UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Building a Psychological Ground Truth Dataset with Empathy and Theory-of-Mind During the COVID-19 Pandemic

Permalink
https://escholarship.org/uc/item/950900w7

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Lee, Yoon Kyung
Jung, Yoonwon
Lee, Inju
et al.

Publication Date
2021

Peer reviewed
Building a Psychological Ground Truth Dataset with Empathy and Theory-of-Mind During the COVID-19 Pandemic

Yoon Kyung Lee (yoonlee78@snu.ac.kr)
Department of Psychology, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

Yoonwon Jung (ywjung@snu.ac.kr)
Department of Psychology, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

Inju Lee (ijlee37@snu.ac.kr)
Department of Psychology, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

Jae Eun Park (dawn2089@snu.ac.kr)
Department of Psychology, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

Sowon Hahn (swhahn@snu.ac.kr)
Department of Psychology, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

Abstract
As the mental health crisis deepens with the prolonged COVID-19 pandemic, there is an increasing need for understanding individuals’ emotional experiences. We have built a large-scale Korean text corpus with five self-labeled psychological ground-truths: empathy, loneliness, stress, personality, and emotions. We collected 19,025 documents of daily emotional experiences from 3,805 Korean residents from October to December 2020. We collected 42,128 sentences with different levels of theory-of-mind. Each sentence was annotated by trained psychology students and reviewed by experts. Participants varied in their ages from the early 20s to late 80s and had various social and economic statuses. The pandemic impacted the majority of daily lives, and participants often reported negative emotional experiences. We found the most frequent topics: responses to confirmed cases, health concerns of family members, anger towards people without masks, stress-relief strategies, change of the lifestyle, and preventive practices. We then trained the Word2Vec model to observe specific words that match each topic from the topic model. The current dataset will serve as benchmark data for large-scale and computational methods for identifying mental health levels based on text. This dataset is expected to be used and transformed in many creative ways to mitigate COVID-19-related mental health problems.

Keywords: COVID-19, Empathy, Theory-of-Mind, Personality, Emotions, Crowdsourcing, Topic Modeling, Word2Vec

Introduction
The outbreak of COVID-19 affected the health of many individuals. Korea was one of the many countries that experienced an upsurge of cases and actively adjusted polices to stop the virus spread since the first confirmed cases on 20th of January, 2020 (World Health Organization, 2020a, 2020b). Due to COVID-19, opportunities for real face-to-face social exchanges have declined sharply. Yet research on the impact of the pandemic situation on people’s emotional experiences is still lacking. Several studies used web-scraped corpus from social media (Chen et al., 2020) or used crowdsourcing to directly collect people’s responses in text. There is a need for documentation of pandemic experiences from various parts of the world, since the impact of the pandemic differs from

Figure 1: Overall procedure of building the dataset. Each sentence was annotated based on levels of theory-of-mind by trained psychology major undergraduate and graduate students (Expert #1) then accepted by reviewers with psychology degree (Expert #2).
country to country. The emotional responses to COVID-19 and public can be different by countries as well. In this study, we collected large dataset with corpus of people’s emotional experiences labeled with empathy, theory of mind, and related psychological traits that are highly associated with many severe mental and physical illnesses (e.g., loneliness and stress).

**Empathetic Experiences in the Pandemic**

Empathy is defined as the ability to understand mental states and emotions of others and therefore employ appropriate social-emotional response (Levenson & Ruef, 1992; Stueber, 2019). Empathy is essential in one’s own emotional experience (Thompson et al., 2021). Empathy can facilitate the compliance with facial masking and social distancing by thinking of others in different perspectives (Pfattheicher et al., 2020). Although there are various viewpoints on what constitutes empathy (Cuff et al., 2016), here we considered sympathy and compassion (also considered ‘affective resonance’; Vachon & Lynann, 2016) as major components of empathy. Another component of empathy, is theory-of-mind (Premack & Woodruff, 1978). Theory-of-mind is also known as perspective-taking which is an ability to take other’s desire or intention in one’s mental state. Previous studies found that different aspects of empathy can have different degree of impact on social relationships. For example, individuals can improve attitudes towards others by understanding others’ perspectives (Shih et al., 2009) and do prosocial behaviors through feelings of personal distress or sympathy (Batson et al., 1987). Ma and Wang (2021) found that the level of perspective-taking and personal distress influences emotional reactivities to COVID-19 pandemic. The degree of perspective-taking of others affected the intention to get vaccinated in COVID-19 pandemic situations (Pfattheicher et al., 2020).

**Impact of Social Isolation to Mental Health**

Many societal impacts of COVID-19 pandemic such as economic recession, unemployment, and public health led to increase in stress levels. Perceived stress entails loss of control, feeling of hopelessness, and fatigue. High levels of distress were found in women, people with lower income, people who live with others (Flesia et al., 2020) and health professionals on critical care services or emergency (Ruiz-Fernández et al., 2020). Loss of social connections and prolonged isolations are positively associated with feeling of loneliness (Pinquart & Sorensen, 2001). It is also negative associated with subjective well-being, life satisfaction, vitality, and positive affect (Elphinston, 2018). Chronic loneliness can exacerbate mental illnesses like major depression, psychosis, and suicide (Hawkley & Cacioppo, 2010).

**Individual Differences in Emotional Responses to the Pandemic**

Personality can be an important factor in identifying differently experienced mental health issues such as stress and anxiety. It can also explain different behavioral outcomes as well. Studies found that higher neuroticism and extraversion experienced a higher level of stress response to the COVID-19 when compared prior to the pandemic (Liu et al., 2021). The degree of compliance with safety rules and social distancing was found to be different by personality types as well. People who were low on agreeableness were also less likely to comply with the social restrictions (Zajenkowski et al., 2020). Low conscientiousness and openness to experience were also associated with greater level of economic anxiety (e.g., loss of jobs; Mann et al., 2020).

**The Present Study**

We conducted an explorative large-scale study and built COVID-19 Emotion Diary with Empathy and Theory-of-Mind Ground Truths Dataset. The dataset contains written daily experiences labeled with emotions theory-of-mind levels, loneliness, perceived stress, empathy, and personality of residents of Korea. The dataset contains real experiences in pandemic situations in written text. The data was collected from Oct. to Dec. 2020. We obtained written daily experiences of people from various ages, education, income levels. Each individual wrote 5 diaries (representing typical 5 days) with self-labeled theme emotions, writing stress and writing efforts. As such, the dataset provides insight on how people express their emotions and opinions regarding current situations by offering their different levels of emotions and psychological states. Also, these psychological ground truth labels were self-annotated by the participants, which reduces the possibility of writer-annotator discrepancy of labels. The data is available for research community.

**Methods.**

**Materials**

A total of 36 questions of demographic information and psychological variables were used. Variables measured are listed in Table 1. We asked whether one is currently working or taking classes from home due to the pandemic (‘remote work’) and whether a person is limiting social contacts less than twice a week due to the pandemic (‘social distancing’). A total of 30 questions were psychological states: personality, loneliness, perceived stress, and empathy.

**Procedure**

Participants were given online informed consent at the beginning of the project. A monetary compensation equivalent to 7 US Dollars was provided. We asked participants to avoid distractions during the participation (e.g., watching videos). Submissions that were considered unoriginal or plagiarized (e.g., copy-and-paste a portion of web page or news articles) were declined. Example guidelines in writing were also provided but were not mandatory. Example topics were transition to remote work or online classes, change in social connections, feelings of others who do not practice social distancing, and coping...
strategies to relieve stress in pandemic. Finally, participants were asked to proofread before submission and not to include any personally identifiable information such as phone number, address, or residential ID number.

In the next step, 5 experts (psychology graduate students) reviewed submissions and developed a guideline for labeling theory-of-mind levels to sentences. We came up with four categories of sentences: no presence of others (‘level 0’), mention of others (but no evidence of inferring other’s intention; ‘level 1’), refuse to see in others’ perspectives (‘level 2’), and perspective-taking (‘level 3’). Example of ‘level 1’ was “I went to go shopping with my friend”, ‘level 2’ was “He is a selfish person not to wear a mask in the middle of pandemic”, ‘level 3’ was “I know it must have been hard for him”. Eighteen trained psychology students labeled sentences according to the guideline, then professional reviewers approved them. Inter-rater reliability was 0.7 based on 300 sample documents.

Results

A total of 3805 participants’ data were analyzed. The average age of participants was 31.55 (SD = 10.27). A total of 36 participants were excluded from analysis. Among participants, female was 76% (N = 2910). Above half (53.3%) reported they are currently doing remote-work or taking online classes school. People who reported that they are practicing social distancing were 64.6%. People in the age range of 20 to 40 were 91% (N = 3452). The lowest income level was 21.6 % of ‘less than 20k’, 18.2 % of 20k-30k, then 16.6% of 30k to 40k. Education level were university (51.6%), high school (20.8%), and college (16.3%). Out of 19,025 documents, self-labeled emotions were sad (7,119), ‘others’ (6,373), angry (5,148), calm (4,602), fear (3,707), disgust (2,754), and surprise (2,603). Top 8 frequently reported emotions within ‘others’ category was frustrated (461), amusing (346), disappointed (271), pleasant (244), irritating (241), proud (192), sympathetic (183), and worry (184).

For loneliness, we used two cutoff points to distinguish 3 levels of loneliness (low, moderate, and high; Neto, 1992). The mean score of accumulated loneliness score was 14.64 (SD = 4.29). ‘moderate’ loneliness (above 15) was 31.6% (1204 out of 3805) and ‘high’ loneliness (above 18) was 26% (1000 out of 3805). The mean score of perceived stress was 8.30 (SD = 2.93). People who scored higher than 12 (high-stress) were 13.56% (516 out of 3805). For empathy, the mean score was 24.57 (SD = 3.77). Thirty-one percent scored higher than a score of 27 were considered as ‘high-empathy’ (1175 out of 3805). The distribution of scores was skewed to the right. For personality, we normalized each score of personality types, then categorized each individual into one personality type with the maximum number of scores. Out of 3508 participants to be categorized, 1012 were agreeable (28.8%), 1010 were neurotic (28.8%), 809 were open (23%), 646 were conscientious (18.4%) and 328 were extraversion (9.3%). There were 13.04 words per sentence on average (total 549,629 words equal to 421,128 sentences). Out of 14,350 documents, there were on average 2.94 (SD=2.15) annotations. A total of 2,513 documents did not have any annotation (‘level 0’). Out of 39,412 sentences annotated, the number of sentences annotated ‘level 1’ was 34,448, ‘level 2’ was 2,513, and ‘level 3’ was 2,451.

Table 1: List of Variables.

<table>
<thead>
<tr>
<th>Name</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathy</td>
<td>0 to 4 (9 items)</td>
</tr>
<tr>
<td>Theory-of-Mind</td>
<td>Level 0 to 3</td>
</tr>
<tr>
<td>Loneliness</td>
<td>0 to 3 (6 items)</td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>0 to 4 (4 items)</td>
</tr>
<tr>
<td>Personality</td>
<td>0 to 4 (12 items)</td>
</tr>
<tr>
<td>Emotions</td>
<td>Happy, fear, disgust, surprise, calm, anger, sad, and *others</td>
</tr>
</tbody>
</table>

*participants wrote their own emotions that best represent the diary content.

Table 2: Topic Models.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>person, mask, corona, think, when, confirm, wear, situation, prevention</td>
</tr>
<tr>
<td>2</td>
<td>job, company, time, remote work, commute, work, think, home, workplace</td>
</tr>
<tr>
<td>3</td>
<td>child, mom, home, husband, family, daughter, son, hospital, brother(sister), mind</td>
</tr>
<tr>
<td>4</td>
<td>work-out, home, corona, stress, time, think, start, day, body, feel</td>
</tr>
<tr>
<td>5</td>
<td>class, school, continue, study, exam, course, online, time, face-to-face, think</td>
</tr>
<tr>
<td>6</td>
<td>people, mask, anger, wear, prevent, understand, uncomfortable, seat, sound, irritate</td>
</tr>
<tr>
<td>7</td>
<td>house, walk, car, feeling, dog, place, road, park, people, neighborhood</td>
</tr>
<tr>
<td>8</td>
<td>friend, corona, person, travel, think, time, gathering, house, situation, plan</td>
</tr>
<tr>
<td>9</td>
<td>house, cafe, food, time, corona, movie, coffee, think, delivery, book</td>
</tr>
<tr>
<td>10</td>
<td>corona, think, work, money, situation, worry, employ, mind, ready</td>
</tr>
</tbody>
</table>

Text Mining We used a python library, Spacy, for tokenizing documents into sentences, then to words (tokens) with Korean model. Among tokens, we extracted tokens with part-of-speech tagger corresponding to nouns and adjectives. We created a tokenized list of words, t (N = 19,021) for topic modeling using LDA (Blei et al., 2003). We ran a total of 15 iterations to train the model. Table 2 shows 10 topics with the highest coherence score. 486. Selected top 10 topics were:
response to confirmed cases (topic 1), ‘transition to remote-work (topic 2), ‘health concerns of family members (topic 3), ‘physical exercise as stress relief (topic 4), ‘transition to online schools (topic 5), ‘feeling unsafe around people with no masks on (topic 6), ‘outdoor activities as stress relief (topic 7), ‘social distancing from friends (topic 8), ‘making new hobbies as stress relief (topic 9),’ and ‘financial difficulties (topic 10).’ We then trained a word embedding model, Word2Vec (Mikolov et al., 2013). We employed the skip-gram model and set the parameters as follows: the dimensionality of embedding vector = 100, window size = 5, the minimum counts of token to use = 5, the number of worker threads = 4. Figure 2 shows a t-SNE two-dimensional map of top 500 frequent (Van de Maaten & Hinton, 2008). We conducted K-means clustering with our trained embedding vectors for grouping the similar embedding vectors. We classified the embedding vectors into 10 clusters. We combined some similar clusters and determined the final 8 clusters. We visualized each cluster with different colors on the plot.

Discussion

In the present study, we collected over 19K documents about the daily experience of people living in pandemic situations. We also collected variables that can be later labeled and provide ground truth for mental health: emotions, loneliness, stress, empathy, and personality. We found that more females than males, younger adults than older adults, people with lower income levels, and higher education levels participated in our study. We collected data from a wide range of age groups (19 to 82), where we found diverse topics related to COVID-19. Sad and angry were the most frequently reported emotions from reported daily experiences. People were generally feeling lonely (above ‘moderate’ level). Empathy showed right-skewed distribution, suggesting people generally regarded themselves as more empathetic. Openness and Neuroticism were the most frequent personality type found in our data.

Limitations and Suggestions for Future Study

The current study leaves several research topics unexplored. First, we did not explore how the daily experience of people differed by specific timelines (weekly or monthly). Future research may explore possible patterns of general sentiment change in the pandemic situation (e.g., from hopelessness to adjustment). The number of people who reported feelings of sadness was twice as many as those who reported fear, which could suggest that as time passed, the immediate response to COVID-19 such as fear and anxiety has subsided and that people nowadays feel more tired and...
We found that internet users may not be representative of average Koreans as they tend to be high in openness and agreeableness. Individuals with these personality types are more likely to participate in online surveys than other types (Marcus & Schütz, 2005; Fan & Yan, 2010; Park et al., 2018). However, our data had a similar number of people with neuroticism to those with openness. This could be due to the fact that more people, in general, are participating in discussions online than offline due to enforced social isolation.

We, however, were able to collect over 353 older adults with ages ranging from 50 to 82. This opens a new stem of research on understanding different experiences among different age groups. For example, since many of the older generations are separated from their children and are likely to have a physical illness, the impact of COVID-19 would be more detrimental. Many studies found higher levels of loneliness and depression in older generations than younger ones (Dykstra et al., 2005). Since different generations experience different aspects of lives, age differences in how stress triggered by prolonged social isolation affect experiences of empathy are to be explored in future studies.

Future studies may use the data for interdisciplinary research on detecting empathy and theory-of-mind in the form of written text. With psychological ground truth embedded in the data such as emotions and personality, one can generate specific data that represents a specific type of personality with a combination of levels of empathy.

### Conclusion

We introduced the first ground truth large text dataset of empathy and theory-of-mind, with psychological traits of people living in the COVID-19 pandemic. The presented data includes labels of emotions, loneliness, stress, personality, empathy, and theory-of-mind, so that each document can be classified by each variable. Our results provide a new direction for scalable and computational methods of predicting empathy and theory-of-mind. We encourage the research community to use the dataset for better understanding of emotional distress and mental health affected by the pandemic.

### Acknowledgments

This project was supported by Data Voucher Project, Korea Data Agency.¹ We appreciate DeepNatural AI² for facilitating data collection.

### References


¹ https://global.kdata.or.kr/en/kdata/

² https://www.deepnatural.ai/
consequences and mechanisms. *Annals of Behavioral Medicine, 40*(2), 218-227.


