Cross-cultural analysis of the Wikipedia community

1. Abstract

This paper reports a cross-cultural analysis of four Wikipedias in different languages and demonstrates their roles as communities of practice (CoPs). Prior research on CoPs and on the Wikipedia community often lacks cross-cultural analysis. Despite the fact that over 75% of Wikipedia is written in languages other than English, research on Wikipedia primarily focuses on the English Wikipedia and tends to overlook the Wikipedias in other languages. This paper first argues that Wikipedia communities can be analyzed and understood as CoPs. Second, norms of behaviors are examined in four Wikipedia languages (English, Hebrew, Japanese, and Malay), and the similarities and differences across these four languages are reported. Specifically, typical behaviors on three types of discussion spaces (Talk, User Talk, and Wikipedia Talk) are identified and examined across languages. Hofstede's dimensions of cultural diversity as well as the size of the community, and the function of each discussion area provide lenses for understanding the similarities and differences. As such, this paper expands the research on online CoPs through an examination of cultural variations across multiple CoPs, and increases our understanding of Wikipedia communities in various languages.

2. Introduction

The proliferation of the social Web and its participatory nature gives rise to many online communities, some of which undertake a common practice that links users together. In this environment, users often become "prosumers," who are consumers and producers at the same time (Tapscott & Williams, 2008). Various companies identified the potential utility of prosumers by soliciting product and research ideas from the prosumers, who are not employed by their organizations. These prosumers are sharing knowledge through massive collaborative efforts. For example, they provide answers to questions posted on Q&A sites such as Yahoo! Answers or write encyclopedic articles on sites similar to Wikipedia. Wikipedia users are engaged in knowledge sharing processes by implementing common practices that create and maintain shared identity. This type of knowledge sharing is pertinent to the idea of "communities of practice" (CoPs) (Lave & Wenger, 1991). The CoP framework offers "a lens by which we can focus our understanding of informal collaborative learning that occurs outside formal classrooms and training environments" (Hara, 2009, p. 2).

Although the practice of informal collaborative learning through knowledge sharing is not limited to organizational settings, a number of studies about CoPs were conducted primarily in organizational environments. Whether it was about employees of Caterpillar (Ardichvili, Page, & Wentling, 2003), coordinators for Adult Learning Councils (Gray, 2004), user interface designers (Veenswijk & Chisalita, 2007), Xerox technicians (Orr, 1996), or insurance claim processors (Wenger, 1990), these studies discussed knowledge sharing practices of CoPs within organizational boundaries. As noted earlier, knowledge sharing practices outside organizations are already occurring. Thus, examination of these open CoPs (that go beyond organizational boundaries) ought to be expanded.

As collaboration across countries increases, a better understanding of cross-cultural interactions is needed. Surprisingly, few cross-cultural analyses of CoPs have been conducted in the past. For instance, Pan and Leidner (2003) studied a knowledge management system within an international organization that is used to connect multiple CoPs, but did not conduct a cross-cultural analysis of these culturally diverse CoPs. In contrast, Ardichvili, Maurer, Li, Wentling, and Stuedemann (2006) conducted qualitative research to examine cultural effects on knowledge sharing in online CoPs and did not find as many national cultural influences as they had expected in the data from three countries (China, Russia, and Brazil). They explained that, since the study was conducted in one company, the organizational culture was possibly a stronger influence than national culture. They also speculated that online behaviors were less affected by national cultures than offline behaviors. As Ardichvili et al. (2006) stated, we need a better understanding of cross-cultural online CoPs.

Another feature of online CoPs that influences the dynamics of participants is the size of the community (Jones, 1997). Although discussions about the size of CoPs are limited, Dubé, Bourhis, and Jacob (2006) included size as one of the categories in their typology of online CoPs. Additionally, Roberts (2006), who identified limits of CoPs, also referred to the sizes of CoPs and further cautioned about the possibility that the larger CoPs become, the weaker the ties among members become. Stvilia, Al-Faraj, and Yi (2009) also speculated that the size of Wikipedias affect how information quality is managed. As such, it is possible that norms of behaviors will differ across Wikipedias of different sizes.

In addition, prior research on online CoPs has not shed light on norms of behaviors in great detail. Genre analysis of online interactions was put forth by Yates, Orlikowski, and Okada (1999). They articulated two types of interactions, which they called "genres," found in a Japanese R&D lab — purpose and form. Coding categorized under "purpose," covers the purposes of communication, such as response, solicitation, and announcement. The "form" coding includes formatting (e.g., greeting, list, ellipsis, etc.) as well as linguistic features (e.g., dialect, emphatic language, informal body, etc.). Hara and Hew (2007) developed a coding scheme to analyze genres of online interactions related to knowledge sharing practices in an online CoP. They identified ten types of activities by expanding the "purpose" categories identified by Yates et al. (1999). Further development of genres of interactions in online CoPs would be useful to understand norms of behaviors in open online CoPs, particularly with the proliferation of such communities on the social web.

Above all, the literature about online CoPs largely uses comprehensive single-case studies (e.g., Barab, MaKinster, & Scheckler, 2004; Gray, 2004; Schlager, M., Fusco, J., & Schank, 2002; Wasko & Faraj, 2005). This tradition of single-case studies originated from its anthropological roots, which the creators of the term, "communities of practice," embraced (Lave & Wenger, 1991). Only a few studies extend investigations of multiple online CoPs (e.g., Dubé, Bourhis, & Jacob, 2006; Hara, Shachaf, & Storeger, 2009; Hew & Hara, 2007). Thus, there is a need for analytical, conceptual, and comparative work regarding online CoPs that transcends common, single-case study approaches, which tend to be confined within organizational boundaries (Hara, Shachaf, & Storeger, 2009). In particular, as organizations become more global and deploy more cross-border CoPs than ever before, it is necessary to conduct cross-cultural analyses of CoPs.

3

The global nature of Wikipedia makes it an interesting case for such cross-cultural analyses; over 75% of Wikipedia is written in languages other than English. Despite the multi-lingual nature of Wikipedia (75,000 active members in more than 260 languages (Wikipedia: About)), prior studies have been predisposed to investigate only the English version of Wikipedia, with two notable exceptions (e.g., Pfeil, Zaphiris, & Ang, 2006; Stvilia, et al., 2009). Pfeil, et al. (2006) examined history pages of the topic "game", from four Wikipedias (Dutch, French, German, and Japanese) and concluded that the dimensions identified by Hofstede (1991) offline are also observed in online environments. Stvilia, et al. (2009) analyzed the information quality of three Wikipedias (Arabic, English, and Korean), contended that these Wikipedias use different models of information quality and suggested that it is difficult to make definite conclusions about the cultural influences over quality models. Based on these two studies one cannot determine what the relationships are between culture and online behaviors on Wikipedia, because one study contends that differences between languages are in alignment with Hofstede's cultural dimensions (Pfeil et al., 2006) and the other study concludes that it is difficult to claim that there are cultural influences over information quality models, although differences among languages exist (Stvilia et al., 2009).

Research of non-English Wikipedias is needed to better understand the larger Wikipedia community. Do differences across various Wikipedia languages reflect similar cultural variations that have been found in the offline environments (e.g., Hofstede, 1991)? As a first step in addressing this issue, this paper focuses on the following research questions: How do Wikipedia norms of behavior vary across languages?; What are the (cultural and size) variations among Talk pages?

Background
Wikipedia as a Community of Practice

There is a need first to examine if Wikipedia possesses characteristics of CoPs before claiming a cross-cultural analysis of Wikipedia as a CoP. Once the Wikipedia online communities are identified as online CoPs, the study can contribute to the discussion of cross-cultural analysis of CoPs by analyzing multiple language versions of Wikipedia. In this paper we use the following definition of CoPs: "Groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis" (Wenger, Dermott, & Snyder, 2002, p. 4). Wenger (1998), in his seminal book, identified four characteristics of communities of practice: practice, community, meaningful learning, and identity. We argue here that all four characteristics of CoPs are prevalent within the Wikipedia community; the manifestation of each of these characteristics in Wikipedia is described next.

1. Practice: Wikipedia users are engaged in knowledge sharing practices, as seen in discussion spaces (e.g., Stvilia, Twidale, Smith, & Gasser, 2008) and policy pages (e.g., Beschastnikh, Kriplean, & McDonald, 2008; Forte, Larco, & Bruckman, 2009). Users write original articles, edit and improve existing articles, provide quality assurance, fight against acts of vandalism, participate in policy setting, and are engaged in community building and maintenance activities. Sharing these practices fosters a sense of community among Wikipedia users.

4

2. Community: Wikipedia users develop a set of community norms and policies. Users are expected to adhere to these norms and the Wikipedia policies (e.g., Beschastnikh, et al., 2008; Riehle, 2006; Shachaf & Hara, in press). New members of the community are informed and encouraged to learn these norms of behaviors by seasoned users who have mastered processes, policies, and practices on Wikipedia. Bryant, Forte, and Bruckman (2005) reported that new users tend to focus on editing individual articles and later become motivated to contribute to the well-being of the Wikipedia community and the Wikipedia project as a whole. This indicates that seasoned Wikipedia users develop a sense of community.

3. Meaningful learning (learning in context): Wikipedia users learn to become Wikipedians (e.g., Bryant, et al., 2005); they learn how to behave, write, and be a member (in good standing) of the community. When examining how new Wikipedia users become Wikipedians, Bryant, et al., (2005) applied a concept of legitimate peripheral participation used in the communities of practice literature (Lave & Wenger, 1991). "Legitimate peripheral participation" is a term used to describe a mode of participation in which newcomers to a community peripherally participate in the practice. Though peripheral, the participation is legitimate in the sense that these new apprentices can observe other members, especially the more experienced, and learn how to become full members of the community. This learning was documented in Bryant, et al.'s (2005) study, focusing on users' participation in Wikipedia over time, as the pattern and type of their contributions and motivations evolved over time.

4. Identity: Shared identity is partially formed around the shared practice of the CoP. The Wikipedia user develops an identity as a Wikipedian (Bryant, et al., 2005). Anthony, Smith, & Williamson (2009) contended that, as shown among open source developers, reputations and group identity are one of the motivators to contribute to public goods. Due to the strong group identity on the Wikipedia community, contributions are flourishing.

Because Wikipedia possesses characteristics of CoPs, we are able to contribute to the discussion of cross-cultural analysis of CoPs by analyzing multiple language versions of Wikipedias.

Hofstede's dimensions of cultural differences

Any cross-cultural analysis is based on an accepted understanding of the term culture. Culture is defined as "the collective programming of the mind which distinguishes the members of one group or category of people from another" (Hofstede, 1991, p. 5). One way that scholars try to understand the nature of cultural differences is to understand a pattern composed of a combination of dimensions (Straub, Loch, Evaristo, Karahanna & Srite, 2002). Each combination is unique to a culture, and a single dimension cannot capture the complexity of a culture; still, some scholars suggest that each dimension can be separately examined. While various such approaches to culture exist (e.g., Hall, 1976; 1983; Hofstede, 1991; Kluckhohn and Strodbeck 1961; Trompenaars & Hampden-Turner, 1998), Hofstede's (1991) five dimensions are the most commonly utilized pattern in cross-cultural research. The dimensions include: low/high power distance, individualism/collectivism, masculinity/femininity, uncertainty avoidance, and long- term/ short-term orientation. The fifth dimension was a later addition to his framework based on the contribution of Chinese scholars (Hofstede, 2008). Empirical studies that utilize the

5

dimensional approaches, however, report inconsistencies in the context of groups that use technology. Most of the studies focus only on 3-4 dimensions, mostly on Hofstede's (1991) dimensions (Mayers & Tan, 2002), and typically, they only focus on the individualism/collectivism dimension.

Power distance (high vs. low) refers to the distance between different people in varied levels of a hierarchy. It is "the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally" (Hofstede, 1991, p.28).

Individualism vs. Collectivism describes the orientation of an individual either first toward the group interest and then to his own, or first to his own and then toward the group interest. The group responsibility toward the individual is evident in a collectivist culture, yet almost non-existent in individualistic cultures where the individual is expected to look after himself. "Individualism pertains to societies in which the ties between individuals are loose: everyone is expected to look after himself or herself and his or her immediate family. Collectivism as its opposite pertains to societies in which people from birth onwards are integrated into strong, cohesive in groups, which throughout people's lifetime continue to protect them in exchange for unquestioning loyalty" (Hofstede, 1991, p.51).

Masculinity vs. Femininity. Masculinity refers to more assertiveness and business-interest orientation, while femininity refers to taking care of quality of life and the individual need, which are external to the business. Hofstede (1991) suggested that "masculinity pertains to societies in which social gender roles are clearly distinct (i.e., men are supposed to be assertive, tough, and focused on material success, whereas women are supposed to be more modest, tender, and concerned with the quality of life); femininity pertains to societies in which social gender roles overlap (i.e., both men and women are supposed to be modest, tender, and concerned with the quality of life)" (p. 82).

Uncertainty avoidance (high vs. low) refers to the notion of taking or avoiding risks, and to the attitude toward uncertainty and ambiguity. Uncertainty avoidance is "[T]he extent to which the members of a culture feel threatened by uncertain or unknown situations. This feeling is, among other things, expressed through nervous stress and in a need for predictability: a need for written and unwritten rules" (Hofstede, 1991, p. 113).

The *Long-term vs. Short term* dimension is a later addition to the four initial dimensions. Hofstede (1991) suggested that

"Short term orientation [societies are characterized by] respect [for] tradition; respect for social and status obligations regardless of cost; social pressure to 'keep up with the Joneses' even if it means overspending; small saving quota, little money for investment; quick results expected; concern with possessing the truth...Long term orientation [societies are characterized by] adaptation of traditions to a modern context; respect for social and status obligations within limits; thrift, being sparing with resources; large savings quotes, funds available for investment; perseverance towards slow results; willingness to subordinate oneself for a purpose; concern with respecting the demand of virtue" (p. 173).

Some Eastern countries are characterized by high scores on Hofstede's Power Distance Index, and high scores on Collectivism. The Western countries, on the other hand, have low scores on the Power Distance Index and high scores on Individualism. Taken together, Hofstede suggested that there is a negative relationship between these two dimensions: "In cultures in

which people are dependent on ingroups these people are *usually* also dependent on power figures" (Hofstede, 1991, p. 55). "In view of the correlation between power distance and collectivism one could consider them as two manifestations of one single dimension of cultural differences" (Hofstede, 1991, p. 56). Nonetheless, Hofstede clarified that these two are distinct, but both correlate with a third factor – a country's economic development. Thus, either dimension will support our interpretations of the differences that are found between Western and Eastern countries. Researchers that focus on cultural differences often distinguish

between Eastern and Western countries (Nisbett, 2003). "Psychologists have conducted a wealth of experiments that seem to support popular notions that easterners have a holistic world view, rooted in philosophical and religious traditions such as Taoism and Confucianism, while westerners tend to think more analytically, as befits their philosophical heritage of reductionism, utilitarianism and so on." (Yong, 2009).

In summary, prior research of online CoPs is driven by single case studies constrained within organizational boundaries, defies the examination of explicit norms of behaviors, lacks extensive use of cross-cultural analysis and can benefit from additional understanding of the influence of size on behaviors. In addition, research on Wikipedia primarily focuses on the English Wikipedia and tends to overlook the Wikipedias in languages other than English. This research examines the variation of norms of behaviors among Wikipedia users in four languages.

Methods

In an effort to answer the research questions, content analysis of 120 Wikipedia talk pages in four languages was conducted. In order to address our research questions and to answer whether size or culture can help understanding of the differences among Wikipedias, we chose four Wikipedias that differed on two dimensions: size and culture. Thirty pages were randomly selected from Wikipedia in each of the four different languages—English, Hebrew, Japanese, and Malay; two languages are spoken in Western cultures (English and Hebrew) and two in Eastern cultures (Japanese and Malay). These Wikipedias also vary in their sizes: two large Wikipedias, with over 100,000 articles each (English, with 2,580,417 articles, and Japanese, with 526,800 articles), and two smaller Wikipedias, with less than 100,000 articles each (Hebrew, with 83,034 articles, and Malay, with 30,890). Table 1 provides additional statistics on Wikipedia in these four languages.
7

Language	Rank	Articles[1]	Share of Wikipedia[2]	Number of Population[3]	Users per article[4]	Article per Edits by size[1]
English	1	2,976,299	22.54%	10,203,869	1,889,458	80.70
Japanese	5	605,534	4.60%	332,402	392,559	36.57
Hebrew	28	94,970	0.72%	88,480	60,190	63.22
Malay	43	44,552	0.27%	40,773	21,897	15.97

Ratio of users to administrators[5]
6,095.50
5,113.88
1,638.82
2,912.36

Table 1. Descriptive information on four Wikipedias [as of August 3, 2009]

Data Collection

Data collection was undertaken in two stages. First, the data collection from the English Wikipedia was conducted in July 2007 and was repeated on the other Wikipedias (Hebrew, Japanese, and Malay) in September 2008.

Wikipedia's massive information is saved under 20 namespaces; the biggest one is the "main" namespace, which includes all encyclopedia articles (Viégas, Wattenberg, Kriss, & van Ham,

2007). Among these 20 namespaces, nine talk namespaces exist. Talk namespaces provide users with places in which they can discuss issues and communicate with each other. Talk namespaces are linked to other namespaces. For example, a talk page is associated with each Wikipedia article, which is called a "Talk" page. Likewise, a talk page is associated with each user page or Wikipedia policies page, which is called a "User Talk" page or a "Wikipedia Talk" page, respectively. In addition, a history page is linked to each Wikipedia page from all namespaces. These history pages provide links to older versions of that page as well as history pages for its respective talk pages; any of these pages may have an archive (in cases of intensive discussions on a specific talk page, some of the older discussions are archived). Among the nine talk namespaces, (1) Talk; 2) User Talk; 3) Wikipedia Talk; 4) Image Talk; 5) MediaWiki Talk; 6) Template Talk; 7) Help Talk; 8) Category Talk; and 9) Portal Talk), the top three most active talk name spaces, i.e., Talk, User Talk, and Wikipedia Talk (Viégas et al., 2007), were chosen for the analysis.

Sampling process

To identify the norms of behaviors on the English Wikipedia, we selected the three most active talk name spaces (Talk, User Talk, and Wikipedia Talk) and randomly selected talk pages from each name space. Using the Wikipedia search capabilities, a list frame for the various talk pages was generated on July 2008 for the random sampling. Random sampling of Wikipedia pages is not a common practice; prior research on Wikipedia norms used purposeful sampling

[1] http://meta.wikimedia.org/wiki/List_of_Wikipedias

[2] http://en.wikipedia.org/wiki/Wikipedia:Multilingual_statistics

[3] http://en.wikipedia.org/wiki/Wikipedia:Wikipedia_articles_per_population [4] http://meta.wikimedia.org/wiki/List_of_Wikipedias_by_edits_per_article

[5] Divide the number of users by the number of administrators found in http://meta.wikimedia.org/wiki/List_of_Wikipedias 8

(e.g., Viégas et al., 2007). As Ortega, Gonzalez-Barahona, and Robels (2008) indicated, contributions to Wikipedia are not equal. As a result, when we randomly chose talk pages, some pages did not have any discussions on talk pages. Whenever this was the case, we did not include these talk pages in the sample. We chose this sampling strategy because highly attracted Wikipedia articles have their unique characteristics (Pfeil, et al., 2006). For example, using topics that attract significant interest, traffic and attention from the media, the general public, and vandals does not create a sample that represents Wikipedia as a whole. Thus, by using random sampling, the findings from this study is more likely to be generalizable to other Wikipedia articles, compared, for example, with Pfeil et al's sample that includes only one topic, "game". Our sample further extends Pfeil et al's study, which limits the analysis to examining revisions people made on articles, by examining three different discussion areas on each of the languages.

The sample in the first step of the pilot study included 30 pages, 10 pages from each of the three name spaces, Talk, User Talk, and Wikipedia Talk (see Table 2). These 30 pages include more than 700 posts.

WIKIPEDIA TALK
HAIKU ABOUT WIKIPEDIA POLICY
SGPEDIANS' NOTICE BOARD/ARTICLE ASSESSMENT
MANUAL OF STYLE (ISLAM-RELATED ARTICLES)/MA MALAKAT AYMANUKUM
FEATURED PICTURE CANDIDATES/ARCHIVE 1 STRESS ALERTS
COPIES OF WIKIPEDIA CONTENT (UNDETERMINED OR DISPUTED COMPLIANCE)
REQUESTS FOR ADMINSHIP/COOL3 (SECOND NOMINATION)
SU
VERIFIABILITY/ARCHIVE 18 DON'T OBJECT TO PROPOSALS
USER TALK
MAIS OUI!
SLIMVIRGIN ANONMOOS

BETACOMMANDBOT SAN DIABLO
BATHROBE
KYLOHK
SKUNKMASTER PEEPHOLE DOODOOBUTTER
TALK
LIST OF CELEBRITIES WHO HAVE BEEN AWARDED THE ORDER OF THE BRITISH EMPIRE
EVERTON FC PEARLAND, TEXAS
TNA X DIVISION TITLE
PETER, PHILIP AND ALEXANDER, PRINCES OF YUGOSLAVIA
VOLGA RIVER
GREAT GATSBY
CARAVAGGIO
DO YOU SEE WHAT I SEE HASSAN (NAME)

Table 2: The sample: List of Talk, User Talk, and Wikipedia Talk pages

During the second step of data collection, the same sampling procedure was utilized to collect pages from the other three Wikipedias in September 2008. The sample at this step included 90 additional pages, from the three other languages (Hebrew, Japanese, and Malay). The final

9

sample used in this study included 120 talk pages, 30 talk pages from each of the four Wikipedias and from three different name spaces, 10 from each type of talk page. In total, this includes over 2,700 Wikipedia posts.

Data Analysis

The coding scheme was developed from the ground up during the pilot study with the English Wikipedia, using interpretive content analysis; the first and second authors coded a subset of the data and suggested a list of codes. These codes were discussed and grouped into broader categories. The discussions with the third author refined individual codes. The final coding scheme includes fifty-six codes under 3 categories: writing norms, information sharing, and community well-being (see the Appendix). This coding scheme was later used to code the pages and to compare norms of behavior among the three types of talk pages.

The data was coded under each scheme, and frequency tables were created for each language. Each of the authors and an independent researcher coded the selected pages in one language, which was his/her native language. All four coders coded 10% of the data of the English Wikipedia to examine the level of inter-coder reliability (number of agreements, divided by the sum of the number of agreements, plus number of disagreements). Inter-coder reliability was high: 95% between the Hebrew and English Wikipedias, 85% between the Japanese and English Wikipedias, and 86% between the Malay and English Wikipedias.

Table 3 presents the total number of posts for the three namespaces in four Wikipedias. Naturally, the English Wikipedia had the largest number of messages in Talk, User Talk, and Wikipedia Talk pages because it is the largest Wikipedia. The size also correlates with the number of messages posted on Wikipedia Talk pages. The larger the Wikipedia becomes, the more it involves administrative tasks. However, if we only consider the size of Wikipedias, the Japanese Wikipedia should have accumulated the second greatest number of messages, then Hebrew, and Malay in all three namespaces. While this logic worked for Talk and Wikipedia Talk pages, the Hebrew and Malay Wikipedias had more messages on User Talk pages than the Japanese Wikipedia. Even in Talk pages, the Japanese and Hebrew Wikipedias had approximately the same amount of messages. The data indicates that the size is not the only explanation of the differences among the various Wikipedias.

**ENGLISH (1264) HEBREW (686) JAPANESE (567)
MALAY (196)
TALK USER TALK (627) (833)
288 421 156 333 162 25
21 54
WIKIPEDIA TALK (1253)
555 197 380
121**

Table 3. Total number of posts in the three namespaces among four languages

10
Because the study was designed to investigate the difference between Eastern and Western cultures and larger and smaller Wikipedias, the data were further aggregated based on culture and size. For each of the four conditions (i.e., Western, Eastern, large, and small Wikipedias) we created an aggregated frequency column for Talk, User Talk and Wikipedia Talk. Subsequently, cross tabulation analysis was conducted on the data using SPSS 17.0.

Findings and Discussion

As a part of the analysis process, some codes were grouped into the following categories: quality and accuracy; courtesy; and conflict and disagreement (see Table 4 for specific codes). Table 5 presents percentages of codes that appeared for specific categories. The percentages were calculated by using code frequencies divided by the total number of posts in order to compensate for size differences among the four Wikipedias.

**QUALITY AND ACCURACY
COURTESY
CONFLICT & DISAGREEMENTS
QUESTION - ABOUT CITATIONS
QUESTION - ABOUT HOW TO DO/ OR ABOUT POLICIES
QUESTION - A REQUEST FOR INFORMATION (I.E., A REFERENCE QUESTION) EXPLANATION - DISCUSSION OF FACTS ACCURACY
SHARING INFORMATION ABOUT HOW TO DO
PERSONAL OPINION
EVALUATION OF QUALITY OF CONTRIBUTIONS
APOLOGIES & REGRETS
APPRECIATION
POLITENESS/COURTESY
GREETING
CONFLICT OVER FACTS**

Table 4. Grouping of codes

As can be seen in Table 5, we aggregated the frequencies of codes by culture (West/East) and size (small/large). English and Hebrew Wikipedias were coded as Western, while Japanese and Malay were coded as Eastern. Large Wikipedias include English and Japanese Wikipedias, and small Wikipedias include Hebrew and Malay Wikipedias.

Culture—Quality and Accuracy				Size—Quality and Accuracy			
	Talk	User Talk	Wikipedi a Talk		Talk	User Talk	Wikiped ia Talk
West	69.67	45.33	37.72	Large	52.27	31.72	28.64
East	45.24	32.02	29.95	Small	62.64	45.63	39.03
Culture—Courtesy				Size—Courtesy			

11

	Talk	User Talk	Wikipedi a Talk		Talk	User Talk	Wikiped ia Talk
West	21.29	120.73	21.54	Large	107.45	254.89	113.92
East	106.87	234.97	114.44	Small	21.33	100.81	22.06
Culture—Conflict and Disagreements				Size—Conflict and Disagreements			
	Talk	User Talk	Wikipedi a Talk		Talk	User Talk	Wikiped ia Talk

West	98.82	26	36.22	Large	22.06	10.69	7.34
East	31.39	11.11	6.54	Small	108.15	26.42	35.42

Table 5. West/East and Large/Small comparison on four categories in percentage

Table 6 presents the results of cross tabulation of 3 namespaces by size and by culture. These are briefly described here and discussed in the following sections. A cross-tabulation analysis was conducted to evaluate whether the frequency of quality and accuracy of posts on the three talk pages differ between the two cultures. The results indicated no statistically significant difference between East and West in this sample (Table 6). Likewise, a cross tabulation analysis to evaluate the frequency of posts about accuracy and quality based on size was done. Findings indicated that these differences were not statistically significant. A cross tabulation analysis that focused on the relationship between conflict and disagreement for both cultures and sizes of talk pages revealed no statistical differences on either (Table 6). The only significant differences that our cross tabulation analysis revealed were based on culture and size in courteous behaviors on the three talk pages.

		Pearson χ2	Cramer's V	p value
East/West	Quality & Accuracy (N=260, df=2)	446	41	.8
	Courtesy (N=620, df=2)	24.415	198	.000
	Conflict & Disagreement (N=210, df=2)	2.062	99	357
Size	Quality & Accuracy (N=261, df=2)	.348	37	.84
	Courtesy (N=620, df=2)	12.413	141	2
	Conflict & Disagreement (N=209, df=2)	3.26	125	.196

Table 6. Cross tabulation results of 3 talk pages by size and 3 talk pages by culture.

Descriptive data indicated that the content of messages varied across types of talk pages and that the pattern of variation was similar across all four languages. For example, postings about accuracy and quality of information were more frequently found on Talk pages than on User Talk pages, or Wikipedia Talk pages. This tendency was evident in both large and small Wikipedias and on Western and Eastern Wikipedias. However, user interactions and their communication style were consistent across talk pages in each language but differed between Wikipedias. These variations correlate with Hofstede's dimensions of cultural diversity, specifically with the power distance index. In the following sections, we describe these similarities and differences across cultures and sizes that we found, and discuss each one of them.

Similarities across cultures & sizes

Postings about accuracy and quality of information were more common on Talk pages than User talk or Wikipedia Talk, in large and small and in Eastern and Western Wikipedias. At the same time, more courteous posts were found in User Talk pages than Talk or Wikipedia Talk pages. We speculate that this pattern reflects the different roles of each type of Talk pages within the Wikipedia community at large. Talk pages are task oriented, as reflected by the emphasis on quality and accuracy, while User Talk pages are more social in nature. This pattern is evident across cultures and sizes.

Another similarity across cultures and sizes is politeness behaviors. Courteous postings (for example, apologies, appreciation, and greetings—see Table 4 for specific codes) were posted more frequently on User Talk than on any other type of talk page, both in large and small and in Eastern and Western Wikipedias (see Figure 1). We speculate that this reflects a higher level of interest in the community well-being on the User Talk pages compared with Talk pages, and higher level of attention to quality and accuracy on the Talk pages compared with User Talk pages. This pattern of differences occurs naturally as users tend to socialize on User Talk pages, and not as much in other task-oriented talk pages. It is clear that users differentiate their behaviors on the various namespaces when they post messages.
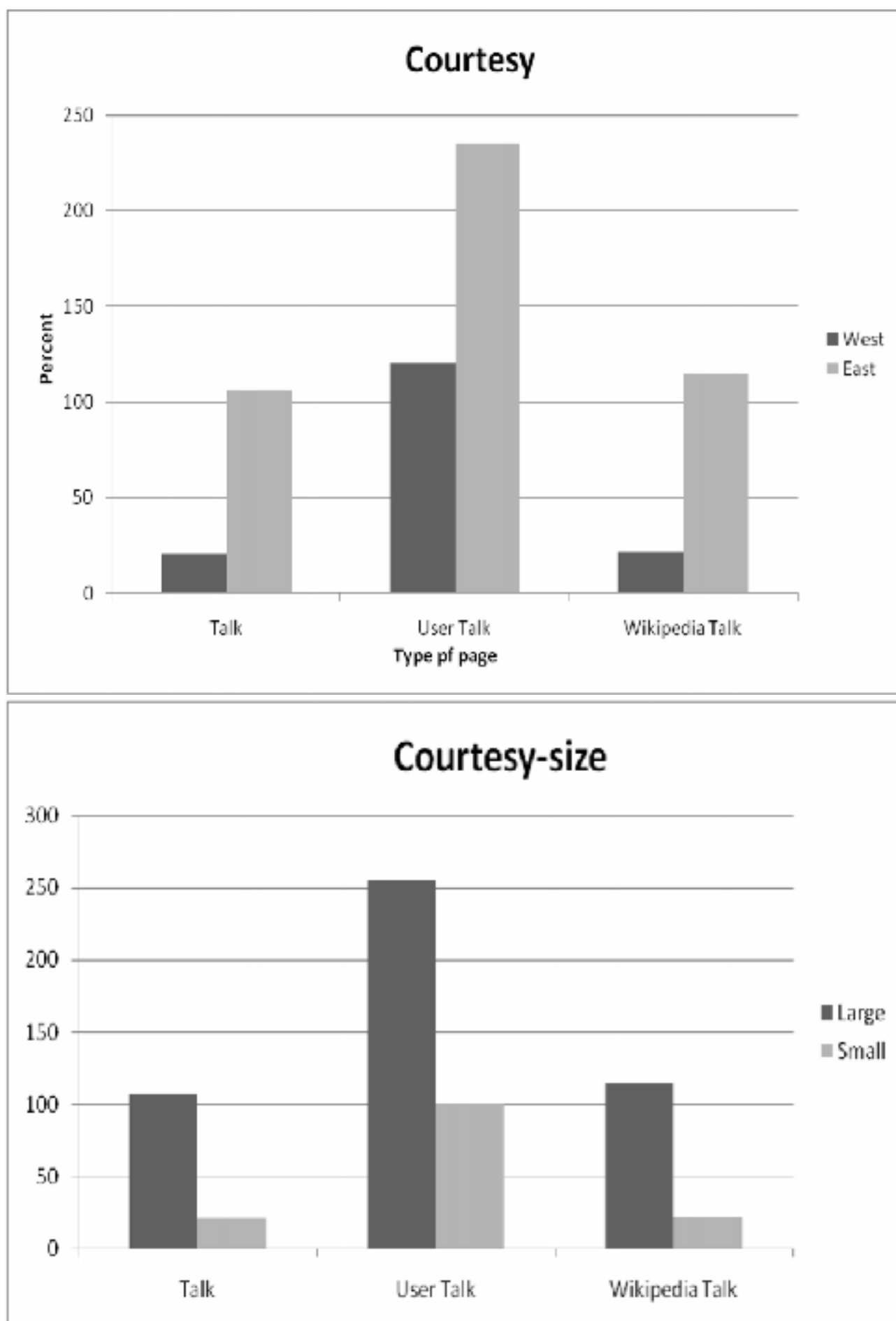
Figure 1. Courtesy for East/West (top) and Size (bottom)

Variations across languages

Variations in communication style were found based on Wikipedia size and culture; these styles were consistent in each language on all three types of talk pages (Talk, User Talk, and Wikipedia Talk). Hofstede's dimensions of cultural diversity can support the interpretations of these variations.

Courtesy behaviors in the large Wikipedias were more frequent compared to the small Wikipedias on all three types of talk pages (see Table 5); these differences in courtesy behavior based on size were found to be statistically significant (Table 6). Eastern Wikipedias had significantly more courteous messages on each type of talk page than the Western Wikipedias (Table 6).

These variations across cultures possibly reflect the difference between high vs. low power distance cultures. Higher levels of politeness are associated with high power distance cultures (e. g., Brown & Levinson, 1987; Spencer-Oatey, 2000; Holmes & Stubbe 2003). Countries differ

14

in the perception of inequalities among their inhabitants in the context of family, school, and work. In high power distance countries (i.e., Japan) there would be more respect toward parents (by their children), teachers (by their students), and bosses (by their subordinates). Even the language systems in high power distance cultures emphasize distinctions based on a social hierarchy. Moreover, in collectivistic/high power distance cultures, relationships prevail over tasks (Trompenaars & Hampden-Turner, 1998). This, again, may be an additional reason for more courtesy behaviors in the East compared with the West. Likewise, the difference between East and West in the amount of conflict and disagreement are similarly rationalized (Trompenaars & Hampden-Turner, 1998). Conflict and disagreement behaviors were more frequently observed in the West (Figure 2), perhaps because tasks tend to prevail over relationships in individualistic countries; these differences, however, were not statistically significant (Table 6). Thus, we can conclude that there is more concern over task – accuracy and quality issues – in the West compared with the East.

Figure 2. Conflict and Disagreement for East/West (top) and Size (bottom).

Small Wikipedias exhibited more conflict behaviors than the large Wikipedias, although the differences were not statistically significant (Table 6). Earlier studies about group size and cooperation would support an opposite behavior pattern (e.g., Hamburger, Guyer, & Fox, 1975). The larger the group becomes, the more conflict could arise due to the stipulation of public

Conflict and Disagreement

Conflict & Disagreements-size
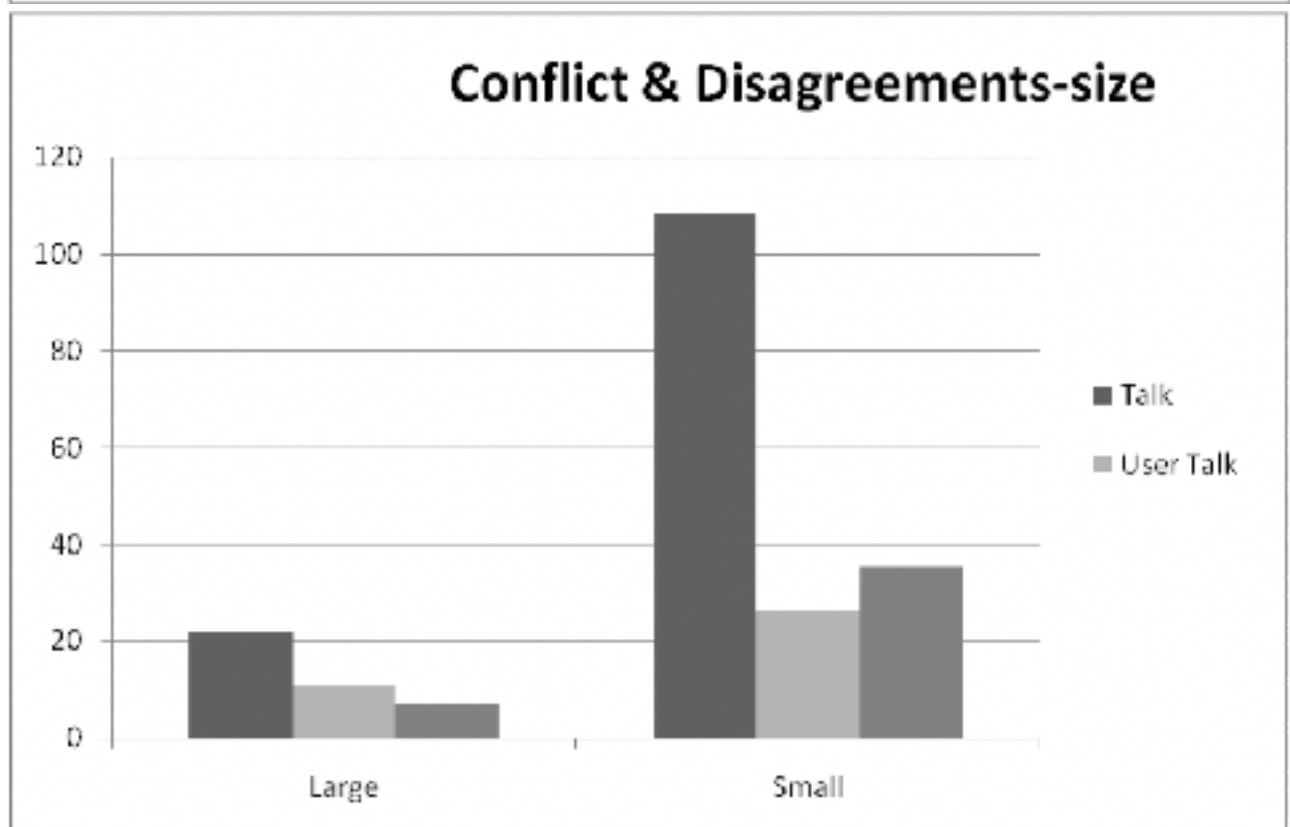
15

goods (Olson, 1965). However, this idea of public goods has been disputed because open online CoPs, such as Wikipedia, offer a milieu for mass contribution to public goods (Anthony, et al., 2009). Ortega, et al. (2008) also found that "in communities with a small number of authors, contributions tend to be concentrated by the group of most active ones." The observation of differences based on sizes might be delineated by the fact that because the small Wikipedias have a small number of contributors and are still under development, every member is greatly invested in the project. This could increase disagreement among members about articles because everyone has a strong attachment to their own contributions. In fact, codes related to quality and accuracy are posted more often in the small than the large Wikipedias. It is possible that smaller

Wikipedias may have more issues with quality control than larger ones. The data also shows that more conflicts occured in Talk pages than the other two namespaces, which indicates that although people disagree on topics of their interests, this conflict is not personal in nature. The behavioral variations based on size differences need further investigation.

Conclusions

This paper identified the Wikipedias as CoPs and examined three types of Wikipedia talk pages in four languages. It discussed similarities and differences among cultures, sizes, and talk pages in light of the function of namespaces and Hofstede's dimensions of cultural diversity.

In general, task-oriented postings, such as quality and accuracy, were found more frequently on Talk pages compared with Wikipedia Talk pages and User Talk pages, in both Eastern and Western Wikipedias and small and large Wikipedias. On the contrary, community well-being postings, such as courtesy, appeared mostly in User Talk pages compared to Talk and Wikipedia Talk pages in both Eastern and Western and small and large Wikipedias. These findings indicated that, in all of these four languages, Wikipedia users similarly differentiate the use of each type of discussion area by posting messages of different natures and purposes on different types of talk pages. We argued that Hofstede's (1991) dimensions of cultural diversity can help us interpret differences between the Eastern and Western Wikipedias that have been found in this study.

One of the limitations of this study that is common to cross-cultural research is that multiple researchers, speakers of various languages, coded the data. Even though the inter-coder reliability on the English pages among all three coders was high, it is possible that some of the variations across languages are partially due to differences between coders. Another limitation is that our analysis is based on the assumption that a Wikipedia in a specific language relates to a national culture of a specific country. The languages used in the Wikipedias, however, may not exactly correspond with a specific country. This assumption is in line with prior cross- cultural research on Wikipedia (Jiang, 2000; Pfeil, et al., 2006). Finally, due to the different size of the four Wikipedias, the frequencies were presented in percentages rather than actual numbers of codes.

Despite these limitations, the implications of this study are as follows: First, based on the clarification that Wikipdias are considered as CoPs, we have developed a coding scheme that examines norms of behaviors in the particular online CoPs in which users engage in knowledge sharing practices. The behaviors we identified focus specifically on "knowledge shaping" (Yates,

16

Wagner, & Majchrzak, in pres) in online CoPs. Knowledge shaping (i.e., edit, integrate, and rewrite content) in online CoPs differs from "knowledge creation" (Yates, et al., in press) and is critical in the context of mass collaboration using wikis that allow edits from multiple contributors.

Second, only a few studies in the past have conducted cross-cultural analysis of CoPs, and the current study is the first cross-cultural analysis of CoPs that exists outside organizational boundaries. The four Wikipedias that we investigated provided a valuable test-bed to conduct such analysis. In the chosen four Wikipedias, members of the communities are all engaged in similar activities with shared goals. However, the data shows that transferability of findings about the English Wikipedia (Western culture) should be made with caution to Eastern cultures. This also means that the overwhelming number of studies conducted in Western countries about CoPs may not be applicable to other parts of the world. There is a dire need for more cross-cultural analyses and studies of CoPs in non-Western countries. Similarly, our study is one of the few that examines more than one Wikipedia, and the first to investigate multiple Wikipedias using two criteria — culture and size.

Third, many previous studies of online CoPs were conducted in organizational settings. A number of organizations began to explore possibilities to extricate from traditional knowledge sharing mechanisms by using conversational tools such as blogs, wikis, and social networking software (e.g., DiMicco, Millen, Geyer, Dugan, Brownholtz, & Muller, 2008; Wagner & Bolluju, 2005). However, knowledge sharing opportunities are increasingly not limited by organizational boundaries. As the practice of knowledge sharing opportunities outside organizations becomes a norm, we need to scrutinize open online CoPs. This study is informative for such open online CoPs.

In conclusion, this research sheds light on how CoPs operate by analyzing norms of behaviors. In particular, the four Wikipedias that we examined provided exemplars of CoPs that exist in different cultural environments. Future research should expand the number of languages. As there are few studies of cross-cultural analysis about online CoPs, the Wikipedias provided nice test-beds to examine variations of norms of behaviors in different cultures. Future research should look into cross-cultural analysis of CoPs in various cultures as well as size differences.

References

Anthony, D., Smith, S. W., & Williamson, T. (2009). Reputation and reliability in collective goods: The case of the online encyclopedia Wikipedia. *Rationality & Society, 21*, 263-306. Ardichvili, A., Maurer, M., Li, W., Wentling, T., & Stuedemann, R. (2006). Cultural influences on knowledge sharing through online communities of practice. *Journal of Knowledge Management, 10*(1), 94-107.

Ardichvili, A., Page, V., & Wentling, T. (2003). Motivation and barriers to participation in online knowledge–sharing communities of practice. *Journal of Knowledge Management*, *7*(1), 64-77.

Barab, S.A., MaKinster, J. G., & Scheckler, R. (2004). Designing system dualities: Characterizing an online professional development community. In S. A. Barab, R. Kling, & J. H. Gray 17 (Eds.), Designing for virtual communities in the service of learning. Cambridge: Cambridge University Press.

Beschastnikh, I., Kriplean, T., & McDonald, D. (2008). Wikipedian self-governmance in action: Motivating the policy lens. In the proceedings of *Web Intelligence Conference*. IEEE CS Press, 45-51.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bryant, S. L., Forte, A., & Bruckman, A. (2005). Becoming Wikipedian: Transformation of participation in a collaborative online encyclopedia. Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work, November 6-9, Sanibel Island, Florida.

DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B., & Muller, M. (2008). Motivations for social networking at work. *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work*, 711-720.

Dubé, L., Bourhis, A., & Jacob, R. (2006). Towards a typology of virtual communities of practice. *Interdisciplinary Journal of Information, Knowledge, and Management, 1*, 69-93.

Forte, A., Larco, V., & Bruckman, A. (2009). Decentralization in Wikipedia Governance. *Journal of Management Information Systems, 26*(1), 49-72.

Gray, B. (2004). Informal learning in an online community of practice. *Journal of Distance Education*, *19*(1), 20-35.

Hall, E.T. (1976). Beyond culture. Garden City, New York: Doubleday.

Hall, E.T. (1983). The dance of life. Garden City, New York: Doubleday.

Hara, N. (2009). Communities of practice: Fostering peer-to-peer learning and information *sharing in the work place*. Berlin: Springer.

Hara, N., & Hew, K.F. (2007). Knowledge sharing in an online community of health-care professionals. *Information Technology & People, 20*(3), 235-261.

Hara, N., Shachaf, P., & Storeger, S. (2009). Online communities of practice typology revisited. Journal of Information Science, 35(6), 740-757.

Hew, K.F., & Hara, N. (2007). Knowledge sharing in online environments: A qualitative case study. *Journal of American Society for Information Science & Technology, 59*(14), 2310- 2324.

Hofstede, G.H. (1991). Cultures and organizations: Software of the mind. London: McGraw-Hill. Hofstede, G.H. (2008). Announcing a new version of the Values Survey Module: the VSM 08. Retrieved January 6, 2009, from http://stuwww.uvt.nl/~csmeets/VSM08.html

Holmes, J., & Stubbe, M. (2003). *Power and politeness in the workplace: A sociolinguistic analysis of talk at work*. London: Pearson.

Hamburger, H., Guyer, M., & Fox, J. (1975). Group size and cooperation. *Journal of Conflict Resolution, 19*(3), 503-531.

Jiang, W. (2000). The relationship between culture and language. *ELT Journal*, *54* (4), 328-334. Jones, Q. (1997). Virtual-communities, virtual-settlements and cyber-archaeology: A theoretical

outline. Journal of Computer Mediated Communications, 3(3). Retrieved July 10, 2005, from http://jcmc.indiana.edu/vol3/issue3/jones.html

Johnson, C.M. (2001). A survey of current research on online communities of practice. *The Internet and Higher Education*, 4(1), 45-60.

18

Kittur, A., & Kraut, R.E. (2008). Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. In the proceedings of *Computer Supported Cooperative Work*, 37-45.

Kluckhohn, C., & Strodbeck, F.L. (1961). *Variation in value orientation*. Evanson, IL: Row and Peterson.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge: Cambridge University Press.

Mayers, M.D., & Tan, F.B. (2002). Beyond models of national culture in information systems research, *Journal of Global Information Systems*, *10*(1)*, 24-32.

Nisbett, Nisbett, R.E. (2003). The geography of thought: How Asians and Westerners think differently and why. New York: The Free Press.

Olson, M. (1965). The logic of collective action: Public goods and the theory of groups. Cambridge, MA: Harvard University Press.

Orr, J.E. (1996). Talking about machines: An ethnography of a modern job. Ithaca, NY: ILR Press. Ortega, F., Gonzalez-Barahona, J.M., & Robels, G. (2008). On the inequiality of contributions to Wkipedia. In proceedings of 41st *Annual Hawaii International Conference on System Sciences*. Piscataway, NJ: IEEE Press. Retrieved January 27, 2009 from http://doi.ieeecomputersociety.org/10.1109/HICSS.2008.333

Pan, S.L., & Leidner, D.E. (2003). Bridging communities of practice with information technology in pursuit of global knowledge sharing. *Journal of Strategic Information Systems, 12*, 71- 88.

Pfeil, U., Zaphiris, P., & Ang, C.S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication*, *12*(1), article 5. Retrieved January 2, 2009, from http://jcmc.indiana.edu/vol12/issue1/pfeil.html

Riehle, D. (2006). How and why Wikipedia works: an interview with Angela Beesley, Elisabeth Bauer, and Kizu Naoko. *Proceedings of the 2006 international symposium on Wiki*s, Odense, Denmark. Retrieved September 29, 2006, from http://doi.acm.org/10.1145/1149453.1149456

Roberts, J. (2006). Limits to communities of practice. *Journal of Management Studies*, *43*(3), 623-639.

Schlager, M., Fusco, J., & Schank, P. (2002). Evolution of an online education community of practice. In K. A., Renninger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (pp. 129-157). Cambridge: Cambridge University Press.

Shachaf, P., & Hara, N. (2010, in press). Beyond vandalism: Wikipedia trolls. *Journal of Information Science.*

Spencer-Oatey, H. (2000). Culturally speaking - Managing rapport through talk across cultures. London: Continuum.

Straub, D., Loch, K., Evaristo, R., Karahanna, E., & Srite, M. (2002). Toward a theory-based measurement of culture. *Journal of Global Information Management*, *10*(1),13-23.

Stvilia, B., Al-Faraj, A., & Yi, Y.J. (2009). Issues of cross-cultural information quality evaluation— The case of Arabic, English, and Korean Wikipedias. *Library & Information Science Research,31*(4), 232-239.

Stvilia, B., Twidale, M., Smith, L.C., & Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science & Technology*, *59*(6), 983–1001.

19

Tapscott, D., & Williams, A.D. (2008). Wikinomics: How mass collaboration changes everything. New York: Portfolio.

Trompenaars, F., & Hampden-Turner. C. (1998). Riding the waves of culture: Understanding cultural diversity in global business. New York: McGraw-Hill.

Veenswijk, M., & Chisalita, C. (2007). The importance of power and ideology in communities of practice. *Information Technology & People, 20*(1), 32-52.

Viégas, F.B., Wattenberg, M., Kriss, J., & van Ham, F. (2007). Talk before you type: Coordination in Wikipedia. *Proceedings of the Forty Hawai'i International Conference on System Sciences* (HICSS-40). Los Alamitos, CA: IEEE Press.

Wagner, C., & Bolluju, N. (2005). Supporting knowledge management in organizations with conversational technologies: Discussion forums, weblogs, and wikis. *Journal of Database Management*, *16*(2), 1-13.

Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly, 29*(1), 35-57.

Wenger, E. (1990). Toward a theory of cultural transparency: Elements of a discourse of the visible and the invisible. Unpublished doctoral dissertation, University of California, Irvine.

Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge: Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W.M. (2002). Cultivating communities of practice: A guide to managing knowledge. Boston: Harvard Business School Press.

Yates, J., Orlikowski, W.J., & Okamura, K. (1999). Explicit and implicit structuring of genres in electronic communication: Reinforcement and change of social interaction. *Organization Science, 10*(1), 83-103.

Yates, D., Wagner, C., & Majchrzak, A. (in press). Factors affecting shapers of organizational wikis. Journal of the American Society for Information Science and Technology.

Yong, E. (March, 2009). East meets west: How the brain unites us all. *New Scientist 2698*. Retrieved February 5, 2010 from: http://www.newscientist.com/article/mg20126981.700-beyond-east-and-west-how-the-brain-unites-us-all.html?full=true

20

Appendix

Table 7. Coding scheme – part1: Writing Norms

Codes	Code Definition
1. Signature	
1.1 signature simple	Plain signature
1.2. signature modified	Add color or include some symbols (i.e. star of david), signs (i.e. chat), or additional links as part of a signature
1.3. signature—later addition	Later addition to indicate who posted the message
1.4. no signature	No signature
1.5. IP addresses as a signature	IP address as a signature
2. Writing style	
2.1. address others by name	Refer to another user, use actual (wikipedia user) names to discuss issues
2.2. mention other user by name	Address another user by his/her name
2.3. expressive writing style	Use certain characters (e.g., !), style (e.g., bold face) or emoticons to express one's feelings.
2.4. Use foreign languages	Insert foreign languages and or foreign characters in the middle of the discussions

3. Playful comments — Sarcasms, joke, humor	Use sarcasms to make a point; insert a joke or make a humorous comment
4. Links	
4.1. link to Wikipedia policies	Make a link to Wikipedia policies
4.2. link to other Wikipedia pages	Make a link to other Wikipedia pages
4.3. link to other webpages	Make a link to pages that are outside of Wikipedia
4.4. link from random and irrelevant words	Make a link to words that are not necessarily important to link (e.g., Well you've found a decent adhoc solution so all is well.)
5. Jargon and Acronyms	
5.1. Wikipedia acronyms	Use Wikipedia acronyms, such as RFC (request for comments)
5.2. Jargons from a specific context	Use jargons form a specific discipline
5.3. Internet acronyms	Use acronyms that are common among regular Internet users, e.g., IMHO (in my humble opinion).
6. Article Ranks	
6.1. ranking of article's quality	Ranking the quality an article based on the quality scale, e.g., FA-Class
6.2. ranking of article's	Ranking the importance of an article based on the importance

importance	scale

Table 8. Coding scheme- part 2: Community Well-Being

Categories and Codes	Code Definition
1. Conflict	
1.1 conflict over style/ format	Disagree about style and/or format of an article.
1.2 spillover conflict to personal attacks/ attribution (trolls)	Intensified conflict including threats (e.g., "And please don't move Scotland back to the top or it will be you getting edit waring"; "So if you keep moving it back without reason surely it is you who should be warned")
2. Agreement	Express agreement with others and/or ideas presented in namespaces
3. Resolution	
3.1. Discussion about mediation	Discuss how to mediate conflicts through various methods embedded within Wikipedia
3.2. Request for mediation	Request for means for mediation, such as survey, polls, etc.
3.3. Declaration of (satisfactory) resolution	Declare a resolution that is either completely or somewhat satisfactory.
3.4. Compromise	Make a reasonable compromise to end conflicts. (e.g., "As you can see, I'm not doing it, although I could quote this FACT.")
3.5. Concern to maintain community well being	Express a concern for the community well-being threatened or a concern for other people's feelings (e.g, "hopefully it can be resolved to people's satisfaction.")
3.6. Coordination efforts to resolve debate	Effort to facilitate discussions, so that debates would be resolved.
3.7. Blocking	Request for blocking, for unblocking, justification for blocking, complaint about blocking.
3.8. Effort to avoid possible future conflict	Make an effort to maintain the community well-being in order to limit future conflict – (e.g., "Okay. As long as we're not going spillover our WT:IAR problems here. I'd hate to be responsible for that.")

4. Vandalism	
4.1. Vandal/ Vandalism Attribution Challenged	A request for justification of naming act as 'vandalism' or a user as a 'vandal'
4.2. Notification of Vandal/ Vandalism	A notification of 'vandalism' or 'vandal'.
4.3. Thanks for Fighting against Vandal/ Vandalism	An appreciation note about anti-vandals acts.
4.4. Comments about User Behavior - possible Vandal/ Vandalism	A comment about a user (deviant) behavior that is posted by a Wikipedian for other Wikipedians to answer/respond/react/ or be informed.
5. Apologies & regrets	Express regrets and/or apologize (e.g., "Technically. I thought I was being so very clever, but at this point it's clear it was just lame soapboxing. I regret my actions.")

22

6. Appreciation	Thanks, value, and appreciate contributions
7. Politeness/courtesy	Use polite words, such as please, would you, regards, etc. (**e.g., "**but please explain why you disagree..."; "Also, please consider..."; "regards,"; Hi, You have just reverted ...").
8. Greeting	Greeting words, such as hi, cheers, etc.
9. Group support	Express support, permission, or appreciation for the group/community (e.g, I am seeking approval before doing this (Everton FC). Also, "I have the link to a question [snip] May I link it here?" (everton)).
10. Request for collaboration	Request for possible collaboration or ask for help

Table 9. Coding scheme – part 3: Information Sharing

Categories and Codes	**Code Definition**

1. Question	
1.1. Question – About Accuracy of Facts	A question about a fact that is posted by a Wikipedian for other Wikipedians to answer/ respond
1.2. Question – About Citations	A question about or request for citation that is posted by a Wikipedian for other Wikipedians to answer/respond
1.3. Question – About How To Do/ or about policies	A question about action that is posted by a Wikipedian for other Wikipedians to answer/ respond
1.4. Question – A Request for Information (i.e., a reference question)	A request for information about a topic/fact that is posted by a Wikipedian for other Wikipedians to answer/respond
2. Explanation	
2.1 Explanation – Discussion of Facts Accuracy	A comment posted by a Wikipedian to convince other Wikipedians of the accuracy of facts.
2.2 Sharing information about how to do	An explanation of how to do [something] on Wikipedia.
3. Suggestion/Call for Action	A suggestion for action posted by a Wikipedian; suggestions for example for rephrasing, expanding, deletion, reference, or adding information.
4. Personal Opinion	A personal opinion posted by a Wikipedian.
5. Evaluation of quality of contributions	Provide evaluative comments regarding the quality of contributions
6. Voting	
6. 1 Announcement of voting on an article	Announcement about voting posted by a Wikipedian
7. Request	

7.1. Request to become	A request posted on a user talk page by a Wikipedian

administrator	asking the user to become an administrator
7.2.Request to join a project	– A request posted on a user talk page by a Wikipedian asking the user to join a project.
7.3. Request to vote on or monitor items for deletion	A request posted on a user talk page by a Wikipedian asking the user to vote on or monitor his/her contributions that are marked for deletion.
7.4. Request to intervene in a conflict	A request posted on a user talk page by a Wikipedian asking the user to intervene in a conflict or participate in mediation.
7.5 Request to Block/ Unblock page	A request posted on a user talk page by a Wikipedian asking the user to Block/ Unblock a page (article)

Hara, N., Shachaf, P., & Hew, K. (2010). Cross cultural analysis of the Wikipedia community. *Journal of the American Society of Information Science and Technology, 61*(10), 2097-2108.

Cross-cultural analysis of the Wikipedia community
Noriko Hara

School of Library & Information Science, Indiana University
nhara@indiana.edu

Pnina Shachaf
School of Library & Information Science, Indiana University shachaf@indiana.edu

Khe Foon Hew
National Institute of Education Nanyang Technological University Email: khefoon.hew@nie.edu.sg