age, gender, grade, location), medical only (e.g., diseases, symptoms, medical history), or both demographic and medical identity information (see Table 1).

Table 1
Example of each amount and type of identity information.

Identity information amount	*Example*
No information	*Does it hurt to have an abortion?*
Medium number	***my daughter is 14** can she take the Morning after Pill?*
High number	*Is this being weird or depressed? Maybe, is it even normal?? During school **I phase out a lot, can't concentrate...so I just sit there and have daydreams about all kinds of things** (I'll make up stories about me, not having so many problems in life, being happy, finding a great husband, travelling, etc.) And when **I get home I sleep a LOT**, I go straight to my bed, because I want to dream some more. I usually kinda recap my day, but in a more "creative" way. I sleep for hours during the day and **don't feel like eating or going out.** But the thing is...**I'm really really sad, and think about my death sometimes, what its like being hurt...and I have nightmares about the time I was raped** This wouldn't be normal would it, or what is it?*
Identity information type	
Demographic only	
Medium	*How do you take a drug addict to the rehabilitation centre?*

---

[5] http://developer.yahoo.com/answers/

	The boy is 20 years old *and will not listen to his mother. how do we help him*
High	*okay my parents are getting separated but not a divorce, and my dad just told me tonight that he already bought a house and at the end of this month hes going to move in. im okay with it-i knew it would happen but i cant stand to see my parents in this condition-sad, depressed, etc. Now i will have to take responsibility to take care of my mother, and especially **my little brother whos only 10. he wont understand. but what do i do? im just 14. i have an older sister who is 17, a senior and next year she will graduate...***
Medical only	
Medium	*how do you know your stressing? like idk if i have or have not been!But it do feel like it..My attitude sucks rite about now && i **feel bi-polar**(but im not) && i nned to know if im stressin or need help..**My head hurts alot** also && i also feel a lil depressed*
High	*I take **valiums** regularly to **fight my insomia**. Am I considered as a drug addict? I have a sleeping sickness called insomnia**. It's hard for me to fall asleep during night-time but I feel sleepy and lethargic during daytime.** My friend told me about Valiums. He said that it can help my sleeping problems. I gave it a try and it worked for me. Now, I found myself relying on Valiums to achieve that well-deserved good sleep. Am I addicted to it?*
Both demographic and medical	
Medium	*What can I do to help my sister who is a drug addict? **My sister is a drug addict. She has been for over 5 years**. She was raped while she was "high" on drugs and now has a son. She takes her son with her when she goes and buys drugs. She lives with our parents. She also has now stolen over $12,000 from our Grandmother, and our Grandmother is not going to press charges. They think it is okay to sweep this under the rug. But I'm fed up. I want to call CPS on her or something. What can I do?*
High	*When **i was 16** i got pregnant and had an abortion now some years have passed and i think i might be pregnant.? **I get irregular periods** like i always have and so ive **taken birth control pills for them since i was 14...** i went off for a year and started again and took em for a couple months but they were making me feel sick so i had to go for a lower dose but through all this me an my boyfreind continue to have sex no big deal but now **im feeling sick when i smell certain things an my nipples are really sore an ive been sleeping alot..** the only thing i can compare this to is my first pregnancy which happened right after i stopped taking my pills but i only had one symptom which **was dark spotting for a day none of this..** any help?*

(*Note* that the highlighted in questions are identifiers that fit within the framework)

Using this framework, content analysis of 400 stigmatized health question was conducted by the coder to identify certain questions that fit in the framework. In the results, 110 of the 400 questions (27.5%) were identified that contain the relevant amounts and types of identity information (see Table 2). These came from across the abortion (n=26), depression (n=39), drug addict (n=30), and STD subcategories (n=15).

Questions with either one or three pieces of identity information were excluded in order to make amount levels more distinct. Moreover, the study also manually collected user profile information for each asker such as user nickname, photo, personal contact information, and the 'about me' page. However, other user profile information such as 'points' and 'levels' (which encourage the users to participate in a variety of activities, e.g., selecting the best answer) were not included in this study since users are not able to show or hide that information for others.

Table 2
Number of each identity information amount and type.

Identity information amount	N
No information	19
Medium number	40
High number	51
Identity information type	
Demographic only	**19**
Medium	10
High	9
Medical only	**24**
Medium	14
High	10
Both demographic and medical	**48**
Medium	16
High	32
Total	*110*

Results

Relating User Profile Identity Information to User Question Identity Information

Users asking stigmatized health questions generally provide minimal identity information in their profile. More specifically, they primarily use 'no name or pseudonym' (n=69, 62.70%) for a user nickname, either 'no photo' or 'fake photo' (n=105, 95.45%), no 'contact information' (n=57, 51.80%), and no details on the 'about me' page (n=91, 82.70%). When relating identity information in user profiles to identity information in user questions, results show there are differences for providing personal contact information in one's profile (see Table 3).

First, there are differences in the amount of personal identity information in user questions relative to profile contact information, χ^2=6.707, df=1, p<.1. More specifically, users with a high amount of personal information in questions are slightly more likely to provide contact information in their profile (53%); users with a medium amount of personal information in questions provide contact information in their profile slightly less often (48%); and users with no personal identity information in questions are even less often including contact information in their profiles (37%).

Moreover, there are differences in the type of personal identity information in user questions relative to profile contact information, χ^2=21.339, df=1, p<.05. More specifically, users who provide only demographic identity information in questions tend to less often provide contact information in their profile (37%); users who use medical identity information in their questions provide contact information in their profiles more often provide personal contact information (54%) and users who include both demographic and medical identity information in their questions provide contact information in their profile even more regularly (58%). Except for contact information, there are no statistically significant differences between

other types of identity information in user profiles and the amount or type of identity information in user questions.

Relating User Question Identity Information to Number of Responses

Although statistical comparisons do not reveal significant differences given some of the small Ns in certain categories, it is worth looking at the descriptive for general trends in the data. Users with a medium amount of information in questions tend to receive slightly higher average number of responses (mean= 11.02, S.D.=10.05); users with a high amount of identity information (mean= 9.40, S.D.=6.87) or no identity information (mean=8.47, S.D.=7.03) received slightly less average number of responses.

More specifically, users with a high amount of demographic identity information in questions without any specific medical identity information received the highest number of responses (mean=13.11, S.D.=9.73), trailed by users who provide a medium number of medical identity information (mean=12.36, S.D.=12.38).

Table 3
Profile Information Relative to Identity Information Amount and Type in Questions

Information amount	User ID	Photo	Contact	About me
No information	No name/pseudonym (n=12, 63.16%) Partial name/full name (n=7, 36.84%)	No photo (n=11, 57.89%) Fake photo (n=7, 36.84%), Real photo (n=1, 5.26%)	No contact (n=12, 63.16%) Contact (n=7, 36.84%)	No about me information (n=17, 89.47%), About me information (n=2, 10.53%)
Medium amount	No name/pseudonym (n=24, 60.00%) Partial name/full name (n=16, 40.00%)	No photo (n=23, 57.50%) Fake photo (n=14, 35.00%), Real photo (n=3, 7.5%)	No contact (n=21, 52.50%) Contact (n=19, 47.50%)	No about me information (n=14, 73.68%), About me information (n=5, 26.32%)
High amount	No name/pseudonym (n=33, 64.70%) Partial name/full name (n=18, 35.30%)	No photo (n=28, 54.90%) Fake photo (n=22, 43.14%), Real photo (n=1, 1.96%)	No contact (n=24, 47.56%) Contact (n=27, 52.53%)	No about me information (n=42, 82.35%), About me information (n=9, 17.65%)
Information type	User ID	Photo	Contact	About me
Demographic only	No name/pseudonym (n=11, 57.89%) Partial name/full name (n=8, 42.11%)	No photo (n=14, 73.68%) Fake photo (n=5, 26.32%), Real photo (n=0, 0%)	No contact (n=12, 63.16%) Contact (n=7, 36.84%)	No about me information (n=14, 73.68%), About me information (n=5, 26.32%)
Medical only	No name/pseudonym (n=16, 66.67%) Partial name/full name (n=8, 33.33%)	No photo (n=12, 50.00%) Fake photo (n=11, 45.83%), Real photo (n=1, 4.17%)	No contact (n=11, 45.83%) Contact (n=13, 54.17%)	No about me information (n=18, 75.00%), About me information (n=6, 25.00%)
Both	No name/pseudonym (n=30, 62.50%) Partial name/full name (n=18, 37.50%)	No photo (n=25, 52.08%) Fake photo (n=20, 41.67%), Real photo (n=3, 6.25%)	No contact (n=20, 41.67%) Contact (n=28, 58.33%)	No about me information (n=42, 87.50%), About me information (n=6, 12.50%)
Total	*No name/pseudonym (n=69, 62.70%) Partial name/full name (n=41, 37.30%)*	*No photo (n=62, 56.40%) Fake photo (n=43, 39.10%), Real photo (n=5, 4.5%)*	*No contact (n=57, 51.80%) Contact (n=53, 48.20%)*	*No about me information (n=91, 82.70%), About me information (n=19, 17.30%)*

Moreover, in terms of information type, users with demographic identity information only received a higher numbers of responses (mean=12.42, S.D.=9.86) than users with only medical identity information (mean=10.58, S.D.=10.22) or with both demographic and medical identity information (mean=9.12, S.D.=6.64). The specific distribution of responses among different amounts and types of identity information that disclose an asker's identity can be seen in Table 4.

Table 4
Average number of responses among different numbers and types of personal information in stigmatized health questions on Yahoo! Answers

Information amount	Number	Average number of response	Std. Deviation
No information	**19**	**8.47**	**7.03**
Medium number	**40**	**11.02**	**10.05**
Demographic	10	11.80	10.46
Medical	14	12.36	12.38
Both	16	9.2	7.79
High number	**51**	**9.40**	**6.87**
Demographic	9	13.11	9.73
Medical	10	8.10	5.82
Both	32	8.79	6.10
Information type			
Only Demographic	19	12.42	9.86
Only medical	24	10.58	10.22
Both	48	9.12	6.64
Total	*110*	*9.82*	*8.18*

Relating User Question Identity Information to Best Answer Selection

In terms of time to best answer selection, users with no identity information in questions, although no significant differences given some of the small Ns in certain categories, generally took the shortest time (in hours) to receive the best answer for stigmatized health questions in Yahoo! Answers (mean=14.78, S.D.=11.02), followed by users with a medium amount of identity information (mean=16.51, S.D.=11.90) and users with a high amount of identity information (mean=17.88, S.D.=11.13).

However, the results show that users with a high number of demographic identity information seem to take the shortest time to select the best answer. Additionally, users with only demographics identity information took the shortest time to the best answer selection (mean=14.67, S.D.=14.31) than users with only medical identity information (mean=17.02, S.D.=8.51) or users with both demographic and medical identity information (mean=17.61, S.D.=10.43) as shown in Table 5.

Table 5
Average time to best answer selection among different numbers and types of personal information in stigmatized health questions on Yahoo! Answers

Information amount	Number	Average time (hour) to best answer selection	Std. Deviation
No information	**19**	**14.78**	**11.02**
Medium number	**40**	**16.51**	**11.90**
Demographic	10	15.75	19.62
Medical	14	16.78	9.13
Both	16	16.75	8.02
High number	**51**	**17.88**	**11.13**
Demographic	9	13.47	4.97
Medical	10	17.35	8.02
Both	32	19.29	12.92
Information type			
Only Demographic	19	14.67	14.31
Only medical	24	17.02	8.51
Both	48	17.61	10.43
Total	*110*	*16.85*	*11.35*

Discussion

The online environment continues to play a vital role in health-related information seeking and sharing activities. For various stigmatized health issues, people prefer to go online to seek health information with anonymous peers rather than meet face-to-face with known people (Suzuki & Calzo, 2004); however, whether sharing more personal identity information online influences the responses one receives has been understudied.

The results indicate that users asking such stigmatized health-related questions in SQA provide minimal identity information in their profile. It seems that online Q&A users who pose these health-related questions would rather provide their personal information in questions as needed than disclose identity in the profile for all to view. It may signify that discursive and visual anonymity that withhold not only personal information but also visual presentation (Qian & Scott, 2007) in online environments is normative—allowing people to participate in a variety of online interactions (Scott & Choi, 2012).

However, the only identity information in one's profile related to amount and type of identity information in questions is personal contact information. This finding revealed that users who do not intend to provide any specific identity information in a question are less likely to provide personal contact information in their profile, whereas users who willingly provide either demographic and/or medical identity information are more likely provide personal contact information. People may consider personal contact information as a less identifiable factor than other profile information (e.g., photo, name, etc) or even expect further communications between an asker and responders via email or instant messaging with regard to an asker's health issues. Yet, we do not know whether people actually communicate between an asker and responders via their personal modes of communication for health questions. Future studies should be done in order to analyze what motivates users include their personal contact information in user profile, or to investigate whether an asker and responders directly interact with each other via their personal contacts beyond question-answering processes in SQA.

Compared to face-to-face interactions, online Q&A environments as a form of computer-mediated communication (CMC) lack social context cues (Walther, 1996), which may hamper understandings of a sender's messages. This indicates that an asker may need to provide more information such as demographic and/or medical information when asking health-related questions in online environments for a better understanding of the current health issues and conditions. However, the results show that users with a medium amount of identity information received a relatively higher number of responses than users with a high amount of identity information with regard to an asker's stigmatized health question. This

signifies that providing excessive amount of identity information when asking a question in SQA would discourage responses rather foster interpretations regarding an asker's information needs. This finding also reflects that one of significant attribute of failure of fact-based questions (Shah et al., 2012) is that questions are complex; thus, questions that are too complicated and/or overly broad or provide excessive identity information are less likely to get a response because of that complexity. Moreover, these findings indicate that providing more identity information when asking stigmatized health questions does not necessarily mean better or clearer information.

One of the implications from these findings is that an asker may revise his/her question by reducing identity information that may be unnecessary and include only major identifiable information that helps responders gain a better understanding of an asker's current health issues and conditions. Another suggestion would be to redirect an asker to other online health services for seeking health information (e.g., PatientsLikeMe, WebMD[6], etc.) where people may interact with more professional health and medical related news and information. Further study should focus on comparisons between SQA and other online health services or forums to investigate how similar stigmatized health questions have either similar and/or different outcomes that satisfy an asker's information needs.

The study also found that users with demographic identity information in a question seem to receive relatively higher numbers of responses and have the shortest time to best answer selection as compared to those who provide medical identity information. This may signify that such demographic identity information is more likely objective and readable than medical information. It also suggests that how an asker describes his/her medical symptoms or conditions in stigmatized health questions is relatively subjective in that that people may describe a similar symptom in various ways by their own terms and expressions. This may hamper better understandings of responders on what an asker intends to ask, which may contribute to fewer responses and longer delays in selecting a best answer. Such subjective and ambiguous forms of information provided by health-related information seekers may limit opportunities to receive more responses from others in SQA environments. This result also supports other findings that ambiguity is one of the significant attributes of failed questions, which impairs better interpretation of the questions (Shah et al., 2012).

Another interesting finding is that users with no identity information (in terms of information amount) have shorter time to best answer selection (mean=14.78, S.D.=11.02). The questions with no identity information may not ask specific personal health issues or conditions; instead, they are more likely to ask the general issues in stigmatized health areas. The following example illustrates questions with no specific personal identity information in stigmatized health topics:
"Does it hurt to have an abortion?"

This question appears to look for more general knowledge about an abortion rather than specific advice or opinion about an asker's health issues, which means that an asker may select any response as the best answer if the response somewhat fulfills an asker's information need. In addition, users with either a medium and/or a high amount of identity information received a relatively higher number of responses than users with no identity information. As one way of accounting for this, Shah (2011) argues that a delayed answer may use earlier posed responses to the given question for a more comprehensive answer for an asker's information needs so that delayed answers could have the likelihood of being selected as the best answer.

Limitations

Even though the study presented here provided insights into the effects of identity information in a user profile and a user question on responses for stigmatized health-related topics, the study is limited in several ways. For example, to provide a distinction between different amounts of identity information, the study did not include questions with either one or three pieces of identity information (although such amounts can, of course, be found on these sites). We do not know exactly how those excluded amounts might be similar or different to the high/medium/low amounts in terms of their influence on the quantity of responses. In addition, because this exploratory study only analyzed 110 question-answer pairs, it is difficult to generalize from this small, non-random sample. For this reason, future study might focus on a

---

[6] http://www.webmd.com/

larger dataset to determine whether there exist significant differences between different amounts of identity information in stigmatized health questions.

This study is also limited in evaluating responses because it only focused on how amount and/or type of identity information in a question relates to the number of responses and the time to get the question resolved in SQA. It is still arguable whether more and faster answers are the better answers to the given question. Future study could focus on evaluating answer quality based on different amount and types of identity information in stigmatized health questions in SQA.

Conclusion

The current study investigated the degree to which an SQA user's identity is disclosed and how that relates to various response characteristics in the question-answering process. Different amounts and types of identity information in a question have been studied regarding stigmatized health questions in SQA. This work may be a useful early step in understanding disclosure of identity information during questioning-answering processes for people seeking professional opinions and advice for such stigmatized health-related questions. Findings from the current study about identity information amount and type shed light on how an asker discloses his/her identity information for a better understanding of his/her health issues and conditions in stigmatized health questions in SQA. An appropriate amount of objective and less ambiguous identity information helps responders understand what an asker looks for and encourages them to provide a response to his/her question.

This study also makes a contribution to anonymity and identity information in online environments by exploring how identifiable information that discloses an asker's identity relates to receiving responses from others and/or getting his/her information needs resolved in SQA. Previous research argues that lack of personal specification causes less social and more information oriented communication among people online (see Azechi, 2005). Yet, the findings in this study suggest that disclosing appropriate amounts and types of identity information should be considered when posing a question in order for better question-answering interactions for an asker's information needs. The current study attempted to focus on identity information in a question and investigated how amount and/or type of identity information relate to responses in SQA. Based on the findings regarding different amounts and types of identity information, further studies should be focused to address information seeking behaviors in various sensitive topics (e.g., income, politics, etc).

References

Aquilino, W. (1994). Interview Mode Effects in Surveys of Drug and Alcohol Use. *Public Opinion Quarterly, 58*, 210-40.

Aquilino, W., & LoSciuto, L. (1990). Effect of Interview Mode on Self-Reported Drug Use. *Public Opinion Quarterly 54*, 362-95.

Azechi, S. (2005). Information humidity model: explanation of dual modes of community for social intelligence design. Social Intelligence Design for Mediated Communication *19*(1), 110–122.

Baker, L., Wagner, T.H., Singer, S., & Bundorf, M.K. (2003). Use of the Internet and e- mail for health care information: Results from a national survey. *Journal of the American Medical Association, 289*(18), 2400- 2406.

Berland, G., Elliott, M., Morales, L., Algazy, J. et al. (2001). Health information on the Internet: Accessibility, quality, and readability in English and Spanish. *Journal of the American Medical Association, 285*(20), 2612-2621.

Bian, J., Liu, Y., Agichtein, E., & Zha, H. (2008). Finding the right facts in the crowd: factoid question answering over social media. In *Proceedings of WWW*. 467-476.

Bock, B., Graham, A., Sciamanna, C., Whiteley, J., Carmona-Barros, R., Niaura, R., & Abrams, D. (2004). Smoking cessation treatment on the Internet: Content, quality, and usability. *Nicotine Tobacco Research, 6*(2), 207-219.

Boldero, J., & Fallon, B. (1995). Adolescent help-seeking: What do they get help for and from whom? *Journal of Adolescence, 18*, 193–209.

Borzekowski, D. L. G., & Rickert, V. I. (2002). Adolescents, the Internet, and health: Issues of access and content. In S. L. Calvert, A. B. Jordan, & R. R. Cocking (Eds.), *Children in the digital age* (pp. 71–82). Westport, CT7 Praeger.

Bouguessa, M., Dumoulin, D., & Wang, S (2008). Identifying authoritative actors in question answering forums: the case of yahoo! answers. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 866–874.

Brubaker, J., Lustig, C., & Hayes, G. (2010). PatientsLikeMe: Empowerment and Representation in a Patient-Centered Social Network. *CSCW'10; Workshop on Research in Healthcare: Past, Present, and Future.*

Caspie, A. & Gorsky, P. (2006). Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior, 9*, 54–59.

Choi, E., Kitzie, V., & Shah, C (2012, October). *Developing a Typology of Online Q&A Models and Recommending the Right Model for Each Question Type.* Poster to be presented in American Society of Information Science & Technology (ASIST) Annual Meeting. Baltimore, Maryland.

Donald, A., Lindenberg, B. & Humphreys, L. (1998). Medicine and health on the Internet: The good, the bad, and the ugly. *Journal of American Medical Association, 280*(15), 1202-1206.

Dubow, E. F., Lovko Jr., K. R., & Kausch, D. F. (1990). Demographic differences in adolescents' health concerns and perceptions of helping agents. *Journal of Clinical Child Psychology, 19*(1), 44–54.

Fox, S. (2006). Online health search 2006. Washington, DC: Pew Internet & American Life Project. Washington, DC: Pew Research Center. Retrieved May 15, 2009, from http://www.pewinternet.org/Reports/2006/Online- Health-Search-2006.aspx

Fox, S. (2008a). New health data. Pew Internet & American Life Project. Retrieved from http://www.pewinternet.org/Commentary/2008/July/New-Health-Data.aspx

Fox, S. (2008b). The engaged e-patient population. Pew Internet & American Life Project. Washington, DC: Pew Research Center. Retrieved May 15, 2009, from http://www.pewinternet.org/Reports/2008/The-Engaged-Epatient-Population.aspx

Frost, J. H., & Massagli M. P. (2008). Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. *Journal of Medical Information and Research 10*(3):e15. doi: 10.2196/jmir.1053.

Frost, J. H., & Massagli M. P. (2009). PatientsLikeMe the case for a data-centered patient community and how ALS patients use the community to inform treatment decisions and manage pulmonary health. *Chronic Respiratory Disease 6*(4) 225-229.

Gazan, R. (2007). Seekers, sloths and social reference: Homework questions submitted to a question-answering community. *New Review of Hypermedia and Multimedia*, *13*(2), 239-248.

Gualitieri, L. N. (2009). The doctor as the second opinion and the Internet as the first. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems (pp. 2489-2489). New York: ACM Press.*

Gross, R., & Acquisti, A. (2005, November 7, 2005). *Information revelation and privacy in online social networks.* Paper presented at the WPES'05, Alexandria, Virginia, USA.

Hesse, B., Nelson, D., Kreps, G., Croyle, R., Arora, N., Rimer, B. et al. (2005). Trust and sources of health information. *Archives of Internal Medicine, 165*(22), 2618–2624.

Kim, S., Oh, J. S., & Oh, S. (2007). Best-Answer Selection Criteria in a Social Q&A site from the User-Oriented Relevance Perspective. In *Proceedings of the ASIST Conference,* 1-15.

Krishnamurthy, B., & Wills, C (2008). Characterizing privacy in online social networks. In *Proceedings of the first workshop on Online social networks*, 37-42.

Locke, S. E., Kowaloff, H. B., Hoff, R. G., Safran, CPopovsky, ., M. A., Cotton, D. J., Finkelstein, D. M., Page, P. L., & Slack, W. V. (1992). Computer-Based Interview for Screening Blood-Donors for Risk of HIV Transmission. *Journal of theAmerican Medical Association 29*, 1301-1305.

London, K., & Williams, L. (1990). *A Comparison of Abortion Underreporting in an In-Person Interview and Self-Administered Questionnaire.* Paper presented at the annual meeting of the Population Association of America, Toronto.

Marx, G. T. (1999). What's in a name? Some reflections on the sociology of anonymity. *The Information Society*, *15*(2), 99–112.

Lee, C. D., & Smith, A. B. (2009). Title of annual review chapter. In Lee, C. D. (Ed.), *Annual Review of Name of Science* (Vol. 17, pp. 172-183). Publisher Location: Publisher Name.

Mosher, W. D., & Duffer, A. P. Jr. (1994). *Experiments in Survey Data Collection: The National Survey of Family Growth Pretest.* Paper presented at the meeting of the Population Association of America, Miami.

Mott, F. (1985). Evaluation of Fertility Data and Preliminary Analytic Results from the 1983 Survey of the National Longitudinal Surveys of Work Experience of Youth. A report to the National Institute of Child Health and Human Development by the Center for Human Resources Research. Columbus, OH: Center for Human Resources Research.

Oh, S. (2012). The characteristics and motivations of health answerers for sharing information, knowledge, and experiences in online environments. *JASIST, 63*(3)*,* 543- 557.

Oh, S., Worrall, A., & Yi, Y. J. (2011). Quality evaluation of health answers in Yahoo! Answers: A comparison between experts and users. In *Proceedings of the ASIST Conference, 48*(1), 1-3.

Pastore, M. (2001, April). *Online Health Consumers More Proactive About Healthcare. CyberAtlas.* Retrieved January 7, 2002 from http://cyberatlas.internet.com/markets/healthcare/print/0,,10101_755471,00.html.

Puskar, K. R., Tusaie-Mumford, K., Sereika, S., & Lamb, J. (1999). Health concerns and risk behaviors of rural adolescents. *Journal of Community Health Nursing, 16*(2), 109–119.

Qian, H., & Scott, C. R. (2007). Anonymity and self-disclosure on weblogs. *Journal of Computer-Mediated Communication, 12*(4), 1-24.

Scott, C. R., & Choi, E. (2012, November). *Anonymity norms in corporate blogs: Relating commenter identity to comment characteristics*. Paper to be presented at the annual convention of the National Communication Association, Orlando, FL.

Shah, C., Oh, J. S., & Oh, S. (2008). Exploring characteristics and effects of user participation in online social Q&A sites. *First Monday, 13*(9).

Shah, C., Oh, S., & Oh, J. S. (2009). Research agenda for social Q&A. *Library & Information Science Research*, *31*(4), 205-209.

Shah, C., & Pomerantz, J. (2010). Evaluating and predicting answer quality in community QA. *SIGIR*, 411–418.

Shah, C (2011). Measuring effectiveness and user satisfaction in Yahoo! Answers. *First   Monday; 16*(2). February 7, 2011.

Shah, C., Radford, M., Connaway, L. S., Choi, E., & Kitzie, V. (2012, October). *"How Much Change Do You Get from 40$?" – Analyzing and Addressing Failed Questions on Social Q&A.* Paper to be presented to at the annual convention of American Society of Information Science & Technology (ASIST). Baltimore, Maryland.

Smedberg, A (2007). How to Combine the Online Community with Ask the Expert System in a Health Care Site. In *Proceedings of the first International Conference on the Digital Society, IEEE Computer Society Press.*

Smedberg, A. (2008). Enabling Cross-Usage of Public Health Systems - A Holistic Approach to Ask the Expert Systems and Online Communities. In *Proceedings of the Second Annual Conference on Digital Society*, 102 – 106.

Strater, K., & Lipford, H. R. (2008). Strategies and struggles with privacy in an online social networking community. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, 111-119

Stutzman, F. (2006). An evaluation of identity-sharing behavior in social network communities*.* Paper presented at the *iDMAa and IMS Code Conference*, Oxford, Ohio.

Suzuki, L. K., & Calzo, J. P. (2004). The search for peer advice in cyberspace: An examination of online teen bulletin boards about health and sexuality. *Journal of Applied Developmental Psychology, 25*(6), 685-698.

Stvilia, B., Mon, L., & Yi, Y. J. (2009). A Model for Online Consumer Health Information Quality. *JASIST*, *60*(9), 1781–1791.

Tatham, M. (2008). U.S. Visits to Question and Answer Websites Increased 118 Percent Year over-Year Yahoo! Answers receives 74 percent of all U.S. visits. Hitwise. Retrieved from http://www.hitwise.com/press- center/hitwiseHS2004/question-and-answer-websites.php

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions. The impact of data collection, mode, question format, and question context. *Public Opinion Quarterly, 60,* 275–304.

Walther , J. B. (1996). Computer-mediated communication: Impersonal, inter- personal and hyperpersonal interaction. *Communication Research. 23*(1) 3–43.

Webber, R., & Wilmot, J. (2012). Social Media and Help Seeking by Young People on Sexual Assault. Somazone Report Retrieved from http://www.acu.edu.au/Somazone_Report_Social-Media-and-Help-Seeking-by-Young-People-on-Sexual-Assault

Wicks, P., Massagli, M., Frost, J., Brownstein, C., Okun, S., Vaughan, T., Bradley, R., & Heywood, J. (2010). Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Information and Research 12*(2):e19. doi: 10.2196/jmir.1549.

Zeng, Q., Kogan, S., Plovnick, R., Crowell, J., Lacroix, E-M., & Greenes, R. (2004). Positive attitudes and failed queries: An exploration of the conundrums of consumer health information retrieval. *International Journal of Medical Informatics, 73*, 45-55.

Zhang, Y. (in press). College students' uses and perceptions of social networking sites for health and wellness information. *Information Research.*

Zhu, Z., Bernhard, D., & Gurevych, I. (2009). A Multi-Dimensional Model for Assessing the Quality of Answers in Social Q&A Sites. In *Proceedings of the 14th International Conference on Information Quality (ICIQ 2009),* 264–265.