



Kent Academic Repository

Litvak, Marina, Otterbacher, Jahna, Ang, Chee Siang and Atkins, David (2016) *Social and Linguistic Behavior and its Correlation to Trait Empathy*. In: *Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*. Association for Computational Linguistics, pp. 128-137. ISBN 978-4-87974-723-5.

Downloaded from

<https://kar.kent.ac.uk/58162/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://peoples2016.github.io/accepted.html>

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Social and Linguistic Behavior and its Correlation to Trait Empathy

Marina Litvak

Department of Software Engineering
Shamoon College of Engineering
Be'er Sheva, ISRAEL
marinal@sce.ac.il

Chee Siang Ang

School of Multimedia and Digital Arts
University of Kent
Kent, UK
C.S.Ang@kent.ac.uk

Jahna Otterbacher

Social Information Systems
Open University of Cyprus
Nicosia, CYPRUS
jahna.otterbacher@ouc.ac.cy

David Atkins

School of Multimedia and Digital Arts
University of Kent
Kent, UK
D.Atkins@kent.ac.uk

Abstract

A growing body of research exploits social media behaviors to gauge psychological characteristics, though *trait empathy* has received little attention. Because of its intimate link to the ability to relate to others, our research aims to predict participants' levels of empathy, given their textual and friending behaviors on Facebook. Using Poisson regression, we compared the variance explained in Davis' *Interpersonal Reactivity Index* (IRI) scores on four constructs (empathic concern, personal distress, fantasy, perspective taking), by two classes of variables: 1) post *content* and 2) *linguistic style*. Our study lays the groundwork for a greater understanding of empathy's role in facilitating interactions on social media.

1 Introduction

Empathy is an important component of social cognition that contributes to one's ability to understand and respond to the emotions of others, to succeed in emotional communication, and to promote pro-social behavior (Spreng, 2009). We explore the correlations between participants' levels of the various types of trait empathy, and their digital traces at Facebook, representing social media activities. To date, empathy has received little attention from social media and human factors researchers. Some work has been done toward understanding "empathic design" of online support communities (Brennan, Moore, & Smyth, 1991), (Tetzlaff, 1997), (Brennan & Ripich, 1994). However, surprisingly, empathy in social media in the context of day-to-day conversations or messaging has not been well studied.

In this work, we conceptualize empathy as a trait, operationalizing it within the context of our study. In the next subsections, we highlight the intimate relationship between empathy, communication, and friendship patterns, and present hypotheses to be tested. Finally, we explain why users' writing patterns are expected to provide a source of information with respect to their underlying levels of empathy, detailing our hypotheses of interest. We test these hypotheses by fitting the Poisson regression model with each IRI score as the outcome variable and a set of explanatory variables suitable for each hypothesis.

1.1 Davis' IRI

Davis' IRI (Davis, 1983) is a measure of trait empathy that considers a set of four distinct but related constructs. Each of the four subscales of the IRI--empathic concern (EC), fantasy subscale (FS), perspective taking (PT) and personal distress (PD)--was assessed with seven items on a five-point Likert

scale (1 = *does not describe me well* to 5 = *describes me very well*). The subscales that pertain to cognitive dimensions of empathy are the FS and the PT subscale. They measure the tendency to get caught up in fictional stories and imagine oneself in the same situations as fictional characters, and the tendency to take the psychological point of view of others, respectively. The EC and PD subscales measure the affective dimensions of empathy. Specifically, the EC measures sympathy and concern for others and is typically considered as an other-oriented emotional response in which attention is directed to the person in distress (Schroeder, et al, 1988). The PD scale considers a self-oriented emotional response in which attention is directed at one's negative emotions of distress and the reduction of these negative emotions. The IRI has demonstrated good intra-scale and test-retest reliability, and convergent validity is indicated by correlations with other established empathy scales (Davis, 1983).

1.2 Empathy in Social Media

Scholars such as Rogers (2003) have noted the abilities of highly empathic individuals to influence the opinions of others. This could also be the case in the context of social media, although it is unclear whether or not measures such as an individual's network size and frequency and types of activities could reflect this. On Facebook in particular, establishing a friendship is a mutual decision, meaning that both sides must confirm it in order to be connected. Intuitively, we can say that the creation of a "friendship" on Facebook is an indication that individuals are open to sharing with others. A previous study performed on a large and diverse dataset of Facebook participants in Bachrach et al. (2012) found significant relationships between their personality traits and the size and density of their friendship network, and their activity online. Kang and Lerman (2015) explored user effort and content diversity in social networks, with a commentary on cognitive constraints in social network activity. Given that little work on empathy in the social media literature, we found it necessary to establish the baseline relationship (if any) between network size, activity, and trait empathy.

H1a: The size of one's friendship network is correlated to her levels of empathy.

H1b: A user's level of activity (i.e., amount of written text) is correlated to her levels of empathy.

While level of activity depends on one's network size, there is also reason to believe that in some cases, people with smaller networks may engage actively with their close friends, thus producing higher levels of activity than those with a larger network. Therefore, we also examine the relationship between level of activity and empathy.

1.3 Empathy and writing patterns

The widespread use of writing therapy by psychologists (Pennebaker, 1997) confirms the tight relationship between writing characteristics and aspects of the self, such as empathy. We believe that the level of one's empathy influences one's writing. Therefore, we analyzed written content (in the form of posts and comments) in order to distinguish between users with different levels of empathy. Previous research (Pennebaker and King, 1999), (Mairesse and Walker, 2008) concluded that linguistic style is an independent and meaningful way of exploring personality and there is a strong correlation between language dimensions, measured by the Linguistic Inquiry and Word Count (LIWC), and personality factors. Our analysis focuses on LIWC Psychological (i.e., content) and Linguistic (i.e., style) measures (see Table 2), to test the following hypotheses:

H2a: Users whose social media texts express more socially oriented content are more empathic.

H2b: Users whose social media texts exhibit linguistic styles that engage others are more empathic.

2 Methodology

2.1 Data collection and preprocessing

We developed a Facebook application (“app”) in order to carry out the following phases of our data collection: capturing each participant’s digital traces during the previous 30 months, and administering a standardized test that measures different types of trait empathy. We analyzed participants’ levels of trait empathy using Davis’ IRI.¹

The app captured participants’ profile (upon their agreement), including the full list of their Facebook friends. In addition, the app tracked users’ recent social activities: with whom and how frequently they interacted through “likes,” “shares,” and “comments” to others’ posts. Then, the participant was prompted to complete the IRI. The participant could also invite friends to complete the survey and become new participants. As such, we employed an opportunistic sampling and snowballing method. Table 1 describes attributes that were collected while Figure 1 depicts the app flowchart.

In order to describe participants’ language behaviors, we considered all textual communication (i.e., posts on one’s own Facebook wall and comments left on the walls of others) that occurred during the previous 30 months. We used LIWC, which analyzes a text by counting word occurrences in psychologically meaningful categories such as negative versus positive emotion, or social versus cognitive processes (Pennebaker, Francis, & Booth, 2001). At the same time, LIWC computes attributes of linguistic style (e.g., the use of punctuation, the extent to which first-, second-, and third-person pronouns are used), otherwise known as stylometric features (Brizan, et al, 2015). Each participant’s set of texts was processed using LIWC, in order to obtain scores on six psychological (i.e., content) measures, described in Table 2, and ten linguistic (i.e., style) measures. The style features included the participant’s total word count and the mean number of words per post. Finally, we considered the proportion of words used belonging to each of the following categories: pronouns, verbs, adverbs, auxiliary (i.e., “helping”) verbs, quantifiers, numbers, swearing, and punctuation.

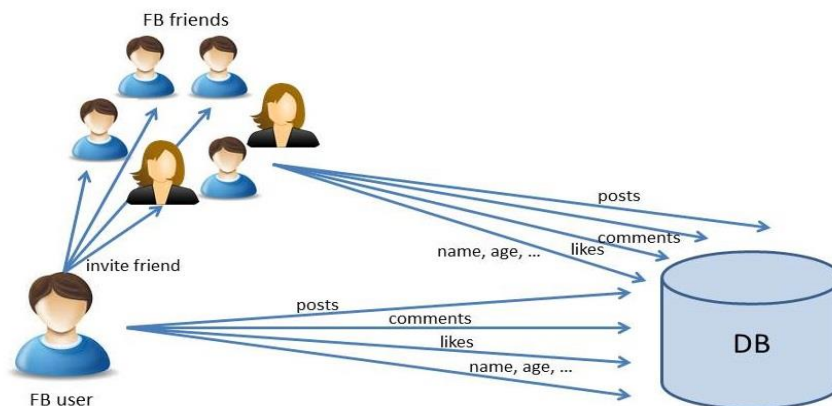


Figure 1: Data collection and system flowchart.

Table 1: Data collected via Facebook application.

Type	Attributes
Profile attributes	Location, Gender, Age
Analyzed profile attributes	Number of friends, Number of likes received from friends, Number of words of the comments received from friends
Trait Empathy Results	IRI Scores: Empathic Concern (EC), Fantasy Scale (FS), Perspective Taking (PT), Personal Distress (PD)

¹ The study was approved by the University of Kent’s research ethics committee.

Table 2: LIWC Categories used to process participants' textual communications.

	LIWC category	Explanation	Key words (examples)
Psychological Processes / Content	Social processes	Communication related to family, friends, people	Daughter, husband, friend, neighbor, baby, boy, talk
	Affective processes	Positive or negative emotions, anger, sadness, anxiety, joy excitement	Love, sweet, happy, cried, ugly, nasty, hate, kill, annoy
	Cognitive mechanisms	Communication related to thought and reasoning	Think, know, consider, cause, should, would, guess
	Perceptual processes	Language describing observations and senses	Hear, feel, view, see, touch, listen
	Biological processes	Communication describing bodily functions	Eat, blood, pain, hands, spit, clinic, love, eat
	Relativity	Language describing motion, space, time	Area, bend, exit, arrive, go, down

2.1.1 Participants

A total of 334 Facebook users participated in the study. In the current analysis, we considered only the users who posted in English, such that their traces could be analyzed via LIWC, and who completed the IRI. We also restricted the dataset to include only individuals whose profiles indicated that they were 65 years old or younger. This was to ensure the integrity of the data. We did not filter short posts and did not distinguish between users using few words and ones using many words.

After applying the aforesaid restrictions, a total of 202 complete profiles were available for the analysis. Of these, 167 participants (82.7%) were female, with mean and median ages of 39.3 and 36.0 years, respectively. This gender imbalance can likely be attributed to the manner by which we incentivized participation. It is well established that there are gender-based differences with respect to empathy. Specifically, women reportedly score higher than men on all four subscales of the IRI (Davis, 1980). Therefore, in our analyses, we included gender as a control variable.

As expected, the distributions of the total number of friends as well as two measures of attention received from others (the number of likes received and the number of words commented on users' posts) were skewed to the right. The mean and median numbers of friends among participants were 304.1 and 238.5, respectively, while the mean and median numbers of likes per post were 16.3 and 13.0, respectively. Participants received a mean of 636, and a median of 257 words, in the comments posted by their friends.

2.1.2 Trait empathy

We considered our participants' scores on the four IRI scales by gender, given that previous studies report salient gender differences. The non-parametric Wilcoxon test reveals that, compared to male participants, females score significantly higher on measures of empathic concern and fantasy. However, no significant gender differences were revealed with respect to perspective-taking and personal distress. These gender differences are somewhat in line with previous research that has reported greater trait empathy overall (i.e., all IRI subscales) among women (Davis, 1980). As mentioned, we retain gender as a control variable in our regression analyses.

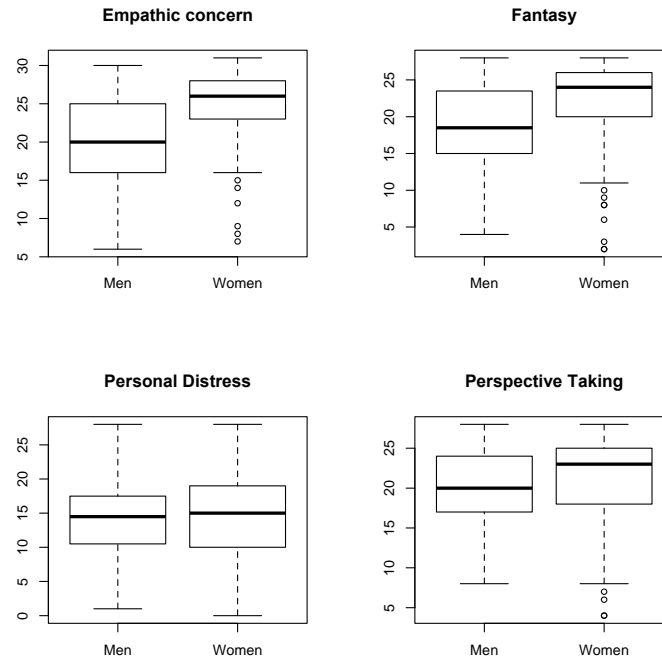


Figure 2: Distribution of IRI scores.

As shown in Figure 2, the median IRI scores for participants is as follows (men / women): EC (20 / 26), FS (18.5 / 24), PD (14.5 and 15), and PT (20 / 23).

2.2 Data analysis

In order to examine the relationship between users' Facebook behaviors and their trait empathy, we used Poisson regression models. Specifically, for each of the four IRI scores, we fit four models, in order to explore the explanatory power of four sets of variables:

- **Control:** Participant gender and age only;
- **Model 1:** The *content* of users' posts, namely, psychological processes exhibited in the text (social, affective, cognitive mechanisms, perception, biological, relativity);
- **Model 2:** The *linguistic style* of posts, namely, linguistic characteristics (total word count, words per post, pronouns, verbs, adverbs, auxiliary verbs, quantifiers, numbers, swearing, punctuation);
- **Model 3:** Measures of users' friendship network (namely, number of total friends and likes).

2.2.1 Poisson regression model

The Poisson regression model is a type of Generalized Linear Model (GLM). Such models use the logarithm link function in order to correlate the model predictors (explanatory variables) to the outcome variable, which is an expected frequency (incidence). In our case, the outcome variable is the IRI score, which ranges from 0 to 28 for each of the four subscales. The estimation of the Poisson models was conducted using the R statistical computing package². For each of the models, we estimate the parameter and statistical significance of each explanatory variable. In addition, we gauged the degree to which the variance in the dependent variable (i.e., the level of empathy as measured by the relevant IRI score) is explained by the set of explanatory variables. To this end, we use Mittlböck's adjusted R^2 , which is appropriate for evaluating Poisson regression models (Mittlböck, 2002).

² <http://cran.r-project.org>

3 Results and discussion

Table 3 shows that gender is correlated to three subscales of empathy. Specifically, female participants display higher levels of empathic concern, fantasy, and personal distress. Age is negatively correlated to the level of personal distress. In all three cases, the control variables do not explain a good deal of variance in the IRI scores. In the case of empathic concern, gender alone explains 13% of the variability (i.e., the R^2 of the model with gender as the only explanatory variable is 0.13).

Table 3: Model with control variables.

	Empathic concern	Fantasy	Perspective taking	Personal distress
Intercept	2.9475***	3.0311***	2.9163***	3.0239***
Gender	0.1884**	0.1864***	0.05642	0.1364**
Age	0.002893*	-0.001945	0.003057	-0.01181***
R^2	0.1573	0.05931	0.03023	0.08097

***p-value < .001; *p-value < 0.1

Focusing on the model that includes the content of users' posts, we see an improvement in the explanatory power of our Poisson model for each of the four subscales of empathy (over the control model), as can be seen from Table 4. However, the most significant improvements are for EC and PT. Gender and social psychological processes in users' text account for just over 20% of the variance in EC score. In particular, female participants and those whose posts are more social (i.e., make references to people, friends and family) tend to score higher on the empathy scale of the IRI. This is expected, given that a user's attentional focus when using social processes within their text is likely to be oriented to others, which is encapsulated in the other-oriented empathic concern subscale. Likewise, the social and perception processes are significant correlates of the PT score. PT is a measure of the dispositional ability to consider the perspective of others.

Table 4: Model 1 - The content of users' posts.

	Empathic concern	Fantasy	Perspective taking	Personal distress
Intercept	2.8755***	2.9714***	2.8797***	3.0160***
Gender	0.1513***	0.1614***	0.02087	0.1375**
Age	0.001361	-0.002870	0.001579	-0.01210***
Social	0.008304*	0.003043	0.007709*	-0.006554
Affect	-0.001940	0.0009739	-0.0007692	0.004469
Cogmech	0.0008126	0.0008076	-0.001549	-0.002789
Percept	0.01296	0.001750	0.03029**	-0.01173
Bio	-0.002502	0.02334*	-0.01211	0.008979
Relativ	0.002555	-0.001853	0.001692	0.007410
R^2	0.2051	0.08630	0.08241	0.09080

***p-value < .001; **p-value < .01; *p-value < 0.05

Thus, using language relating to social processes would likely enable users to take the perspective of others; they need information about their communication partner to take their perspective. Unsurprisingly, language referring to perceptual processes significantly correlates to trait PT. In order to take the perspective of another (i.e., to understand the issues, thoughts, and feelings of others) one needs to use perceptual processes. That said, the effect sizes are rather small to draw any definitive conclusions regarding the *H2a*.

Table 5 shows that for all four subscales of empathy, adding the stylistic characteristics of users' texts increases the proportion of variance explained. Interestingly, the use of auxiliary verbs significantly correlates to both greater cognitive and affective empathy (i.e., all IRI subscales), while pronouns significantly correlate to FS and PT. The fact that the number of words per post is significantly correlated to greater PD could refer to a need to express oneself (e.g., an opinion, complaint or need). However,

the fact that the word counts of the participant's posts or comments are not correlated to other empathy measures approves that, in general, *H1b* is not supported.

A possible explanation can be found in Bachrach et al (2012)'s work, where all Big Five personality traits have demonstrated correlation to the activity level of users in Facebook, expressed by number of likes, uploaded photos, statuses, and more. It was found that extroverts are more likely to reach out and interact with other people on Facebook. Given that word counts of posts or comments are also self-generated content, it is arguably possible that a greater volume of written words is reflective of a more extroverted personality trait. However, there is no direct correlation between empathy and extraversion trait (Magalhães et al., 2012).

Table 5: Model 2 - The linguistic style of posts.

	Empathic concern	Fantasy	Perspective taking	Personal distress
Intercept	2.962***	2.969***	2.867***	2.990***
Gender	0.1291**	0.1355**	0.005643	0.1482**
Age	0.002012	-0.002920*	0.002114	-0.01200***
Word count	-0.00003474	0.00004009	0.000007792	-0.00009535
Words per post	0.0005015	-0.000005146	-0.00002968	0.002625**
Pronouns	0.004271	0.007317*	0.006648*	-0.005947*
Verbs	-0.007095	-0.01249*	-0.002170	-0.003509
Adverbs	-0.008531	0.007475	-0.002882	-0.001027
Aux verbs	0.01662*	0.01543*	0.01538*	0.02764**
Quantifiers	0.002737	-0.0007620	-0.03161**	-0.04031*
Numbers	0.002383	0.02999*	0.006492	0.05199*
Swearing	-0.008739	-0.008929	-0.04189	0.01032
Punctuation	-0.0006776*	-0.0003195	0.000008210	-0.00004612
R ²	0.2540	0.1146	0.1009	0.1376

***p-value < .001; **p-value < .01; *p-value < 0.1

We observed a number of correlations between linguistic styles and empathy measures. Importantly, the use of pronouns is positively correlated with PT and FS, but negatively correlated with PD. The regular use of pronouns might indicate that a user is switching perspectives frequently within a session, which would in turn exercise perspective taking skills. Although communication partners are real, as opposed to a character in a novel, there is still a barrier between the user and his or her communication partner because they are not physically face-to-face. In a sense, this type of communication is surreal and may require some fantasy. Interestingly, the use of auxiliary verbs such as am, will, or have, is positively correlated with all empathy measures. Auxiliary verbs add functional meaning to the clause in which the auxiliary verb appears and thus can express tense and emphasis among other meanings. Therefore, these words function to create a more vivid sense of an action, which consequently would exercise more mental imagery. The use of our mental imagination capacities is inherent in the FS and PT subscales.

Further, greater words per post significantly correlating to greater personal distress could refer to a need to express oneself. Feelings of personal distress are uncomfortable and someone who is distressed has a reason that has evoked negative feelings in the first place. In the digital domain, one way of alleviating the distress would be to write about one's feelings as a cathartic exercise, or perhaps to express their point if in an argument or debate; both would likely require more words to achieve.

In sum, it is safe to say that users with varying levels of empathy do exhibit different linguistic styles when communicating in social media. Therefore, the results of our analysis support *H2b*.

The number of available cases to fit model 3 was reduced to 169 participants (from 202), because some of the data regarding friends of participants could not be collected (most likely because it was protected by the respective Facebook users).

The total number of likes and comments received on posts, as well as number of friends were used in order to examine the relationship between these measures and empathy. As can be seen in Table 6, the number of likes on one's posts and total number of friends were not correlated to any of the four types of empathy. The volume of comments (measured as the total number of words of the comments) on one's posts is weakly correlated to PD, although the direction of the relationship is negative. Therefore, we can conclude that *H1a* is not supported; network size alone is not a clear signal of an empathic personality.

A possible explanation of this finding might be that participants use Facebook to manage a large number of "weak ties" (people from different social circles) while still maintaining closer relationship with a smaller number of friends (see (Marsden 1987) and (Putnam 2001) for details). However, without quantifying the nature (weak vs strong) of each Facebook friend, we cannot test this explanation. Still another possibility is the link between network size and narcissism, which is negatively correlated to empathy (see (Mehdizadeh, 2013) and (Buffardi and Campbell, 2008) for more explanations).

Table 6: Model 3 - Measures of user activity and interactions with friends.

	Empathic concern	Fantasy	Perspective taking	Personal distress
Intercept	2.9638***	3.0123***	2.9443***	2.9117***
Gender	0.1953***	0.1792***	0.06689	0.08796*
Age	0.001214	-0.001493	0.001394	-0.006859***
Friends	-0.003352	-0.002132	0.002311	-0.014941
Likes	0.007440	-0.009567	-0.004216	0.01942
R ²	0.1375	0.05622	0.01861	0.05250

***p-value < .001; ** p-value < 0.01; *p-value < 0.05

4 Conclusions and future work

Given the unprecedented scale of human connectivity realized through social media, with unforeseeable consequences on a global scale, it is timely to study the relationship of online interactions with such an important human characteristic as empathy. In this paper, we explored correlations between multiple behavioral cues on social media and empathy. We considered a snapshot of a user's Facebook data, collected over a given time interval, to understand how different behavioral cues correlate to the user's levels of empathy. In other words, we explore how other Facebook users might form impressions about someone's level of empathy based on his or her behavior. The main focus and novelty of our study was to explore whether the writing characteristics can describe the user in terms of empathy.

We learned that the relationship between participants' social media behaviors, friendships and interactions with others, and their levels of trait empathy is rather complex. While we began with hypotheses grounded in previous literature, we observed some unexpected correlations. In particular, it appears to be the case that not all interactions are equal; it is likely that simple traces of interaction such as "likes" and "commenting" may tell us different things about an individual's willingness and ability to engage others. Future work could probe deeper in order to understand how and why users exhibiting relatively high and low levels of empathy engage "the other".

We also generated some ideas for future work, including experimenting with more targeted linguistic features (such as modal and hypothetical verbs); using syntactic structure for a more complex measurement of style; building a cross-validated predictive model; analysis of other traits, e.g. narcissism, big five (BF) personality (and considering empathy as an aspect of agreeableness from the BF personality traits); distinguishing between friendship, acquaintances, and incidental/semi-random FB connections in our model; considering how empathy relates to the care/harm dimension of moral foundations theory; and exploring trolling (Buckels et al., 2014) as an opposite of empathy.

In summary, this paper has highlighted a few interesting research directions: the relationship between social media activities, communication patterns, and the human characteristic of empathy. Future work must focus on recruiting a larger sample of participants in order to obtain a more balanced representation of different cultural groups as well as gender representation. In addition, the study can be extended to inter-group interactions based on social classes, religions, nationality, and so on. An in depth understanding of inter-group interaction online and its relationship to empathy is an important direction of research, and would potentially provide insights to those who design social technology that would facilitate positive intergroup interactions, thus creating a more empathic online environment.

Acknowledgements

The authors would like to thank the efforts of the three anonymous reviewers in providing thorough feedback on this work. In particular, we would like to acknowledge their suggestions for several interesting future directions, which are mentioned in the conclusion of the paper.

References

- Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., and Stillwell, D. (2012). Personality and Patterns of Facebook Usage. In Proceedings of ACM Web Sciences 2012.
- Brennan, P. F., Moore, S. M., & Smyth, K. A. (1991). ComputerLink: Electronic support for the home caregiver. *Advances in Nursing Science* , 13 (4), 14-27.
- Brennan, P. & Ripich, S. (1994). Use of a homecare computer network by persons with AIDS. *International Journal of Technology Assessment in Health Care* , 10 (2), 258-272.
- Brizan, D. G., Goodkind, A., Koch, P., Balagani, K., Phoha, V., & Rosenberg, A. (2015). Utilizing linguistically enhanced keystroke dynamics to predict typist cognition and demographics. *International Journal of Human-Computer Studies* , 82, 57-68.
- Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, 67, 97-102.
- Buffardi, L. E. & Campbell, W. K. (2008). Narcissism and social networking websites. *Personality and Social Psychology Bulletin* , 34, 1303-1324.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology* , 85.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology* , 44 (1), 113-126.
- Kang, J-H & Lerman, K. (2015). User Effort and Network Structure Mediate Access to Information in Networks. arXiv preprint arXiv:1504.01760.
- Mairesse, F. & Walker, M.A. (2008). Trainable Generation of Big-Five Personality Styles through Data-driven Parameter Estimation. In Proceedings of ACL-08: HLT, 165–173
- Marsden, P.V. (1987). Core discussion networks of Americans. *American Sociological Review*, 122–131.
- Magalhães E, Costa P, Costa M. J. (2012). Empathy of medical students and personality: evidence from the Five-Factor Model. *Med Teach*. 34(10), 807-812.
- Mehdizadeh, S. (2013). Self-Presentation 2.0: Narcissism and Self-Esteem on Facebook. *Cyberpsychology, Behavior, and Social Networking* , 13 (4), 357-364.
- Mittlböck, M. (2002). Calculating adjusted R2 measures for Poisson regression models. *Computer Methods and Programs in Biomedicine* , 68, 205-214.
- Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science* , 8 (3), 162-166.
- Pennebaker, J. W., King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6), 1296-1312.

- Pennebaker, J., Francis, M., & Booth, R. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC 2001*. Mahwah, NJ, USA: Erlbaum.
- Putnam, R.D. *Bowling alone: The collapse and revival of American community*. Simon and Schuster, New York, 2001.
- Rogers, E. (2003). *The Diffusion of Innovations*. New York: Free Press.
- Schroeder, D. A., Dovidio, J. F., Sibicky, M. E., Matthews, L. L., & Allen, J. L. (1988). Empathy and helping behavior: Egoism or altruism. *Journal of Experimental Social Psychology* , 24, 333-353.
- Spreng, R. M. (2009). The Toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment* , 91 (1), 62-71.
- Tetzlaff, L. (1997). Consumer informatics in chronic illness. *Journal of the American Medical Informatics Association* , 4 (4), 285-299.