User Based Question Answering: An Exploratory Study of Community Generated Information Exchange in Yahoo!Answers

Rong Tang
Graduate School of Library and Information Science
Simmons College
rong.tang@simmons.edu

Sheila Denn
Graduate School of Library and Information Science
Simmons College
sheila.denn@simmons.edu

**Final Report
2008 OCLC/ALISE Library and Information Science
Research Grant Project**
5 September 2011

I. Overview

This project was funded by OCLC/ALISE Library & Information Science Research Grant Program Grant and by the Simmons President Fund for Faculty Excellence in early 2008. The major endeavor of this exploratory research study included data processing and coding, and conducting content analysis of user community generated question-answer-comment sets through the platform of Yahoo!Answers.

Online user-based Q&A systems present an excellent opportunity to investigate how users interact with one another through a virtual social environment; however, the research in this area is still limited. There is a lack of systematic, structured, and in-depth content analysis on users' question-answer data. Consequently, the complexity and nuances involved in users' Q&A exchange processes, the association between questions and answers, the interactive creation and selection of answers, and the social value bestowed upon a public-shared Q&A community have not yet been fully revealed.

The research project reported here addresses the research problem by employing a question taxonomy, an answer classification scheme (based on sentence typology and the features/elements contained), and a classification criteria set for comments to examine 3,000 randomly selected question-answer-comment sets from 25 question categories in Yahoo!Answers. In addition to the analysis of question types, investigation into the answers provided for a given question, the linguistic properties of question-answer sets, and how an answer has been chosen as the best, have enabled a thorough understanding of information exchange in a social Q&A process.

The research questions for this study are:

1. What are different types of questions found in Yahoo!Answers datasets?
2. What are different types of answers in responding to the questions found in Yahoo!Answers datasets?
3. What is the association between the best answer and its linguistic characterization as well as the presence of certain element such as the inclusion of an URL?
4. What are most notable criteria emerged from askers' comments with regard to why the best answer was chosen?
5. What specific impact do answers to research questions 1-4 have on the provision of library and information services?

Over the course the project, three large datasets were created based on 3,000 randomly sampled question sets from the entire Yahoo!Answers categories. The datasets consisted of:

1. Questions and question-descriptions (sometime called question content), with 120 items per category
2. Answers, best answers, and the scores for the best answers to the questions from 1.
3. Comments on the best answer, mostly from the asker, as he/she gave score to the best answer from 2.

Due to the sheer quantity of the datasets and the nature of the coding which required at least two coders coding for each data element then meet to double check and cross validate, the project has been a huge undertaking. In the duration of the project, both the Principal Investigator and the Co-PI on multiple occasions had to delay the project due to their health

problems. The Co-PI Sheila Denn had to take a medical leave starting March 2009, and was unable to return to the project. The PI Rong Tang has been able to work on the project with the aid from research assistants, however, her own health problem has also caused the further delay of the project completion.

Overall, this has been a very interesting, intellectually rewarding, and valuable research project. The findings are very rich, insightful, and meaningful to the advancement of LIS research and professional practice. It is the hope that in the near future, this research study will be written up as research papers, to be published in scholarly journals or presented at professional conferences.

I. Innovative Significance of the Study

The purpose of this project is to conduct content analysis of randomly selected question/answer/comment sets extracted from the Yahoo!Answers system through its API services. This funded research is one of the first studies to perform a multi-level linguistic analysis of the question/answer/comment sets on a user-based Q&A system. Our work is innovative in that the results will extend the theories of questions and answers beyond the content-oriented perspective of reference services. The identification of how the best answers are constructed in Yahoo!Answers will help librarians to provide pertinent responses to users' requests. Additionally, the results from the sociolinguistic analysis of the community based Q&A process will further our knowledge of user needs and patterns of online communication in a digital reference environment. By undertaking a thorough analysis of the questions and answers submitted to Yahoo!Answers, we are taking a first step toward understanding how user community-based Q&A services fit in the Web 2.0 (and Library 2.0) information landscape, and how we as information professionals can improve all facets of the information-seeking process for our users.

II. Related Research

III.1. Theories of Question and Answer Classification

The study of questions, as pointed out by Freed (1994) and Pomerantz (2005), has been carried out for decades by researchers from a variety of domains. Speech act theorists, conversation analysts, psycholinguists, computational linguists, information scientists, and reference librarians have all been involved in the research on question categorization or the process of question answering. According to Freed (1994), there is general agreement among researchers about the essential characteristics of questions; that is, "questions have an identifiable syntactic form, a characteristic prosodic nature, a semantic or propositional content, which is separate from their pragmatic and social function" (p. 622). On the other hand, both Freed (1994) and Pomerantz (2005) indicate that there are considerable disagreements on what approach to take to analyze and categorize questions. One controversial issue is whether questions can be understood without looking at the responses and feedback that they receive. A number of researchers (Schegloff & Sacks, 1973; Clark & Wilkes-Gibbs, 1986; Roberts &

Bavelas, 1996) have asserted that questions must be analyzed along with their answers as a part of "adjacency pairs" in a collaborative communication setting.

A remarkable amount of research has been conducted on question typology, from the perspectives of linguistic and discourse analysis, automated Q&A systems, and digital reference services. Several influential classification schemes of questions have been developed, including Robinson & Rackstraw's (1972) wh-question classification, Lehnert's (1978) conceptual categories of questions, Graesser, Mcmahen & Johnson's (1994) information content-based question classes, Freed's (1994) taxonomy of question functions, and Pomerantz's (2005) groupings of reference questions. Some of these taxonomies are conceptually associated with one another and have been empirically tested and used for conversational analysis, TREC Q&A studies, and IR research (Keyes, 1996; Lytinen & Tomuro, 2002; Harabagiu, Moldovan, Pasca, Mihalcea, Surdeanu, Bunescu, Girju, Rus, & Morarescu, 2002; Tomuro, 2004). As an example, the analytical model of question-asking patterns developed by Graesser and his colleagues (1994) focuses on the content of information sought by the questioner and classifies the questions based on the length of the answers they will receive, instead of looking at the linguistic characteristics of question stems such as why, who, or where. Questions that invite short answers, such as verification, disjunctive, or concept completion, require a response merely in a word or a phrase. The long-answer questions, such as causal antecedent, enablement, or judgmental, typically elicit responses that are several sentences long. The authors indicate that such a typology is adequate for empirical testing in that it both exhaustively represents a wide range of questions and can be reliably applied by trained judges with a high level of coding consistency. Pomerantz (2005) indicates that this taxonomy is indeed useful in classifying real-world questions. He also points out that because the taxonomy is based on the function of expected answers, it is more a classification of question-answer pairs than of questions alone.

In analyzing over 1,000 question tokens occurring in dyadic conversations, Freed (1994) proposes to classify questions according to the type of information sought or conveyed, i.e., the function of questions. Freed suggests that there are 16 categories of questions that range on a continuum from public information, to shared information, to phatic information, and finally to rhetorical information. Freed also maps the syntactic form of the questions to the functional groups and finds that certain function categories are closely associated with a particular syntactic type of question, such as yes/no questions, wh-questions, or tag questions.

Both Graesser et al. (1994) and Freed's (1994) classifications of questions present useful approaches to the analysis of question-answer sets. While Graesser and colleagues focus on question functions of expected answers, Freed's taxonomy deals with pragmatic and social aspects of questions in a paired conversational setting. Because the two grouping models address question categories at different levels, a multi-faceted taxonomy combining both schemes may prove to be fruitful.

In the area of library information services, many attempts to categorize reference questions have been made (Rothstein, 1964; Seng, 1978; White, 1981; Phenix, 1983; Fogarty, 1995; Carter & Janes, 2000). Some of these were derived based on standardized definitions of reference transaction types occurring in the Library General Information Survey (U.S. Department of Education, 1981). In a comprehensive review of the literature on question

taxonomy, Pomerantz (2005) presents 11 types of reference questions based on his synthesis of previous research. This particular classification of questions considers the forms of expected answers and the type of "task that the librarian must perform to formulate an appropriate answer" (p. 721). The types of reference questions included range from directional, ready reference, and readers' advisory, to research, analysis, and critique.

Thus far only limited research has been done to investigate or classify answers. In describing her computational model of question answering, QUALM, Lehnert (1978) articulates three selection rules for answers. They are: (1) an answer that conveys information by inference is preferable to one that spells out such inference explicitly; (2) a good answer does not tell the questioner something he already knows; and (3) when a question asks about a single conceptualization, good answers do not require the questioner to infer the missing parts of an incomplete causal chain. Another study that contributes to the understanding of why an answer is chosen to be the best is Kim and her colleagues' (2007) recent work on comments made about chosen answers in Yahoo!Answers.

III.2. Studies on Automated Q&A Systems

Since the mid-1970s, automated Q&A system research has had two primary foci: 1) to develop computational models of human question answering that can be tested; and 2) to provide tools to be used as decision support mechanisms. Computational models of question answering have been built on three main activities: parsing of the original question into its conceptual components; retrieval of corresponding concepts from memory; and generation of a usable answer (Lehnert, 1978).

Clearly there is a great deal of complexity embedded in these three activities. For this reason, research in automated Q&A in the 1980s and 1990s focused on the relatively straightforward tasks involved in understanding and generating answers for simple, closed-ended questions (such as "In what city is Fenway Park located?"). At the beginning of the twenty-first century, a renewed focus on tackling the challenges of open-ended questions has been articulated in two white papers released by the National Institute of Standards and Technology as part of the ongoing TREC series of conferences in information retrieval (Carbonell et al., 2000; Burger et al., 2001).

III.3. Studies on Digital Reference Services

The new wave of digital reference services, whether via email or online chat, is perhaps closest to the kind of Q&A we see in Yahoo!Answers. The primary difference is that the questions posed to a digital reference service are asked in the particular context of the library, and the answerer is an information professional assumed to have particular expertise in question answering.

Over the years, a great deal of work has been done to categorize the types of questions that are received by reference services, both traditional and digital (see, for example, Arnold & Kaske, 2005; Braxton & Brunsdale, 2004; Brown, 1985; Bunge, 1990; Diamond & Pease, 2001;

Katz, 2002; Kibbee, Ward, & Ma, 2002; Lee, 2004; and Pomerantz, 2005). These categorizations tend to emphasize the purpose of the question, such as questions about library policies and procedures, subject questions, holdings questions, ready reference, and directional questions (Arnold & Kaske, 2005).

A remarkable series of studies conducted by Radford and Connaway (2008) on evaluating virtual reference services, investigated in part (during Phase II of the project) 850 randomly sampled chat transcripts. Researchers employed and developed a good number of classification coding schemes for analyzing transcripts, ranging from question classification, relational content coding schemes, and query clarification coding scheme, and ready reference accuracy coding scheme. The question classification scheme used, combining Katz (1997), Kaske and Arnold (2002), and Radford's (2006) work, categorizes questions into the classes of: directional, ready reference, subject search, research, policy/procedural, holdings/do you own?, no question, inappropriate, and reader's advisory. Several question types included in the KKAR could also be used to describe the question content in a social question-answer environment. Radford and Connaway's (2008) found that the success of reference encounter depend on the positive interaction between the librarian and the user in both the relational dimension and content dimension. Their investigation into user/librarian relational content yields valuable results in learning the impact of a good interaction in the opening and closing of the chat references such as rapport building and abrupt ending on the success/failure of the entire reference transaction. Another interesting area was the study on query clarification. Radford, Connaway, Confer, Sabolcsi-Boros, and Kwon (2011) discovered that further questions were asked or information was provided "to clarify topic, background, search history, extent/depth of information needed and type of resources sought" (p. 264). It was discovered that librarians and users engage in different patterns and have different goals in their query clarification. Librarians tend to ask clarification questions prior to conducting search where as users tend to clarify during the search process. The results on the nature and process of query clarification are helpful and relevant to understanding the question answer exchange in a social QA setting.

A 2010 study by Harper and his colleagues (Harper, Weinberg, Logie, & Konstan, 2010) extended Aristotle's rhetoric framework to classify the questions from three social QA systems including answerbag, metafilter, and Yahoo!Answers. The questions were categorized into the classes of deliberate, epideictic, and forensic, and within each class, two subclasses were included, in a manner similar as what is listed below in Table 1. Harper et al. (2010) further examined how important terms, the length, number of responses are associated with each question type. They also studied compound questions, and the linkage between different question type and the type of qualitative value they carry.

Table 1 Rhetoric Framework of Question Type

Deliberate (future-focused)	Epideictic (present-focused)	Forensic (past-focused)
• Advice • Identification	• (Dis)Approval • Quality	• Prescriptive • Factual

III.4. Studies on User-based Q&A Systems

Despite all the theories developed for question analysis and automated Q&A, existing question taxonomies have never been applied to examine the question-answer data in a social Q&A environment. During the formative years of social Q&A systems and within recent years, a good number of empirical studies (Gazan, 2006, 2007; Kim et al., 2007; Kim & Oh, 2009; Shah & Pomerrantz, 2010, Harper, et al., 2010; Shah, 2011) have presented results related to these systems. As one of the original developers of the Answerbag system, Gazan conducted several exclusive investigations on the question-answer behavior of the user. His 2007 paper examines users who submit homework questions to the system. Through analyzing 992 homework questions in terms of the number of answers received, ratings obtained, and with further examination of question text, answers, and comments, the author identifies two kinds of questioners: "seekers" and "sloths." While the former typically interact with people by engaging in rounds of conversations about their questions, the latter copy their actual assignment questions verbatim and do not communicate with people who respond to their questions. Evidence from the Answerbag data suggests that seekers have been preferred by the user community, and that the shared responsibility of rating, answering, and commenting on questions has shown to be a very powerful mechanism to control the level of services that users provide to one another in such an environment.

In another study, Gazan (2006) focuses on the behavior of users who answer others' questions. In sampling nearly 10,000 answers, the author groups the answerers into two categories: specialists and synthesists. The major difference between the two types of answer is that synthesists' answers always contain one or more references to external sources. Overall, the synthesists' answers were voted to be more useful than the answers by specialists; however, there are specific categories such as parenting, divorce law, criminal law, Mormon religion, or relationships where specialists' answers were preferred over synthesists' ones.

In a study by Kim and her colleagues (2007), 1,200 samples of user comments on best answers in Yahoo!Answers were analyzed. Based on the questioners' own comments on why a given answer was chosen as the best to their questions, the researchers performed content analysis to uncover users' criteria for answer relevance. The final set of criteria consists of 25 individual items in seven general categories, with "general statement" added as a one-item category. The criteria categories range from content value, cognitive value, socio-emotional value, extrinsic value, information source value, and utility to general statement. The data show that users applied "socio-emotional value" related criteria most frequently in presenting their reasons for selecting an answer. In a follow-up study of a more extended sample, Kim and Oh

(2009) revised their best answer selection criteria to include new elements of rationality, length, taste, and humor. They also revised their grouping of the criteria to some degree. One of the conclusions the authors reached is that "the relative importance of criteria varies according to the topic categories" (p.725).

In a study that examined one month of Yahoo!Answers activities, which included over 1 million questions and over 8.4 million answers, Adamic, Zhang, Bakshy, and Ackerman (2008) also found that the Q&A dynamics vary by topics. Based on thread length, content length, and asker/replier overlap, the authors performed k-means cluster analysis and grouped the datasets into categories of interactive discussion, everyday advice and factual questions. In terms of best answers, Adamic et al. found that reply length, the number of competing answers, and the track record of the user were most significant indicators for an answer being selected as the best.

A more recent case study by Bouguessa, Wang and Dumoulin (2010) focused on developing an automatic procedure in identifying knowledge-sharing communities, using Yahoo!Answers as the case. Such a procedure first identifies authoritative users and then discovers "the communities that form around them"(p. 46). The algorithm developed by the researchers calculates the authority score then used a transactional clustering measure named TRANCLUS to detect communities through askers' interaction with authoritative users. The approach was tested through data from six Yahoo!Answers categories that recorded one-year activities. The results showed that Yahoo!Answers user community was more of a "dense" community structure "within each of which askers are closely clustered around one dominant authoritative user" (p.46).

A recent study by Shah and Pomerantz (2010) looked into how to automatically predict the answer quality of Yahoo!Answers based on a predefined set of criteria. The researchers used the crowd sourcing technique through Amazon Mechanical Turk workers to rate the answer quality based on 13 criteria. Then the authors used automatically extracted features to compare automatic assessment with human assessment. They concluded that the automatic procedure was very robust and achieved over 80% accuracy. In his most recent study, Shah (2011) focuses on properties associated with the best answer, including distribution of answer ratings, the time lapse between a question and best answer, as well as what position the best answer is in a list of answers. Shah (2011) found that while some of the best answers would appear within the first 15 minutes after the question was posted, the majority of the best answer took more than an hour to appear. Shah also found that the sixth answer or the first answer tend to be selected as the best answer.

While these studies provide valuable, first-hand reports of user community based Q&A systems, an in-depth, systematic analysis of randomly selected question-answer sets is needed to develop a holistic understanding of user information exchange behavior. Furthermore, none of these studies has used the theories of question taxonomy in the investigation of the questions and corresponding answers. While Kim et al. (2007) and Kim & Oh (2009) look at the criteria people use to choose best answers as well as the relationship between question types (operationally defined by the researchers into three groups: information, suggestion, and opinion) and the criteria, the data they use are only the comments made by the questioner on the chosen answer. It is questionable whether these comments may be examined as stand-alone units, detached from

the actual question posted and answers provided. Additionally, the focus of their investigation is on the relevance criteria of the answers, rather than the characteristics of questions and answers themselves. In addition, even though Adamic et al. (2008) revealed that reply length and other external factors can predict an answer's possibility of becoming the best answer, these factors are not necessarily applicable in the environment of library and information services where the focus is on the construction of best answers based on some internal characteristics of the answer itself. As a result, research gaps remain in the empirical investigation of user-based Q&A systems through treating question-answer sets as "adjacency pairs," fully exploring the linguistic, discourse, and functional properties of questions and answers, and further defining the impact of social interaction on the shaping of how a question is posed, a response is generated, and a best answer is selected.

III.5. Sociolinguistic Analysis of Questions

While most of the theories and taxonomies of questions and answers discussed above concern the subject content of the questions, the social and pragmatic purposes of questions and answers can be analyzed as well. Graesser et al. (1994) describe these purposes as question generation mechanisms. They represent states that can exist in a questioner's mind that spur the question to be asked. The four question generation mechanisms identified by Graesser and his colleagues include 1) questions that seek to fill in knowledge gaps; 2) questions that establish and monitor common ground; 3) questions that coordinate action of a social nature; and 4) questions that direct the conversation and control attention. Categorizing questions according to these purposes provides a more comprehensive picture of not only the content being sought by the questioner, but also the social motives that, in many cases, are more important to the questioner than the content of the answers themselves.

Similarly, Freed's (1994) taxonomy takes into account the social function of questions within conversations as well as their specific informational content. In particular, questions in the relational categories (conversational focus, shared information, phatic information, and elaboration) and the expressive style categories (didactic function, rhetorical function, humor, self-directed function, and reported speech) convey as much (or more) information to the hearer as they actually seek from the hearer.

The sociolinguistic focus in both Graesser et al. (1994) and Freed (1994) is on questions as used in naturalistic conversation. We must recognize that the questions submitted to and answered in Yahoo!Answers are certainly not part of naturalistic conversation, just as reference inquiries (whether made face-to-face or virtually) are not part of truly naturalistic conversations either. However, adding a sociolinguistic component to our analysis of questions will give us a richer account of those factors beyond pure information content that are important both to questioners and answerers.

III. 6. Taxonomy Used for This Study

In summary, the previous research has provided a rich body of theories, frameworks and empirical knowledge about the nature of questions and answers, which serve as the foundation

for this research study. The coding schemes employed by the present study were developed primarily based on the following three classification mechanisms from the existing literature:

(1) For questions, the "wh-" question taxonomy by Robinson & Rackstraw (1972) was used, which classifies questions into the categories of "who," "which," "what," "when," "where," "why," and "how."

(2) For answers, since a description of the composition of an answer would be derived from the examination of sentences that made up the answer, the coding scheme is guided in part through the sentence classification theories. The following sentence typology, modified and synthesized from schemes articulated in Curme (1947) and Sadock & Zwick (1985), was used as the first part of the answer coding:

Sentence Type
- Declarative
- Imperative
- Exclamatory
- Fragmentary
- Interrogative

(3) For comments, the criteria presented by Kim, Oh, & Oh (2007) was applied to code and analyze comments. Note that the criteria and grouping published in Kim & Oh's (2009) paper are somewhat different from their original work. Since the coding scheme was originally developed based on authors' 2007 paper, the grouping was also done according to the Kim et al. 2007's structure.

Note that modifications and expansions have been made to the original schemes. A detailed description of the coding processes is provided in the next section.

III. Scope, Methodology, and Data Coding Processes

The first step of the project involved setting up a project management tool by using a password-protected wiki page that documents the relevant literature and project progression, with extensive meeting notes on coding decisions. A network drive was set up for storage of datasets. All team members (PI and Co-PI, technological consultant, and research assistant) have access to the wiki page and the network drive.

The second step involved data collection and sampling during which a stratified random sample of 3,000 question/answer sets was generated, with those sets stratified based on the 25 top-level categories assigned by Yahoo!Answers. The datasets were generated by the end of May of 2008 and, from June to August 2008, multiple word counting on questions, answers, and comments were performed. During this step the PIs worked closely with the technological consultant to discuss and work around the emerging problems/issues with API performance. The

research assistant was instructed to generate coding sheets based on the extracted datasheet. Meanwhile, more in-depth reading and literature reviews were performed on question types and answer taxonomy. Starting June 2008, both PIs began performing syntactic coding of questions. Once the questions coding was completed, the coding continued on question descriptions. Then start in November 2009, the coding for answers were conducted, with the PI and a research assistant coding on each answer by the type of sentences and the presences of certain elements such as reference and URL. In early 2010, the coding of comments started also with the PI and a research assistant working together to code the comments according to the criteria specified by Kim, et al. (2007). Additional criteria and their definitions were developed during the process of coding.   While data analysis for questions and their descriptions was completed in 2009, the data analysis for answers and comments was performed in July 2011 after the data processing for both coded datasets was completed.

IV. 1. Question and Question Description Coding

The coding of questions and question contents were based on the "wh-" question taxonomy (Robinson & Rackstraw, 1972). In addition to the regular categories of "who," "which," "what," etc., a category called "inversion" was added. Inversion questions are the questions that used an inversion of the word order such as "are you…" "do you.." to signal the format of a question. For example, the question "are tomatoes a fruit or veggie?" was coded as an inversion question. In cases such as "do you know where the capital of China is?" the question was double coded as both an "inversion" and a "where" question. Examples of double coding and special cases were discussed and documented on the project wiki. When the question does not fit into either the "wh-" categories or the "inversion" type, it was grouped under the "other" category.

The process of coding of questions and question content involved two coders performing coding individually and then reviewing the coding together to reach consensus on the final coding of each question. The use of informal language presents a challenge for coding; however, the coding itself went on very diligently and efficiently. Figure 1 is a sample coding sheet for questions in the category of "Family & Relationships." Table 1 shows the coding results after two coders consolidated and arrived at the final coding.

		Yahoo!Answers Data Collection									
		Family & Relationships									
Question	**Multiparts**	**Who**	**Which**	**What**	**When**	**Where**	**Why**	**How**	**Inversion**	**Others**	**Notes**
62 MY PARENTS ARE GETTING MARRIED AFTER 28 YEARS. SHOULD THEY THROW A CEREMONY										1	
63 How did you propose to your love?							1				
64 Advice girl problems?										1	
65 DOES this make me a HO?										1	
66 Should I give my friend another chance?										1	
67 Guys, I have a question for you?										1	
68 Is it a bad sign if....?										1	
69 Trying again! I've got this kid at school(high school) who keeps trying to goat me into an arguemen										1	
70 My mother is having an affair?										1	
71 Stay, or Leave? ...after 2 years	2									2	
72 i am in love with a muslim married guy and he lovesme al		1									
73 Ways to do this?										1	
74 Which photos?			1								
75 Guy friend wanted sex now he acting different...what did i		1									
76 i need a chick..?										1	
77 I have a huge phopia of not having a girlfriend. Advice?										1	
78 Help me understand??? Ok, so I am trying to be open minded, and not be a pain in the a** husband....?										1	
79 Name Change Question?										1	
80 i want to be with him but he has a girlfriend? what should		1									
81 who would you do?		1									
82 Anwser please guys and girls!!!?										1	
83 How do you get a girl to go on a date with you?							1				
84 what is the meanest thing u have ever done to your family		1									
85 Guys, is this true? Please Answer!!!?										1	
86 Why do husbands feel they ha	2				1		1				
87 Blondie or brunet what's the better girlfriend?			1								
88 made me moan just by touching my back?										1	
89 Girls, if you boyfriend told you he wanted you to dominate him, would you be happy to do it?										1	
90 Does the counting method work?										1	
91 hellp lol 2?										1	
92 serious help!!!!?										1	
93 guy I like says I'm like a sister?										1	
94 Talking to a Freshman Girl?										1	
95 How to "dump" a girlfriend?							1				
96 what would you tell your friend?				1							
97 MY GIRLFRIEND PASSED OUT dURING SEX!?!?										1	
98 Please Girls make your comments.?										1	
99 I know this is wrong but......?										1	

Figure 1 Sample Coding Sheet for Questions

Table 2 Coding of Question Types

	Who	Which	What	When	Where	Why	How	Inversion	Others
artshumanities	5	4	37	2	1	8	13	29	27
beautystyle	1	4	23	1	5	4	24	26	46
businessfinance	2	1	24	1	8	5	29	35	25
carstransportation	1	2	19	2	8	7	17	25	48
computerinternet	1	3	30	0	3	6	22	26	38
consumerelectroni	0	3	23	1	14	4	26	30	35
diningout	2	2	41	0	16	4	11	26	39
educationreferenc	7	4	35	2	4	5	24	18	33
entertainmentmusi	11	1	32	4	4	4	15	32	35
environment	2	2	18	2	5	11	18	32	43
familyrelationships	3	3	31	1	1	8	17	29	36
fooddrink	0	2	37	1	7	4	21	30	28
gamrecreation	2	1	19	1	7	6	24	26	51
health	0	1	28	1	4	5	12	36	49
homegarden	0	1	30	1	6	1	36	21	39
localbusiness	1	1	22	0	20	1	16	37	37
newsevents	4	0	28	1	4	12	13	48	32
pets	2	2	25	0	4	8	19	37	29
politicsgov	6	5	22	2	2	11	10	45	24
pregancyparenting	0	1	22	0	3	7	25	38	34
sciencemath	0	2	32	1	4	8	21	20	43
socialscience	4	2	32	0	0	14	18	35	26
societyculture	1	2	31	1	9	17	23	38	19
sports	16	3	28	2	6	9	18	27	25
travel	1	3	36	1	13	4	13	23	33

The coding of question content descriptions (users were given the opportunity when posting a question to provide a detailed description of the question context) were completed by December 2008. Figure 2 is a sample coding sheet for the description of the questions using the same categories as the ones for questions. Note that not all the questions had descriptions, so the coding sheet reflects coding for the questions that had descriptions/contents.

	YAQID	Que	Descriptio	Interrogat	who	Which	What	When	Where	Why	How	Inversion	Others
	Yahoo!Answers Data Collection: Question Content												
	Family & Relationships												
62	######	MY I	AFTER 28	2							1	1	
63	######	How did you propose to your love?											
64	20080326	Advi	I broke up	2								2	
65	20080326	DOE	ive cheated	2								1	1
66	20080326	Shou	My friend	1			1						
67	20080326	Guy:	Well I am	2								1	1
68	20080326	Is it	Ok, here&	2								2	
69	20080326	Tryin	what shoul	1			1						
70	20080326	My r	I know this	2									2
71	20080326	Stay	All he see	1									1
72	20080326	i am	am 22 and	1			1						
73	20080327	Way	Any sugge	2									2
74	20080327	Whi	reading										
75	20080327	Guy	guy friend	2								2	
76	20080327	i nee	who can c	1	1								
77	20080327	I hav	years old,										
78	20080327	Help	My wife an	1									1
79	20080327	Nam	I plan on c	2								2	
80	20080327	i want to be with him but he has a girlfriend? what should i do?											
81	20080327	who f you could do anyone in the whole world…celebrity, friend, politician etc.											
82	20080327	Anw	my guy frie	3							1	2	
83	20080327	How	I dont wan	1							1		
84	20080327	what is the meanest thing u have ever done to your family members?											
85	20080327	Guy:	was	5						1		3	1

Figure 2 Sample coding sheet for question content (descriptions)

IV.2. Answer Coding

The coding of answers took place between late 2008 to late 2009. The coding was done by the first PI and a research assistant, with each answer coded separately then confirming/verifying through cross checking. The process of coding answers involved multiple steps as follows:

(1) Create a coding sheet for each question;
(2) Record the number of sentences in each answer; also record the score for the best answer;
(3) Code each sentence in an answer by its question type, which include declarative, imperative, exclamatory, fragment, and imperative.
(4) If a sentence contains multiple parts, it was also recorded through a "1" for presence. The frequency of the sentences containing multiple parts was not coded;
(5) Finally each answer was coded by whether it contained scientific notation, URL, emoticon, reference, or e-abbreviations.

Below is an example for answers containing scientific notations. This particular answer was in the "Science and Mathematics" category:

The question (20080329031744AAUzgpG) was: can sum1 help me with dis quadratic prob.?

The best answer was: *q1   1. h(x) = a(x-b)+c = ax - (ab-c)   It's just a linear equation that form a line.   2. It is clear that that is a parabolic equation   3. h(x) = ax(x-b)(x-c) = (ax^2-abx)(x-c) = ax^3-acx^2-abx^2+abcx   it will form a line also.   So, the equation that perform a shape is the second equation...* [the reminder of the answer is excluded here]

If an answer contained symbol used for online expression such as ";)", it would be coded as having "emoticon." The e-abbreviations would be the combination of first letter of each word such as "LOL," "TTYL," or self-invented spelling abbreviations such as "r u…," "abt," "athg," or "digi" etc. Whether an answer contained URL or reference was also noted. Figure 3 below is a screen shot of the answer coding sheet in a question in the "Sports" category.

	Answer Categories	#Sentences	Multi-Parts	Declarative	Imperative	Exclamatory	Fragment	Interrogat	Scientific	URL	Other	Email	Emoticon	Reference	E-Abbr.
1	Yahoo!Answers Data Collection: Answer Analysis														
2	Sports!!														
3	Answer Categories	#Sentences	Multi-Parts	Declarative	Imperative	Exclamatory	Fragment	Interrogat	Scientific	URL	Other	Email	Emoticon	Reference	E-Abbr.
4	Question ID	20080326085218AAmCgQk													
5	Question Nam	Should the NBA have "challenges" troughout entire games??													
6	Question Cont	Like in Tennis and Football. Should the NBA let players have a chance to argue a call? Guve players a chance to argue a call. Not neccesay fouls but something like out of bounds calls and deflections a
7	UserNick	andy p													
8	Num Answers	7													
9	(UserNick) An:	(pinoyplaya2	4		2		1		1						
10	(UserNick) An:	(Din) No bec	2		2										
11	(UserNick) An:	(michael d) r	2	1	2										
12	(UserNick) An:	(greg) no, itll	2	1	2										
13	(UserNick) An:	(chinamigarc	4		4										
14	(UserNick) An:	(Captain Ten	2	1	2										
15	Best	5													
16	(UserNick) An:	(jhendricks3	5	1	5										

Figure 3 Sample coding sheet for answer coding

IV.3. Comments Coding

The actual coding of comments took place between January and May of 2010. The comments are typically askers' comments once they gave a score to the best answer. However, not all askers gave a comment and in some cases, people other than askers made some comments. For all the questions that were sampled, there are a total of 1536 that contained comments (about 51% of the questions). Among the comments, there are 1481 made by the askers, which counts for 96% of the comments in the sample. Across the 25 Yahoo!Answers categories, the average number of comments is 61 for total comments, and 59 for comments made by askers. Table 2 gives detailed about the number of comments per category. Note that categories such as "Family Relationship" "Health" "Society and Culture" and "Beauty & Style" had the highest number of comments, while categories such as "Local business," "Home Garden" and "Cars and Transportation" had the lowest number of comments.

Table 3 Number of Comments in the Sample

Question Categories	No. of Comments	No. of Comments Made by Askers
Arts&Humanities	55	53
Beauty&Style	80	80
Business&Finance	47	47
Cars&Transportation	44	42
Computer&Internet	71	69
ConsumerElectronic	47	44
Diningout	47	44
Education&Reference	59	57
Entertainment&Music	76	73
Environment	46	46
Family&Relationships	86	85
Food&Drink	58	56
Game&Recreation	60	54
Health	84	84
Home&Garden	43	42
LocalBusiness	35	29
News&Events	46	45
Pets	57	55
Politics&Government	74	72
Pregnancy&Parenting	72	70
Science&Math	61	58
SocialScience	72	69
Society&Culture	87	82
Sports	73	70
Travel	56	55

Comments were coded based on the 26 criteria developed by Kim et al (2007). Table 3 lists these criteria and the grouping. Note that for most of the analysis, these criteria were treated as individual items, only in one section of the report the criteria were put into different classes for the purpose of comparison. While most of the criteria were given some definitions or examples, these definitions/examples were not necessarily specific or operationalizable. In the process of coding, we developed more concrete definitions of the criteria and also provide specific examples for a criterion. All the coding process was documented on the project wiki page. Figure 4 is a screenshot of the coding examples for one of the criteria – "Writing Style." Additional criteria such as "Answer Extension," "Symbols" created to cover instances where none of the 25 criteria

describe. For a complete set of definitions of all criteria, see Appendix A. Figure 5 displays a sample coding sheet for comments made to the answers of a question in the category of "Beauty & Style."

Table 4 Criteria for Coding Comments

Type	Criteria Label	Grouping
Original Criteria	Accuracy	Content
	Scope & specificity	
	Clarity	
	Writing style	
	Completeness	
	Existence of particular detail	
	Characteristics	
	Content novelty	Cognitive
	Perspective novelty	
	Understandability	
	Questioner's background/experience	
	Emotional support	Social-emotional
	Answerer's attitude	
	Answerer's effort	
	Answerer's experience	
	Agreement	
	Confirmation	
	Affect	
	External verification	Extrinsic
	Available alternatives	
	Quickness	
	Reference to external sources	Information Source
	Source quality	
	Effectiveness	Utility
	Solution feasibility	
	General Statement	
Additional Criteria	Answer Extension	Grouped to Content
	Symbols	
	Other	
	Questioner's Excitement	Grouped to Social-emotional

Figure 4 Documented coding examples for the criterion "writing style."

Figure 5 Sample coding sheet for comments

IV. Data Analysis

Data analysis of question and question content was performed both at the descriptive level and inferential level. Descriptive statistics included: frequency of answers provided per question; average length of time to first answer; distribution of subject categories; distribution of question types; and the top categories where a particular type of questions got asked. Factor analysis was performed based on types of questions each category received. Comparisons were also made between data patterns for questions alone and for questions and question description combined.

The analysis of answers focused on the characteristics of best answers. The descriptive summaries included (1) the average score for best answers (2) the average number of sentences that made up the best answer; (3) the proportion of the best answers contained sentences that consisted of multiple parts; (4) the proportion of the best answers that are of different types of sentences; and (5) the proportion of the best answers contained elements such as scientific notations, URLs, email addresses, references, emoticons, abbreviations, etc. Correlation analysis was performed at multiple levels: (1) the association between the score and various sentence types and the presences of certain elements. (2) the association among sentence types based on their frequencies, (3) the association among elements contained based on their presences in the best answers (4) the association between sentence type and element contained. Furthermore, analysis was also performed on a smaller sample data which consisted on top two best score categories (Home&Garden and News&Events), bottom two best score categories (Beauty&Style, Computer&Internet), and the middle one best score category (Food&Drink).

The analysis of comments was centered on examining the comments made by askers. Descriptive summaries were produced on the frequency of criteria mentioned by the askers when commenting on why he/she choose the answer as the best answer. A summary was also provided to list the Yahoo!Answers categories that contained the highest proportions of the comment type. Correlation analysis was performed to ascertain significant associations among various comment criteria.

V. Results

The report of the results is sequenced by the order of (1) question and question description analysis; (2) answer analysis, and (3) comment analysis.

VI.1. Analysis of Questions and Question Descriptions

VI.1.1 Average Time Length Taken to Receive the First Answer. Table 4 lists the average number of minutes for a user to post a question and receive the first answer from the user community. The top three categories that took the longest time to get a response were "Local Business," "Entertainment and Music," and "Business and Finance." The categories that received

quickest answers were "Family and Relationships," "Pregnancy and Parenting," and "Food and Drink."

**Table 5 Average length (minutes elapsed) to receive the first answer**

	Average length (min.) to receive first answer
Family&Relationships	10.8
Pregnancy&Parenting	41.78
Food&Drink	59.86
Beauty&Style	74.83
SocialScience	87.4
Home&Garden	90.52
Science&Math	157.47
Health	163.28
News&Events	163.67
Society&Culture	171.22
Arts&Humanities	182.07
Environment	197.31
Sports	277.68
Pets	286.07
Politics&Gov	302.37
Cars&Transportation	326.75
Diningout	346.44
Computer&Internet	370.91
Education&Reference	402.5
Travel	411
ConsumerElectronic	463
Game&Recreation	485.2
Business&Finance	635
Entertainment&Music	660.77
LocalBusiness	1635.04

VI.1.2. Average Number of Answers by Categories. In terms of number of answers received for each question within a given category, categories such as "pregnancy and parenting" and "dining out" receive more than seven answers per question, while categories such as "local business" and "business and finance" received less than three answers per question. Figure 6 illustrates the average number of answers received per question for the 25 Yahoo!Answers categories.

Average Number of Answers
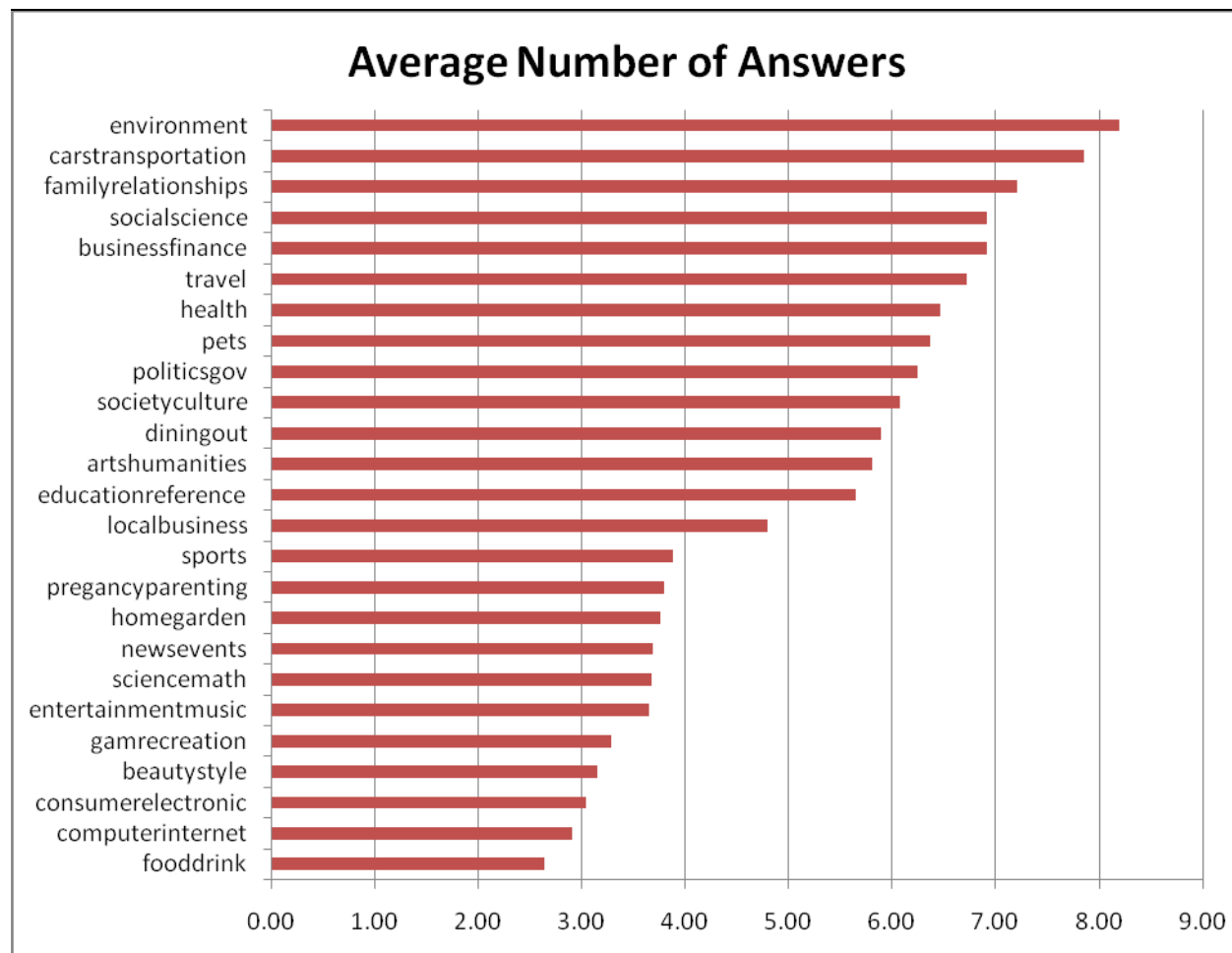
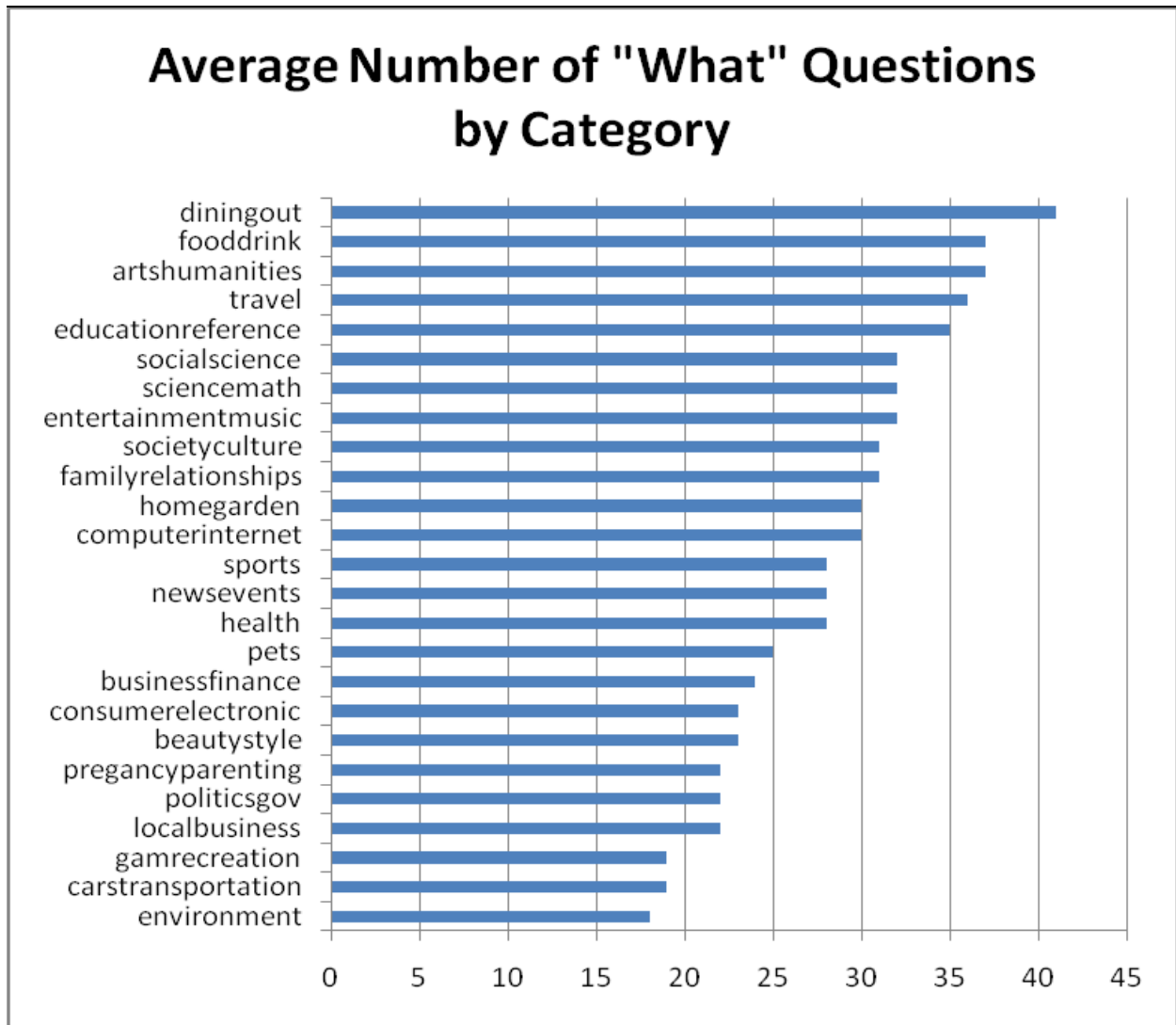**Figure 6 Average number of answers by category**

VI.1.3. Question Type. In terms of the type of questions, the frequencies vary based on the categories. For instance, the "what" questions occurred most frequently in the category of "dining out,""food and drink," and "arts and humanities." The "how" questions were most frequently seen in the category of "home and garden,""business and finance" and "consumer electronics." The categories of "society and culture," "social sciences," and "news events" contain the highest number of "why" questions. Note that the average number of "why" questions across categories is much smaller than the "what" or "how" questions. For inversion questions, the top three categories were "news and events," "politics and government," and "society and culture." Figure 7 provides the average frequency of question type across 25 categories, whereas Table 5 lists only the Yahoo!Answers categories that have the highest frequency by question type. As it is shown from Table 5, "Sports" had the highest number of "who" questions, whereas "which" was popular in asking "politics and government" questions. Understandably, people mostly asked "where" questions when concerning local business, and they asked "when" regarding "entertainment and music." In looking at the most popular type of questions get asked in what Yahoo!Answers categories, the data patterns seemed reasonable and interesting. Figures 8-11 display the actual distributions of those "wh-" questions.
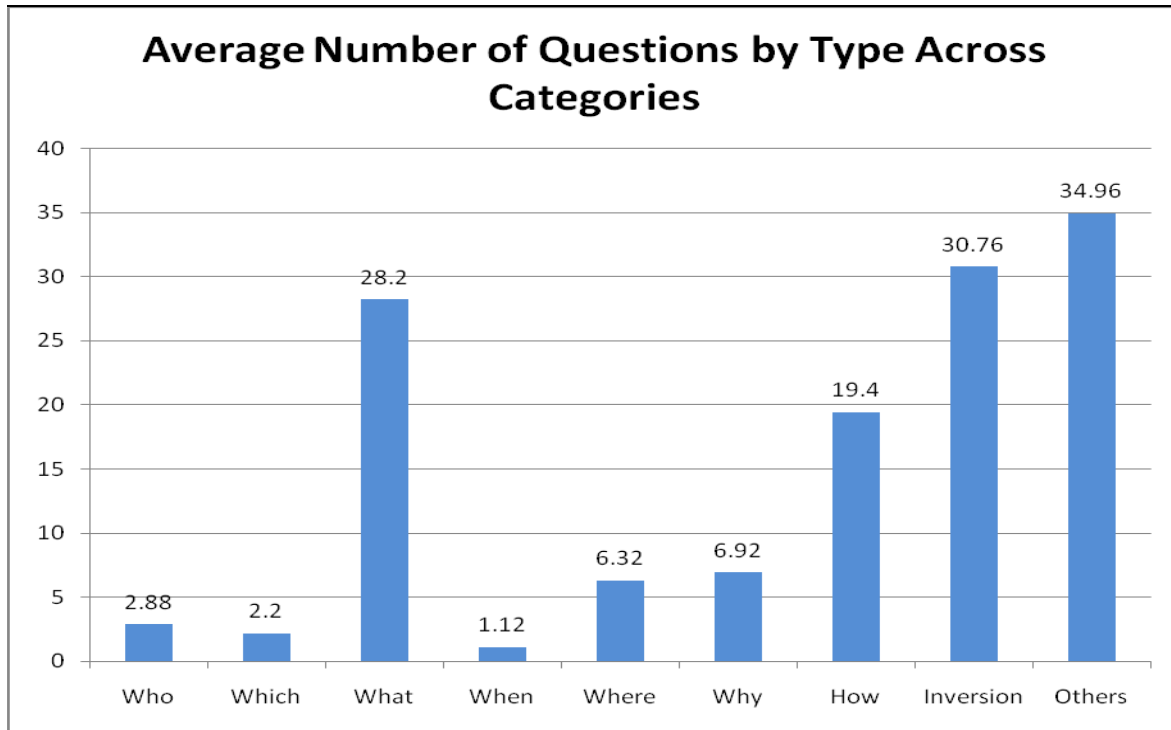
Figure 7 Average number of question by types

Table 6 Yahoo!Answers Categories Containing Highest Counts of Question Type

Top Category	Question Type	Number of Questions
Sports	Who	16
Politics&Government	Which	5
DiningOut	What	41
Entertainment&Music	When	4
LocalBusiness	Where	20
Society&Culture	Why	17
Home&Garden	How	36
News&Events	Inversion	48
Game&Recreation	Other	51

Average Number of "What" Questions by Category

Category	Value
diningout	41
fooddrink	37
artshumanities	37
travel	36
educationreference	35
socialscience	32
sciencemath	32
entertainmentmusic	32
societyculture	31
familyrelationships	31
homegarden	30
computerinternet	30
sports	28
newsevents	28
health	28
pets	25
businessfinance	24
consumerelectronic	23
beautystyle	23
pregancyparenting	22
politicsgov	22
localbusiness	22
gamrecreation	19
carstransportation	19
environment	18

Figure 8 "What" questions frequency

Average Number of "How" Questions by Categories

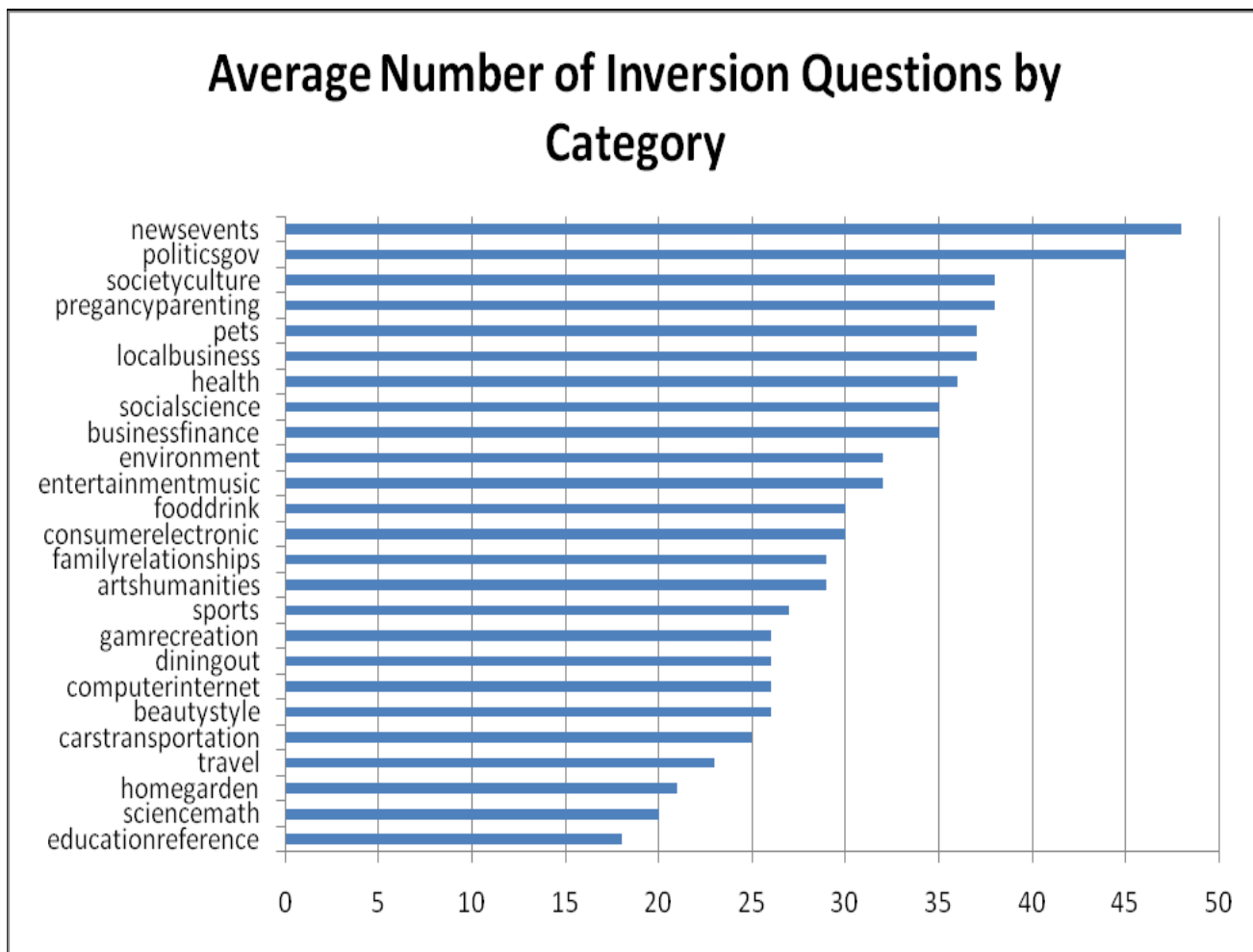| Category	Value
homegarden	36
businessfinance	29
consumerelectronic	26
pregancyparenting	25
gamrecreation	24
educationreference	24
beautystyle	24
societyculture	23
computerinternet	22
sciencemath	21
fooddrink	21
pets	19
sports	18
socialscience	18
environment	18
familyrelationships	17
carstransportation	17
localbusiness	16
entertainmentmusic	15
travel	13
newsevents	13
artshumanities	13
health	12
diningout	11
politicsgov	10

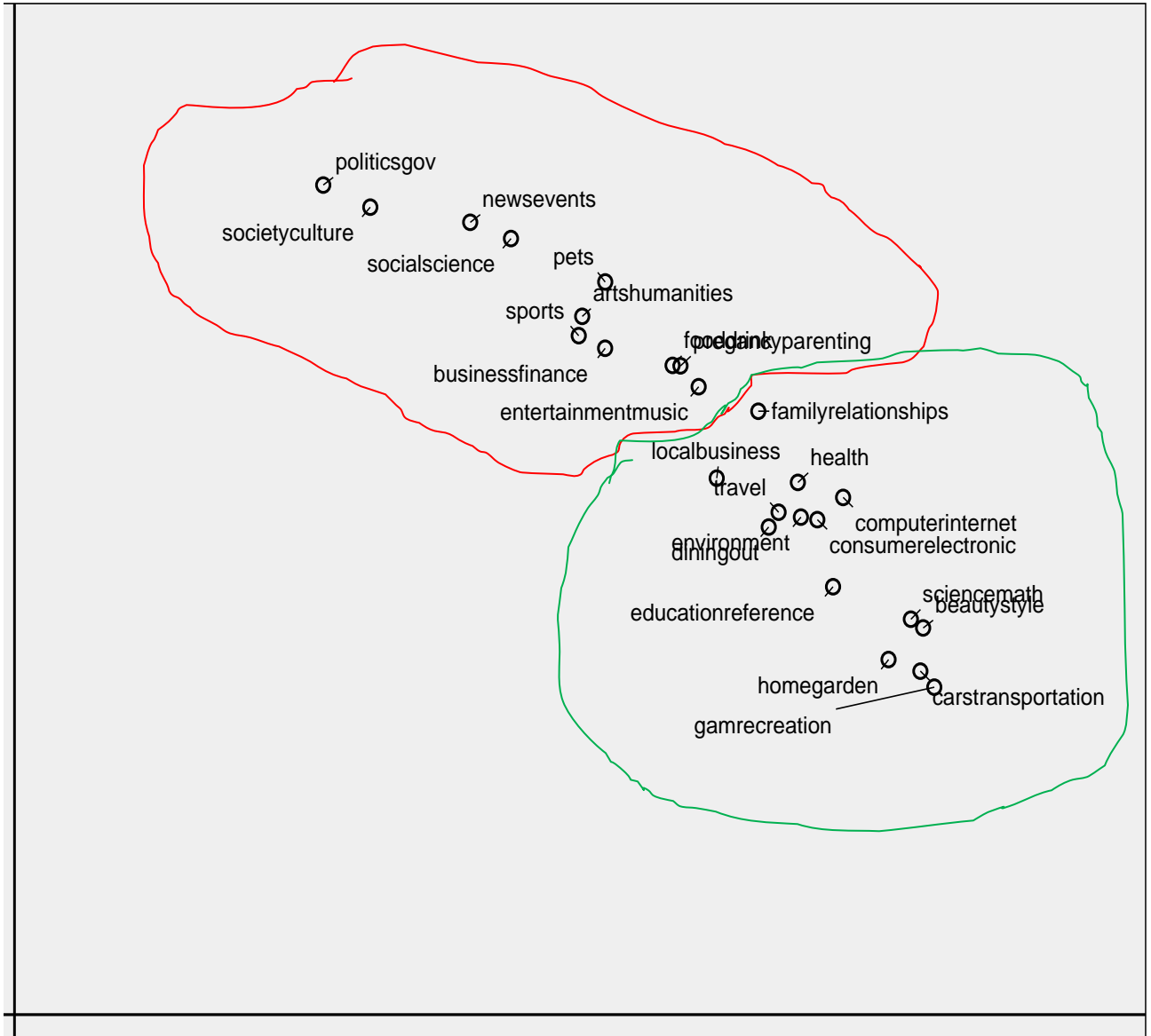**Figure 9 "How" questions frequency**
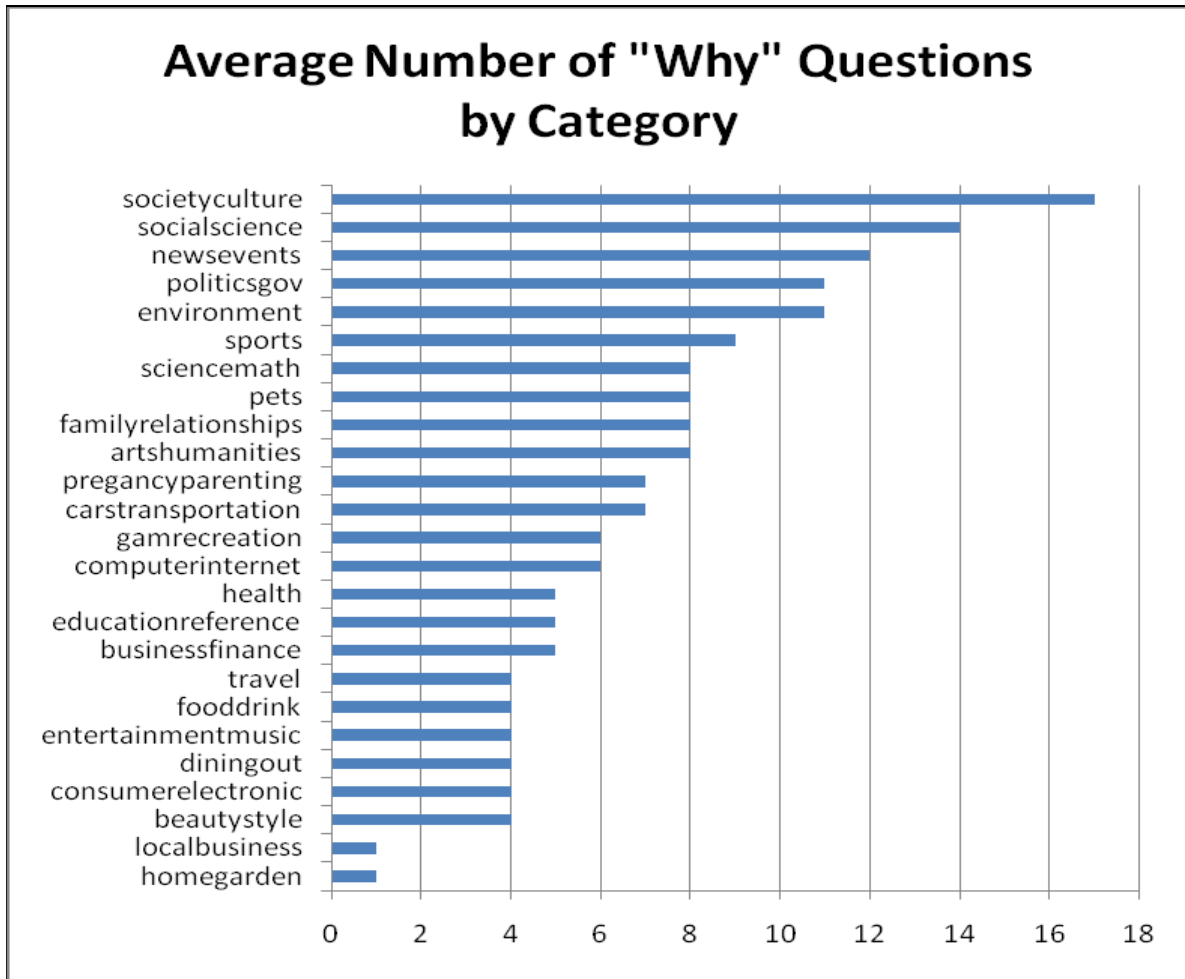
Figure 10 "Why" questions frequency

Figure 11 Inversion questions frequency

VI. 1.4. Analysis of Differences. Based on the frequency of kinds of questions, a factor analysis and it resulted in two components among 25 categories. Figure 12 shows the clustering of the 25 categories. For the two components, categories such as "Politics and Government," "News and Events," "Society & Culture," "Pets," "Sports," etc., belonged to the component one, whereas "Family & Relationships," "Home & Garden," "Environment," and "Health," etc., belonged to a different component.

Figure 12 Factor analysis based on frequency of kinds of questions

Using the components for grouping, MANOVA revealed significant difference for **Who** (p=0.04), **Why** (p=0.01), **Inversion** (p=0.001), and **Other** (p=0.00) questions. The number of answers per question was significantly different between the two groups of categories (p=0.02).

VI.1.5. Question Descriptions. For each question, the user is prompted to elaborate in the text box underneath the one-line question box. Our coding showed that the questions in question description followed similar patterns as the questions themselves. Figure 13 shows the average number of combined question and question description counts across different question type. In comparison with Figure 7, the only difference was a higher number of inversion questions than

the "other" question type. In Figure 7, the number of "other" question type is higher than inversion questions.
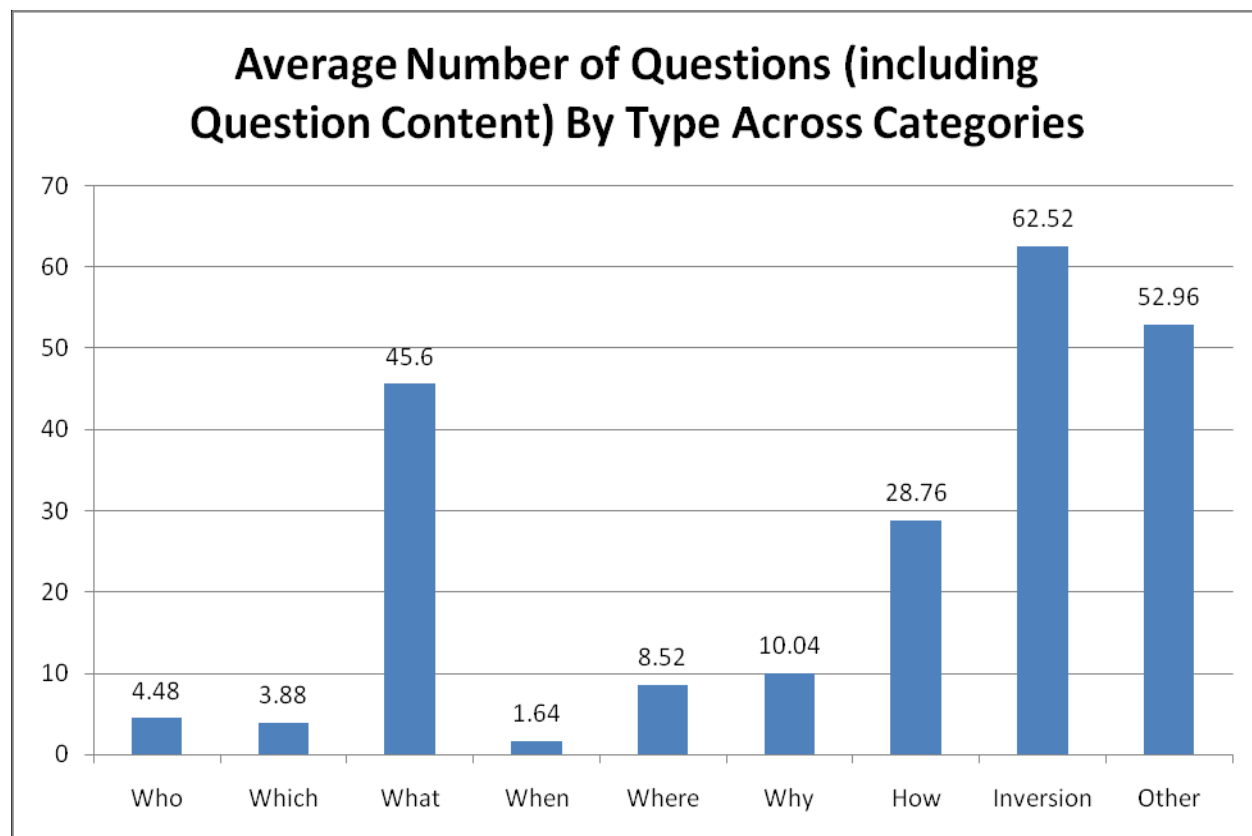
Figure 13 Average Number of Questions (with Question Description Included) by Type

In terms of the most popular question type asked in which Yahoo!Answers categories, there were some changes with regard to the overall distributions when the questions in the question description field were added. Table 6 provides the total counts. In comparison to Table 5, the differences include

(1) "Beauty & Style" had the highest number of "which" questions
(2) "Arts & Humanities" tied with entertainment music of having the highest number of "when" questions
(3) "Education & Reference" became the top category for "what" questions
(4) "Health" became the number one category for inversion questions as well as "other" question.
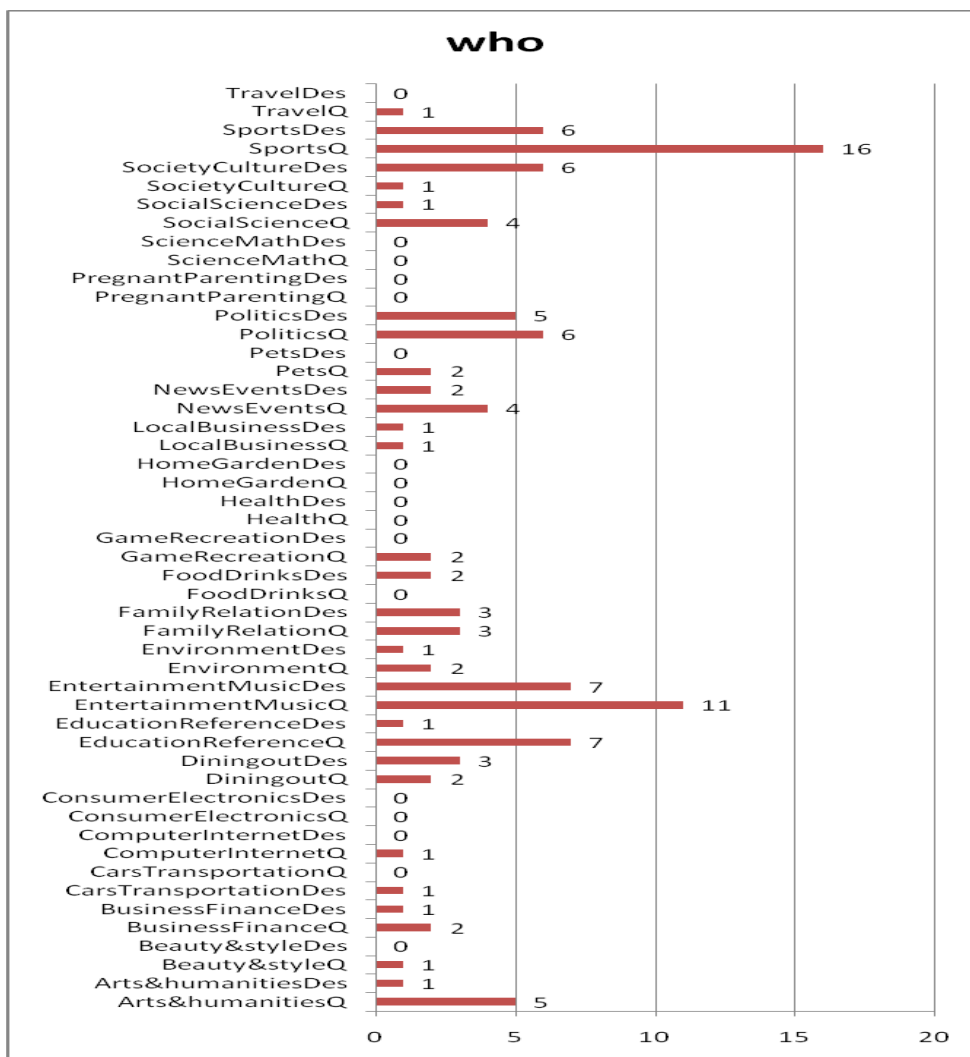
The following patterns are the same:

(1) "Sports" contained the highest number of "who" questions

(2) "Local Business" had the largest group of "where" questions
(3) "Home & Garden" is where people asked the most "how" questions
(4) "Society &Culture" is the top category for "why" questions.

Table 7 Yahoo!Answers Categories Containing Highest Counts of Question Type (Questions and Question Descriptions Combined)

Top Categories	Question Type	Number of Questions
SportsTotal	Who	22
Beauty&Styletotal	Which	9
Education&ReferenceTotal	What	76
Arts&Humanitiestotal	When	4
Entertainment&MusicTotal	When	4
LocalBusinessTotal	Where	25
Society&CultureTotal	Why	27
Home&GardenTotal	How	46
HealthTotal	Inversion	94
HealthTotal	Other	92

Figure 14 below presents the comparative counts of "who" questions in questions and question descriptions by categories, where category label and a "Q" stands for question proper, and category label with a "Des" stands for question descriptions.

who

Category	Value
TravelDes	0
TravelQ	1
SportsDes	6
SportsQ	16
SocietyCultureDes	6
SocietyCultureQ	1
SocialScienceDes	1
SocialScienceQ	4
ScienceMathDes	0
ScienceMathQ	0
PregnantParentingDes	0
PregnantParentingQ	0
PoliticsDes	5
PoliticsQ	6
PetsDes	0
PetsQ	2
NewsEventsDes	2
NewsEventsQ	4
LocalBusinessDes	1
LocalBusinessQ	1
HomeGardenDes	0
HomeGardenQ	0
HealthDes	0
HealthQ	0
GameRecreationDes	0
GameRecreationQ	2
FoodDrinksDes	2
FoodDrinksQ	0
FamilyRelationDes	3
FamilyRelationQ	3
EnvironmentDes	1
EnvironmentQ	2
EntertainmentMusicDes	7
EntertainmentMusicQ	11
EducationReferenceDes	1
EducationReferenceQ	7
DiningoutDes	3
DiningoutQ	2
ConsumerElectronicsDes	0
ConsumerElectronicsQ	0
ComputerInternetDes	0
ComputerInternetQ	1
CarsTransportationQ	0
CarsTransportationDes	1
BusinessFinanceDes	1
BusinessFinanceQ	2
Beauty&styleDes	0
Beauty&styleQ	1
Arts&humanitiesDes	1
Arts&humanitiesQ	5

Figure 14 Comparison of "Who" question frequency in question proper and question description

VI. 2. Analysis of Answers

The in-depth analysis of answers was geared mainly towards syntactic characteristics for the best answers. Note that best answers were typically given by the asker, with a rating ranging from 1 to 5, with 5 as the highest score.

VI.2.1. Best answer by sentence type. Table 7 includes the proportion of best answers by particular sentence type in each category, with the highest proportions marked in dark green. It is clear that "declarative" sentences made up the highest proportion of the best answer sentences across all 25 categories. The category of "Politics & Government" contained the highest proportions of declarative sentences. For a majority of the categories, the second largest sentence type for best answers was "imperative" sentences. The category of "Food & Drink" had the

highest proportion of imperative sentences in its best answers. With the remaining sentence types, the proportions were all lower than 10%, with "Family & Relationships" containing highest percentages in exclamatory sentences, "Science & Mathematics" in "fragment" sentences, and "News and Events" in interrogative sentences. Figure 15 displays the average proportion of sentence types in best answers for all categories.

Table 8 Best Answer Sentence Type across Categories

Categories	Declarative	Imperative	Exclamatory	Fragment	Interrogative
Arts&Humanities	84.38%	7.26%	2.52%	1.42%	4.42%
Beauty&Style	62.71%	25.97%	7.24%	0.56%	1.67%
Business&Finance	69.61%	22.18%	3.49%	1.85%	3.08%
Cars&Transportation	71.81%	19.27%	2.64%	0.00%	6.29%
Computer&Internet	79.51%	18.03%	1.64%	0.00%	1.64%
ConsumerElectronic	80.73%	11.78%	4.50%	0.00%	2.78%
Diningout	82.09%	6.05%	6.05%	3.26%	3.26%
Education&Reference	81.15%	10.19%	4.81%	0.19%	2.69%
Entertainment&Music	84.71%	4.71%	6.18%	0.29%	1.47%
Environment	81.44%	7.48%	2.88%	0.00%	3.74%
Family&Relationships	68.71%	18.06%	7.42%	0.16%	5.65%
Food&Drink	67.03%	26.78%	4.44%	0.13%	2.02%
Game&Recreation	77.46%	12.88%	7.04%	0.20%	3.62%
Health	76.90%	15.51%	4.38%	0.00%	2.19%
Home&Garden	70.40%	21.32%	3.31%	0.18%	4.04%
LocalBusiness	74.39%	13.93%	5.53%	1.23%	6.35%
News&Events	77.90%	9.82%	6.92%	0.22%	7.81%
Pets	75.11%	14.63%	7.09%	0.60%	3.47%
Politics&Government	87.76%	4.48%	2.93%	1.90%	3.28%
Pregnancy&Parenting	78.53%	11.39%	7.33%	0.52%	3.14%
Science&Math	83.26%	4.43%	2.29%	8.87%	1.14%
SocialScience	83.79%	6.10%	4.82%	0.48%	5.30%
Society&Culture	85.28%	5.25%	6.09%	0.68%	5.92%
Sports	80.43%	9.79%	5.53%	1.91%	2.98%
Travel	83.66%	8.17%	6.04%	0.89%	2.84%
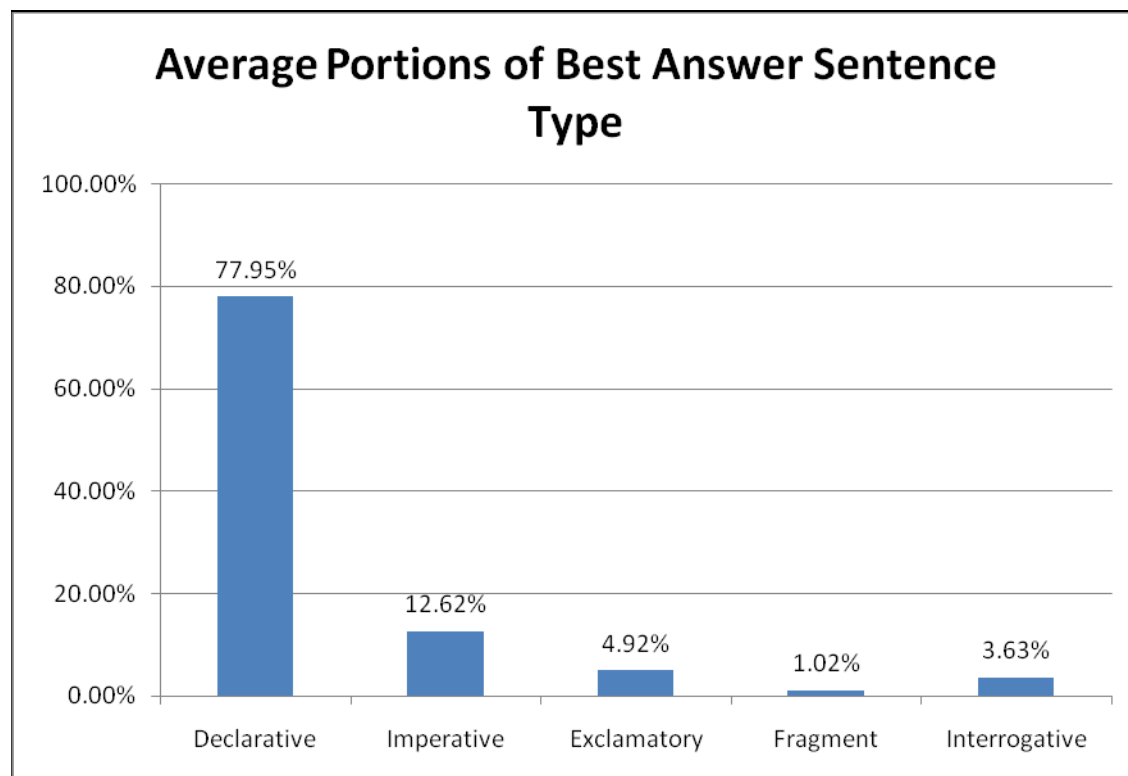
Figure 15 Average portions of the best answer sentence types

     VI.2.2. Correlation between Best Score and Sentence Type. Correlation analysis was performed to ascertain the association between the scores received for best answers and the proportion of the answer sentence types. Statistically significant negative correlation was identified between the proportion of declarative sentence and the score (r=-0.44, p=0.028), and there was significant positive correlation between proportions of imperative sentences in the best answer and the score (r=0.46, p=0.022). This is quite interesting in that while the presence of declarative sentence leads to a lower score, the use of imperative sentence would gain higher points. It appears that confidence and authoritative tone conveyed through an imperative sentence would be an indicator for a better answer. The level tone delivered through declarative sentences does not seem to be believable to askers as the best answers, although the majority of the best answers are made of declarative sentences.

     VI. 2.3. Correlation between Best Score and Elements Contained. Correlation analysis was also performed to examine the association between the scores received for the best answer and the percentage of the answers containing elements such as URL, reference, email, emoticon, etc. No significant correlations were found.

     VI.2.4. Correlation among Best Answer Sentence Type. Based on the proportions of answer sentences occurring in best answers across the categories, correlation analysis among various sentence types was performed. It was found that there was a significant negative

correlation between imperative and declarative sentences (r=-0.941, p=0.000). No other pair-wise correlations were significant.

VI.2.5. Correlation among Best Answer Elements Contained. Based on the proportions of certain features or elements appearing in best answers across the categories, correlation analysis among various feature types was performed. It was found that there was a significant negative correlation between existence of emoticon and the presence of a URL in the best answers (r=-.438, p=0.028). It was also found that the presence of URLs in the best answers correlate positively with the inclusion of reference in the best answer (r=0.726, p=0.000).

VI.2.6. Correlation between Sentence Type and Elements Contained. Three elements hold significant association with the frequency of certain best answer sentence types. These elements are:

(1) Emoticon. The presence of emoticon is negatively associated with the frequency of the declarative sentences in the best answer (r=-0.578, p=0.002); The presence of emoticon is positively correlated with the proportion of the imperative (r=0.492, p=0.012); and exclamatory sentences that made up the best answer(r=0.474, p=0.017);
(2) Reference. The presence of a "reference" element in the best answer is negatively associated with the fragment sentence type (r=-0.418, p=0.037);
(3) Scientific notation. The presence of scientific notation in the best answer is negatively associated with exclamatory sentence type (r=-0.543, p=0.005).

The above correlations suggest interesting characteristics of the online social communication. The fact that emoticon is positively associated with imperative sentences and exclamatory sentences sounds reasonable in that both sentence types require certain emotional elements that would be expressed through an "emoticon" symbol.

VI.2.7. Correlation Analysis of Representative Sample Data. During the process of correlation analysis, it was decided that while it was not feasible to analyze the entire answer sample of 15346 answers, it would be a meaningful exercise to take a partial sample data that contained individual scoring and counts of the sentence types and elements contained. A smaller sample data that included answers to categories that had top two average best score (Home&Garden, News&Events), the middle point (Food&Drink), and bottom two (Computer&Internet, Beauty&Style) was taken to run some further correlation analysis. No significant correlations were found between the number of answers and best score, as well as between the number of sentences and best score; No significant correlations were found between the best score and number of different sentence types. There were significant positive correlation between the best scores and the presences of URLs in the best answers (r=0.083, p=0.042) and significant negative correlation between the best score and the presence of emoticon (r=-0.112, p=0.006).
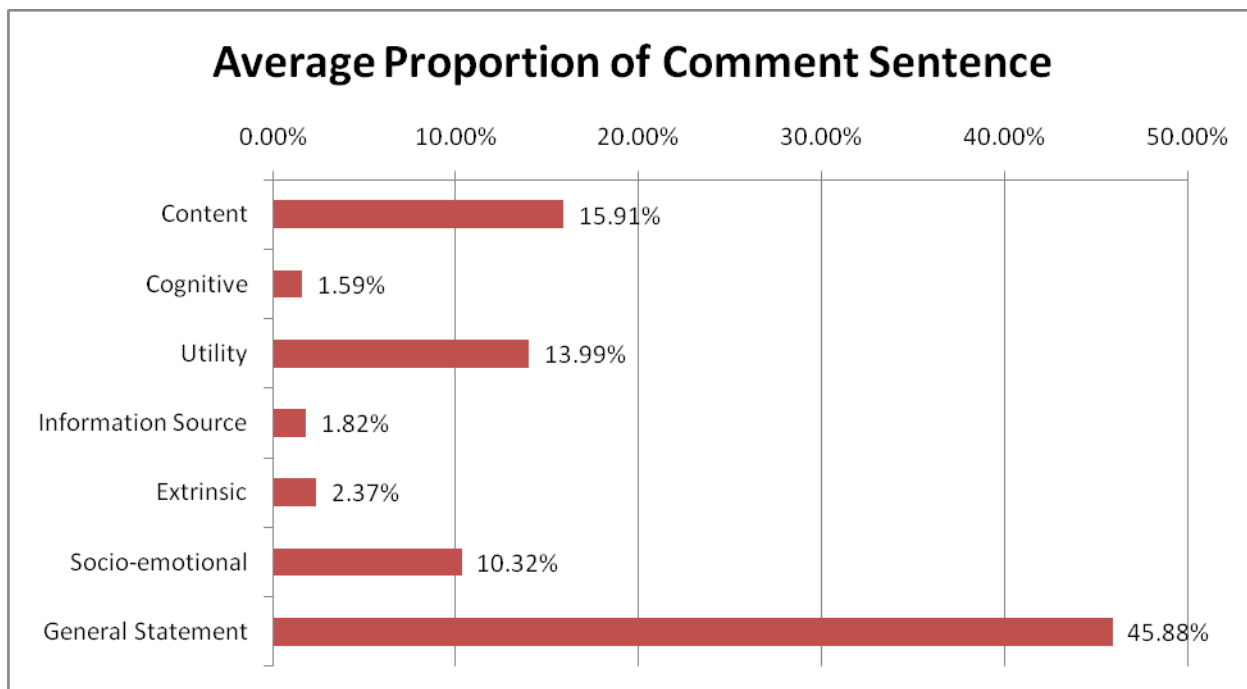
VI.3. Analysis of Comments (criteria)

VI.3.1. *Average Proportions of Comments Type across Categories*. A total of 1481 comments by the asker were analyzed based comments criteria. It is found that almost half of the comments were "general statement" (46%). General statements are comments made by askers to describe his/her overall impression of the answer and his/her gratitude to the answers, which typically contains a simple "Thanks." (For a complete list of the comments criteria definitions and examples, see Appendix A). The remaining categories of comments are all less than 10%. Specifically, "effectiveness," "solution feasibility," "accuracy," "characteristics, agreement, and scope & specificity are above 2%. Figure 16 reflects the proportions of all categories in the original coding. Note that other than the "general statement," the top three reasons that the askers gave their answerers the best score included "effectiveness," "solution feasibility," and "accuracy."
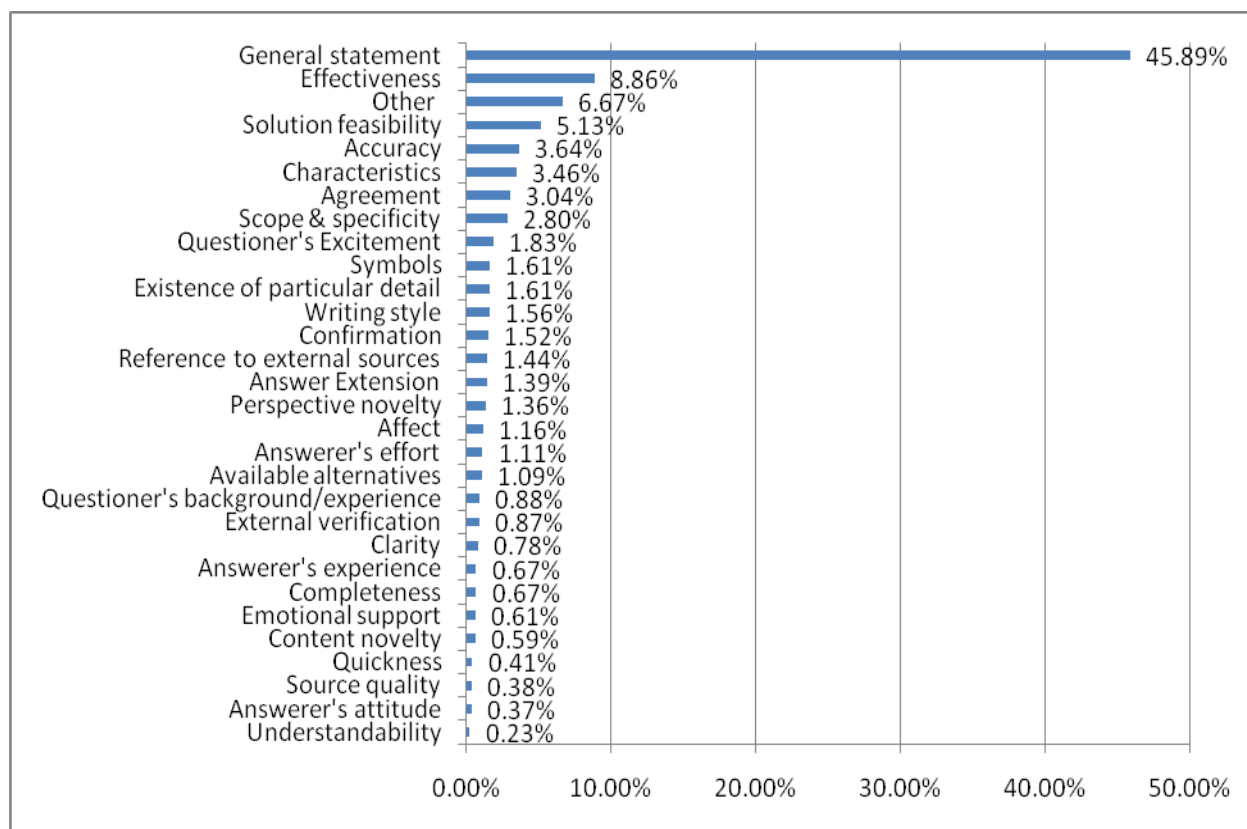
Figure 16 Average Proportion of Comment Type across Categories

An attempt was made to group the comments based on the criteria classes outlined by Kim et al. (2007). For additional criteria, "answer extension" was grouped to the "content" category, and "questioner's excitement" was grouped under the socio-emotional category. The symbol and other categories were discarded in the grouping. Figure 17 shows the proportions of comment sentences that reflect the selection criteria by class. Note that when excluding general

statements, the top three groups of comments were concerned about answer's content (16%), utility (14%) and its socio-emotional value (10%). This is somewhat different from Kim & Oh's (2009) results where the number one class was "social-emotional" (30%), then "content" (26%) and thirdly "utility" (23%). It is important to note that general statement did not factor into Kim and Oh's (2009) percentages.

Figure 17 Average Proportion of Comment Criteria

VI.3.2. Maximum Proportion of Comments Category by Criteria. The analysis of the comment type and the category that contained the highest portion of best answers in a given type revealed interesting patterns. Table 8 lines up the information about the Yahoo!Answers category that had the highest portion of each of the comment type. It is shown that for "cars and transportation," "external verification" stood out as one of the key criteria to determine whether the answer was the best one. For "Consumer Electronics," answerer's experience is important and the asker tends to get very excited in their comments to the content of the best answer. For advices on "dining out," reference to external sources was essential, whereas quickness to respond was critical in the best answer to questions related to music and entertainment. The fact that the answers contained "particular detail" was weighted significantly for questions on the topic of health; whereas whether the answers were effective was a measure of success in the "games & recreation" category. In order to be chosen as the best answer to questions in government and politics, the answer needed to be accurate, but also allow people debate on issues they agree or disagree. Interestingly, for answers concerning news and events, content novelty is crucial, whereas the best answers for pregnancy and parenting should lend users some emotional support. For sports, people were very affective in their best answers, whereas confirmation was the key to indicate the best answer to questions in the subject matter of science and mathematics.

All of these top categories seem very logical in their association with unique criteria determining the quality and characteristics for a best answer.

Table 9 Categories Containing the Highest Proportions of Comment Type

Top Categories	Criteria	Proportion
Arts&Humanities	Source quality	5.63%
	Answer Extension	4.23%
	Understandability	1.41%
Beauty&Style	General statement	66.40%
Cars&Transportation	External verification	3.08%
Computer&Internet	Available alternatives	5.38%
ConsumerElectronic	Questioner's Excitement	6.90%
	Answerer's experience	3.45%
Diningout	Reference to external sources	6.56%
	Answerer's attitude	3.28%
Education&Reference	Answerer's effort	3.95%
Entertainment&Music	Questioner's background/experience	5.56%
	Quickness	1.85%
Environment	Scope & specificity	8.86%
Family&Relationships	Other	20.83%
Game&Recreation	Effectiveness	21.84%
Health	Existence of particular detail	5.00%
Home&Garden	Clarity	3.23%
	Completeness	3.23%
LocalBusiness	Solution feasibility	12.50%
	Symbols	5.00%
News&Events	Content novelty	6.94%
Politics&Government	Agreement	13.89%
	Accuracy	8.33%
	Writing style	6.48%
Pregnancy&Parenting	Emotional support	5.56%
Science&Math	Confirmation	4.76%
SocialScience	Perspective novelty	5.05%
Society&Culture	Characteristics	14.01%
Sports	Affect	4.72%

VI.3.3. Correlation Among Comment Criteria. Based on the proportion of the comment type occurring across the Yahoo!Answers categories, a correlation analysis was performed to examine the relationship among comment criteria. Table 9 includes all the significant correlations among a variety of pair-wise variables. In Table 9, a correlation is only reported once. For instance, the correlation between accuracy and scope specificity only occurs once in the accuracy row. "Other" as a criteria category is not reported here. As a result, later criteria such as "content novelty" do not show up in the first column of criteria list, due to the fact that the correlation with an early listed criterion was already there. Note that while the majority of the correlations were positive, there were five negative associations (marked in yellow). Specifically, "accuracy" is negatively associated with "scope specificity," "completeness" is negatively associated with "affect" and "quickness." While higher level "perspective novelty" leads to a lower level of "solution feasibility," the more experience that an answer has, the less likely he/she will include "reference to external sources" in the answer.

Similar to the correlation results shown in previous section, the correlations revealed in Table 8 also seem sensible. For example, one would imagine that clarity is associated with completeness whereas the quicker one receives an answer, the less likely the answer would be complete. It is quite understandable that understandability has to do with source quality, and the more the answer has experiences, the more likely his/her answer would be effective.

Table 10 Correlations among Comment Criteria Based on Proportion of Comment Type

Criteria	Correlated Criteria	Coefficient and p Value
Accuracy	Scope Specificity	r=-.501, p=.011
	Agreement	r=.582, p=.002
Scope Specificity	Understandability	r=.531, p=.006
	Questioner's Excitement	r=.513, p=.009
Clarity	Completeness	r=.550, p=.004
	External Verification	r=.499, p=.011
Writing style	Characteristics	r=.517, p=.008
	Perspective Novelty	r=.449, p=.024
	Solution Feasibility	r=.524, p=.007
Completeness	Affect	r=-.529, p=.007
	Quickness	r=-.477, p=.016
Characteristics	Content Novelty	r=.468, p=.018
	Perspective Novelty	r=.609, p=.001
	Affect	r=.482, p=.015
	Solution Feasibility	r=.582, p=.002
Perspective Novelty	Affect	r=.529, p=.007
	Available Alternatives	r=.423, p=.035
	Solution Feasibility	r=-.574, p=.003
Understandability	Source Quality	r=.455, p=.022
Questioner background experience	Agreement	r=.399, p=.048
	External Verification	r=.419, p=.037
Answerer's Attitude	Agreement	r=.490, p=.013
	Reference to External Sources	r=.403, p=.045
	Answer Extension	r=.470, p=.018
Answerer's Effort	Confirmation	r=.402, p=.047
	Answer Extension	r=.420, p=.036
Answerer's Experience	Reference to External Sources	r=-.434, p=.030
	Effectiveness	r=.414, p=.040
Confirmation	Quickness	r=.488, p=.013
Available Alternatives	Quickness	r=.403, p=.046
Quickness	Reference to External Sources	r=.408, p=.043
Source Quality	Answer Extension	r=.398, p=.049
Effectiveness	Questioner's Excitement	r=.615, p=.001
Solution feasibility	General Statement	r=.527, p=.007

VI.3.3. Correlation between Best Score and Comment criteria. Correlation analysis of the best score the comment type showed that the score has a significant negative correlation with comments that were affective (r=-0.457, p=0.022) or had unique characteristics (r=-0.435, p=0.030). Significant positive correlations were identified between the score and the comments that provided "solution feasibility" (r=.513, p=0.009), or had "external verification", (r=0.464, p=0.019) or contained "general statements" (r=0.679, p=0.000), or expressed "questioner's excitement" (r=0.401, p=0.047).

As before, the correlations revealed here make sense. For instance, it is easy to see the connection between an answer that gives the asker some solution, or is creditable and being scored higher.

VI.4. Practical Implications of the Results

The analysis of questions, question descriptions, answers, and comments produced a magnitude of meaningful patterns and interesting results. While the results concerning question type revealed that certain types of questions tended to be used to inquire a particular subject matter or in a specific context, it is fascinating that there were similarities and differences between an initial question when it is considered alone, and combined counts when the explanation of the question which included additional questions were examined together with the initial question. In a reference interview setting or a typical information exchange process, people tend to start with an initial question, followed by a more in-depth description of what the question is about. Knowing what kinds of "wh-" questions tend to be asked to start an inquiry for what subject matter would be helpful to prepare for potential responses. This particular finding would not only be helpful for library professionals who provide both the face-to-face and digital reference services but also for those who design automatic QA systems, in their creation of answer templates to respond to users' spontaneous questions. It is also fruitful to examine the content of the question to pinpoint the progression between an initial inquiry and the sense making process shown through the question content.

A significant part of the results concerns with the characterizations of best answers. Using the sentence type to code the answers, it is interesting to see which sentence type got used more often as the best answers for different topics. While declarative sentence made up 78% of the best answer sentences, the best answer score was found to be negatively correlated with the frequency of declarative sentences. Imperative sentences, through their positive correlations with the score, may be suggesting that certain manifestation of creditability and authoritativeness is needed in the delivery of the best answer. Various correlation results demonstrated the associations between various sentence types, elements contained as well as between sentence type and element type. Several statistically significant associations seemed logical and meaningful; for example, emoticon is associated positively with imperative and exclamatory sentence types while negatively associated with declarative sentences. The partial sample data analysis also provided meaningful results that correlate the scores positively with the presence of URLs in the best answer, and negatively with the emoticon in the best answer. All these findings help to establish a multi-faceted understanding of the syntactic makeup and features associated

with the best answers. Such an understanding is greatly valuable to reference services or any services that involve interactions with users.

Data patterns emerged from the analysis of comments provided magnificent insights into how askers considered as important when they chose the best answer to their questions. On a general level, most of the best answers provided good content, practical utility, and had high socio-emotional value. At the individual criteria level, besides general statement, the effectiveness of an answer, being able to provide some feasible solutions in the answer, as well as the accuracy of the response, seemed to be the top standards in assessing the quality of an answer. More noteworthy from the results is that different criteria were weighted as carrying different levels of importance for evaluating the answers to questions of different topics, which is consistent with the findings from both Adamic et al.'s (2008) study and Kim & Oh's (2009) research. For instance, understandability is important for questions in "Arts & Humanities," whereas emotional support is crucial in answering questions concerning pregnancy and parenting. With health related questions, it is helpful for the answer to contain certain level of detail, while answer's experience is a determining factor for the best answer to questions on consumer electronics. While content novelty is important to answer news and events questions, perspective novelty is the key for a good discussion in matters concerning social science. Many of these connections are interesting and valuable for the development of a good answer to users' inquiries.

In addition, several significant correlations among the comment criteria found in the study sample are useful in understanding the complexity associated with providing best answers. For example, understandability is associated with better source quality whereas the higher the perspective novelty an answer presents, the lower it is for the answer to contain feasible solutions. These are meaningful fruits for thought in our provision of best and appropriate answers. Meanwhile, it is equally useful to pay attention to some of the correlations between best score and comment criteria. While solution feasibility and external verification are positive indicators for higher scores, affective answers are actually the opposite. This also reminds us the importance for our responses to provide solutions to users' problems and to be credible and verifiable.

VII.   Research Outputs and Future Research

Multiple presentations were given to report the preliminary results of the study. It is in the plan that research papers will be written to submit for publications in scholarly journals or at professional conferences.

As discussed in previous sections, the present project has produced rich and valuable results. The 3,000 datasets are goldmine for a series of future studies. Further analysis will be carried out to examine the semantic and pragmatic aspects of the questions and answers. It is certain that a complete understanding to the information exchange in a social Q&A platform will be achieved through a multi-dimensional, micro-and macro- level's examination of patterns and associations that occur as individual instances as well as in an aggregated form. Ultimately, research studies such as the one reported here enrich and advance our empirical knowledge on

user community behaviors in a social Q&A process, and also help to improve the practice of information services by offering guidance on providing the best possible answers to users' questions.

References

Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008). Knowledge sharing and Yahoo Answers: Everyone knows something. *Proceedings of the 17th International Conference on the World Wide Web (WWW '08)* (pp.665-674). New York: ACM.

Arnold, J., & Kaske, N. (2005). Evaluating the quality of a chat service. *Portal, 5*, 177-193.

Braxton, S.M., & Brunsdale, M. (2004). E-mail references as substitute for library receptionist. *Library Journal, 130* (1), 46-47.

Brown, D.M. (1985). Telephone reference questions: A characterization by subject, answer format, and level of complexity. *RQ, 24*(3), 290-303.

Bouguessa, M., Wang, S., & Dumoulin, B. (2010). Discovering knowledge-sharing communities in question-answering forums. *ACM Transactions on Knowledge Discovery from Data, 5* (1), Article 3. 1-46.

Bunge, C.A. (1990). Factors related to output measures for reference services in public libraries: Data from thirty-six libraries. *Public Libraries, 29*, 42-47.

Burger, J. et al. (2001). Issues, tasks and program structures to roadmap research in question & answering. National Institute of Standards and Technology.

Carbonell, J. et al. (2000). Vision statement to guide research in Question & Answering (Q&A) and Text Summarization.

Carter, D.S., & Janes, J. (2000). Unobtrusive data analysis of digital reference questions and service at the Internet public library: An exploratory study. *Library Trends, 49*(2), 251–265.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*(1), 1-39.

Diamond, W., & Pease, B. (2001). Digital reference: A case study of question types in an academic library. *Reference Services Review, 29*, 210-219.

Curme, G. O. (1947). *English Grammar.* New York: Barnes & Noble Books.

Freed, A. F. (1994). The form and function of questions in informal dyadic conversation. *Journal of Pragmatics, 21,* 621-644.

Fogarty, E. (1995). Reference questions: Who, what, where, when, how, and why? *New Jersey Libraries, 28,* 19–21.

Gazan, R. (2006). Specialists and synthesists in a question answering community. *Proceedings of 2006 Annual Meeting of the American Society for Information Science and Technology.* Retrieved August 30, 2007, from: http://eprints.rclis.org/archive/00008433/01/Gazan_Specialists.pdf.

Gazen, R. (2007). Seekers, Sloths and Social Reference: Homework Questions Submitted to a Question-Answering Community. Paper submitted to *New Review of Hypermedia and Multimedia.*

Graesser, A.C., McMahen, C.L., & Johnson, B.K. (1994). Question asking and answering. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 517–538). San Diego: Academic Press.

Harabagiu, S., Moldovan, D., Pasca, M., Mihalcea, R., Surdeanu, M., Bunescu, R., Girju, R., Rus, V., & Morarescu, P. (2002). Falcon: Boosting knowledge for answer engines. *Proceedings of the* 9th Text REtrieval Conference. Retrieved on August 30, 2007, from: http://www.iccs.inf.ed.ac.uk/~s0239548/qa-group/papers/harabagiu.2001.trec.ps.

Harper, F.M., Weinberg, J., Logie, J., & Konstan, J. A. (2010). Question types in social Q&A sites, *First Monday*, *15*(7), retrieved on September 5, 2011, from: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2913/2571.

Kaske, N., & Arnold, J. (2002). An unobtrusive evaluation of online real time library reference services. American Library Association, Annual Conference, Atlanta, GA, June 15, 2002.

Katz, W. A. (2002). *Introduction to reference work* (8th ed.)*, Vol.1; Vol. 2*. Boston: McGraw-Hill.
_____. (1997). *Introduction to reference work*. (7th ed.). New York: McGraw-Hill.

Keyes, J.G. (1996). Using conceptual categories of questions to measure differences in retrieval performance. Proceedings of the American Society for Information Science, *33,* 238-242.

Kibbee, J.Z., Ward, D.H., & Ma, W. (2002). Virtual service, real data: Results of a pilot study. *Reference Services Review, 30*(1), 25-36.

Kim, S., Oh, J. S., & Oh, S. (2007). Best-Answer selection criteria in a social Q&A site: From the user-oriented relevance perspective. Paper accepted to appear in 2007 Annual Meeting of American Society for Information Science and Technology.

Kim, S., & Oh, S. (2009). Users' relevance criteria for evaluating answers in a social Q&A site. *Journal of the American Society for Information Science and Technology, 60*(4), 716-727.

Lee, I.J. (2004). Do virtual reference librarians dream of digital reference questions? A qualitative and quantitative analysis of email and chat reference. *Australian Academic and Research Libraries, 35*, 95-110.

Lehnert, W.G. (1978). *The process of question answering: A computer simulation of cognition.* Hillsdale, NJ: Erlbaum.

Liddy, E. D. (1998). Enhanced text retrieval using natural language processing. *Bulletin of the American Society for Information Science and Technology, 24*(4), 14-16. Retrieved on August 30, 2007, from: http://www.asis.org/Bulletin/Apr-98/liddy.html.

Lytinen, S. L, & Tomuro, N. (2002). The use of question types to match questions in FAQFinder. *Proceedings of AAAI Spring Symposium on Mining Answers from Texts and Knowledge Bases.* Retrieved on August 30, 2007, from: http://condor.depaul.edu/~ntomuro/papers/spring02.pdf.

Phenix, K. (1983). *Analysis of the 1981/82 HEGIS/LIBGIS Responses of Illinois Academic Libraries (Illinois Library Statistical Report No. 9).* Springfield, IL: Illinois State Library.

Pomerantz, J. (2005). A linguistic analysis of question taxonomies. *Journal of the American Society for Information Science and Technology, 56*(7), 715-728.

Radford, A. (2004). *English syntax: An introduction.* Cambridge, UK: Cambridge University.

Radford, M. L. (2006). Encountering virtual users: A qualitative investigation of interpersonal communication in chat reference. *Journal of the American Society for Information Science and Technology, 57*(8), 1046-1059.

Radford, M. L., & Connaway, L. S. (2008). *Seeking synchronicity: Evaluating Virtual Reference Service from User, Non-User, and Librarian Perspective: IMLS Final Performance Report.* Retrieved September 5, 2011, from: http://www.oclc.org/research/activities/synchronicity/reports/20080626-final.pdf

Radford, M.L., Connaway, L. S., Confer, P. A., Sabolcsi-Boros, S., & Kwon, H. (2011). "Are we getting warmer?" Query clarification in live chat virtual reference. *Reference & User Service Quarterly, 50*(3), 259-279.

Roberts, G.L., & Bavelas, J.B. (1996). The communicative dictionary: A collaborative theory of meaning. In J. Stewart (Ed.), Beyond the symbol model: Reflections on the representational nature of language
(pp. 135–160). Albany, NY: State University of New York Press.

Robinson, W.P., & Rackstraw, S.J. (1972). *A question of answers* (Vol. I). Boston: Routledge & Kegan Paul.

Rothstein, S. (1964). The measurement and evaluation of reference service. *Library Trends, 12*(3), 456–472.

Sadock, J. M., & Zwicky, A. M. (1985). Speech act distinctions in syntax. In Shopen, T. (Ed.) Language typology and syntactic description. Cambridge, UK: Cambridge University Press.

Schegloff, E., & Sacks, H. (1973). Opening up closings. *Semiotics, 7*(4), 289-327.

Seng, M. (1978). Reference service upgraded using patrons' reference questions. *Special Libraries, 69*(1), 21–28.

Shah, C. (Feburary 7, 2011). Effectiveness and user satisfaction in Yahoo!Answers. *First Monday, 16(2).* Retrieved September 5, 2011, from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3092/2769.

Shah, C., & Pomerantz, J. (2010). Evaluating and predicting answer quality in community QA, *SIGIR'10.*

Tomuro, N. (2004). Question terminology and representation for question type classification. *Terminology, 10*(1), 153-168.

U.S. Department of Education. (1981). *LIBGIS—Public library universe.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

White, H.D. (1981). Measurement at the reference desk. *Drexel Library Quarterly, 17*(1), 3–35.

Appendix A: Comments Criteria List and Definitions

Original Criteria (Developed by Kim et al., 2007):

- **Accuracy:** In the comments, questioners described based on their own judgment, whether the answers were correct or truthful
- **Scope & specificity:** how specifically the answers covered the question topics
- **Clarity:** How clearly the answers were explained
- **Writing style:** 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."
- **Completeness:** *our definition (not from the article)*: Questioners commented on how the answer covers all aspects of the question, or the information provided is complete, not partial.
- **Existence of particular detail:** Sometimes, questioners elaborated their fondness of a particular fact/detail illustrated in the best answers 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"
- **Characteristics:** Questioners evaluated the answers' historical, religious or factual aspects and the related comments were grouped into the 'Characteristics' criterion.
- **Content novelty:** Questioners evaluated whether the best answers provided new information they did not know e.g., "huh!!!! Never heard that name"
- **Perspective novelty:** Or whether they dealt with new approaches that questioners had not thought of 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."
- **Understandability:** The answers were easily understandable.
- **Questioner's background/experience:** Sometimes, the answers which reminded questioners of or reflected on their previous experiences or backgrounds were chosen. 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
- **Emotional support:** Questioners responded to the 'Emotional supports' from answerers, saying "That makes me feel better." or "Your words really helped"
- **Answerer's attitude:** Questioners also sensed whether answerers were sincere, honest or open-minded
- **Answerer's effort:** Questioners appreciated the 'Answerer's efforts' in putting answers together, as stating like "She took enough time to try and feed me some info."
- **Answerer's experience:** Sometimes, questioners pointed out that they appreciated answers sharing their personal experiences. 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."
- **Agreement:** Whether the questioners agreed with the answerers "Finally, someone who agrees with me"
- **Confirmation:** Whether the answerers confirmed what the questioners originally expected (Confirmation), was also reflected in the comments "This guy thinks like me so he must be alright."

- **Affect:** Sometimes, questioners chose the best answers solely based on what they felt about it e.g., "That was hilarious," or "You answer was too funny"
- **External verification:** 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."
- **Available alternatives:** Questioners picked the only available answer e.g., "Well there was just one answer."
- **Quickness:** The first response "You were the first one to answer"
- **Reference to external sources:** Specified that questioners showed their appreciation of the external source offerings in the comments.
- **Source quality:** The perceived quality of the sources, whether worthwhile to review or useless, were also evaluated in the comments 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")
- **Effectiveness:** The questioners stated that they chose the best answer because that answer had actually helped to solve their original problems, either partially or entirely 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"
- **Solution feasibility:** The questioners showed their intention to try out what they had been instructed in the answers 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"
- **General statement:** 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."

Added Criteria/Categories

- **Answer Extension:** Questioners or otherID users did not address the quality of the answers, instead, they provide additional information for answering the question
- **Symbols:** The comments contain symbols instead of complete sentences. In some cases, the questioners were prompted by the system to put comments in, and they simply put a symbol like "." to avoid giving comments.
- **Other:** The comments do not fit into a understandable category or are irrelevant to the question, for instance, comments written in other languages
- **Questioner's excitement:** The comments reflect obvious excitement of the questioner about the answerer's ability or the quality of the answer
  - Example:
  - Entertainment & Music: 20080328161422AAAyjGE HAHAHAHAHA omgggggg thanx sooooo freakin much gurlyy!!! HAHAHAH i have ur butt ointment, sir. im sooooo doing tht my next sleepover!!!!! LMFAOO xD