Cultural Norms and Interpersonal Relationships: Comparing Disclosure Behaviors on Twitter

Anju Punuru
Indian Institute of Technology, Kharagpur
Kharagpur, India
anju.punuru@iitkgp.ac.in

Tyng-Wen Cheng
Brigham Young University
Utah, USA
tschen01@byu.edu

Isha Ghosh
Rutgers University
New Jersey, USA
isha.ghosh@rutgers.edu

Xinru Page
Brigham Young University
Utah, USA
xinru@cs.byu.edu

Mainack Mondal
Indian Institute of Technology, Kharagpur
Kharagpur, India
mainack@cse.iitkgp.ac.in

ABSTRACT

This study performs an initial exploration of cultural differences in social media disclosure behaviors. We focus on what U.S. and India users disclose about interpersonal relationships on Twitter, a popular social networking platform that has gained enormous traction outside the U.S.

We developed a taxonomy of words representing interpersonal relationships and then collected all tweets containing these words (~4.5 million tweets) uploaded from India and the U.S. over a one-month period of time. We found that Indian tweets about others tend to be statistically significantly more positive and uncover differences in how they tweet about various relationships (family, friends, others) in comparison to U.S. users. Drawing on theories of collectivism and individualism, we discuss how different cultural attitudes may explain these behaviors. We present implications for research and for designing to support cultural norms.

INTRODUCTION

In recent years, Twitter has seen a substantial increase in users from different countries [6]; Now only 21% of its users reside in the United States. Research shows that Twitter is used not only as a broadcast medium, but also commonly as a personal communication channel [10, 11]. Prior studies have by and large focused on U.S.-based users [10, 11] for understanding usage attitudes and behaviors, leaving out a large portion of Twitter users. Culture-based differences in privacy attitudes have been uncovered and shown to impact behavior in prior research [9, 14], suggesting that culturally-based differences may also impact how people prefer to use social media. Understanding these preferences can help us better support different users. This study is an initial exploration of these differences and takes a data-driven approach to analyze disclosure behavior surrounding interpersonal relationships. Specifically, our study addresses the following research question: ***How do social media users from India and the U.S. differ in their online disclosures about interpersonal relationships?***

Prior research has identified cultural norms as a key mechanism in shaping the creation and maintenance of interpersonal relationships. Most relevant for our study, culture has been shown to influence the content and style of communication between people in different social contexts [12]. Researchers investigating the link between culture and interpersonal relationships have used Fiske's (1999) Relational Model Theory (RMT) [7, 8]. According to RMT, social relationships can be classified as friends, family, and organizations [7]. Table 1 provides a detailed description and most frequent words for each category included in the dataset. We use the RMT classification to investigate cultural differences between India and U.S. users in how they tweet about their social relationships on Twitter. The main contributions of this work are: (i) Creating a systematic approach to identify posts where social media users mention interpersonal relationships; (ii) Uncovering cultural differences in relationship disclosure behavior of India and U.S. users.

METHOD

Creating a Saturated Lexicon of Interpersonal Relationships: The first goal of this project was to create a saturated list of words representing social relationships in two different cultures (U.S. and India). Both these countries vary widely in terms of cultural norms; most significantly, India has been characterized as a collectivist society and the U.S. as an individualistic culture [9]. First, three authors created a lexicon of English words representing interpersonal relationships (e.g., mother, boss, colleague). For each word, the first author searched for related words from multiple sites (https://relatedwords.org/, https://www.thesaurus.com/, https://www.dictionary.com/, https://www.urbandictionary.com/) and added to the list any new words generated from the search. This process was repeated until no new words were found from the related word search. To further saturate this list, we asked a native English speaker to read through this list of relationship words and

Relationship	Definition	Frequent words
Family	People with whom the individual feels kinship and has a specific established tie or relationship.	family, kid, mom
Friend	People with whom the individual feels kinship but is not related to by blood or social ties.	love, friend, buddy
Organization	Membership within any group with a common purpose is treated as an organization.	teacher, boss, manager
Mixed	This category is used in cases where a term might be used to refer to people either from Family or Friend.	bro, baby, brother

Table 1. Description and examples of each relationship category.

	U.S.	India
Family	899,517	115,955
Friend	639,431	89,349
Organization	135,335	49,684
Mixed	261,371	33,762

Table 2. Number of tweets falling into each category.

add any synonyms as well as additional new words. Finally, all authors crosschecked the list to verify no new relationship words could be added.

Data Collection: Tweets were collected during January 1 – 31, 2020 using a Twitter data collection tool called Twint [13]. We set geographic perimeters to specifically collect tweets from within India and the U.S. This initial data collection leveraged our previously developed lexicon and resulted in ~4.5 million tweets from which we excluded repeated tweets and removed urls, user mentions (e.g., @cscw), and punctuations, then converted each tweet to lower case. Furthermore, we kept only English words in each tweet for our analysis. After the data cleaning process, we had a total of 2,095,792 tweets from 604,895 unique users in India and the U.S. There were 1,823,037 (86.99%) tweets from U.S. and 272,755 (13.01%) tweets from India (Table 2 shows the total number of tweets collected in each category). We randomly sampled ten tweets from U.S. for each relationship word and manually verified that for majority of the relationship words (more than 57%) at least 70% of these tweets are indeed about relationships (our main findings remain unchanged even when we considered only those words where at least 70% of tweets are about relationships). Tweets were then divided into relationship categories based on the RMT theory. The categorization of interpersonal relationships into friends, family, and organization has previously been used in sociological research investigating relationship groups [4, 7, 8]. While following this categorization, we realized that some words (e.g., "bro") could be used to refer to both family and friend relationships (e.g., "bro" could be a friend or a family member) and so these words were categorized as Mixed. Table 1 shows the most frequent words classified into each relationship. For the full list of relationship words see: https://drive.google.com/file/d/101SeXf8RhEgznR8Ulr43y6NQ4g_WEy2r/view?usp=sharing

RESULTS AND DISCUSSION

To investigate the differences in Twitter disclosure behaviors of India and U.S. users, we leveraged sentiment analysis. We used iFeel2.0 [2], a tool that applies 18 sentiment detection algorithms to a given tweet and outputs 18 sentiment scores. We assigned the average of these scores as a tweet's sentiment score; sentiment scores varied between -1 (negative sentiment) and +1 (positive sentiment). For our analysis, we used the average sentiment score of tweets for each relationship group in a given country and performed pairwise Mann-Whitney tests between all pairs of inter- (U.S. vs. India for a given relationship category) and intra- (e.g., family vs. friend in the same country) cultural relationship groups, applying Bonferroni correction for multiple comparisons. The p-value was significant at $p < 0.001$ level for all of our pairwise comparisons. Table 3 shows the average sentiment scores (i.e., the average positivity of tweets) as well as standard deviation of these scores for each of the India and U.S. relationship categories. Furthermore, in Table 4 we focus on the inter-cultural differences in sentiment scores for tweets from different categories. This result shows that, in each relationship category, tweets from India are more positive than tweets from the U.S. These results have implications for both theory and practice.

Category	Sentiment Mean	Sentiment std. dev.
India – Family	0.16	0.36
India – Friend	0.48	0.33
India – Organization	0.22	0.32
India – Mixed	0.24	0.34
U.S. – Family	0.10	0.37
U.S. – Friend	0.44	0.35
U.S. – Organization	0.17	0.36
U.S. – Mixed	0.14	0.37

Table 3. Sentiment score (and standard deviation) for each category. Higher number is more positive sentiment.

	U.S. sentiment	India sentiment
Family	0.10	0.16
Friend	0.44	0.48
Organization	0.17	0.22
Mixed	0.14	0.24

Table 4. Sentiment score differences in each category between U.S. and Indian tweets. All differences are significant at **p < 0.001***** level.

First, when examining the expression of different relationships online, it is important to consider the diverse priorities of individualistic and collectivist cultures. Research has previously shown that the culture in which an individual is raised and socialized has a direct effect on their communication practices [1, 14]. For instance, members of individualistic cultures prioritize the major values of their culture (e.g., independence, achievement) and therefore tailor their communications to present themselves as unique individuals, standing out from the herd. Members of individualistic society may therefore be more concerned with how they themselves are perceived by their audience and be less reserved in their online communications as they attempt to show authenticity.

Members of collectivistic cultures, on the other hand, are taught to prioritize different values (e.g., harmony, solidarity) and think of themselves as interconnected with the larger society. For these members, presentation of self is linked with the perception of the overall communities. These members might, therefore, choose to be more diplomatic in their choice of language or tweet construction as they would be concerned with how their overall family or community is being perceived based on their tweets. Our results show that these differences can play a significant role in how they tailor their online communications.

From a design perspective, our results uncover an opportunity for designers to integrate culture-specific norms into the design of social network sites. For instance, members from collectivist cultures might appreciate a feature showing the overall positive versus negative score of their tweet before tweeting. This feature would allow the user to consider if their tweet might result in them "loosing face" in their community and therefore help users better manage their online communications. Another feature could be to detect the overall sentiment score for a user's tweets for each relationship type and to personalize the sentiment of auto-word suggestions, or to prompt the user and warn when they seem to be writing an unusually negative tweet for that relationship type. Yet another possible feature is to use the relationship lexicon developed in this work to classify tweets into relationship categories. Users could "pin" tweets for relationship categories where it is important to keep informed and respond immediately, avoiding regret.

CONCLUSION AND FUTURE DIRECTIONS
By examining differences in online disclosures of Twitter users from India and the U.S., we show that the cultural norms of a user can have a significant influence on their online communication behaviors. We also developed a lexicon of interpersonal relationship words that can be used for future research. While this study focused only on tweets written in the English language, we have followed a similar process to develop a taxonomy for Hindi and Telugu and will conduct a subsequent study collecting tweets written in these languages. We will also expand this research to examine tweets from other individualist and collectivist cultures to determine if our findings are generalizable. As we have become an increasingly global and connected society, cross-cultural differences in communication will only become more important to understand and support.

REFERENCES

[1] Norah Abokhodair & Sarah Vieweg. (2016). Privacy & Social Media in the Context of the Arab Gulf. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (pp. 672–683).

[2] Matheus Lima Diniz Araújo, João Paulo Diniz, Lucas Bastos, Elias Soares, Manoel Junior, Miller Ferreira, Filipe Nunes Ribeiro, & Fabrício Benevenuto (2016). iFeel 2.0: A Multilingual Benchmarking System for Sentence-Level Sentiment Analysis. In *Proceedings of the Tenth International Conference on Web and Social Media* (pp. 758-759)

[3] Nancy K. Baym (2015). *Personal connections in the digital age*. John Wiley & Sons.

[4] Rodrigo Brito, Sven Waldzus, Maciej Sekerdej, & Thomas Schubert (2011). The contexts and structures of relating to others: How memberships in different types of groups shape the construction of interpersonal relationships. *Journal of social and personal relationships*, 28(3), 406-432.

[5] Valeria Castillo (2019). Which Languages Are Spoken In India? https://www.babbel.com/en/magazine/what-languages-are-spoken-in-india

[6] J. Clement (2020). Leading countries based on number of Twitter users as of April 2020, https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/.

[7] Alan Page Fiske (1991). *Structures of social life: The four elementary forms of human relations*. New York: Free Press.

[8] Askan F. Koerner (2006). Models of relating-not relationship models: Cognitive representations of relating across interpersonal relationship domains. *Journal of Social and Personal Relationships*, 23(4), 629-653.

[9] Ponnurangam Kumaraguru & Lorrie Faith Cranor (2005). Privacy in India: Attitudes and awareness. In *International workshop on privacy enhancing technologies* (pp. 243-258). Springer, Berlin, Heidelberg.

[10] Alice E. Marwick, & danah boyd, (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, 13(1), 114-133.

[11] Mor Naaman, Jeffrey Boase, & Chi-Hui Lai, (2010). Is it really about me? Message content in social awareness streams. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 189-192).

[12] L.A. Samovar, R.E. Porter, & N.C. Jain, (1981). *Understanding intercultural communication*. Wadsworth Publishing Company.

[13] TWINT - Twitter Intelligence Tool: https://github.com/twintproject/twint

[14] Blase Ur and Yang Wang. 2013. A cross-cultural framework for protecting user privacy in online social media. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 755–762).