

A Comparison of Quality, Speed, Scope and Usability between English and Chinese CQAs

Alton Y. K. Chua and Snehasish Banerjee

Abstract—This paper compares CQAs from two popular languages used on the internet, English and Chinese, on the basis of four dimensions, namely, answer quality, answer speed, corpus scope and interface usability. To enhance granularity, answers are compared separately in terms of their content quality, cognitive quality and socio-emotional quality. Results indicate that the English CQAs fared better than the Chinese CQAs in terms of answer quality and answer speed, while the Chinese CQAs outperformed the English counterparts in terms of corpus scope and interface usability. However, the English CQAs and the Chinese CQAs did not differ significantly in terms of content quality of answers. The implications of the findings are discussed. Finally, the paper concludes with notes on limitations and directions for future research.

Index Terms—community question answering, quality, speed, scope, usability, hierarchical ANOVA

I. INTRODUCTION

THE ability of search engines to return thousands of web pages for any given query in a matter of seconds had long made them the gateways for online information seeking. However, their inability to parse queries in natural language [1] and offer succinct answers to users' questions [2], coupled with continual advancements in web technologies have paved the way for the emergence of more interactive information seeking channels. Known as community question-answering sites (CQAs), these refer to social media applications where “any user can pose a question, and in turn other users – potentially many of them – will provide an answer” [3, p 759]. Besides, CQAs allow users to browse the corpus of previously contributed questions and answers, rate the quality of answers, offer comments and provide votes in favor of superior answers [4].

Since the inception of the first CQA in 2002 in the shape of Knowledge iN, these sites have been growing at a prolific pace in multiple languages. For example, Yahoo! Answers, one of the most widely used English CQAs, reported 62 million unique visitors per month on average in the United States alone in 2010 [5]. Similarly, Chinese CQAs such as

Baidu Knows were accessed more than 2.26 billion times solely in April, 2010 [6]. Such a trend is expected to continue on its upward trajectory in the near future. With the growing popularity of CQAs among users, it is no wonder that scholarly interest in CQA research is also on the rise.

Yet, two research gaps can be identified in extant CQA literature. First, most CQA research tends to draw datasets from English CQAs such as Yahoo! Answers [7], [8]. However, given the growing popularity of CQAs in multiple languages, confining datasets to CQAs of a single language may severely constrain the generalizability of findings.

Second, not much scholarly inquiry has delved into the evaluation of multiple CQAs from an encompassing perspective. Most studies focus on quality of answers (eg. [9] [10]). Yet, they usually lack the granularity to shed sufficient light on answers' content richness, cognitive value, and socio-emotional adequacy, the three important facets of answer quality [9], [11], [12]. Moreover, other research-worthy dimensions of CQAs that have not been adequately explored thus far include speed of receiving answers, scope of corpora, and usability of interfaces.

To address the aforementioned research gaps and to augment the limited prior research along these themes (eg. [6], [13], [14], [15]), this paper compares CQAs from two popular languages used on the internet, English and Chinese, on the basis of four dimensions, namely, answer quality, answer speed, corpus scope and interface usability. To enhance granularity, answer quality is measured separately in terms of content quality, cognitive quality and socio-emotional quality. Specifically, the following research questions are investigated:

RQ 1: How do English and Chinese CQAs differ based on (a) content, (b) cognitive, and (c) socio-emotional quality of answers?

RQ 2: How do English and Chinese CQAs differ based on speed of attracting answers?

RQ 3: How do English and Chinese CQAs differ based on scope of corpora?

RQ 4: How do English and Chinese CQAs differ based on usability of interfaces?

The significance of the paper is two-fold. On the research front, this paper attempts to provide an encompassing perspective for evaluation of CQAs by incorporating the dimensions of answer quality, answer speed, corpus scope and interface usability. Furthermore, a comparative study of multiple CQAs of different languages has been relatively uncharted thus far. Given that behavior patterns of social media users are influenced by cross-lingual effects [16], [17], this study could shed light on differences between

Manuscript received April 29, 2013. This work is supported by Nanyang Technological University Research Grant AcRF Tier 1 (RG58/09).

A. Y. K. Chua is with Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (phone: +65-6790-5810; e-mail: altonchua@ntu.edu.sg).

S. Banerjee is with Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (e-mail: snehasis002@e.ntu.edu.sg).

users of English CQAs, and those of Chinese CQAs.

On the practical front, this paper sheds light on nuances in answer quality, answer speed, corpus scope and interface usability between the English CQAs and the Chinese CQAs. An understanding of such issues has implications on the effectiveness of CQAs in meeting users' information needs, which in turn, may determine their sustainability. Furthermore, with the availability and easy access of translation portals such as Google Translate, users of English CQAs could make sense of answers contributed in Chinese CQAs, and vice-versa. Hence, users might lean on the findings of this paper to choose between English CQAs and Chinese CQAs to adequately meet their information needs.

The remainder of the paper is structured as follows. The following section describes the literature which revolves around four central themes, namely, answer quality, answer speed, corpus scope, and interface usability. Next, the procedures of data collection and analysis are explained. This is followed by the results and the discussion. Finally, the paper concludes with notes on limitations and directions for future research.

II. LITERATURE REVIEW

A. Answer Quality

Quality of an answer in CQAs refers to the extent to which it adds value for users [8]. In most CQAs, users act as volunteers to post answers to self-selected questions without any gate-keeping process. As a result, answer quality can fluctuate drastically from excellent to pathetic [10]. While seeking answer for a particular question, users are thus presented with a list of answers with varying quality. Finding the appropriate answer among them is not trivial as users may not be knowledgeable enough to assess the quality of answers to questions asked by themselves [7]. This makes their task of sieving the grain from the chaff grueling [14].

For the purpose of this paper, answer quality is operationalized by the degree to which answers show richness in terms of content quality, cognitive quality and socio-emotional quality [9], [11], [12]. Content quality refers to the overall content richness of answers and can be explained by three factors: reasonableness, soundness and dependability. Reasonableness is the extent to which an answer is consistent and believable [12]. Soundness refers to the extent to which an answer is error free, complete and coherent. A dependable answer is one which is current, secure, and temporally coherent [18]. Cognitive quality refers to an answers' ability to stimulate the cognitive cues of users' knowledge. It can be explained by two factors: understandability and novelty. Understandability is the extent to which an answer is easily comprehended [9], [11]. Novelty refers to the extent to which an answer invokes creative thinking among users [19]. Socio-emotional value refers to the social aspect of CQAs, denoting interpersonal relationships and emotions as reflected through answers. Gratitude, appreciation and empathy are some forms of emotions commonly expressed in CQAs to thank others for

sharing their knowledge or providing emotional support [9], [11].

B. Answer Speed

Limited scholarly attention has been trained into the speed of answers in CQAs thus far. In one of the few studies, a comparative analysis was done between Yahoo! Answers and Google Answers to identify the factors that could contribute to the former's success and the latter's failure [15]. The findings revealed that users could obtain answers much faster from Yahoo! Answers than Google Answers. In another related study, Yahoo! Answers was found to attract fastest answers in a matter of few minutes, followed by satisfactory answers in a matter of few hours [20]. The on-going widespread support for Yahoo! Answers could be attributed to the speed of attracting answers.

A similar study on Stack Overflow, a closed domain CQA site for developers and programmers, revealed that most answering activities take place within the first hour after a question is being posted [21]. In general, users seem to value promptness, and lack the proclivity to wait too long for good answers. In fact, users who cannot obtain prompt answers from CQAs turn to alternative sources of information. They never bother to return to CQAs in order to check if their questions have been answered [22]. Thus, if a CQA site does not offer good answers fast, its sustainability will be called into question. For the purpose of this paper, answer speed is operationalized based on the time elapsed in minutes between posting a question and receiving an answer.

C. Corpus Scope

CQAs provide an avenue to accumulate growing corpora of questions and answers [3]. Scope of CQA corpora refers to their breadth, comprehensiveness and coverage to satisfy the information needs of users. Supported by efficient information retrieval and clustering techniques, CQA corpora consist of questions and their associated answers that had been previously contributed to the sites. Besides, they also include other data items such as timestamps, ratings to answers, comments attracted by answers, and best answers selected by users [23].

When a question is posted, a list of related questions available in the CQA corpora is also presented to users. Such a list of related questions can be of utmost importance because past responses to some of those questions could also be a potential answer to the newly posted question [24]. The relevance of the related questions with respect to the newly posted question greatly depends on the scope of CQA corpora. Wider the corpus scope, the related questions would be more likely relevant to the new question [14]. Hence, for the purpose of this paper, corpus scope is operationalized in terms of the relevance of the related questions presented by CQAs in response to a newly posted question.

D. Interface Usability

Interface usability of CQAs refers to the efficiency, effectiveness and satisfaction associated with its use [13]. It is influenced by the way the sites are navigated, organized and labeled to facilitate unobtrusive browsing, searching

and ease of use [25]. In CQAs, users perform various tasks such as asking questions, contributing answers, browsing past-contributed questions and answers, offering comments and votes, as well as engaging in social interactions. It is conceivable that users' willingness to perform these tasks would be positively associated with interface usability of CQAs [26].

For the purpose of this paper, interface usability is operationalized in terms of the efficiency and ease with which it facilitates information sharing and information organization tasks [27]. Information sharing tasks typically involve bilateral information flow between users who play the role of askers and answerers. These include asking and answering questions, supplementing answers by sharing external resources, and offering votes or comments. On the other hand, information organization tasks typically comprise those that allow users retrieve the information sought easily and effectively. These include easy access to question categories such as entertainment, sports and science, indication of best answer to a given question, and a summary of users' profile to recognize the trusted contributors of answers [13].

III. METHODOLOGY

A. Selection of Language and CQAs

This paper draws data from six CQAs, three English and three Chinese. These two languages were chosen for being the top two languages used on the internet, supported by some 536.6 (26.8%) and 444.9 (24.2%) millions users respectively [28]. Moreover, much scholarly attention has been trained on the investigation of both English CQAs (eg. [3], [16]) and Chinese CQAs (eg. [29], [30]).

The English CQAs selected for analysis include Yahoo! Answers, WikiAnswers and Answerbag. On the other hand, the Chinese counterparts comprise Baidu Knows, Tencent Soso Wenwen and Sina iAsk. These six CQAs were chosen given that all of them have been established for more than three years, and attract at least 20,000 unique visitors per month on average. The statistics on unique visitors for the English and the Chinese CQAs were drawn from Compete Site Analytics and ChinaRank respectively as of July, 2011. Such longevity and popularity allows for comprehensive data collection.

B. Data Collection

The data collection period, which lasted from July to December, 2011, involved three steps, gathering questions from the CQAs, cross-posting the gathered questions across the CQAs, and harvesting answers attracted by the cross-posted questions. The first step involved gathering some 100 questions from each of the six CQAs. Specifically, questions were garnered from four categories, namely, entertainment, sports, computer/internet, and science/mathematics. These categories were chosen as they are commonly available and allow for comparison across all CQAs. Questions that were culturally-specific were omitted as they were deemed inappropriate for cross-posting across the CQAs from the two languages. A few iterations were

made to the question gathering process to produce six sets of 100 questions, each set comprising 25 questions from each of the four categories.

The second step involved cross-posting of the gathered questions across the six CQAs. Specifically, 500 questions gathered from the other five CQAs were posted into each CQA. This meant that questions extracted from the English CQAs had to be translated to Chinese, whereas those drawn from the Chinese counterparts had to be translated to English. Translations were done by three research associates (henceforth, known as coders) who were effectively bilingual, possessed graduate degrees in Information Science with minimum two years of professional experience, and were well knowledgeable about the four chosen categories. The coders back-translated and pre-tested the questions before cross-posting. When needed, they also conferred among themselves to ensure the accuracy of the translated questions. Questions were posted throughout the day, seven days a week with randomized timings as much as possible to minimize any confounding effects of different time zones, time of the day, and day of the week on attracting answers from users across the globe.

The third step involved harvesting answers attracted by the cross-posted questions from the six CQAs. A window of four days was given to each question to solicit answers. In response to the 3,000 questions (500 questions per CQA site), a total of 5,356 answers were harvested. Along with the questions and their respective answers, related data items such as time-stamp and URLs were also archived.

C. Coding and Measurements

For measuring answer quality, content analysis was used to identify characteristics within the answer text [31]. The coders evaluated answers on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) to indicate the extent to which content quality, cognitive quality, and socio-emotional quality could be observed. The mean pair-wise Cohen's kappa for inter-coder reliability among the three coders was found to be 0.83, indicating non-chance level of agreement in the quality scores [32].

For measuring answer speed, system timestamps were used. The timestamp of posting a question to a particular CQA was recorded. Thereafter, the timestamps of receiving all answers within the window of four days were also recorded. Finally, the speed of each answer was computed as the difference in minutes between the timestamp of receiving the answer and that of posting the question.

When questions are posted in CQAs, a list of four or five related questions available in their corpora is presented to users. For measuring corpus comprehensiveness, the relevance of such related questions in response to a newly posted question was considered. For each question, the individual related questions were assigned a relevance score ranging from 1 to 5. For instance, if a related question was assigned a score of 1 (5), it was deemed completely irrelevant to the parent question (highly relevant to the parent question). Scores were assigned by the three coders and the mean pair-wise Cohen's kappa was found to be 0.81, indicating non-chance level of agreement in relevance scores [32]. Finally, the score of corpus scope for each

question was calculated as the arithmetic mean of relevance scores of all related questions presented by the CQAs.

For measuring interface usability, the three coders evaluated the six CQAs on the extent to which the sites facilitated efficient and easy information sharing and information organization tasks on a five-point Likert scale (1=strongly disagree, 5=strongly agree). Specifically, the sites were evaluated with respect to efficiency of information sharing tasks, efficiency of information organization tasks, ease of use of information sharing tasks, and ease of use of information organization tasks. The mean pair-wise Cohen’s Kappa for inter-coder reliability among the three coders was found to be 0.81, indicating non-chance level of agreement in the usability scores [32].

D. Data Analysis

To address RQ 1 and RQ 2, the 5,356 harvested answers were the units of analysis. On the other hand, the 3,000 cross-posted questions were the units of analysis for addressing RQ 3. To address RQ 4, the individual websites were the units of analysis.

The first three RQs were analyzed using a combination of independent samples t-test (t-test) and hierarchical analysis of variance (HANOVA). The result of the t-test indicated if there exists significant difference between the English and the Chinese CQAs. Next, HANOVA was employed to delve deeper because it is appropriate when multiple categorical levels are nested hierarchically within some higher order independent variable [33], [34]. In this case, language is the higher order independent variable with three CQAs nested within English, and three nested within Chinese. For RQ 4, the means and standard deviations of interface usability are reported for the six CQAs to indicate the differences.

IV. RESULTS

Of the 3,000 cross-posted questions across the six CQAs from the two languages (500 questions per CQA site), 2,065 (68.83 %) were answered resulting in a collection of 5,356 answers. The three English CQAs attracted 2,231 answers in response to 894 questions (59.60 %) with an average of 2.50 answers per question. On the other hand, the three Chinese CQAs received 3,125 answers in response to 1,171 questions (78.10 %) with an average of 2.67 answers per question.

Among the three English CQAs, Yahoo! Answers had the most questions answered (88.4 %). It also attracted the highest number of answers (1,357) with 3.07 answers per question on average. On the other hand, WikiAnswers was found to lie at the other end of the spectrum. It had the least questions answered (27.40 %) and also attracted the least number of answers (140) with 1.02 answers per question on average.

Among the three Chinese CQAs, Baidu Knows had the most questions answered (81.60 %). It also attracted the highest number of answers (1,227) with 3.01 answers per question on average. On the other hand, Sina iAsk was found to lie at the other end of the spectrum. It had the least questions answered (76.00 %) and attracted the least number of answers (766) with 2.02 answers per question on

average. The general descriptive statistics of the six CQAs from the two languages in terms of the number of questions and answers are summarized in Table I.

Table I
General Descriptive Statistics

CQA	Questions answered # (%)	Answers received (#)	Answers/question
Yahoo Answers	442 (88.40)	1,357	3.07
Wiki-Answers	137 (27.40)	140	1.02
Answer-bag	315 (63.00)	734	2.33
Baidu Knows	408 (81.60)	1,227	3.01
Tencent Soso Wenwen	383 (76.60)	1,132	2.96
Sina iAsk	380 (76.00)	766	2.02
English CQAs	894 (59.60)	2,231	2.50
Chinese CQAs	1,171 (78.10)	3,125	2.67
Aggregate	2,065 (68.83)	5,356	2.59

The descriptive statistics of the six CQAs from the two languages based on the evaluation dimensions: content quality (CON), cognitive quality (COG), socio-emotional quality (SOC), answer speed in minutes (SPE), corpus scope (SCO), and interface usability (USA) are summarized in Table II. It appears that the English CQAs fared better than the Chinese CQAs in terms of answer quality and answer speed. However, in terms of corpus scope and interface usability, the Chinese CQAs seem to outperform the English counterparts.

Table II
Descriptive Statistics based on the Evaluation Dimensions

CQA	CON (1-5)	COG (1-5)	SOC (1-5)	SPE (min.)	SCO (1-5)	USA (1-5)
Yahoo Answers	3.90 ± 0.67	3.86 ± 0.63	3.96 ± 0.64	351.11 ± 1013.59	1.37 ± 0.77	3.47 ± 0.65
Wiki-Answers	3.88 ± 0.96	3.95 ± 0.90	3.96 ± 0.93	721.90 ± 1200.57	2.08 ± 1.44	3.31 ± 0.58
Answer-bag	3.83 ± 0.65	3.80 ± 0.58	3.86 ± 0.62	464.33 ± 850.57	1.88 ± 1.50	3.77 ± 0.27
Baidu Knows	3.71 ± 1.02	3.67 ± 0.98	3.73 ± 1.00	550.88 ± 1009.75	2.59 ± 1.73	4.11 ± 0.43
Tencent Soso Wenwen	3.91 ± 0.75	3.83 ± 0.66	3.92 ± 0.69	451.00 ± 1044.69	2.34 ± 1.74	4.03 ± 0.64
Sina iAsk	3.95 ± 0.63	3.92 ± 0.56	3.98 ± 0.56	975.57 ± 1303.18	1.52 ± 1.08	4.21 ± 0.32
English CQAs	3.88 ± 0.69	3.84 ± 0.64	3.92 ± 0.66	411.62 ± 980.58	1.78 ± 1.32	3.52 ± 0.50
Chinese CQAs	3.84 ± 0.85	3.79 ± 0.79	3.86 ± 0.81	618.80 ± 1119.98	2.15 ± 1.61	4.12 ± 0.52
Aggregate	3.86 ± 0.78	3.81 ± 0.73	3.89 ± 0.75	532.50 ± 1068.93	1.97 ± 1.47	3.81 ± 0.52

With respect to content quality of answers, the English

CQAs (3.88 ± 0.69) performed better than the Chinese CQAs (3.84 ± 0.85). Among the English CQAs, Yahoo! Answers (3.90 ± 0.67) attracted answers with the highest content quality, while Answerbag (3.83 ± 0.65) was at the other end of the spectrum. Among the Chinese CQAs, Sina iAsk (3.95 ± 0.63) received the most content rich answers, while Baidu Knows (3.71 ± 1.02) lagged behind in the rear. The results of the t-test for RQ 1(a) suggest no statistically significant difference in content quality of answers between the English CQAs and the Chinese CQAs. However, the results of HANOVA indicate a statistically significant nested effect of the CQAs within the two languages [$F(4, 5350) = 15.60, p < 0.001$].

With respect to cognitive quality of answers, the English CQAs (3.84 ± 0.64) performed better than the Chinese CQAs (3.79 ± 0.79). Among the English CQAs, WikiAnswers (3.95 ± 0.90) attracted answers with the highest cognitive quality, while Answerbag (3.80 ± 0.58) was found to lie at the other end of the spectrum. Among the Chinese CQAs, Sina iAsk (3.92 ± 0.56) received answers with the highest cognitive value, while Baidu Knows (3.67 ± 0.98) lagged behind in the rear. The results of the t-test for RQ 1(b) reveal a significant difference in cognitive quality between answers obtained from the English CQAs and those harvested from the Chinese CQAs [$t(5274.90) = 2.87, p < 0.05$]. The results of HANOVA further indicate a statistically significant nested effect of the CQAs within the two languages [$F(4, 5350) = 16.89, p < 0.001$].

With respect to socio-emotional quality of answers, the English CQAs (3.92 ± 0.66) performed better than the Chinese CQAs (3.86 ± 0.81). Among the English CQAs, both Yahoo! Answers (3.96 ± 0.64) and WikiAnswers (3.96 ± 0.93) attracted answers with higher socio-emotional elements compared to Answerbag (3.86 ± 0.62). Among the Chinese CQAs, Sina iAsk (3.98 ± 0.56) received answers with the highest socio-emotional quality, while Baidu Knows (3.73 ± 1.00) was found to lie at the other end of the spectrum. The results of the t-test for RQ 1(c) reveal a significant difference in socio-emotional quality between answers retrieved from the English CQAs and those obtained from the Chinese CQAs [$t(5269.03) = 3.19, p < 0.05$]. The results of HANOVA further indicate a statistically significant nested effect of the CQAs within the two languages [$F(4, 5350) = 18.92, p < 0.001$].

With respect to answer speed, the English CQAs (411.62 ± 980.58) fared better than the Chinese CQAs (618.80 ± 1119.98). Among the English CQAs, Yahoo! Answers (351.11 ± 1013.59) received the fastest answers, while WikiAnswers (721.90 ± 1200.57) was found to lie at the other end of the spectrum. Among the Chinese CQAs, Tencent Soso Wenwen (451.00 ± 1044.69) received the most responsive answers, while Sina iAsk (975.57 ± 1303.18) lagged behind in the rear. The results of the t-test for RQ 2 reveal a significant difference in answer speed between the English CQAs and the Chinese CQAs [$t(5137.22) = -7.18, p < 0.001$]. The results of HANOVA further indicate a statistically significant nested effect of the CQAs within the two languages [$F(4, 5350) = 35.21, p < 0.001$].

With respect to corpus scope, the Chinese CQAs (2.15 ± 1.61) performed better than the English CQAs (1.78 ± 1.32). Among the English CQAs, WikiAnswers (2.08 ± 1.44) offered the most comprehensive corpus scope, while Yahoo! Answers (1.37 ± 0.77) was found to lie at the other end of the spectrum. Among the Chinese CQAs, Baidu Knows (2.59 ± 1.73) offered the best corpus scope, while Sina iAsk (1.52 ± 1.08) lagged behind in the rear. The results of the t-test for RQ 3 reveal a significant difference in corpus scope between the English CQAs and the Chinese CQAs [$t(2883.06) = -6.935, p < 0.001$]. The results of HANOVA further indicate a statistically significant nested effect of the CQAs within the two languages [$F(4, 2994) = 52.56, p < 0.001$].

With respect to interface usability, the results of RQ 4 suggest that the Chinese CQAs (4.12 ± 0.52) performed better than the English CQAs (3.52 ± 0.50). Among the English CQAs, Answerbag (3.77 ± 0.27) scored the highest, while WikiAnswers (3.31 ± 0.58) was found to lie at the other end of the spectrum. Among the Chinese CQAs, Sina iAsk (4.21 ± 0.32) offered the most usable interface, while Tencent Sosos Wenwen (4.03 ± 0.64) lagged behind in the rear.

The results from the four RQs are summarized as follows. For RQ 1(a), answers harvested from the English CQAs did not significantly differ from those retrieved from the Chinese CQAs with respect to content quality. For RQs 1(b) and 1(c), answers from the English CQAs significantly outperformed those from the Chinese CQAs in terms of cognitive quality and socio-emotional quality. For RQ 2, the English CQAs attracted significantly faster answers compared to the Chinese CQAs. For RQ 3, the Chinese CQAs significantly offered broader corpus scope vis-à-vis the English CQAs. For RQ 4, the Chinese CQAs appeared to fare better than the English counterparts in terms of interface usability.

V. DISCUSSION

Four main findings could be culled from this study. First, contrary to prior research [16], [35], [36], there seems to be no English-Chinese language divide on the internet insofar as content quality of answers between the English CQAs and the Chinese CQAs are concerned. Though it was once conceived that “*if you want to take full advantage of the Internet there is only one real way to do it: learn English*” [35, p 226], this may no longer be the case. Users of both the English CQAs and the Chinese CQAs appear equally proficient in providing content rich answers in their respective languages. This could be vestige of the gradual trends of globalization in web usage and ICT across the globe. Nonetheless, the English CQAs outperformed the Chinese CQAs with respect to cognitive quality and socio-emotional quality of answers.

Second, despite attracting more answers per question (as revealed from Table I), the Chinese CQAs significantly lag behind in terms of answer speed compared to the English CQAs. Perhaps, users of Chinese CQAs are more dependent on cues from initial answers compared to those of the English CQAs, who tend to be more self-reliant and hence,

more prompt in offering answers [37], [38]. However, once the first answer to a given question is submitted in the Chinese CQAs, it seems to pique the interests of other users to contribute more answers [14]. This reflects the collectivist social orientations among Chinese CQAs' users, who appear highly motivated by expectations of social reciprocity [38]. Consistent with prior research, the findings thus confirm that behavior patterns of asking and answering among CQA users are affected by cross-lingual differences [16], [17].

Third, irrespective of language, all CQAs can be improved further with respect to corpus scope. With the growing popularity of CQAs [3], [7], their corpora are expected to become more comprehensive over time. Also, it is feasible to rank all past questions and answers based on relevance through inter-question, inter-answer, and question-answer similarity [24], [39], [40]. Yet, the dearth of relevant related questions in response to newly posted questions from most CQAs suggests that CQA corpora are perhaps not well equipped with efficient information retrieval and clustering techniques. Given that innumerable questions posted in CQAs are recurrent [24], users' dependency on the goodwill of answerers could be reduced considerably if all CQA corpora are supported with better retrieval and clustering techniques.

Fourth, the efficiency and ease of use for information organization and information sharing tasks were generally inadequate for most CQAs. As a result, they indicated scope for improvement in terms of interface usability. Most CQAs did not cater for the different needs of novices and expert users [13]. Moreover, there was limited scope for customization in all the six CQAs. In the modern web-blitzed society, users have become sophisticated and tech-savvy in their information seeking behaviors [41], [42]. Being spoilt for choices, they increasingly look for applications that allow customization. Hence, substantial room for users' customization could be essential for the sustainability of CQAs.

VI. CONCLUSION

This paper compared CQAs from two popular languages used on the internet, English and Chinese, on the basis of four dimensions, namely, answer quality, answer speed, corpus scope and interface usability. To enhance granularity, answer quality was measured separately in terms of content quality, cognitive quality and socio-emotional quality. The English CQAs studied comprise Yahoo! Answers, WikiAnswers and Answerbag. On the other hand, the Chinese counterparts include Baidu Knows, Tencent Soso Wenwen and Sina iAsk. Results indicate that the English CQAs fared better than the Chinese CQAs in terms of answer quality and answer speed, while the Chinese CQAs outperformed the English counterparts in terms of corpus scope and interface usability. However, the English CQAs and the Chinese CQAs did not differ significantly in terms of content quality of answers.

The findings of the paper should be interpreted in light of two constraints. First, selection of the six CQAs from the two languages was based on longevity and popularity of the

sites. Consistency in system features and efficacies across the CQAs was not taken into account. For example, Answerbag is the only CQA among the six that allows images to be posted as answer content. Second, the relevance of the related questions was constrained by the period of data collection. Given the ever growing corpora of questions and answers, the related questions presented by CQAs in response to a newly posted question are likely to change dynamically. Hence, the scores for corpus scope are likely to vary over time.

This paper has implications for both theory and practice. On the theoretical front, this paper presents an encompassing perspective for evaluation of CQAs by incorporating the dimensions of answer quality, answer speed, corpus scope and interface usability. Furthermore, it extends the territories of CQA research through a comparative study of multiple CQAs from more than one language. Given that behavior patterns of social media users are influenced by cross-lingual effects [16], [17], this study sheds light on differences between users of English CQAs, and those of Chinese CQAs. As a result, it also raises concerns about the generalizability of findings when datasets are restricted to a single platform [7].

On the practical front, this paper offers insights into nuances with respect to content quality, cognitive quality, socio-emotional quality of answers, as well as other dimensions such as answer speed, corpus scope and interface usability between the English CQAs and the Chinese CQAs. An understanding of these issues has implications on the effectiveness of CQAs in meeting users' information needs. This is pertinent for the CQAs, which in turn, may determine their sustenance. Moreover, with the availability of translation portals such as Google Translate, users of English CQAs could make sense of answers contributed in Chinese CQAs, and vice-versa. Hence, users might lean on the findings of this paper to choose between English CQAs and Chinese CQAs to effectively meet their information needs.

This study offers some potential directions for future CQA research. One possible area of investigation could be to evaluate CQAs in terms of their system features that promote users' engagement. After all, an actively engaged community is essential for CQAs to flourish. A second area of research could cater to examining whether answer quality, answer speed, and corpus scope vary with question categories such as entertainment, sports, computer/internet, and science/mathematics. Such an investigation could offer insights into the extent to which the performance of CQAs vary across question categories. A third area of research could adopt a user study to investigate the degrees to which individuals find the usability of CQAs to vary across different platforms. Such an approach of evaluation could shed richer insights on the usability of different CQA interfaces.

REFERENCES

- [1] M. J. Blooma, A. Y. K. Chua and D. H. L. Goh, "A predictive framework for retrieving the best answer," in *Proceedings of the 23rd*

- Annual ACM Symposium on Applied Computing, Fortaleza, Brazil, 2008, pp. 1107-1111.
- [2] M. J. Blooma, A. Y. K. Chua, D. H. L. Goh and L. C. Keong, "A trend analysis of the question answering domain," in *Proceedings of the 6th International Conference on Information Technology: New Generations*, Las Vegas, United States, 2009, pp. 1522-1527.
 - [3] F. M. Harper, D. Moy and J. A. Konstan, "Facts or friends?: Distinguishing informational and conversational questions in social Q&A sites," in *Proceedings of the International Conference on Human Factors in Computing Systems*, ACM, 2009, pp. 759-768.
 - [4] B. Qu, G. Cong, C. Li, A. Sun and H. Chen, "An evaluation of classification models for question topic categorization," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 5, pp. 889-903, Feb. 2012.
 - [5] R. Gazan, "Social Q&A," *Journal of the American Society for Information Science and Technology*, vol. 62, no. 12, pp. 2301-2312, Dec. 2011.
 - [6] D. Wu and D. He, "A study on Q&A services between community-based question answering and collaborative digital reference in two languages," in *Proceedings of iConference*, 2013, pp. 326-337.
 - [7] P. Fichman, "A comparative assessment of answer quality on four question answering sites," *Journal of Information Science*, vol. 37, no. 4, pp. 1-11, Oct. 2011.
 - [8] M. J. Blooma, A. Y. K. Chua and D. H. L. Goh, "Selection of the best answer in CQA services," in *Proceedings of the 7th International Conference on Information Technology: New Generations*, Las Vegas, United States, 2010, pp. 534-539.
 - [9] S. Kim and S. Oh, "Users' relevance criteria for evaluating answers in a social questions and answers site," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 4, pp. 716-727, Jan. 2009.
 - [10] M. A. Suryanto, A. Sun, E. Lim and R. H. L. Chiang, "Quality-aware collaborative question answering: Methods and evaluation," in *Proceedings of the 2nd ACM International Conference on Web Search and Data Mining*, Barcelona, Spain, 2009.
 - [11] S. Kim, J. S. Oh and S. Oh, "Best-answer selection criteria in a social Q&A site from the user-oriented relevance perspective," in *Proceedings of the American Society for Information Science and Technology*, vol. 44, 2007, pp. 1-15.
 - [12] M. J. Blooma, A. Y. K. Chua and D. H. L. Goh, "What makes a high-quality user-generated answer?," *IEEE Internet Computing*, vol. 15, no. 1, pp. 66-71, Jan.-Feb. 2011.
 - [13] A. Y. K. Chua and R. S. Balkunje, "Comparative evaluation of community question answering websites," in *The Outreach of Digital Libraries: A Globalized Resource Network*, pp. 209-218, Springer Berlin Heidelberg, 2012.
 - [14] A. Y. K. Chua and S. Banerjee, "English versus Chinese: A cross-lingual study of community question answering sites," *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2013*, 13-15 March, 2013, Hong Kong, pp. 368-373.
 - [15] C. Shah, J. S. Oh and S. Oh. (2008, September 1), "Exploring characteristics and effects of user participation in online social Q&A sites," *First Monday*, vol. 13, no. 9. [Online]. Available: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2182/2028>
 - [16] S. A. Hale. "Net increase? Cross-lingual linking in the blogosphere," *Journal of Computer-Mediated Communication*, vol. 17, no. 2, pp. 135-151, Jan. 2012.
 - [17] J. Yang, M. R. Morris, J. Teevan, L. Adamic and M. Ackerman, "Culture matters: A survey study of social Q&A behavior," in *Proceedings of the International Conference on Weblogs and Social Media*, Barcelona, Spain, 2011.
 - [18] B. K. Kahn, D. M. Strong and R. Y. Wang, "Information quality benchmarks: product and service performance," *Communications of the ACM*, vol. 45, no. 4, pp. 184-192, Apr. 2002.
 - [19] D. Kelly, N. Wacholder, R. Rittman, Y. Sun, P. Kantor, S. Small and T. Strzalkowski, "Using interview data to identify evaluation criteria for interactive, analytical question-answering systems," *Journal of the American Society for Information Science and Technology*, vol. 58, no. 7, pp. 1032-1043, May 2007.
 - [20] C. Shah. (2011, February 7), "Effectiveness and user satisfaction in Yahoo! Answers," *First Monday*, vol. 16, no. 2. [Online]. Available: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3092/2769>
 - [21] L. Mamykina, B. Manoim, M. Mittal, G. Hripcsak and B. Hartmann, "Design lessons from the fastest Q&A site in the west," in *Proceedings of the International Conference on Human Factors in Computing Systems*, ACM, 2011, pp. 2857-2866.
 - [22] V. Kitzie and C. Shah, "Faster, better, or both? Looking at both sides of online question answering coin," in *Proceedings of the American Society for Information Science and Technology*, vol. 48, 2011, pp. 1-4.
 - [23] C. Shah and J. Pomerantz, "Evaluating and predicting answer quality in community QA," in *Proceedings of the International SIGIR conference on Research and development in information retrieval*, ACM, 2010, pp. 411-418.
 - [24] A. Shtok, G. Dror, Y. Maarek and I. Szpektor, "Learning from the past: answering new questions with past answers," in *Proceedings of the International Conference on World Wide Web*, ACM, 2012, pp. 759-768.
 - [25] J. Nielsen, "Designing web usability: The practice of simplicity," New Riders Publishing, Indianapolis, 1999.
 - [26] C. W. Phang, A. Kankanhalli and R. Sabherwal, "Usability and sociability in online communities: A comparative study of knowledge seeking and contribution," *Journal of the Association for Information Systems*, vol. 10, no. 10, pp. 721-747, 2009.
 - [27] B. Detlor, "Information management," *International Journal of Information Management*, vol. 30, no. 2, pp. 103-108, Apr. 2010.
 - [28] Internet World Stats, "Internet world users by language," retrieved Jun., 2012 from <http://www.internetworldstats.com/stats7.html>
 - [29] S. Deng, Y. Liu and Y. Qi, "An empirical study on determinants of web based question-answer services adoption," *Online Information Review*, vol. 35, no. 5, pp. 789-798, 2011.
 - [30] Y. Z. Wu, J. Zhao, X. Duan and B. Xu, "Research on question answering and evaluation: a survey," *Journal of Chinese Information Processing*, vol. 19, no. 3, pp. 1-13, 2005.
 - [31] K. A. Neuendorf, *The content analysis guidebook*. Beverly Hills, CA: Sage Publications, 2002.
 - [32] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37-46, 1960.
 - [33] C. Dytham, *Choosing and Using Statistics: A Biologist's Guide*. West Sussex, UK: Wiley-Blackwell, 1999.
 - [34] R. M. Walczuch and R. T. Watson, "Analyzing group data in MIS research: Including the effect of the group," *Group Decision and Negotiation*, vol. 10, no. 1, pp. 83-94, Jan. 2001.
 - [35] D. Crystal, *Language and the Internet*. New York, US: Cambridge University Press, 2006.
 - [36] L. Lu, "Digital divide: Does the internet speak your language?," in *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2010, pp. 4022-4025.
 - [37] R. E. Nisbett, K. Peng, I. Choi and A. Norenzayan, "Culture and systems of thought: Holistic versus analytic cognition," *Psychological Review*, vol. 108, no. 2, pp. 291-310, Apr. 2001.
 - [38] M. E. W. Varnum, I. Grossmann, S. Kitayama and R. E. Nisbett, "The origin of cultural differences in cognition: The social orientation hypothesis," *Current Directions in Psychological Science*, vol. 10, no. 1, pp. 9-13, Feb. 2010.
 - [39] D. Carmel, M. Shtalhaim and A. Soffer, "eResponder: Electronic question responder," *Lecture Notes in Computer Science*, vol. 1901, pp. 150-161, Springer Berlin Heidelberg, 2000.
 - [40] J. Jeon, W. B. Croft and J. H. Lee, "Finding semantically similar questions based on their answers," in *Proceedings of the International SIGIR Conference on Research and Development in Information Retrieval*, ACM, 2005, pp. 617-618.
 - [41] A. Y. K. Chua and S. Banerjee, "Aggregated community question answering," *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2013*, 13-15 March, 2013, Hong Kong, pp. 397-402.
 - [42] B. Jingyuan and G. Xingwu, "Situation and development countermeasures of Chinese government website," in *Proceedings of the International Conference on Artificial Intelligence, Management Science and Electronic Commerce*, 2011, pp. 5777-5780