

Why Users of Yahoo! Answers Do Not Answer Questions

David Dearman and Khai N. Truong

Department of Computer Science

University of Toronto

Toronto, Ontario, M5S 3G4, Canada

{dearman, khai}@cs.toronto.edu

ABSTRACT

Posing a question to an online question and answer community does not guarantee a response. Significant prior work has explored and identified members' motivations for contributing to communities of collective action (e.g., Yahoo! Answers); in contrast it is not well understood why members choose to not answer a question they have already read. To explore this issue, we surveyed 135 active members of Yahoo! Answers. We show that top and regular contributors experience the same reasons to not answer a question: subject nature and composition of the question; perception of how the questioner will receive, interpret and react to their response; and a belief that their response will lose its meaning and get lost in the crowd if too many responses have already been given. Informed by our results, we discuss opportunities to improve the efficacy of the question and answer process, and to encourage greater contributions through improved design.

Author Keywords

Question and answer, Q&A, community, motivation

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors

INTRODUCTION

Yahoo! Answers is an online question and answer (QA) community comprised of 120 million members and a searchable archive of over 400 million answers [9] – it is a place where people “*get real answers from real people*” [15]. Despite the size and diversity of its membership, posing a question to the community does not guarantee a response. Depending on the type of QA service, between 5-53% of questions do not receive a response [1, 6, 7, 12, 13]. Even so, a response does not guarantee an adequate answer [6, 7]. Members are free to answer questions at their own discretion and generally free to post any response they want.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.

Copyright 2010 ACM 978-1-60558-929-9/10/04...\$10.00.

A significant body of accumulated knowledge identifies a breadth of motivations for contributing to *communities of collective action* [4, 8, 11, 14]. Helping others is generally attributed to altruism, in that people “*enjoy helping others just as they enjoy being helped*” [10]. Kollock posits that community members can be motivated by: anticipated reciprocity; reputation building; a sense of efficacy in that their contribution has a positive effect; and a personal or community need for the information [8]. A framework of motivational factors for contributing to virtual communities by Moore and Serva [11] highlights the diversity of motivations observed in previous literature, citing: altruism, belonging, collaboration, egoism, egotism, emotional support, empathy, knowledge, power, reputation, self-esteem, self-expression, and wisdom. Although it is believed that an understanding of these reasons can be used to improve the efficacy of the question and answering process, many questions remain unanswered. It can be assumed that members of Y!A do not answer questions when they are not interested in the subject or are unable to provide an answer. However, these surface considerations do not provide a complete picture.

In this paper, we report on the results of an online survey asking active members of the Yahoo! Answers community why they choose to not answer a question they have read. Our results show that:

- Top and regular contributors do not answer questions for the same reasons;
- Questions that receive too many responses are less likely to be answered because it is likely the response “*will get lost in the crowd and won't be read*”;
- The respondent's perception of how the asker will receive, interpret and react to their response is important. Respondents do not want to get reported for abuse and potentially lose access to the community.

Informed by our results, we discuss opportunities to improve the efficacy of the question and answer process for communities of collective action, and to encourage greater user contribution.

EXPERIMENTAL METHOD

We conducted an online survey exploring why members of the Yahoo! Answers (Y!A) question and answering community choose not to answer a question they have read. Over a 15 week period (May 5, 2008 to August 11, 2008), we identified the top weekly contributors (determined by

points) posted on the public leader boards for the localized United States and Canada versions of the Y!A website. Each week we logged these top contributors and regular contributors who responded to the same questions, but are not top contributors themselves. Using the log, we randomly sampled 731 members (349 top; 382 regular) and contacted them using the email functionality built into Y!A. Members' email addresses are not publicly listed; as such we had to contact them through the community interface. Limitations imposed by Y!A limited us to contacting 10 members a day, until all the potential participants were emailed. The email invited the person to participate in the study and provided a personalized URL to the survey. Of the 731 potential respondents, 152 attempted the survey and 135 (100 top; 35 regular) completed it – a response rate of 18.5%. Participants were not compensated.

One hundred and thirty-five participants completed the questionnaire – 62 female and 73 male. The age of respondents varied: 18-25 (25), 26-35 (19), 36-45 (30), 46-55 (24), 55+ (34) and unreported (3). The majority reported reading (111) and answering questions (105) more than once a day or daily. No observed significant differences existed across the participant demographics or locality. Per week, respondents answer an average of 86.53 questions ($SD=119.08$).

Using an online questionnaire allowed us to gather data from a large number of participants who are difficult to research otherwise. The survey consisted of 14 short questions – eight closed-ended and six open-ended. The types of questions asked of participants included:

- demographics (e.g., gender, age, location, site usage)
- five reasons “*why [they] may not respond to a question posted on Yahoo! Answer after having read it*”
- how frequently (7-point Likert scale: 7-‘Always’; 1-‘Never’) when responding to a question do they “*know the answer*”, “*know part of the answer*”, and “*do not know the answer, but have the ability to find or otherwise provide an answer*”
- how frequently (7-point Likert scale: 7-‘Always’; 1-‘Never’) they “*answer a question that has already received responses*”, “*read the majority of responses to the question before I post my answer*”, and why

RESULTS

No significant difference was observed in the responses from the United States and Canada community members sampled. Similarly, there was no observable difference in the responses from the top and regular contributors. As such, we present the results together as a single community.

State of knowledge when question is first read

Presented in Table 1 are the participants' responses with respect to their state of knowledge when answering a question, immediately after reading it; a nonparametric Friedman Test revealed a significant main effect, $\chi^2_{(2, N=134)} = 57.19$, $p < .001$. Post-hoc pairwise comparisons using the Wilcoxon Test with Holm's sequential Bonferroni correction showed that participants most frequently answer a question

Table 1. The mean (M) and standard deviation (SD) for how frequently (1-‘Never’; 7-‘Always’) each statement reflects the participants' state of knowledge to answer a question, immediately after reading it.

States of Knowledge	M (SD)
I know the answer.	5.80 (0.78)
I know part of the answer, but not the full answer.	4.66 (1.40)
I do not know the answer, but I have the ability to find or otherwise provide and answer.	4.70 (1.45)

when they know the answer, rather than having partial, $p < .001$, or no knowledge, $p < .001$, of the answer.

Reasons not to answer questions

The analysis of the 595 open-ended responses for “*why [they] may not respond to a question posted on Yahoo! Answer after having read it*,” was completed using grounded theory affinity clustering [2, 5]. Two researchers created an affinity diagram by clustering similar responses in three iterations until consensus was reached on the seven primary categories (see Table 1). Each researcher then independently re-categorized all responses using the seven categories with a high level of observed agreement (0.987) and inter-rater reliability (Cohen's Kappa: 0.983, $p < .005$).

Participants reported overwhelmingly that their greatest reason for non-response is the **nature or content** of a question (M1). For example, the question may not be sincere, but rather “*touches [upon] inflammatory subjects where the asker seems to be stirring things up...*” The question may violate community guidelines by discriminating against particular groups or individuals, or involve illegal activities. Participants perceive such questions as “*a waste of time*,” and not worth their effort.

It is common for participants to “*answer a question that has already received responses*” ($M=5.21$, $SD=1.13$) and in doing so “*read the majority of responses to the question before I post my answer*,” ($M=4.47$, $SD=1.95$). They do so because they feel “*... other answers are incomplete or inaccurate*” in which case they “*can provide more detail to a solution ... or at least add more information*” that will be useful to the asker. They also believe that they often “*... understand the question better and have a better response*” or that the question is open to interpretation and their “*... answer may be different than someone else[s]*.” In contrast, the fact that a question has **already received responses** (M3) from community members is a strong reason not to respond, especially if “*someone has already answered [the question] correctly and completely*”, or if “*the question has been [asked and] answered a million times before*.” One participant commented that: “*When there are already good answers. I read all the answer [sic] as well as the questions and if I feel it has been answered ...[with] great answers I give them the thumbs up and move onto the next question.*” If they respond to a question that has received too many responses then their response “*will lose it's [sic] power or meaning*”, or “*get lost in the crowd and won't be read*.” They want to be recognized if they take the time to respond to a question.

Table 2. The participants' reasons for not responding to a question they have read. The responses were grouped into seven high-level categories using grounded theory affinity clustering. The counts (N) are intended to show the prevalence of comments, not to provide statistical evidence of their frequency. The percentages (%) are calculated for each category within a group.

Reason Categories	Top Contributors		Regular Contributors		Total (N)
	N	(%)	N	(%)	
M1. The nature or content of a question	177	(38.7)	53	(39.0)	230
M2. The responder's time, effort and expertise	79	(17.2)	28	(20.6)	107
M3. The question has already received responses	72	(15.7)	20	(14.7)	92
M4. The syntax of the question	69	(15.0)	21	(15.4)	90
M5. How the answer will be received and interpreted	24	(5.2)	6	(4.4)	30
M6. The questioner's time, effort and expertise	23	(5.0)	6	(4.4)	29
M7. The poster's characteristics, behaviour and history	15	(3.3)	2	(1.5)	17
Total	459		136		595

Participants indicated that they do not respond to a question if they are unsure **how the answer will be received and interpreted** by the asker (M5): *"Certain questions I don't reply to because I am afraid that if I express my personal opinion, I might offend someone."* Responders fear that the asker *"will just report [their] answer because [they] have a different opinion."* When a question is reported as abuse, Yahoo can refuse further access to the community or impose reprimands on the reported member. In addition, they perceive that some questioners lack common sense, such that *"no matter what you suggest by way of experience, they won't listen,"* the *"answer will fall on deaf ears/blind eyes."*

The **time, effort and expertise** required of the **responder** (M2) are important. Responders are unable and for the most part unwilling to answer a question when they *"do not know the answer"* or are *"not qualified to answer."* However, as supported by Table 1, the majority of participants indicated that they do answer some questions when they do not know the answer, but that others simply require *"too much effort to answer properly"*, or that they do not have the time to do the research for the questioner. Alternatively, the **time, effort and expertise** (M6) that the questioner has exerted or has the potential to exert to answer the question themselves is equally important. Some questions are trivial enough that the information can be easily found online by searching previous questions or a search engine. Participants commented that they expect users to have made some effort to find the answer on their own, *"people who are too lazy to do they[sic] own research are difficult to take seriously."*

The **syntax** of a question (M4) is important. A question cannot be answered if it is *"not understandable as written"* or *"makes no sense."* For example, questions may be written using *"TXT SPK"* or not grammatically and orthographically correct. Additionally, it can be the case that *"the question does not list enough information"* for an answer to be given or that *"the question is way to[o] long and too detailed,"* overwhelming the reader.

Interestingly, a minority of the participants have a strong understanding of the **characteristics, behaviors** (M7) and interaction **history** of certain other community members. For example, some users cheat, using *"numerous profiles to vote for themselves and give other answers thumbs down,"*

while other simply have *"shown a propensity to choose poorly thought out answers ... that are blatantly wrong."*

RESEARCH AND DESIGN OPPORTUNITIES

Community-based question and answer services provide access to diverse expertise that anyone with an Internet connection can access. The result of our survey of members of Yahoo! Answers (Y!A) reveals a number of research and design opportunities that we present in this section.

The Question Has Received Many Responses Already

Responders want to be recognized. Many believe that a new response to a question with many existing responses will simply *"get lost in the crowd."* To address this concern, visualizations should be developed to display accumulated responses creatively. For example, responses from top contributors could be elevated, prominently distinguishing these responses. Alternatively, responses can be grouped to give the perception of fewer answers while emphasizing agreement between responses. Grouping responses can also benefit the questioner because they can quickly differentiate between themes in the responses rather than reading them all.

Ensure the Response is Received Well

Responders are wary of answering questions where their response can be misinterpreted or misconstrued. A response could be interpreted as a passionate expression of a personal opinion by one member and as a personal attack by another. This is particularly true for questions that deal with sensitive subject matter and zealous ideals. Fear of losing access to the community is a strong reason not to respond. Community dialog and responses are typically accessible to all members, but a questioner's reactions to a response (e.g., reports of abuse) are not observable within the UI. Knowledge of the actions a questioner performs in response to an answer is valuable and can help a responder choose when to voice her opinion. If the potential responder knows a member flags a lot of questions as abuse, then she might avoid answering the person's questions or construct and present her response differently. The number of responses flagged as abuse or the question/responses themselves could be archived with the member's profile, giving other members the ability to make an educated decision as to how the member will react. Arguably this will require changes to how communities manage abuse and complaints, but it is important to ensure that crying 'wolf'

does not impact the community's efficacy or ostracize specific members.

Ensure Question Diversity While Targeting Interest

Targeting a responder's interests and expertise is important for eliciting a response and is the locus of attention with expertise systems. However, it is equally important to ensure diversity in the questions; participants do not like answering questions that "are repeatedly asked over and over again." In combination with targeting expertise and interest, community systems should ensure a consistent diversity in questions. Suggesting questions to a user that are dissimilar to what she has answered in the past, but appropriate to her. Additionally, there is benefit to showing a questioner while they are composing their question if similar questions have been asked recently and if so, the corresponding answers. Simple feedback can show the questioner there is no need to ask the question; highlight the need to differentiate their question from previous ones; or help him realize that the question is problematic because it has been asked by others and not answered properly or at all.

Question Complexity & Length

Questioners should craft their questions to include all relevant information so the reader can make an informed choice; however, they must do so intelligently. Participants avoided questions that are too long and complex: "I do not respond to postings that are over supplied with questions." Poorly written questions convey sloppiness on the part of the questioner. This issue could be addressed by providing feedback about the length and complexity of a question before it is posted, allowing the questioner to restructure appropriately. For example, the interface could disclose how many questions of a similar length or complexity are answered, and how many responses they receive. If the question is long or complex, the interface could suggest breaking it into individual questions and asking them separately, or perform this action on the questioner's behalf.

CONCLUSION

Online communities such as Yahoo! Answers (Y!A) provide access to a diverse set of knowledge and expertise that may not be available to an individual otherwise. We surveyed 135 members of Y!A to identify why they choose to not answer a question they have read. Top and regular contributors choose to not answer questions for the same reasons: the nature and composition of the question; their perception of how the asker will receive and interpret their response; and the belief that their response will simply get lost in the crowd if there have been too many responses already. Informed by our results, we discussed opportunities to improve the efficacy of the question and answering process. The goal of this work is to identify challenges preventing every question from being answered. However, as discussed in this paper, getting more responses is not necessarily a desired effect. Instead, an interesting research and design challenge is to come up with ways that ensure questions are positioned to receive at least one useful answer.

ACKNOWLEDGEMENTS

We thank Gillian Hayes and Elaine Huang for their thoughtful feedback on earlier versions of this paper.

REFERENCES

1. Adamic, L.A., Zhang, J., Bakshy, E., and Ackerman, M.S. Knowledge sharing and yahoo answers: everyone knows something. In *Proc. WWW 2008*, ACM Press (2008), 665-674.
2. Beyer, H. and Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*. 1998, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
3. Engelbart, D.C. Augmenting human intellect: A conceptual framework. Stanford Research Institute Summary Report on Contract AF 49(638)-1024, 1962.
4. Forte, A. and Bruckman, A. Why do people write for wikipedia? Incentives to contribute to open-content publishing. In *Proc. GROUP 2005 workshop: Sustaining community*, ACM Press (2005).
5. Glaser, B.G. *Basics of Grounded Theory Analysis: Emergence Vs. Forcing*. 1992, Mill Valley, CA, USA.
6. Harper, F.M., Raban, D., Rafaeli, S., and Konstan, J.A. Predictors of answer quality in online Q&A sites. In *Proc. CHI 2008*, ACM Press (2008), 865-874.
7. Janes, J., Hill, C., and Rolfe, A. Ask-an-expert service analysis. *Journal of American Society for Information Science and Technology* 52, 13 (2001), 1106-1121.
8. Kollock, P. The economies of online cooperation: Gifts and public goods in cyberspace, in *Communities in Cyberspace*, M. Smith and P. Kollock (1999), Routledge. London, 220-239.
9. Leibenluft, J. A librarian's worst nightmare: Yahoo! Answers, where 120 million users can be wrong, *Slate Magazine* 2007.
10. McMillan, D.W. and Chavis, D.M. Sense of community: A definition and theory. *Journal of Community Psychology* 14, 1 (1986), 6-23.
11. Moore, T.D. and Serva, M.A. Understanding member motivation for contributing to different types of virtual communities: A proposed framework. In *Proc. SIGMIS-CPR 2007*, ACM Press (2007), 153-158.
12. Rafaeli, S., Raban, D.R., and Ravid, G. How social motivations enhances economic activity and incentives in the Google answers knowledge sharing market. *International Journal of Knowledge and Learning* 3, 1 (2007), 1-11.
13. Su, Q., Pavlov, D., Chow, J.-H., and Baker, W.C. Internet-scale collection of human-reviewed data. In *Proc. WWW 2007*, ACM Press (2007), 231-240.
14. Wasko, M.M. and Faraj, S. "It is what on does": why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems* 9 (2000), 155-173.
15. Yahoo, Yahoo! Answers, <http://answers.yahoo.com>