Empathy between Neighboring Nations: Distance Matters

Avik Chakrabarti[1], Clay H. Yoo[2], and Ashiqur R. KhudaBukhsh[3]

[1] UW-Milwaukee (UWM), 3203 N Downer Ave, Milwaukee, WI 53211. chakra@uwm.edu
[2] Jona, 2299 Summer Street · Suite 1095 · Stamford, CT 06905. houseyoo@gmail.com
[3] Department of Software Engineering, Rochester Institute of Technology, 1 Lomb Memorial Dr, Rochester, NY 14623, USA. axkvse@rit.edu

Abstract. Empathy in computational linguistics has a rich literature in mental health support, disability support, and online toxicity. With the advent of large language models (LLMs), empathy in LLM responses has received considerable research attention. Since the outbreak of COVID-19, there has been a renewed interest among economists in empathy. In this paper, we meld expertise from natural language processing and econometrics to answer an intriguing research question: Is empathy, between the residents of neighboring nations, associated with distance? Via a unique dataset of geo-tagged tweets collected during India's healthcare collapse during the second wave of COVID-19, we study empathy in tweets from Pakistan. Our analyses reveal a significant positive association between empathy and distance of the geo-tagged tweets' locations from the India-Pakistan border.

Keywords: Empathy, Natural Language Processing, Econometrics, India, Pakistan, COVID-19

1 Introduction

The origin of the expression einfujhlung (feeling oneself into), later translated to empathy, can be traced back to the late 19 th century [26]. In the new millennium, since the outbreak of the severe acute respiratory syndrome coronavirus 2 (COVID-19), there has been a renewed interest among economists in empathy to the extent that it "enables people to put themselves into others' shoes" [15]. Economic models have typically relied on agents' ability to predict others'; actions: for example, the notion of backward induction implicitly assumes that each player is able to view the game from the rivals' perspectives [41]. While economists assume that altruism is mainly driven by fairness norms, psychologists consider empathy to be a key motivator for altruistic behavior and provide convincing evidence revealing that empathy induction substantially increases altruistic behavior [24]. As Robin [35] reminds us, Adam Smith was among the earliest economists and psychologists to recognize the role of empathy in human behavior: *An Inquiry into the Nature and Causes of the Wealth of Nations* opens

with pleasant stories of workers cooperating in a pin factory and quick-witted boys scrambling over steam engines, inventing labor-saving devices for pistons and boilers, so they can leave the machines unattended and rush off to play with their mates [42]." Smith's conviction, "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest" need not be construed as a mere statement of egoism or selfishness but as an "injunction to orient ourselves toward others" to the extent that "market requires us to talk to other participants not of our own necessities but of their advantages." With this backdrop, measuring empathy by a unique dataset on the average proportion of supportive tweets that originated from distinct districts of Pakistan for the population residing in India affected by the most recent pandemic [1], we pose the following research question:
RQ: *Is empathy, between the residents of neighboring nations, associated with distance?*

We meld the disciplines of natural language processing and econometrics to answer this intriguing research question. Via a unique dataset of geo-tagged tweets collected during India's healthcare collapse during the second wave of COVID-19 [49], we study empathy in tweets from Pakistan. Our analyses reveal a significant positive association between empathy and distance of the geo-tagged tweet's locations from the India-Pakistan border. Our study contributes to a rich line of economics literature that investigated the role of distance driving human behavior (see, e.g., [8,31,36])

The remainder of the paper is organized as follows. In Section 2, we provide the historical context to India-Pakistan conflict spanning decades. Next, we present a brief description of related research (Section 3). In Section 4, we describe the data used in our search for an answer to our research question, and Section 5 we present our modeling results. In Section 6, we present our empirical model and inferences. Based on our findings, we conclude in the last section.

2 Context

India and Pakistan are two nuclear-armed adversaries with a contentious shared history, marked by four major wars and several skirmishes. A major source of ongoing unrest in South-East Asia, the Kashmir issue has long captured the attention of political scientists [27,7,39]. The origin of this issue can be traced back to India's struggle for independence and the subsequent partition of the subcontinent into India and Pakistan in 1947. Overall, an estimated 27,650 soldiers were killed and thousands wounded in the four full-fledged war India and Pakistan fought with the 1971 war being the goriest (11,000 killed from both sides) which resulted in the largest number of prisoners of war (90,000 POWs) since the Second World War [1].

The 2019 conflict, triggered by the Pulwama terror attack, is the most recent edition of this ongoing bilateral tension spanning decades. The Pulwama terror

[1] https://www.bbc.com/news/world-asia-india-57683808

attack brought the two countries precariously close to declaring a full-fledged war (a recent study projects that a full-fledged conflict between these two nuclear adversaries could inflict more than 100 million deaths [44]). Although the imminent crisis was averted by the leaders of both countries, the web manifestation of this conflict received sustained attention from computational social scientists [32,45,21,18,22,14].

During the second wave of COVID-19 in India, India suffered a massive healthcare crisis with accute oxygen shortage in hospitals [2]. While a considerable chunk of social web interactions from Pakistan was supportive, negative tweets expressed politically motivated hate or schadenfreude, or demonstrated a war-mongering attitude. Hashtags such as #IndiaNeedsOxygen, #PakistanStandsWithIndia, and #(I)EndiaSaySorrytoKshmir heavily trended [49]. Yoo *et al.* [49] observed that (1) empathetic tweets considerably outnumbered negative tweets; and (2) empathetic tweets garnered more engagement (in terms of likes and shares) than negative tweets.

3 Related Work

The social media dynamics of India and Pakistan during conflicts has been studied primarily during the Pulawama 2019 terror attack. Palakodety *et al.* [32] studied web manifestation of this conflict on YouTube while Tyagi *et al.* [45] focused on Twitter (currently known as X). Along the broad line of literature on counter speech, [6] Palakodety *et al.* [32] proposed a line of research to detect peace-seeking, hostility-diffusing content, dubbed *hope speech*, to combat online toxicity during warlike scenarios. On the substantive front, both studies' findings were aligned as both studies reported that de-escalating content dominated during the time period the two countries came precariously close to declaring a full-fledged war. Follow-on research on *hope speech* focused on extending linguistic resources to low-resource languages through efficient sampling methods [21,23]. Tyagi *et al.* [45] focused on the hashtags and used label propagation to study polarization. Closest to our current line of research is the work conducted by Yoo *et al.* [49] that investigated how empathetic content was shared during the COVID-19 healthcare collapse in India. While we share the dataset with Yoo *et al.* [49], our work contrasts with all prior research studying social web interactions between India and Pakistan as our research question is novel: we study how empathy is associated with distance from the border.

In computational linguistics, there has been a sustained research focus on empathy in diverse domains that include mental health support, online toxicity [48], and disability support applications. How to operationalize empathy for improved mental health support [40] or disability support [29] to ensuring LLM generations are less toxic and more empathetic [25,34] – there are several rich lines of computational linguistics research that focus on empathy from diverse perspectives and applications. The NLP community has stimulated research on empathy through shared tasks [16] and novel benchmarks [10]. Our work considers an empathy dataset grounded in psychological and behavioral research [9,3,4,43,17,28]. Our work complements NLP literature in furthering our understanding on how empathy in conflict regions associates with the distance from the border.

To the best of our knowledge, we are the earliest to explore any role that cross-border distance per se (as distinct from social distancing), may play in affecting empathy. A plausible explanation for such a void stems from the challenges involved in the measurement of empathy. A thorough review of the extant literature reveals that several distinct yet related elements (e.g. taking others' perspectives; imagining into others' actions and/or reactions; orienting concern for others etc.) add multiple facets to the challenges involved in the process of measuring empathy [46]. As the science of measuring empathy continues to evolve, with relatively new measures aligning more closely with theoretical, methodological as well as experimental advances in understanding the various components of empathy, we take a step forward by leveraging top trending hashtags on Tweets coupled with LLM-powered natural language understanding during the most recent pandemic [49]. In comparison, there is a long tradition of using distance "as the single most important force of spatial organization in the human realm" [11]. There is mounting evidence on the role of distance in driving human behavior notwithstanding perceptions that waves of globalization should have led to the "death of distance." For instance, the magnitude of the estimated elasticity of bilateral trade with respect to distance ranges from 0.8 to 1.3: even more interesting, somewhat paradoxically, is the fact that the absolute value of this elasticity almost always increases over time [8]. In hindsight, for centuries, distance has been the "anchor point" of the dismal science for analyzing spatial-economic interactions ranging from transportation, international trade, migration, commuting, to tourism: the primary role of distance in economics stems from costs of bridging remoteness [31]. The wide range of impact that distance may have remains under radar, even more so since the information technology revolution, in light of dramatic transformations associated with electronic communication [36]. Our study is a timely contribution in this direction of research.

4 Data

4.1 Tweets

We consider the same dataset used by Yoo *et al.* [49][2]. The original data set considers 309,394 Tweets collected between 21 April 2023 and 04 May 2021 using top trending hashtags in Pakistan during the COVID-19 crisis. They include supportive hashtags (e.g., #IndiaNeed(s)Oxygen, #PakistanStand(s)WithIndia) and non-supportive hashtags (e.g., #I(E)ndiaSaySorryToKashmir).

We consider two subsets of data from the original data set. The first is the annotated data set of 3,903 tweets, where 2,273 are labeled empathetic and 1,630 not empathetic. The annotation of this dataset is grounded in psychological and behavioral research [9,3,4,43,17,28]. We use this subset to train the empathy

[2] The authors shared this dataset with us upon our request.

classifier. The second subset consists of tweets with geographical metadata. For those with longitude and latitude, we found 1,543 tweets identified as being in Pakistan. Upon preprocessing the tweets using the code the original authors shared, the second subset contained 672 unique tweets. It is worth noting the power law distribution in the number of times people posted the same Tweet. The most frequent tweet was shared by 508 users and the second most frequent tweet by 186 users.

may allah almighty keep you all safe from this pandemic your pain is our pain we are brothers by origin no borders can divide us
i am pakistani and my prayers and love are with you india in this pandemic situation due to covid 19 as humanity we all stand together
the countries were divided not human beings in this hour of need the entire pakistani nation sincerely prays for the people of india that allah almighty may deliver them from this worst time
india must end illegal occupation on territory world witnessing of india s atrocities and caged siege of innocent civilians of can learn from today cursed condition might resulted of kashmiris execration
u have kept kashmris locked up 650 days u made them hold their breath to live now god has taken revenge on innocent kashmris and stopped breathing for you look at the condition of your people and understand how much oppression you have done to kashmris
the greatest oppression in the world is inflicted on the people of the world s most oppressed subjugated nation the state of kashmir by endian infidels this is the of allah almighty in the form of

Table 1: Positive (highlighted in blue) and negative examples (highlighted in red) of tweet texts exhibiting empathy.

4.2 District Level Demographic Information

We source district level observations on literacy rate, sex ratio, as well as the proportion of urban population from 2019 from the Pakistan Statistical Year Book, an annual publication of the Pakistan Bureau of Statistics.

5 Modeling Empathy

We use the annotated dataset provided by the authors of Yoo *et al.* [49] and train multiple baselines. Yoo *et al.* [49] considered `BERT`, which was one of the most widely used and high performance models in 2021. We consider `Llama 3` [13]; `Mistral` [19]; `ModernBert` [47]; and `BERT` [12]. Table 2 summarizes our

Model	Precision	Recall	F-1
BERT	83.3	81.0	81.1
ModernBERT-base	$93.2_{0.1}$	$93.0_{0.2}$	$93.0_{0.2}$
ModernBERT-large	$93.5_{0.6}$	$93.0_{0.4}$	$93.0_{0.4}$
Llama-3.2-3B	$93.7_{0.5}$	$93.5_{0.5}$	$93.5_{0.6}$
Mistral-7B-Instruct-v0.3	$93.8_{0.9}$	$93.6_{1.1}$	$93.6_{1.1}$

Table 2: Weighted precision, recall, and F-1 score of empathy classification task. We trained with three seeds for every model except the original model (BERT) and reported the mean and standard deviation (in subscript). None of the new models are statistically different from each other.

if the this pendamic can unite the people of india and pakistan why not our leaderships are unable to promote peace friendly relation and resolve the political conflicts
devastated over the loss of lives in as hospitals are crowded amp oxygen supplies are dwindling please donate in every possible manner i am cancelling my eid shopping amp will donate to help patients in india please step forward amp donate

Table 3: Example empathetic tweets that `BERT` could not correctly classify whereas all the new models (`Llama3`, `Mistral`, and `ModernBERT`) could. `BERT` has a well-documented limitation in handling out-of-vocabulary words [30]. The word *pandemic* is misspelled in the first example and the word *and* is misspelled in the second.

results. We observe that compared to `BERT`, all the more-recent, finetuned models performed considerably better. We notice that much of these newer models' performance improvement stemmed from their robust handling of spelling and grammar disfluencies typical to South-East Asian social media discourse [38,37]. Table 3 lists few examples that `BERT` missflcassified but the newer LLMs were able to correctly classify.

Since `Llama3` and `Mistral` are generative models, it is not theoretically guaranteed to output a fixed, closed set of labels when trained via supervised fine-tuning[3]. In contrast, computing the prediction probability from `ModernBERT` (fine-tuned with a classification head) is deterministic. Given that the performance of all three models (`ModernBERT`, `Llama3`, and `Mistral`) are within the error bound, we use `ModernBERT` empathy probabilities for our analyses. Table 1 lists a few positive and negative example geo-tagged tweets from Pakistan as classified by `ModernBERT`.

[3] When we used `Llama3` or `Mistral` with a classification head, it performed considerably worse than the corresponding models trained with supervised fine-tuning (SFT).

6 Empirics

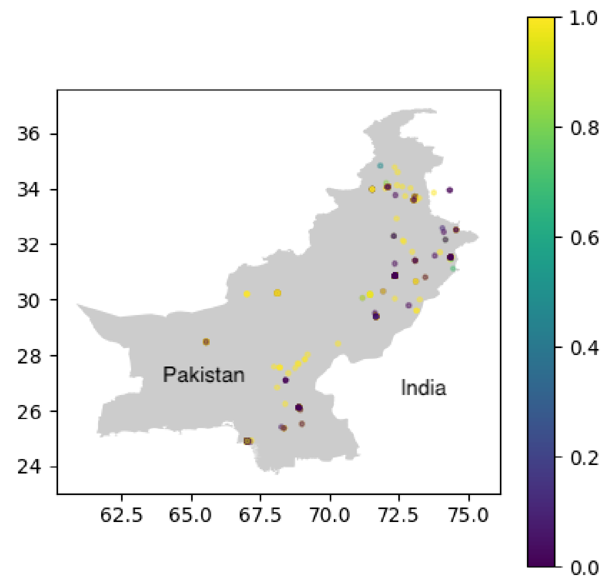


Fig. 1: Empathy: Spatial Distribution. The color bar represents model's (`ModernBERT`) predicted probability that a tweet is empathetic.

At the outset, for exposition, we place empathy on a map and visualize its spatial dispersion: the plot (Figure 1) reveals significant variation in empathy between the districts of Pakistan.

Our regression analysis is based on variants of the model:

$$empathy_i = \alpha + \beta distance_i + X_i \Gamma + \epsilon_i \tag{1}$$

where i indexes districts; our variable of interest is the Euclidean distance between a district in Pakistan, from which a tweet originated, and the nearest district in neighboring India; X_i lists each district's a) literacy rate; b) sex ratio; and c) proportion of urban population; and captures errors due to measurement and/or omission of potential explanatory variable(s). District level observations on literacy rate, sex ratio, as well as the proportion of urban population from 2019 are obtained from the Pakistan Statistical Year Book, an annual publication of the Pakistan Bureau of Statistics. Table 4 summarizes the descriptive statistics corresponding to each of the observed variables. Figure 2 presents a scatter plot and an Ordinary Least Squares (OLS) fit reflecting a positive association between empathy and distance.

Figure 3 presents the average marginal effect of the explanatory variables (literacy rate; sex ratio; and proportion of urban population) on empathy along

	Empathy	Distance	Literacy Rate	Sex Ratio	Proportion of Urban Population
Minimum	0	38.343	40.930	93.660	2.120
Maximum	1	712.803	82.450	111.130	100
Mean	0.671	261.660	59.769	104.152	33.632
Median	0.698	234.183	58.590	105.060	29.440
Standard Deviation	0.293	145.765	11.781	4.279	21.394

Table 4: Descriptive Statistics. The *empathy* column denotes the prediction probability output by `ModernBERT` that a tweet is empathetic. District level observations on literacy rate, sex ratio, as well as the proportion of urban population from 2019 are obtained from the Pakistan Statistical Year Book, an annual publication of the Pakistan Bureau of Statistics.

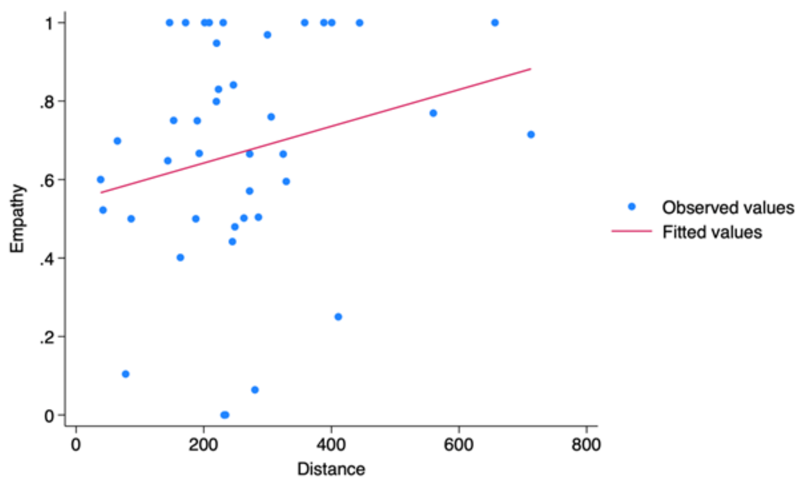
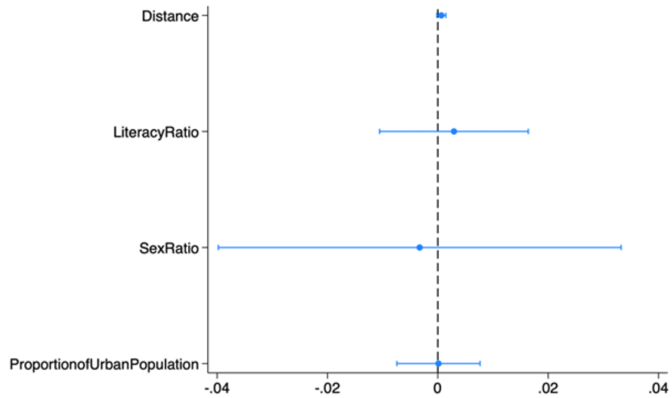
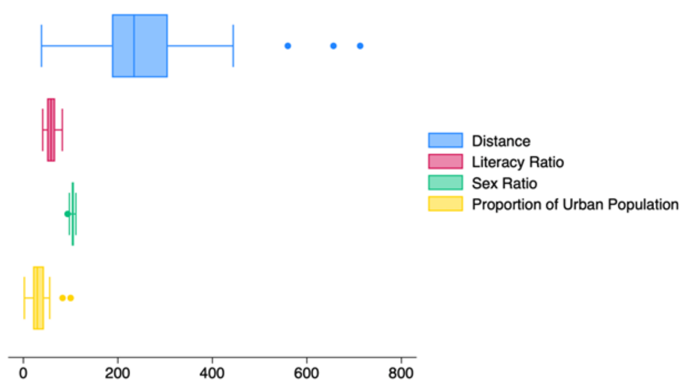


Fig. 2: Empathy and Distance: Scatter Plot

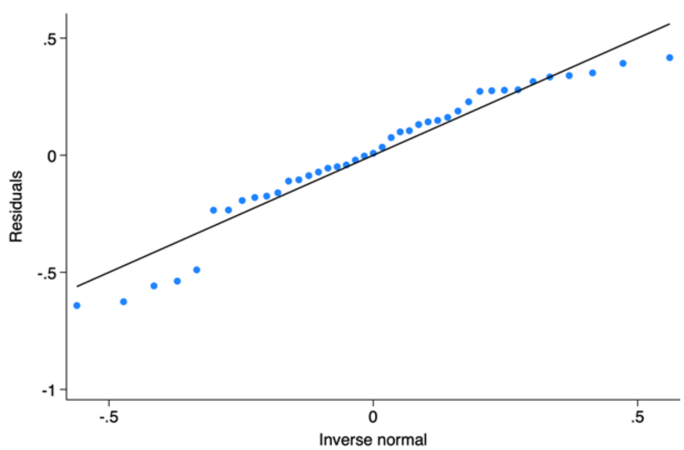
with the 95% confidence intervals, box and QQ plots. However, as detected from the box and QQ plots, the presence of outliers can skew the distribution of the error term violating the underlying assumptions of OLS: an outlier with twice the error magnitude of a typical observation contributes four times as much to the squared error loss and, therefore, has more leverage over the regression estimates.



(a)



(b)



(c)

Fig. 3: Average Marginal Effects with 95% Confidence Intervals, Box and QQ Plots .

Explanatory Variable	1	2	3	4	5	6	7	8
Distance	0.0005**	0.0006**	0.0007**	0.0006**	0.0006**	0.0006**	0.0005**	0.0005**
Literacy Ratio		0.0045	0.0031	0.0029	0.0036			
Sex Ratio			-0.0028	-0.0033				
Proportion of Urban Population				0.0001	-0.0004	0.0012		0.0004
Constant	0.5485**	0.3432	0.6199	1.7335	0.3752	1.3001	0.9520	0.5349**
Sample Size	41	41	41	41	41	41	41	41
R-squared	0.05	0.07	0.07	0.07	0.07	0.06	0.06	0.05
F-statistic	4.34	2.46	1.92	1.42	1.68	1.68	2.86	4.72

Table 5: Empathy and Distance: Robust Regression Results. Dependent variable is empathy. ** indicates a regression coefficient is significant at 5% level.

To address this issue, Table 5 presents statistical estimates of the coefficient of our variable of interest, the vector of coefficients of controls, and the constant term, obtained from running robust regressions with all possible combinations of explanatory variables. In columns 1 through 8 of this table, we validate that the significant positive association between empathy and distance after controlling for distinct combinations of literacy rate, sex ratio, as well as the proportion of urban population. Collectively, our results provide convincing evidence of a significant positive partial correlation between empathy and distance, after controlling for demographic characteristics, that is not sensitive to alterations in the conditioning information set to the extent that the sign as well as statistical significance of the partial correlation do not change in the face of changing combinations of controls. We infer, with 95% level of confidence, that a rise in the point estimate of the likelihood of expressing empathy ranges from 0.05 to 0.07 for every 100 kilometers increase in distance from those affected.

But for any consideration of distance making the heart grow fonder, at first sight, our finding may come across as counterintuitive. However, recurring tensions along the shared borders between India and Pakistan may have contributed to a relative lack of empathy among those residing in and around the eastern districts of Pakistan for those of the neighboring nation. Yet another plausible explanation, not too far-fetched, may be sought through the strengthening ties between India and Afghanistan [20] that could have nurtured empathy for the Indian populace among those in Pakistan relatively close to their bordering nation on the west.

7 Conclusion

In this paper, we present an interdisciplinary study at the intersection of natural language processing and econometrics investigating an intriguing research question: *Is empathy, between the residents of neighboring nations, associated with distance?* We consider a geotagged dataset of tweets collected during the second wave of COVID-19 that devastated India. We consider multiple widely known and used large language models as baselines and train empathy classifiers on this dataset. We next use this empathy classifier to study the association between empathy and distance of the geotagged tweets' locations from the India-Pakistan border.

We observe a significant positive association between empathy, as reflected in supportive tweets originating from Pakistan for the residents of India in the aftermath of COVID-19, and distance. Our results remain robust to the extent that the sign and statistical significance of the partial correlations between empathy and distance are not sensitive to alterations in the conditioning information set. Indeed, notwithstanding the credibility of the sources that allow access to relevant data, any inference drawn from our analysis must be interpreted with caution since the quality and reliability of data may be sensitive to reporting accuracies. A natural extension of our work would allow room for spatial dependence in an econometric model where spatial weight matrices depend on distance decay parameters.

8 Ethical Statement

In this work, we use an existing dataset of geo-tagged tweets collected during the COVID-19 pandemic [49]. This dataset has been collected using publicly available Twitter API. We conduct aggregate analyses and do not focus on individual users. We thus see no major ethical concern. We use an empathy classifier trained on large language models. While we use multiple widely used LLMs to corroborate our findings, large language models (LLMs) are known to have a wide range of biases due to the train data [5]. Our empathy classifier is trained on an existing dataset. We are aware of the vast literature on annotation subjectivity [33].

References

1. Ali, T.: Can Pakistan survive?: the death of a state. Penguin Books London (1983)
2. Baheti, A.D., Nayak, P.: Covid-19 in india: oxygen shortages and a real world trolley problem (2022)
3. Batson, C.D., Fultz, J., Schoenrade, P.A.: Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. Journal of personality 55(1), 19–39 (1987)
4. Batson, C.D., Shaw, L.L.: Evidence for altruism: Toward a pluralism of prosocial motives. Psychological inquiry 2(2), 107–122 (1991)
5. Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: ACM FaccT. pp. 610–623 (2021)
6. Benesch, S., et al.: Counterspeech on Twitter: A field study. Dangerous Speech Project (2016)
7. Bose, S.: Kashmir: Roots of conflict, paths to peace. Harvard University Press (2009)
8. Brun, J.F., Carrère, C., Guillaumont, P., De Melo, J.: Has distance died? evidence from a panel gravity model. The World Bank Economic Review 19(1), 99–120 (2005)

9. Buechel, S., Buffone, A., Slaff, B., Ungar, L., Sedoc, J.: Modeling empathy and distress in reaction to news stories. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. pp. 4758–4765. Association for Computational Linguistics, Brussels, Belgium (Oct-Nov 2018), https://www.aclweb.org/anthology/D18-1507
10. Chen, Y., Yan, S., Liu, S., Li, Y., Xiao, Y.: EmotionQueen: A benchmark for evaluating empathy of large language models. In: Findings of the Association for Computational Linguistics: ACL 2024. pp. 2149–2176. Association for Computational Linguistics (Aug 2024), https://aclanthology.org/2024.findings-acl.128/
11. Couclelis, H.: The death of distance (1996)
12. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. In: NAACL-HLT 2019. pp. 4171–4186. Association for Computational Linguistics (2019)
13. Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., et al.: The llama 3 herd of models. arXiv preprint arXiv:2407.21783 (2024)
14. Dutta, A., Ali, S.M.S., Naseem, U., KhudaBukhsh, A.R.: Towards a bipartisan understanding of peace and vicarious interactions. In: Proceedings of the 34th International Joint Conference on Artificial Intelligence, IJCAI 2025. p. To Appear (2025)
15. Fleischacker, S.: Being me being you: Adam Smith and empathy. University of Chicago Press (2020)
16. Furniturewala, S., Jaidka, K.: Empaths at WASSA 2024 empathy and personality shared task: Turn-level empathy prediction using psychological indicators. In: Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis. pp. 404–411. Association for Computational Linguistics (2024)
17. Goetz, J.L., Keltner, D., Simon-Thomas, E.: Compassion: an evolutionary analysis and empirical review. Psychological bulletin 136(3), 351 (2010)
18. Hussain, S., Shahzad, F., Saud, A.: Analyzing the state of digital information warfare between india and pakistan on twittersphere. SAGE Open 11(3), 21582440211031905 (2021)
19. Jiang, A.Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D.S., Casas, D.d.l., Bressand, F., Lengyel, G., Lample, G., Saulnier, L., et al.: Mistral 7b. arXiv preprint arXiv:2310.06825 (2023)
20. Kaura, V.: India-afghanistan relations in the modi-ghani era. Indian Journal of Asian Affairs 30(1/2), 29–46 (2017)
21. KhudaBukhsh, A.R., Palakodety, S., Carbonell, J.G.: Harnessing code switching to transcend the linguistic barrier. In: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020. pp. 4366–4374 (2020)
22. KhudaBukhsh, A.R., Palakodety, S., Mitchell, T.M.: Harnessing unsupervised word translation to address resource inequality for peace and health. In: Social Informatics - 13th International Conference, SocInfo 2022. Lecture Notes in Computer Science, vol. 13618, pp. 159–180. Springer (2022)
23. KhudaBukhsh, A.R., Palakodety, S., Mitchell, T.M.: Harnessing unsupervised word translation to address resource inequality for peace and health. In: Social Informatics: 13th International Conference, SocInfo 2022, Glasgow, UK, October 19–21, 2022, Proceedings. p. 159–180. Springer-Verlag (2022)
24. Klimecki, O.M., Mayer, S.V., Jusyte, A., Scheeff, J., Schönenberg, M.: Empathy promotes altruistic behavior in economic interactions. Scientific reports 6(1), 31961 (2016)

25. Lahnala, A., Welch, C., Neuendorf, B., Flek, L.: Mitigating toxic degeneration with empathetic data: Exploring the relationship between toxicity and empathy. In: Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 4926–4938. Association for Computational Linguistics, Seattle, United States (Jul 2022), https://aclanthology.org/2022.naacl-main.363/
26. Lipps, T.: Zur Psychologie der Suggestion: Vortrag, Gehalten am 14. Januar 1897, in der" Psychologischen Gesellschaft" zu München. JA Barth (1897)
27. Malik, I., Wirsing, R.G.: Kashmir: Ethnic conflict international dispute. Oxford University Press Oxford (2002)
28. Mikulincer, M.E., Shaver, P.R.: Prosocial motives, emotions, and behavior: The better angels of our nature. American Psychological Association (2010)
29. Mishra, K., Burja, M., Ekbal, A.: ABLE: personalized disability support with politeness and empathy integration. In: Al-Onaizan, Y., Bansal, M., Chen, Y. (eds.) Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024. pp. 22445–22470. Association for Computational Linguistics (2024)
30. Nayak, A., Timmapathini, H., Ponnalagu, K., Gopalan Venkoparao, V.: Domain adaptation challenges of BERT in tokenization and sub-word representations of out-of-vocabulary words. In: Proceedings of the First Workshop on Insights from Negative Results in NLP. pp. 1–5. Association for Computational Linguistics (2020)
31. Nijkamp, P.: The death of distance. In: Economic ideas you should forget, pp. 93–94. Springer (2017)
32. Palakodety, S., KhudaBukhsh, A.R., Carbonell, J.G.: Hope speech detection: A computational analysis of the voice of peace. In: ECAI 2020, pp. 1881–1889. IOS Press (2020)
33. Pavlick, E., Kwiatkowski, T.: Inherent disagreements in human textual inferences. Transactions of the Association for Computational Linguistics 7, 677–694 (2019)
34. Qian, Y., Zhang, W., Liu, T.: Harnessing the power of large language models for empathetic response generation: Empirical investigations and improvements. In: Findings of the Association for Computational Linguistics: EMNLP 2023. pp. 6516–6528. Association for Computational Linguistics (2023)
35. Robin, C.: Empathy & the economy. New York Review of Books, December 8, 43–46 (2022)
36. Rosenthal, S.S., Strange, W.C.: How close is close? the spatial reach of agglomeration economies. Journal of economic perspectives 34(3), 27–49 (2020)
37. Sarkar, R., Mahinder, S., KhudaBukhsh, A.R.: The non-native speaker aspect: Indian english in social media. In: Proceedings of the Sixth Workshop on Noisy User-generated Text, W-NUT@EMNLP 2020 Online. pp. 61–70. Association for Computational Linguistics (2020)
38. Sarkar, R., Mahinder, S., Sarkar, H., KhudaBukhsh, A.R.: Social media attributions in the context of water crisis. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020. pp. 1402–1412. Association for Computational Linguistics (2020)
39. Schofield, V.: Kashmir in conflict: India, Pakistan and the unending war. Bloomsbury Publishing (2010)
40. Sharma, A., Miner, A., Atkins, D., Althoff, T.: A computational approach to understanding empathy expressed in text-based mental health support. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 5263–5276. Association for Computational Linguistics, Online (Nov 2020), https://aclanthology.org/2020.emnlp-main.425/

41. Singer, T., Fehr, E.: The neuroeconomics of mind reading and empathy. American Economic Review 95(2), 340–345 (2005)
42. Smith, A.: In: An inquiry into the nature and causes of the wealth of nations: Volume One. London: printed for W. Strahan; and T. Cadell, 1776. (1776)
43. Sober, E., Wilson, D.S.: Unto others: The evolution and psychology of unselfish behavior. No. 218, Harvard University Press (1999)
44. Toon, O.B., Bardeen, C.G., Robock, A., Xia, L., Kristensen, H., McKinzie, M., Peterson, R.J., Harrison, C.S., Lovenduski, N.S., Turco, R.P.: Rapidly expanding nuclear arsenals in Pakistan and India portend regional and global catastrophe. Science Advances 5(10), eaay5478 (2019)
45. Tyagi, A., Field, A., Lathwal, P., Tsvetkov, Y., Carley, K.M.: A computational analysis of polarization on indian and pakistani social media. In: International Conference on Social Informatics. pp. 364–379. Springer (2020)
46. Vieten, C., Rubanovich, C.K., Khatib, L., Sprengel, M., Tanega, C., Polizzi, C., Vahidi, P., Malaktaris, A., Chu, G., Lang, A.J., et al.: Measures of empathy and compassion: A scoping review. Plos one 19(1), e0297099 (2024)
47. Warner, B., Chaffin, A., Clavié, B., Weller, O., Hallström, O., Taghadouini, S., Gallagher, A., Biswas, R., Ladhak, F., Aarsen, T., et al.: Smarter, better, faster, longer: A modern bidirectional encoder for fast, memory efficient, and long context finetuning and inference. arXiv preprint arXiv:2412.13663 (2024)
48. Wright, M.F., Wachs, S.: Does empathy and toxic online disinhibition moderate the longitudinal association between witnessing and perpetrating homophobic cyberbullying? International journal of bullying prevention 3, 66–74 (2021)
49. Yoo, C.H., Palakodety, S., Sarkar, R., KhudaBukhsh, A.: Empathy and hope: Resource transfer to model inter-country social media dynamics. In: Proceedings of the 1st Workshop on NLP for Positive Impact. pp. 125–134. Association for Computational Linguistics, Online (2021), https://aclanthology.org/2021.nlp4posimpact-1.14/