Best-Answer Selection Criteria in a Social Q&A site from the User-Oriented Relevance Perspective

Soojung Kim[*]

College of Information Studies, University of Maryland, College Park, MD 20742. iam@umd.edu.

Jung Sun Oh*

School of Information and Library Science, University of North Carolina at Chapel Hill, CB# 3360, 100 Manning Hall, Chapel Hill, NC 27599-3360. ohjs@email.unc.edu.

Sanghee Oh*

School of Information and Library Science, University of North Carolina at Chapel Hill, CB# 3360, 100 Manning Hall, Chapel Hill, NC 27599-3360. shoh@email.unc.edu.

As an attempt to better understand how people seek, share, and evaluate information in a social Q&A environment, this study identifies the selection criteria people employ when they select best answers in Yahoo! Answers in the context of relevance research. Using content analysis, we analyzed the comments people left upon the selection of best answers to their own questions. From 1,200 samples of comments, only 465 mentioned the specific reasons for their selection, thus becoming eligible for analysis. Through an iterative process of evaluating the types of comments, the best-answer selection criteria were inductively derived and grouped into seven value categories: Content value, Cognitive value, Socio-emotional value, Information source value, Extrinsic value, Utility, and General statement. While many of the identified criteria overlap with those found in previous relevance studies, the Socio-emotional value was particularly prominent in this study, especially when people ask for opinions and suggestions. These findings reflect the characteristics of a social Q&A site and extend our understanding of the relevance of an electronic environment where people bring their every day problem-solving and decision-making tasks.

Introduction

Contemporary views of users hold that they are not passive information receivers but active producers and distributors. New research paradigms pay close attention to users' voluntary participation and interaction in the on-line environment. Since Tim O'Reilly introduced the term web 2.0 in 2005, it has become a conceptual framework under which a variety of Internet applications and online activities are labeled as "participatory Web" (Madden & Fox, 2006).

Among the various activities that characterize web 2.0, this study focuses on a social Q&A site. Distinguished from traditional Q&A systems or services in library settings, social Q&A sites provide a venue where people voluntarily ask of and answer questions from fellow users, without the intervention of experts or authorities in the field of interest. Through the process of asking and answering a wide range of questions from philosophical inquiries to miscellaneous problems in daily lives, people not only seek information but share their experiences, opinions, advice, and fun. Thanks to ordinary people who bring their own knowledge, experiences, and opinions into the site, information exchanged tends to be relative and unstable rather than objective and absolute. It is the questioner who has the responsibility to evaluate the answers posted by other people and filter the ones that best fit his/her information needs.

In Yahoo! Answers, the social Q&A site examined in this study, a questioner is supposed to select the most appropriate answer and leave comments on that answer, which becomes the 'best-answer,' so that other users

---

[*] Each author made an equal contribution to this work. Authors are listed in alphabetical order.

can continue to consult and evaluate the resolved question and the associated best-answer.

To better understand how people seek, share, and evaluate information in such a social Q&A site, this study poses two research questions:

1) What criteria do questioners employ in selecting the best answer to their question?

2) What, if any, correlation exists between these criteria and the types of questions people ask?

Yahoo! Answers was the social Q&A site studied [or something like that]. Yahoo! Answers was selected for its dominant position in the social Q&A market, which accounts for 96% as of the week ending December 23rd, 2006 (Cashmore, 2006).

Literature Review

User-Oriented Relevance

User-oriented relevance is viewed as a subjective, cognitive, situational, multidimensional, dynamic, and measurable concept (Schamber, 1994). The earliest studies in this line of research are Cuadra & Katter (1967) and Rees & Schultz (1967). These studies show that human relevance judgments are affected by a number of non-topical variables, which were grouped into five categories: (1) the kind of document representation, (2) the way the request is expressed, (3) features of the judge, such as his/her knowledge of the subject, (4) the mode of expressing the judgment, such as?, and (5) the situation/context in which the judgment is expressed. Although both studies examined experts' judgments rather than actual users, their findings are still widely accepted.

In the 1980s and '90s, a plethora of empirical studies were conducted to elicit relevance criteria directly from the users. To highlight a few, the study by Park (1993) identified variables that affect eleven graduate students' relevance judgment as they prepared research proposals for their master's theses. The variables were grouped into three categories: (1) Internal context, associated with the user's own prior experience, (2) external context, associated with the current search and research, and (3) problem context, associated with the motivations underlying the intended use of a citation. Barry (1994) similarly identified 23 user-defined criteria in an academic environment where 18 students and faculty prepared class assignments, doctoral dissertations, and professional presentations and publications. The identified criteria were grouped into one of seven categories as follows: (1) the information content of documents, (2) the user's previous experience and background, (3) the user's beliefs and preferences, (4) other information and sources within the information environment, (5) the sources of documents, (6), the document as a physical entity, and (7) the user's situation. These studies and others (e.g., Cool et al., 1993; Wang & Soergel, 1998) confirmed the existence of factors beyond topicality affecting the user's relevance judgment and those criteria significantly overlap with those of the earliest studies (Mizzaro, 1997). Saracevic (1996) further categorized the types of non-topical relevance into cognitive relevance or pertinence, situational relevance or utility, and motivational or affective relevance. Cognitive relevance is concerned with the relation between the state of knowledge and cognitive information need of a user and texts retrieved. Cognitive correspondence, informativeness, and novelty are examples of this category of relevance criteria. Situational relevance is concerned with the situation/task/problem at hand and usefulness and appropriateness are criteria that fall into this category. Motivational relevance is concerned with a user's intents/goals/motivations. Such relevance criteria as satisfaction, success, and accomplishment belong to this category.

Schamber (1994) also categorized previously suggested relevance criteria into six classes based on an extensive literature review. She stated that a finite set of criteria existed for users regardless of information problem situations. It should be noted, however, that a majority of relevance studies have examined a typical task, namely students and faculty using bibliographic databases to find information in an academic setting. In this setting, printed documents are often the medium against which relevance is judged. Interestingly enough, when people seek interpersonal resources, not documents, in serious life- and health-related situations, love, friendliness, and other affective factors come into play in relevance judgment (Nilan et al., 1988). 'Love' has yet to be found elsewhere in the literature of relevance except in Nilan et al.'s study (Schamber, 1994) probably because emotions or affective factors are more involved when judging people than documents. Since the social interaction taking place on the site we chose to study is similar to the interpersonal interaction in Nilan et al.'s study even though people on that site judge other people's answers, not the people themselves, we

expected to see the socio-emotional factors in relevance judgments in that online community environment.

Social Q&A

The widespread use of the Web and personal computers has motivated people to consult Web resources for their information seeking in daily lives. In response to this phenomenon, various types of web-based question-answer services have appeared in recent years to provide personalized answers to users who are frequently overwhelmed by the number of search results returned by search engines. Since LookSmart Live!, a pioneer of web portal based question-answer services/search engines, launched in 1999 by LookSmart (available at http://www.looksmart.com), many expert question-answer services/search engines such as Ask Jeeves have followed (Lee, 2006). With AnswerBag, Yahoo! Answers, and other social Q&A sites launched over the last a couple of years, however, non-expert and free services in which ordinary people voluntarily ask their questions to the fellow users are particularly gaining popularity at a rapid rate. Because social Q&A websites are still in their infancy, there are only a few studies that examine how people seek and evaluate information in this environment, although some studies announced the appearance of social Q&A phenomenon or compared the features of different social Q&A sites (e.g., Elsa Wenzel, 2006).

Among the few studies that exist, Lee (2006) administered a survey to Korean-speaking Web users to understand their perceptions and assessments of Naver Knowledge-IN, the most popular social Q&A site in Korea. According to her findings, the users do not passively adopt the knowledge from answers: instead they adopt or reject the answer using their own judgment. The source of the answer plays a major role in evaluating the trustworthiness and usefulness of answers in the social Q&A sites. Gazan (2006), who examined AnswerBag, characterized answerers into two categories: specialists and synthesists. Specialists are those who claim expertise in a given topic and answer questions without referencing other sources. Synthesists are those who include one or more references to external sources in their answers. In the AnswerBag community as a whole, the answers of synthesists tended to be rated more highly than those of specialists, though specialists' answers were rated more highly within certain categories such as Health.

Methodology

In Yahoo! Answers, questioners are requested to leave comments when they select the best answer to their questions. Under the assumption that these comments, to some extent, would indicate the reason why a particular answer was chosen over others, we collected and analyzed the comments. Since the question and the best answer together provide the crucial context for each comment being analyzed, we collected the respective questions and best answers along with the comments. Consequently, our data record consists of 1) a question, 2) a best answer, and 3) a questioner's comment on the best answer. Additionally, the subject categories to which the question belongs were included to check if any patterns exist between the subject categories and the selection criteria. The distribution of the criteria among subject categories will be discussed in the following section.

Between December 22nd and 26th, 2006, a total sample of 1,200 comments (with questions and best answers) were selected such that 50 came from each of the 24 subject categories in Yahoo! Answers. Under each subject category, Yahoo! Answers presents separate tabs for open questions, resolved questions, and undecided questions. Among those questions, resolved questions with the best answer chosen by the questioner are of our interest, because of the presence of the comments. Each page of 'resolved question' tab displays 20 questions in a retrospective order, beginning with the most recent one. We selected the first question of each page, and collected the best answer chosen by the questioner and the associated comment. Starting from the first page, the process was repeated until we obtained 50 samples from each category. In this way, we tried to avoid potential biases, either by subjects (topics) or by recent events/activities.

We conducted an inductive content analysis (Krippendorf, 1980). The primary focus of the analysis was on the comments, although we also examined associated questions and answers for establishing contexts. The comments supposedly contain what the questioners liked about the chosen answer or why they liked it. In our analysis, we first identified such expressions or accounts on 'selection criteria' within each comment. These selection criteria were the basic coding unit. Out of 1,200 comments, 465 included one or more criteria. The total number of criteria identified from 465 comments was 493. The average number of criteria mentioned per comments was 1.06. The other 735 comments were excluded from the analysis since the comments were filled

with meaningless symbols, or simple expression of appreciations such as "Thank you" "Great," "Good," or "Nice." In some cases, there are comments that contained the questioners' follow-up questions, not mentioning any evaluative remarks on the best answers. Although it was an interesting feature of social Q&A in that the questioners and answerers seek follow-up interactions in order to solve their problems, we dropped these comments from the analysis because they are beyond the scope of this study.

Although only about 40% of the 1,200 samples were actually used in the analysis, given the fact that there is no control or direction over the way people use this commenting function in the site, we believe that the number of samples containing some expressions of value was reasonably large. While acknowledging potential biases toward or against individuals with particular characteristics (e.g., general expressiveness), we firmly believe that the comments present invaluable data for understanding people's judgments, precisely because people made those comments upon making their own choices, in their own words, and in the real situations.

The analysis process was explorative and iterative. In the first phase, the initial set of codes were developed by examining comments we collected, also with reference to various relevance criteria uncovered in previous studies. With the initial set, each author analyzed 50 samples, meticulously making notes. As a result of discussions on noted issues, the coding scheme was revised several times. A framework for the best-answer selection criteria (See, Table 1) was developed out of this process. In the second phase, the framework was used to classify all the comments collected. The authors analyzed data in parallel and constantly consulted with one another during the entire process of the analysis. Overall, coding was not conducted independently.

Additionally, in order to identify the relationships between question types and the selection criteria, we divided the questions into three groups, information, suggestion, and opinion, based on what the questioners wanted to receive from the answers.

- Information: finding specific facts or understanding phenomenon (e.g., "How long after purchase can I cancel a plane ticket from Orbitz?")
- Suggestion: seeking advice, recommendations, or viable solutions (e.g., "What is the best way to store unopened bags of flour for long term storage?")
- Opinion: surveying other people's thoughts or tastes, or initiating discussions about social issues (e.g., "What is the meaning of Xmas to the beggars and homeless?")

There were questions that did not fit any of these three categories, like a factual survey, for example, "How many languages do you speak?" or a math quiz. We included these questions in the "Others" category in the analysis. The following section will explain the best answer selection criteria in detail and show the findings including the distribution of the criteria among different types of questions.

Results

Best-Answer Selection Criteria Framework

The framework of the best answer selection criteria is composed of the 7 value categories and 24 individual criteria. Each value represents what aspects of the answers people considered in choosing the best ones, and the related criteria indicate the specific reasons to the selection. As can be seen in Table 1, the resulting framework has considerable overlaps with many relevance criteria uncovered in the other settings then the social Q&A sites of the previous studies (e.g., Barry, 1994; Cool et al.,1993; Park, 1993).

The most notable difference in our framework, however, is that 'Socio-emotional value' (33.3%) is turned out to be the biggest category, both in terms of the number of criteria (7 criteria) and the number of samples coded under that category (33.3%).. As for the frequency of individual criteria, 'Agreement' (15%) under 'Socio-emotional value' ranks the top. 'Utility' is also important. Even though 'Utility' accounts only for 14.8% of the total criteria identified, each of two criteria ('Effectiveness' and 'Solution feasibility') under this category is in the top-tier. When individual criteria from across different categories are compared; 'Solution feasibility' (8.7%) is the second most frequent one following 'Agreement' (15%) and 'Effectiveness' (6.1%) is the forth. 'Affect' under 'Socio-emotional value' (6.7%) ranks the third, as a whole, in between 'Solution feasibility' and

'Effectiveness.' The detail explanation of the framework and the distribution of criteria are as follows.

Table 1. Best-Answer Selection Criteria

Category	Criteria	Total No	%
Content value	Accuracy	18	3.7%
	Scope & specificity	12	2.4%
	Clarity	14	2.8%
	Writing style	9	1.8%
	Completeness	13	2.6%
	Existence of particular detail	16	3.2%
	Characteristics	6	1.2%
	Subtotal	88	17.8%
Cognitive value	Content novelty	14	2.8%
	Perspective novelty	11	2.2%
	Understandability	3	0.6%
	Questioner's background/experience	2	0.4%
	Subtotal	30	6.1%
Socio-emotional value	Emotional support	17	3.4%
	Answerer's attitude	10	2.0%
	Answerer's effort	7	1.4%
	Answerer's experience	2	0.4%
	Agreement	74	15.0%
	Confirmation	21	4.3%
	Affect	33	6.7%
	Subtotal	164	33.3%
Extrinsic value	External verification	4	0.8%
	Available alternatives	7	1.4%
	Quickness	6	1.2%
	Subtotal	17	3.4%
Information source value	Reference to external sources	9	1.8%
	Source quality	8	1.6%
	Subtotal	17	3.4%
Utility	Effectiveness	30	6.1%
	Solution feasibility	43	8.7%
	Subtotal	73	14.8%
General Statement		104	21.1%
	Subtotal	104	21.1%
	Total	493	100.0%

'Content value' (17.8%) is the second most frequent value in selecting best answers, excluding general statement. This value indicated that questioners evaluated the content quality of the answers. In the comments, questioners described based on their own judgment, whether the answers were correct or truthful (Accuracy), how specifically the answers covered the question topics (Scope/specificity), or how clearly the answers were explained (Clarity). In addition, questioners evaluated the 'Writing styles' of the answers, for instance, making statements like "You have just summed it all up very neatly indeed." or "I choose this answer as best because it was worded perfectly." Sometimes, questioners elaborated their fondness of a particular fact/detail illustrated in the best answers (Existence of particular detail, e.g., "Great ideas! And I really like the tip about the family name," or "All of the answers are good, but Pearlsawme's explanation was the best in stating that all radicals have two values"). Questioners evaluated the answers' historical, religious or factual aspects and the related comments were grouped into the 'Characteristics' criterion.

The criteria which belong to 'Cognitive value' (6.1%) indicated that questioners selected the best answers in regard with how the answers stimulated the cognitive cues of their knowledge or views in understanding concepts or situations. Questioners evaluated whether the best answers provided new information they did not know (Content novelty, e.g., "huh!!!! Never heard that name") or whether they dealt with new approaches that questioners had not thought of (Perspective novelty, e.g. "Very interesting perspective" or "I never thought about using a shimmer eye shadow to go over the beige. It sounds like a good idea"). Whether the answers were easily understandable for them (Understandability) was also assessed. Sometimes, the answers which reminded questioners of or reflected on their previous experiences or backgrounds were chosen (Questioners' personal experience). For example, a questioner who asked about a favorite sweet potato pie recipe commented that he/she picked the recipe that had the word "Southern" because it matched with his/her native background in the South.

In this study, the researchers were particularly attentive to how 'Socio-emotional value' would be represented in the data, since we assumed that the participatory setting of the social Q&A site, integrated with the technical advantage that users can easily create, modify and search questions and answers, enhances the interaction among users and contributes to building social environments, which encourage users to voluntarily share their feelings, emotions, and experiences as well as information. From the data, we confirmed this assumption and we even found that 'Socio-emotional value' (33.3%) was the most frequently referred value among others. Questioners responded to the 'Emotional supports' from answerers, saying "That makes me feel better." or "Your words really helped". Questioners also sensed whether answerers were sincere, honest or open-minded (Answerers' attitudes). Questioners appreciated the 'Answerer's efforts' in putting answers together, as stating like "She took enough time to try and feed me some info."

Sometimes, questioners pointed out that they appreciated answers sharing their personal experiences (Answerers' experience). For example, a questioner asked about what was "Hypothyroidism." While there were several answers about thorough reviews of definitions and symptoms of the disease, the questioner chose an answer from someone who illustrated his/her own case of the disease. The questioner annotated in the comment that "Thanks to the other person who posted the big list of symptoms, but I wanted a more personal statement."

Whether the questioners agreed with the answerers (Agreement), or whether the answerers confirmed what the questioners originally expected (Confirmation), was also reflected in the comments (e.g., "Finally, someone who agrees with me" or "This guy thinks like me so he must be alright."). Especially 'Agreement' (15%) was an influential criteria. Sometimes, questioners chose the best answers solely based on what they felt about it (Affect, e.g., "That was hilarious," or "You answer was too funny"), and 'Affect' was the third most common criteria (6.7%) across the value categories. The top first and third criteria, 'Agreement' and 'Affect' were affiliated with 'Socio-emotional value.'

'Extrinsic value' (3.4%) indicated that questioners assessed the answers, not according to the content of the answers or their reflections about the answerers, but according to practical matters of the situations surrounding the questions and answers. The 'External verification' criteria meant that questioners chose the best answers because the answers verified what they heard from other sources (e.g., "I've been told to wait by a lot of people; I just hope it works!" or "I found out that some people did exactly what you have recommended."). Sometimes, questioners picked the only available answer (Available alternatives, e.g., "Well

there was just one answer.") or the first response (Quickness, e.g., "You were the first one to answer") as the best answer.

In 'Information source value' (3.4%), the 'Reference to external sources' criteria specified that questioners showed their appreciation of the external source offerings in the comments. Most answers which fit this criterion provided URLs of specific websites related to the topics of the questions or answers. On top of that, the perceived quality of the sources, whether worthwhile to review or useless, were also evaluated in the comments (Source quality, e.g., "Perfect answer from the best source"). A comment about the expertise of answers was also grouped in this criterion as we considered the expertise as one type among sources of information (e.g., "You are quite the expert on gaming/consoles")

'Utility,' (14.8%) was about whether answers counted as useful in solving questioners' problems. Questioners selected best answers before or after they tested the answers. In the former case, the questioners showed their intention to try out what they had been instructed in the answers ('Solution feasibility,' e.g., "I like your answer. I will have to do that." or "That will help me so much. I am going to try this next time I am feeling down"). In the latter case, the questioners stated that they chose the best answer because that answer had actually helped to solve their original problems, either partially or entirely ('Effectiveness,', e.g., "Lots of good answers. I did use your idea. They turned out good.", or "All the suggestions where great and after trying them out I found this one worked the best").

Questioners, sometimes, described their overall impression of the answers in the comments, like "this is exactly what I was looking for" or "That makes sense thank you." These are comments which the questioners expressed somehow the reasons why they chose the best answers, but the comments didn't provide enough information to figure out which particular aspects of the answers they liked the best. We separated them into the general statement category, which accounted for 21.1% of the total criteria mentioned.

## Relevance Criteria Distribution by Question Types

In this section, the questions associated with the sampled 465 comments were categorized into three question types – information, suggestion, and opinion – and the frequency distribution of the selection criteria by these question types were investigated.

According to the distribution of the numbers of the questions in our sample, 'Opinion' (39%) was most frequent, followed by 'Information' (35%) and 'Suggestion' (23%). Since we classified the types of questions of the 465 selected sample data, instead of the original 1,200 samples, it is hard to say whether this distribution of the questions represents the distribution of questions in the social Q&A site in general. The distribution of data across the question types, however, provides interesting patterns of the criteria, suggesting the relative significance of each category for different question types. Table 2 shows the distribution of the 7 value categories across the question types and Figure 1 illustrates it with a bar graph. The full description of the frequency distribution of the data can be found in the following URL: http://www.unc.edu/~shoh/2007_ASIST.htm.

Table 2. Relevance Criteria Distribution by Question Types

Category	Information		Suggestion		Opinion		Others		Total	
	No	%	No	%	No	%	No	%	No	%
Content values	44	25%	16	14%	26	13%	2	22%	88	18%
Cognitive values	14	8%	6	5%	10	5%	0	0%	30	6%
Socio-emotional values	22	13%	22	19%	114	58%	6	67%	164	33%
Extrinsic values	7	4%	6	5%	4	2%	0	0%	17	3%
Information source values	11	6%	5	4%	1	1%	0	0%	17	3%
Utility	22	13%	39	34%	11	6%	1	11%	73	15%
General Statement	55	31%	20	18%	29	15%	0	0%	104	21%
Total	175	100%	114	100%	195	100%	9	100%	493	100%

Figure 1. Relevance Criteria Distribution by Question Types
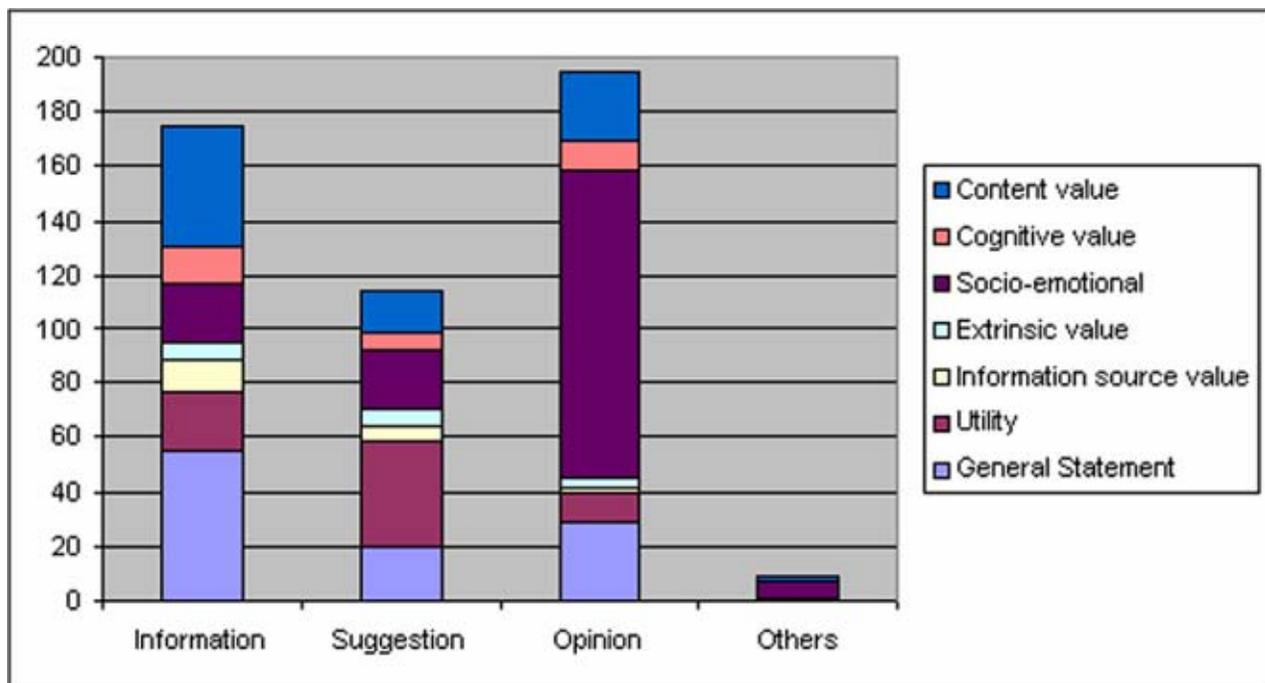
For the information type questions, 'Content value' turned out to be the most frequent category, followed by 'Utility' (excluding 'General statement'). Although it is not specifically illustrated in Figure 1, we found that certain characteristics of the content of the answer such as 'Accuracy' or 'Clarity' were often mentioned in the 'Content' value.' Interestingly, in many cases, questioners specifically preferred the answers related to the 'Existence of particular detail' criteria (e.g., "The answer told me what kind of barrel was for sell and who to contact.").

Questions classified as 'suggestion' in this study are mostly for seeking advice on personal matters or for getting solutions/tips for a problem at hand. Therefore, it is expected to see 'Utility' most often with regard to this type of questions. Perhaps, a more interesting observation would be the high frequency of the 'Socio-emotional value' category. As illustrated in the following example, "I think this is the best answer because the person that answered this is going through the same thing that I will be going through. He offers an example by

sharing a personal story all the while being supportive of how I feel and how my friend feels," people often prefer to have something beyond a technical solution in the social Q&A environment.

For the opinion type questions, 'Socio-emotional value' clearly stands out, covering 58% of cases in total. Moreover, according to our data, the 'Agreement' criterion alone takes 33%. In some cases, the answer was chosen because the questioner agreed on some points that the answerer made, even if the answer per se did not resolve their question, as shown in this comment, "Not really as answer, but I agree with you." Not surprisingly, 'Affect' and 'Emotional support' turned out to be also important. In addition, almost all of the comments categorized as 'Answerer's attitude' were observed in opinion type questions (e.g., "This person didn't try and 'shame' me like some of the other people and I really appreciated that").

Relevance Criteria Distribution by Subject Categories

Another additional finding has to do with the difference in selection criteria according to the subject categories. Even with the small number of samples in each subject, some patterns were noted (See, Figure 2).

Figure 2. The Distribution of the Selection Criteria in the Subject Categories of Yahoo! Answers[**]
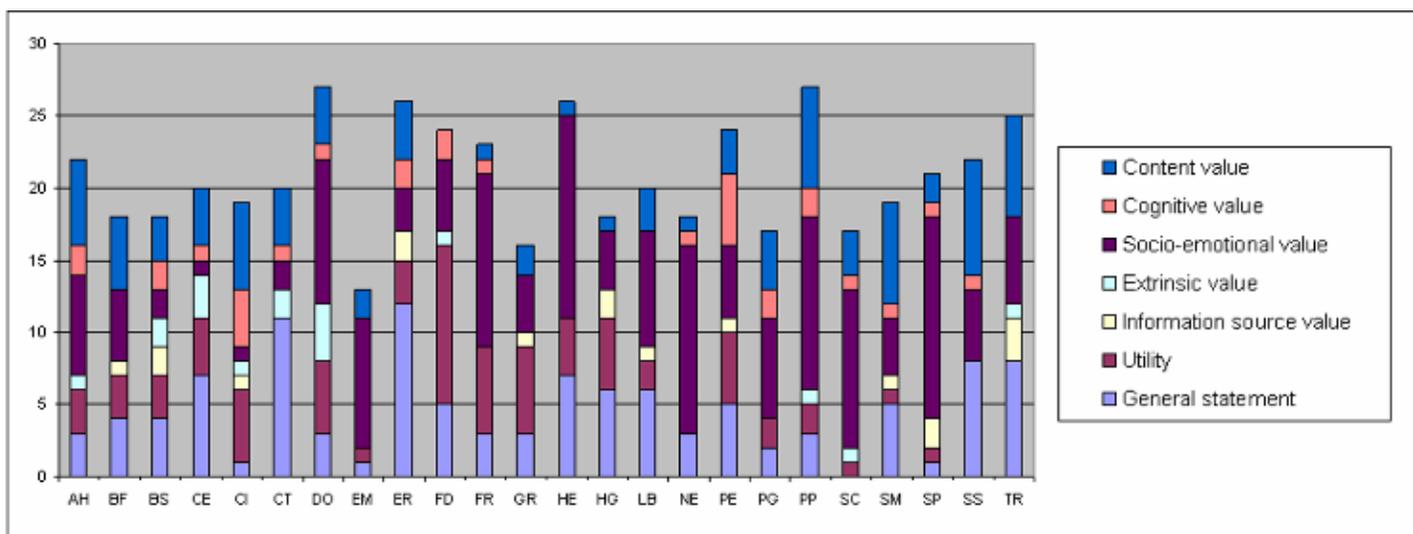
In areas such as Health (HE), Sports (SP), News & Events (NE), Family & Relationship (FR), and Pregnancy & Parenting (PP), the 'Socio-emotional value' category stood out. Also, Content value was frequent in Social Science (SS) and Travel (TR), and Utility was frequent in Food & Drink (FD).

Some specific criteria appeared often in certain subjects but not in others. For example, a provision of further references to other sources ('Information source value') was highly valued in subjects such as 'Beauty & Style' (BS), 'Education & Reference' (ER), 'Home and Garden' (HG), and 'Travel' (TR). With a larger sample, it would be possible to verify those small patterns observed in this study.

Discussion

[**] The full names of the subject categories in Figure 2 are Art & Humanities(AH), Business & Finance (BF), Beauty & Style (BS), Consumer Electronics (CE), Computers & Internet (CI), Cars & Transportation (CT), Dining Out (DO), Entertainment & Music (EM), Education & Reference (ER), Food & Drink (FD), Family & Relationships (FR), Game & Recreation (GR), Health (HE), Home & Garden (HG), Local Business (LB), News & Events (NE), Pets (PE), Politics & Government (PG), Pregnancy & Parenting (PP), Society & Culture (SC), Science & Mathematics (SM), Sports (SP), Social Science (SS), and Travel (TR).

A social Q&A site presents a new information seeking environment. Its various features, including a wide range of questions, the relative and unstable nature of information being exchanged, and the active constructive role of the users, render a social Q&A site distinct from other information systems or services. It is important to gain a better understanding of people's information seeking behaviors/judgments in this particular environment in that it allows us to obtain invaluable data created in a natural setting.

As noted in the result section, the set of selection criteria in our study has considerable overlaps with many relevance criteria uncovered in the previous studies (e.g., Barry, 1994; Cool et al.,1993; Park, 1993), enabling us to place our study in the context of relevance research. The overlapping categories suggest that there indeed exists a certain fundamental commonality in people's relevance judgment and there is a limited set of criteria regardless of particular contexts (Schamber, 1994). However, the kind of information people seek in a social Q&A site and/or their expectations of the answers they will receive are different in many ways from those typically observed in other studies, which mostly were carried out in academic settings. The dominance of the 'Socio-emotional value' category in our data reflects the differences.

The relative proportion of different types of questions in our data set is in and of itself interesting. Out of 465 questions associated with categorized comments in our study, opinion type questions, asking other people's opinions, evaluations, or interpretations about a given situation or fact, appeared most frequently (almost 40%). In the context of social Q&A, people seem to be interested in gathering rather 'subjective' takes on a matter. Considering social motivations, it is not surprising that the 'Socio-emotional value' category accounts for 58 % of best answer selection criteria for this type of questions. On the other hand, 23% of the questions are asking for suggestions, mostly to come up with a viable solution to a problem at hand. Here people could describe their problem situations in detail and thus seek for 'customized' solutions. Again, the 'Socio-emotional value' category takes a large proportion, following right after 'Utility'. The information that the questioner needed in order to resolve the problem situation may or may not be subjective, but people highly appreciated those answers that provided answerers' personal experiences along with suggested solutions. Taken together, we believe the large proportion of opinion or suggestion type questions (63%) and the dominance of the 'Socio-emotional value' category illustrate what people seek from the social Q&A site.

This study was the first attempt to analyze the relevance criteria in the distinct online environment of the social Q&A sites. Thus, it mainly focused on understanding the nature of the data and developing the relevance criteria framework. The framework was evaluated with the pilot study of the small scale of the data samples. The research to characterize the relevance criteria in the social Q&A setting will be continued. In the next study, the relevance criteria framework will be reexamined with the expanded number of the data collection in order to validate/modify the current framework and find the additional relevance patterns of the best answer selection. In the current study, we were not able to collect enough data to illustrate the relevance criteria distribution across the subjects. In addition, the data in the 'General statement' category was excluded from the further analysis since the general expressions of the comments in that category vary and it was hard to group them into separate criteria. It is expected that the patterns of those would be recognizable if we collect more number of data. Comparing to the current approach to code data based on the collaborative reviews among researchers, the next study will be based on the data analysis with independent coding. It would allow to ensure the reliability of the data.

Conclusion

Social Q&A sites allow people to ask questions to fellow users, and assess the relevance of given answers on their own ways. To identify the criteria people employ when selecting best-answers in a social Q&A site, this exploratory study analyzed 465 pairs of questions and comments left on the best answers in Yahoo! Answers. Although high overlap between the identified criteria in this study and those found in previous ones confirms that there exists a set of common criteria frequently used in people's relevance judgment regardless of contexts, the prominence of 'Socio-emotional value' is notable because it reflects the characteristics of the social Q&A environment: people not only seek specific information but share subjective opinions and suggestions on their everyday problem-solving and decision-making tasks. The findings of this study expand the knowledge of relevance to the social computing environment.

References

Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science,* 45, 149-159.

Cool et al. (1993). Characteristics of texts affecting relevance judgments. *Proceedings of the 14th National Online Meeting*, 77-84.

Cuadra, C. A. & Katter, R. V. (1968). Opening the black box of relevance. *Journal of Documentation,* 23, 291-303.

Krippendorf, K. (2004). *Content analysis: An introduction to its methodology* (2nd). Thousand Oaks, CA: Sage.

Gazan, R. (2006). Specialists and Synthesists in a Question Answering Community, *ASIST Annual Meeting*, Austin, Texas, November 3-9, 2006.

Lee, Y. (2006). *Toward a new sharing community: Collective intelligence and learning through Web-portal-based intelligence. Communication*, Culture, and Technology Master of Arts Program. Georgetown University.

Madden, M. & Fox, S. (2006). *Reading the waves of "Web 2.0": More than a buzzword, but still not easily defined.* Pew Research Center. Retrieved January 5, 2007, from http://www.pewinternet.org/pdfs/PIP_Web_2.0.pdf.

Cashmore, P. (2006). Yahoo answers dominate Q&A – No wonder Google gave up. *Mashable!, Social networking 2.0.* Retrieved January 5, 2007, from http://mashable.com/2006/12/28/yahoo-answers-dominates-q-a-no-wonder-google-gave-up/.

Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science*, 48, 810 - 832.

Nilan et al. (1988). A methodology for tapping user evaluation behaviors: An exploration of users' strategy, source, and information evaluating. *ASIST Annual Meeting*, 1988.

Park, T. (1993). The nature of relevance in information retrieval: Empirical study. *Library Quarterly*, 63, 318-351.

Rees, A. M., & Schultz, D. G. (1967). *A field experiment approach to the study of relevance assessments in relation to document searching: Final report: Volume 1*. Cleveland, OH: Case Western. Reserve University, School of Library Science, Center for Documentation and Communication Research.

Saracevic, T (1996), Relevance reconsidered, in Ingwersen, P., Pors, N.O (Eds), *Information Science: Integration in Perspectives*, Royal School of Librarianship, Copenhagen, 201-18.

Schamber, L. (1994). Relevance and information behavior. *Annual Review of Information Science and Technology*, 29, 3-48.

Wang, P., & Soergel, D. (1998). A cognitive model of document use during a research project. Study I. Document selection. *Journal of the American Society for Information Science*, 50 (2), 98-144.

Wenzel, E. (2006). *Dropping knowledge: question-and-answer sites.* Retrieved January 5, 2007, from http://reviews.cnet.com/4520-9239_7-6634568-1.html.