Taylor & Francis
Taylor & Francis Group

Technical Note

Seekers, sloths and social reference: Homework questions submitted to a question-answering community

R. GAZAN*

Department of Information and Computer Sciences, University of Hawaii at Manoa, 1680 East West Rd. POST 314D, Honolulu, HI 96822, USA

An increasing number of students are seeking homework help outside library structures and systems, such as on social reference sites, where questions are answered by online community members who rate one another's answers and provide collaborative filtering in place of traditional expertise. This paper reports the preliminary results of a participant observation and content analysis of homework questions submitted to Answerbag, a social reference site with over one million unique visitors per month. The results suggest that members of the online community are able to distinguish between questions submitted by Seekers—those who interact with the community and engage in conversation about their questions—and Sloths, those who post their homework questions apparently verbatim and interact no further. How the community reacts to these distinct types of questioners reflects values similar to those of professional reference providers, and the community structure also allows members to educate questioners about community standards and the ethics of information seeking.

1. Introduction

The goal of educational technologies is to help students learn. But what constitutes appropriate help? While there has been considerable attention in the information science literature focused on the relationship between digital technologies and educational outcomes, educational technology research generally takes as its object of study a library-centric or classroom-centric system, be it a digital library, online catalogue, or courseware suite. This paper will explore two interrelated issues that warrant closer investigation: how some students broaden the definition of educational technologies to encompass collaborative, social reference websites outside the purview of the library, and the normative question of how much help educational technologies should provide. These questions will be addressed via the preliminary results of a case study of homework questions submitted to Answerbag, a Web-based question-answering community, and how the community's aggregate reactions to homework questions compare with standards of professional reference service.

---

*Email: gazan@hawaii.edu

At many levels of education, evidence of student understanding is indistinguishable from simply getting the right answer on the page. While educational technologies hold the promise of enabling a deeper and more productive journey to those answers, their costs go well beyond the financial, and for the less motivated student, effectively restrict use. There are access paths, interfaces, and vocabulary to learn, time spent forming at least a serviceable mental model of how it works, as well as the important yet often overlooked fact that engaging with the system is almost always in the context of work, as opposed to discovery or entertainment. If there is a more time-efficient alternative that yields a sufficiently good answer, people will use it, even if the alternative wasn't designed for the task. Student use of Wikipedia for research papers is the current canonical example. From the perspective of the information seeker, a less restrictive information resource, one that can be used for information as well as entertainment and social interaction, may be an attractive supplementary—or primary—alternative for homework help.

2. Background

Despite being on the front lines of providing homework help to students, librarians do not have an overarching ethical policy on what constitutes too much help (ALA 2007). Some researchers (see, for example, Wengwert 2001) argue that librarians have operated in an environment of limited information resources and limited access pathways for so long that the ethical focus of the profession has been on minimizing restrictions, whereas the modern information environment requires that hard decisions about appropriate conditions for filtering often must be made. In the absence of a consistent ethical framework, standards and best practices can reveal and shape the professional conception of what constitutes appropriate homework help. At a higher level, librarians are trained to approach any reference transaction by considering both the request (what they want) and the user (why they want it) in tandem (Saxton and Richardson 2002).

Walter and Mediavilla (2005) analysed transcripts of an online homework help service for teenagers, and found that the reference transactions they studied were "severely lacking in the qualities required for effective reference service," when measured against standards from the ALA's Reference and User Services Association. They suggest meeting users halfway and engaging in informal chat conventions, de-emphasizing their professional expertise to create a more hospitable virtual reference environment. By so doing, librarians would not be breaking new ground, but rather catching up to how users are already most comfortable seeking information: from peers.

Train (2005) discusses how school library media specialists should prepare students to use live Internet reference services, and it is by no means a cost-free interaction:

> Students should have a clear idea of what they are looking for and should be prepared to tell the librarian where else they have looked and what they need that isn't on the library's Web page, in their textbooks, and in other books

they've consulted. They should be expected to compare what they receive from
the librarian with what they already have. (Train 2005, p. 20)

However, this ideal virtual reference interaction is rarely reached, even with
systematic preparation of professional librarians. Some literature suggests
that neither reference librarians (Gross 2000) nor students (Meyers 1999) feel
that the other group is particularly engaged in the 'homework help' reference
transaction.

The phrase 'communities of practice' refers to groups of people who share a
concern or a passion about a topic, and who deepen their knowledge and
expertise by interacting on an ongoing basis (Wenger *et al*. 2002). Learning
has long been conceptualized as social in nature, with newcomers engaging in
legitimate peripheral participation (Lave and Wenger 1991)—or entirely
passive 'lurking' in the Web world (Barnett 2000)—before fully engaging with
the community. How these processes play out in online communities has been
widely studied (see, for example, Wasko and Faraj 2000), and sheds light on
the intrinsic value people derive from participating, which has not been
commonly emphasized in system-centred educational technology research.

Van Scoyoc and Cason (2006) found that undergraduates relied primarily
on Internet sites for research, rather than university-funded resources.
Further, while the authors expected to find that upper-division under-
graduates would use library-based resources more frequently as they grew
more familiar with library offerings and had more demanding assignments,
no significant difference between the use of Internet resources for research
between lower division and upper division undergraduates was found.

Similarly, Tenopir (2003) reports that college and high school students use
the Internet more than the library for research, and that even graduate
students and faculty members tend to value electronic resources that
are convenient, and support their 'natural work patterns.' From the
perspective of an individual's total information environment, what constitutes
a natural work pattern changes with technology, familiarity, and custom.
What once called for a personal letter could later be handled with a phone call
or fax, then an e-mail, and now perhaps a text message. A journal article
database that has been designed and accessed in roughly the same way for a
decade or more might seem slow and clunky compared with information
exchange in a chat window or handheld device. Social reference websites fit
more smoothly into students' natural use patterns, and have far lower barriers
to entry than digital libraries or other educational technologies.

Though the familiar echoes of the principle of least effort (Zipf 1949) are
present here, Xu *et al*. (2006) have challenged the dominance of the
accessibility variable in users' choice of information sources, and argue for
a more integrative view: that while accessibility may be the dominant factor in
how users choose between different information source types (e.g. documents,
people, media), once a source type is chosen, source quality is the dominant
factor. A question-answering site that contains a collaborative rating and
filtering mechanism—essentially the aggregate opinion of large numbers of
strangers—might provide a satisfactory assurance of quality.

Throughout the literature, there is a common thread of library-centric educational collections and technologies as just one part of users' overall information environments, and often not the primary part. Also, while general reference standards exist, there are few clear guidelines for appropriate homework help, let alone how these norms should be communicated and enforced. As users so often take existing technologies and use them in ways never intended by their designers, an investigation of homework questions submitted to a Web-based question answering community is warranted.

3. Setting and method

Answerbag (http://www.answerbag.com) is an online question answering community of roughly 100 000 registered users, which at this writing draws over one million unique visitors per month. Users submit questions, answers and answer comments in a nearly limitless variety of categories, and rate the questions and answers of others positively or negatively on a sliding points scale. The more points a user has accumulated, the more they can award or penalize, from plus or minus one point for beginners to plus or minus six points for the most advanced users. Multiple answers to a question are permitted, and the highest-rated answers in terms of overall points are listed first, providing collaborative filtering while still allowing users to browse the range of different answers. Participants with the highest ratings in various categories have their screen names and statistics posted on the site. Answerbag is both a public website and a research testbed, and administrator-level access to all real-time and legacy site data was readily available.

An increasing number of websites use a similar collaborative architecture, where content and value are created by aggregate participation and opinion—these sites are usually described with a term such as social computing or Web 2.0. With the explosion of user-tagged content found in sites like flickr (http://www.flickr.com/) for photographs and del.icio.us (http://del.icio.us) for websites, there is now a controlled term in the Library and Information Science Abstracts (LISA) database for Social bookmarking (i.e. aggregate peer cataloguing). However, at this writing, there is not yet a corresponding term for Social reference (aggregate peer question answering).

In Answerbag's first few years, the volume of submissions was low enough that a few site moderators could review every question and answer submitted. But as Answerbag grew, the number of daily submissions increased a hundredfold. The only way to vet content was to rely even more on the Web 2.0 model, and give users the tools not only to submit questions and answers but to rate and flag the content of other community members for review by site moderators as well. The social reference model has been echoed by Yahoo! Answers (http://answers.yahoo.com), and intrigued enough users that question answering now takes place virtually in real time. Answerbag now features a dynamic timer on the home page, tracking how quickly the last submitted question drew an answer, usually less than 1 min. Users can post a

question and be alerted via e-mail or RSS feed when an answer is posted, freeing them to multitask.

The motivation of questioners to use social reference sites like Answerbag is clear. But why would people provide answers for free? Those who have strong experience or interest in a topic have a forum in which to share it, in an environment where value is determined on the strength of content—not a job title or formal credentials—and expressions of approval and appreciation are central to the function of the site. Some users have commented that the answer ratings and comments they receive from a few hours a week on Answerbag outstrip those they receive over months at their full-time jobs. The points system creates virtual social capital for individuals' online identities, and even though they are most commonly masked behind a handle or other pseudonym, people nurture and defend them, take pride in their growth, and value being part of a community of answerers.

So, with an enthusiastic, worldwide question-answering community, where good contributions are rewarded, it creates a situation for homework-burdened students that is reminiscent of Tom Sawyer (Twain 1876) feigning interest in whitewashing a long, long fence to entice other boys into helping him. A risk inherent in any commons is free-riding, and while Answerbag was not created to be a library-quality reference service, a free, speedy, and anonymous question-answering site has proven tempting for many lost or lazy students. Though perhaps not as craftily as Tom Sawyer in convincing others of the attractiveness of doing their work for them, questioners can offer the implicit promise of good ratings and comments to answerers—the coins of the Answerbag realm—as well as the intrinsic satisfaction of a question well answered.

4. Identifying homework questions

For the purposes of this study, a homework question was initially defined heuristically, as a submission to Answerbag that, by its content, format, category placement, or submitter's comments, appeared to be taken directly from a school assignment, whether from grade school, Sunday school, or graduate school. Candidate homework questions were identified by reviewing user flags and feedback over the course of 15 months, performing site searches on common homework question terms such as "discuss," "analyse," and "explain," as well as mathematical symbols and expressions, and browsing categories such as Education, where homework questions are often submitted. Candidate homework questions were marked for further analysis, and the ratings, comments, and flags attached to these questions were tracked.

Answerbag employs paid and volunteer moderators to review content that has been flagged by other users as potentially miscategorized, nonsensical, spam, offensive, or generally inappropriate for the site. Roughly 500–800 flags are set per day. Users can also submit e-mail feedback to report inappropriate ratings, submission patterns or other rogue behaviours (Gazan 2007) that do

not fit into the flag structure. Combining all the channels, 1317 candidate homework questions were identified over a 15-month period. After further individual analysis, the sample was refined to include only those questions which were most clearly derived from homework assignments. Some examples of these high-confidence homework questions include the following, and are presented exactly as submitted, including grammatical errors:

- Explain Activities involved in Logistic management
- Who is the inspiring sister who won the Nobel Peace proze for her work with the poor of India?
- Discuss the major process of educational administration and illustrate your discussion with actual secondary school activities
- At 25 MPH, it will take you about _____ feet to stop your car A. 25 B. 62 C. 144
- The Sun's average density is almost exactly the same as the average density of?
- I have to analyse the strategies used in arguments regarding current events. Is there any website with a collection of current events and supporting arguments of the opposing views?

A total of 325 candidate homework questions were removed from the sample after closer analysis yielded no strong evidence that they had been taken directly from homework assignments—even though they may have been. Examples of these lower-confidence homework questions include:

- Do mathematical objects (circles, sets, numbers, vector spaces, fractals, etc.) exist independently of the human mind? Or do they only exist as a result of humans inventing them?
- Words that contain little words inside

The remaining 992 high-confidence homework questions were analysed in terms of number of answers received, ratings received, and content analysis of question text, answers, and comments. User profiles, postings, and site logs were also analysed to determine the subsequent activities of the question submitters. Some data, such as the time of posting, were ignored because the site draws users from around the globe, and often users in different time zones interact on the site in real time, asynchronously, or both.

5. Seekers and sloths

The analysis yielded eight homework question characteristics, and a distinction became apparent between questions submitted by "Seekers"—those who interact with the community about their question—and (with all due respect to the animal) "Sloths," who post their question apparently verbatim and interact no further. It is important to note that the terms Seeker and Sloth are used to describe apparent user intent *for individual questions*: users might well submit some questions coded Seeker and others coded Sloth.

Also, some individuals have multiple user accounts, and some single accounts are shared by multiple people; hence the individual question was chosen as the unit of analysis.

Seekers and Sloths are not exclusive categories, just different ends of the same continuum of apparent homework effort and engagement. While there is no way to be certain about the intent of an anonymous user via a few electronic transactions and some text on a screen, and no way for the precise content (or existence) of their homework assignments to be verified, a list of observable indicators of homework questions submitted by Seekers and Sloths emerged from the content analysis, and is provided in table 1.

6. Results and discussion

The evaluation metric is based on the rating, flagging, and comment functions native to Answerbag, which come with several caveats. Users can edit their submissions and add clarifying comments, and other users can adjust their ratings accordingly. Offering recognition on the site in the form of points increases site traffic, but also invites unscrupulous gaming of the rating system, for example by vindictively downrating the answers of a competitor. However, as with any collaborative filtering system, the more answer ratings there are, the more likely the impact of rogue users will diminish. Therefore, the data in this study are best understood as snapshots, inherently fluid, and no statistical analysis is attempted.

The results suggest that Answerbag users distinguish between homework questions submitted by Seekers and those submitted by Sloths, and demonstrate standards of reference service similar to those employed by professionals—gauging the questioner's intent along with the content of the query, and providing resources and explanation accordingly, usually stopping short of providing 'the answer.' They use ratings, flags, answers, and comments to enforce standards of conduct for appropriate homework questions, to educate and encourage the legitimately confused, and to berate the slackers. Table 2 summarizes the findings of the analysis of the 992 homework questions in the sample:

Table 1. Indicators of homework questions submitted by seekers and sloths.

Seekers	Sloths
Appear to summarize homework questions	Appear to quote homework questions directly
Post questions once	Post questions multiple times
Post questions in the appropriate category	Post questions in random or inappropriate categories
Provide context within the question (e.g. "I have an assignment on …")	Provide no context
Include information already obtained	Provide no additional information
Describe point of confusion	Provide no additional information
Interact with other users via ratings or comments	Do not interact
Return to the site and join the community	Never return to the site

Table 2. Numerical results of homework-question analysis.

	Seekers	Sloths	Site average
No. of questions	436	556	na
Average question rating	+7.4	−0.8	+4.5
Average number of answers per question	1.9	0.7	1.3
Average number of comments per answer	3.2	1.2	1.6
No. of questions flagged for administrator review	4 (0.9%)	38 (6.8%)	0.05%

The numerical data suggest that Answerbag users greatly preferred questions that fit the Seeker profile over that of the Sloth. These questions drew much higher ratings, more answers, and more answer comments. Sloth questions flagged for review usually carried the 'Nonsense' explanation, though some users preferred to flag these questions as 'Spam/Offensive'. Some users who opened new accounts only to flood the site with apparently verbatim homework questions drew the wrath of veteran users, who wrote several long and scathing e-mails to site administrators about the appropriateness of "using Answerbag in lieu of one's own mind" for this kind of homework help.

A content analysis of answers received by Seekers and Sloths yielded evidence of a normative emphasis on the part of the community. There were many instances of homework questions receiving answers like the following:

> A. This sounds like a school assignment to me. My suggestion is to get out a map. I won't be doing your schoolwork for you.

Answers can be rated just like questions, and when the 'do your own homework' type of answer was given to a Sloth question, it was almost always the most highly rated answer. Another common way Sloths were upbraided was to give completely false answers:

> Q. Locate and list all bodies of water touching the confederate states
> A. Arctic Ocean
>
> Q. We know about the interior structure of the Sun from . . .
> A. From examining a scaled down model . . . Put that answer on your test—I'm sure your teacher will give you an A+

It is important to emphasize again that Seekers and Sloths are not exclusive categories. One admittedly flustered newbie submitted a raft of astronomy homework questions in an inappropriate category, then made the effort to clarify the questions after being sternly lectured by one longtime member of the community. The newbie's persistence and participation apparently convinced the veteran that perhaps the newbie was not a complete Sloth after all, and deserved a response:

> Q. Two stars have the same luminosity and spectral type. Star A is at a distance of 5 light years from us, and star B is at a distance of 15 light years from us. How bright does star B appear compared with star A?
> A. For this one, I will tell you how to get the answer. Light follows the inverse

square law ($x = 1/\text{distance}\hat{\ }2$) just as gravity and many other things do. So, we have two identical stars. The only difference between them is their distance from us. One is 3 times as far away as the other. So, that is the number you plug in to get your answer.

However, this questioner's experience demonstrates that the process from lurking to legitimate peripheral participation to full community membership is by no means in lockstep. This questioner, apparently satisfied with the quality of service provided by the community, submitted two more batches of related homework questions over the next few months, but in late May— suspiciously close to the traditional end of spring term—this person's participation abruptly ceased.

Seekers' questions are often examples of the essence of homework help: a person who is only missing a resource or concept they need to arrive at the answer themselves, where Sloths are missing the motivation or engagement necessary to do any more than the minimum required to get the answer on the page. This last example demonstrates how an answerer can teach a Seeker to fish, instead of handing them a fish:

> Q. Math help please! I need to evaluate the series to the nearest whole number ... and then there is this series ... $8 + 17 + 26 + \ldots + 305$.
> A. This is an arithmetic series where each term equals the previous term plus 9: $8 + 9n$ where $n = 0, 1, 2, \ldots 33$. 34 terms in all. So, the sum $S = 34 * 8 + 9 * (0 + 1 + \ldots + 33)$ There is a well-known formula for the sum in parentheses, which you will need to determine in order to complete the calculation. Hope that helps to get you started. Good luck!

One of the most common requests new Answerbag users make is for a guidebook to becoming a good Answerbagger—essentially, the rules of the community. However, just as the site is built on a one-question, multiple-answers architecture, the results of this study suggest that no single person could write such a guidebook, because standards for acceptable behaviour are themselves created in the aggregate. Studying patterns of participation in online communities can help reveal these shared standards, and inform the development of homework help services and educational technologies more generally.

These preliminary results also suggest that though students may use social reference sites to get quick homework answers, the Answerbag community demonstrates some of the same standards as professional reference librarians in deciding how best to respond: evaluating each question and questioner as a unit, and preferring to guide people toward an answer rather than providing it outright. When a user approached the community in the best tradition of virtual reference, prepared with some information and willing to engage in a two-way conversation, the Answerbag community tended to provide answers that were more highly rated. 'Drive-by' questioners seeking quick answers often found themselves on the receiving end of a lesson in educational ethics and self-reliance—something most librarians and educational technologies cannot or do not provide. In that sense, online question-answering communities may provide a unique and necessary element of social indoctrination, and serve as an effective supplemental resource for homework help. Ongoing

research will address this question, as well as the larger processes by which the virtual social capital of rating points and perceived expertise on Answerbag influences user behaviour.

7. Conclusion

This study has shown that a social reference community, where answering and rating questions is a shared responsibility, can distinguish between homework questions submitted by Seekers and Sloths, provide appropriately different levels of service, and even educate questioners about appropriate information-seeking behaviour for homework assignments. The engaged Seeker is the user for whom systems and services should be designed, while the Sloth is less likely to meet builders and providers of information services halfway. Engaging Seekers, and transforming Sloths into Seekers, requires supporting people's natural work patterns, understanding the social aspects of learning, and broadening the definition of educational technology.

References

ALA, "Code of ethics of the American Library Association", 2007. Available online at: http://www.ala.org/ala/oif/statementspols/codeofethics/codeethics.htm (accessed 17 September 2007).

G. Barnett, "Information exchange in virtual communities: a typology", *Inform Res*, 5(4), Available online at http://informationr.net/ir/5-4/paper82.html

R. Gazan, "Understanding the rogue user", In *Information & Emotion: The Emergent Affective Paradigm in Information Behavior Research and Theory*, D. Nahl and D. Bilal, Eds., Medford, NJ: Information Today, 2007, pp. 177–185.

M. Gross, "The imposed query and information services for children", *J Youth Serv Libr*, 13(2), pp. 10–17, 2000.

J. Lave and E. Wenger, *Situated learning: Legitimate peripheral participation*, Cambridge University Press: New York, 1991.

E. Meyers, "The coolness factor: Ten libraries listen to youth", *Am Libr*, 30(10), pp. 42–44, 1999.

M.L. Saxton and J.V. Richardson, *Understanding Reference Transactions: Translating an Art into a Science*, San Diego, CA: Academic Press, 2002.

C. Tenopir, "Use and users of electronic library resources: An overview and analysis of recent research studies", in *Council on Libraries and Information Resources Report*, 2003. Available online at: http://www.clir.org/pubs/reports/pub120/pub120.pdf (accessed 11 May 2007)

M. Train, "Answers at your fingertips: Live Internet reference for teachers, students, parents and school library media specialists", *Knowl Quest*, 33(3), pp. 19–21, 2005.

M. Twain, *The Adventures of Tom Sawyer*, Hartford, CT: American Pub. Co, 1876.

A.M. Van Scoyoc and C. Cason, "The electronic academic library: Undergraduate research behavior in a library without books", *Libr Acad*, 6(1), pp. 47–58, 2006.

V.A. Walter and C. Mediavilla, "Teens are from Neptune, librarians are from Pluto: An analysis of online reference transactions", *Libr Trends*, 54(2), pp. 209–227, 2005.

M.M. Wasko and S. Faraj, "It is what one does: Why people participate and help others in electronic communities of practice", *J Strat Inform Syst*, 9, pp. 155–173, 2000.

E. Wenger, R. McDermott and W.M. Snyder, *Cultivating Communities of Practice*, Cambridge, MA: Harvard Business School Press, 2002.

R.G. Wengwert, "Some ethical aspects of being an information professional", *Libr Trends*, 49(3), pp. 486–510, 2001.

Y. Xu, C.Y. Tan and L. Yang, "Who will you ask? An empirical study of interpersonal task information seeking", *J Am Soc Inform Sci Technol*, 57(12), pp. 1666–1677, 2006.

G.K. Zipf, *Human Behaviour and the Principle of Least Effort*, Cambridge, MA: Addison-Wesley, 1949.