Healthcare Innovation Program: Pandemic Proposal

Community-Focused Assessment of the Experiential Impact of COVID-19

Introduction

Background

The outbreak of coronavirus 2019 (COVID-19) was declared a global emergency by the World Health Organization (-WHO), on January 30, 2020 (Sohrabia, et al., 2020), and a global pandemic on March 11, 2020 (Bedford, et al., 2020). Evidence from medical, social and behavioral science-informed investigations of the effects of pandemic, including COVID-19-specific effects, on individuals and communities is continuing to be reported through scholarly research, professional practice activities, and media discussions. Standards and recommendations have rapidly been released from the public health community in an effort to summarize best practices for the general public and health providers in their daily operations during a time of rapidly elevating workload and occupational/personal stress, as well as maintaining a measure of self-care through this time (CDC, 2020; SAMHSA, 2020; SAMHSA, 2020; IASC, 2020). One recent study of the psychological impact of COVID-19 on healthcare workers in Wuhan, China, has identified that this community is one particularly at risk for experiencing mental health issues due to their direct exposure (Kang, Ma, Chenb, Yangb, & Wang, 2020, in press). Effects of such aspects as quarantine and social distancing; drastic fluctuations in workload, work process, or income; and physical or mental exposure to COVID-19 through direct/indirect illness or professional responsibilities, have significantly impacted people around the globe (Brooks, et al., 2020). A number of community interventions and research priorities should be considered as we navigate through experiences related to the COVID-19 pandemic (Holmes, et al., 2020; Pfefferbaum & North, 2020; Torales, O'Higgins, Castaldelli-Maia, & Ventriglio, 2020; Van Bavel, 2020).

As communities continue to experience the impacts of COVID-19, and attempt to resume various activities such as returning to work, school, or other daily activities, it is critical that community leaders strive to understand the full range of what their people are experiencing in their daily lives, the level of understanding that the public has related to events occurring or public actions taken, and how they are reacting to and coping with their diverse pandemic experiences. This should include reactions to public health measures taken to reduce the spread of disease, as well as the actions, communications, and attitudes of leaders, supervisors, businesses, and other members of the general public. Particular members of the public depending on various individual factors, such as how well someone tolerates a reduced sense of control, may be more susceptible to stress and mental health effects (Bedford, et al., 2020; Taha, Matheson, Cronin, & Anisman, 2014). It is suggested by studies such as these, that some people depending on their underlying characteristics may be more at-risk for serious illness if they experience events associated with pandemic such as, prolonged isolation, exposure to pandemic-related death, loss of income/career, increased workload, lack of pertinent information related to public health risk and actions, or others.

Instruments developed recently and for the purpose COVID-19, as well as established, validated instruments that can be applied or adapted to pandemic-related assessment, are available in the literature and can be compiled into a brief assessment battery that can be delivered and completed online for remote data collection (Grasso, Briggs-Gowan, Ford, & Carter, 2020) (Boyd & Pennebaker, 2017; Bedford, et al., 2020; Taha, Matheson, Cronin, & Anisman, 2014). Grasso et al. (2020) have developed the Epidemic/Pandemic Impacts Inventory (EPII) specifically to assess personal impact of pandemic. The EPII, comprised of 92-items across 10 categories, assesses a range of experiences in multiple life areas potentially affected by COVID-19, including: employment, education, social, economic, physical health, physical distancing and quarantine, infection history, and positive change. The EPII appears to be unique in that it recognizes that it captures both negative and positive experiencs. Additionally, prior work of the team in digital psychiatry has revealed that social media data, when paired with computational and machine learning techniques, can provide unobtrusive insights into the mental health states of individuals, communities, and populations, ranging from major depressive disorder and anxiety disorders, to suicidal ideation (De Choudhury, Counts, & Horvitz, 2013) (DeChoudhury, Kiciman, Dredze, Coppersmith, & Humar, 2016) (Birnbaum, et al., 2019) (Bagroy, Kumaraguru, & De Choudhury, 2017) (De Choudhury & Kiciman, 2017). For instance, Co-I De Choudhury's research has been able to successfully predict, based on machine learning methods (De Choudhury, Gamon, Counts, & Horvitz, 2013), risk to depression, prior to reported onset, from Twitter data of individuals with 71% accuracy. Various measures spanning language, social media engagement and activity were significant predictors: decreases in social activity, increases in negative affect, and greater religiosity.

Aims and Hypotheses

We will leverage current COVID-19 research, mental health pandemic consulation guidelines developed by PI Kaslow and Co-I Druss, and prior research by co-PI Crooks, Co-I De Choudhury, and Co-I Muchlinski to create and conduct an online, anonymous, experiential interview, and guided social media analysis, in order to assess how communities have been impacted by the COVID-19 pandemic (Crooks, 2020; Crooks, et al., 2020) (De Choudhury, Morris, & & White, 2014) (De Choudhury & De, 2014). We will conduct an online interview over a number of weeks using Qualtrics survey software licensed through Georgia Tech. We will then analyze the online interview data to understand first-person experience of the impact of COVID-19 for each community assessed. Further, drawing upon Co-I De Choudhury's prior research that has revealed the potential to understand people's mental health needs via social media (De Choudhury, Morris, & & White, 2014) (De Choudhury & De, 2014), we will focus our analysis on the COVID-19 related experiences of healthcare workers and first responders by conducting an exploratory social media analysis to discover key experiences related to psychological impact of COVID-19 on these professional groups. Our data collection methodology will be comprised of a range of specialized, fielded and validated instruments, and designed to collect relevant data for immediate exploration of community needs as a result of COVID-19 impact. This research will lead to a greater understanding of the psychological and other impacts of COVID-19 on community members from a range of demographics, occupations, and lived experiences.

Specific Aim 1: Demonstrate a method for collecting pandemic-related experiential data from targeted communities of interest using first-person experiential online interviews, narrative language analysis of interview content, and iterative machine learning-based social media analysis for discovery.
Specific Aim 2: Utilize our survey and social media findings to inform return-to-work planning, communications, training, and resources, based on direct assessment of community needs.

Method

Phase 1 (Months 1-6): Online Survey

Participants

A 6-month online survey study will be conducted with participants from Georgia Tech and Emory who meet study inclusion/exclusion criteria (estimated N=500 survey respondents for each community). The survey protocol was approved as "Exempt" by the Georgia Institute of Technology (Georgia Tech) Institutional Review Board (IRB) on 5/11/2020. Recruitment and informed consent for Georgia Tech and Emory University participants will follow standard Human Use procedures in accordance with each university's IRB policy. The survey will be launched by co-PI Crooks using the Georgia Tech Qualtrics program, with the Georgia Tech community as an initial comparison cohort. Additional participants for the healthcare/first responder survey cohort will be obtained in collaboration with PI Kaslow from Emory University. A protocol will be submitted to the Emory IRB in keeping with IRB guidelines. The survey will take approximately 1-hour to complete on a voluntary basis, and participants will not be compensated for their participation in this study. Data collection through Qualtrics will permit electronic capture of de-identified survey data, accessible to and exported only by the research team through a study code.

Materials

Background measures. A semi-structured experiential interview will be created by the Investigator(s) for use in this study and administered by the online survey program. The interview will include basic demographic data such as age, gender, and occupation, in a demographic questionnaire based on structure provided in McAdams (1988) (see Appendix A). The interview will also include an open-ended interview question derived from narrative research methodology described in McAdams (1988) and Cramer (1996), which will ask the participant to describe aspects of their personal experience with consequences of COVID-19 pandemic (see Appendix B).

Targeted questionnaire measures. Subsequent to the background questionnaire, the following three questionnaires will be administered by the online survey program, in the order presented below (see Appendices C, D, and E):

1) The Epidemic – Pandemic Impacts Inventory (EPII) (92 items) (Grasso, Briggs-Gowan, Ford, & Carter, 2020)
2) Stress Appraisal (28 items) (Peacock & Wong, 1990)
3) The Brief COPE (28 items) (Carver, 1997)

Procedure and Schedule

The online survey program will administer and collect electronic data from the background questionnaires and targeted questionnaires included in the battery described in the preceding section. Capacity to provide informed consent will be evaluated through acceptance of the consent statement provided by the online survey program prior to beginning the questionnaire battery. Individuals unable to give consent will not be included. Following consent, all participants will complete the full online questionnaire battery in the order presented under the "Materials" section of this protocol. Following completion of the online questionnaire, the survey program will indicate to the participant that the survey has concluded, and thank them for their participation. For social media data collection, we will use an in-house automated web data scraping infrastructure based on the social media platforms' official Application Programming Interfaces (APIs) and the Open Authorization protocol; this infrastructure was developed in Co-I De Choudhury's prior work and has enabled a secure, privacy-preserving mechanism for social media data collection among consented clinical and non-clinical populations (Saha, et al., 2019). The infrastructure allows both retrospective and prospective social media data collection, depending on specific project needs. It only relies on a single point of consent, and does not require participants to actively contribute data.

Analysis

A variety of questions will be asked through the battery of five questionnaires, addressing participant demographics and COVID-19 related experiences. Data collected by the online survey program will be downloaded by authorized study team members, and organized by participant code into our electronic spreadsheet and statistics program for subsequent data analysis. To evaluate Aim 1, researchers will compare general data relationships across questionnaires using, as appropriate, parametric/non-parametric correlational analyses, structural equation modeling, and associated data transformation techniques. To evaluate Aim 2, qualitative topic and language content analysis will be conducted to assess specific linguistic features, emotionality, and meaning of the open-ended question posed within aim 2. Subsequent discussion of these analyses will generate information about specific community-related COVID-19 experiences. From this data, we will assess the sufficiency of planning, communications, training, and resources available to these communities.

Phase 2 (Months 7-24): Social Media Analysis

Social Media Data Harmonization, Cleaning, and Processing

For harmonizing the diverse social media data of participants, first, we will group them into the following categories: 1) Unstructured data from text, tags, captions, and descriptions associated with textual or multimedia posts/status updates shared on social media. 2) Temporal data such as timing of posts shared on the social media. 3) Activity and engagement data, include Facebook "likes", "retweets" and "favorites" on Twitter; Instagram "likes" made and received; Facebook comments; Twitter @-replies made and received; Instagram, comments made and received. And, 4) Network data includes Facebook social graph; Twitter follower and following graph; Instagram social graph, constructed using the friendship links on these platforms, as well as based on interpersonal interactions via private messaging on Facebook and Instagram, or @-replies and @-mentions on Twitter.

Our next task will focus on cleaning, processing and analysis of the above data types acquired above. First, we will filter for duplicate content and employ a spam checker on all of our data to eliminate irrelevant content. Then to overcome the challenges of noise and polysemy of unstructured data, we will apply an integrative approach for data cleaning and normalization. We will disregard stopwords in our data, then perform stemming, lemmatization and morphological analysis to accurately identify the lemma for each word in textual social media data. We will also convert emoticons, phrasal abbreviations, expressively lengthened words, and slangs into standardized language forms using string and distributional similarity based sequence labeling techniques as well as other hashmap techniques. To do this work, we will draw on Co-I De Choudhury's extensive prior experience in taming social media data.

Theory- and Data-Driven Attribute Extraction from Social Media

We will then extract a number of theoretically- and empirically grounded behavioral and psychosocial attributes from these voluntarily shared social media data of participants.

Mood. To assess mood in participants' social media data, we will use a maximum entropy affect classifier, developed by Co-I De Choudhury (De Choudhury, Gamon, Counts, & Horvitz, 2013), to infer a distribution of emotional states in the unstructured data (e.g., 'fear', 'sadness', 'fatigue'). This classifier characterizes human mood on Twitter via 11 affective classes, and uses 200 explicit mood words (validated through a crowdsourcing approach) to be supervising signals for inferring affect in textual data such as social media posts.

Psychological Arousal. Psychological arousal is a key aspect of mental health experiences -- negative valence and high arousal together are an indicator of stress (Coyne, 1991). Therefore, to further capture the psychological arousal of mood in a participant's social media data, extending Co-I De Choudhury's prior work (De Choudhury, Gamon, & Counts, 2012), we will use the widely used word embedding technique (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013) and the Affective Norms for English Words (ANEW) lexicon (Nielsen, 2011).

Stress and Anxiety. Since stress and anxiety are both causes, mediators, as well as effects associated with mental health states, we will use linguistic attributes that identify commonly associated mental health symptomatic expressions of stress and anxiety. To estimate this, we will replicate the transfer learning classifiers built in the team's recent research (Saha, et al., 2019). These classifiers (Support Vector Machines) are trained on those data from Reddit communities that are most closely associated with each of these mental health conditions/symptoms and show a mean accuracy ranging between ranging between 0.81 and 0.91. We will use these classifiers to machine label the textual content of our social media archives with the presence (or absence) of expressions corresponding to the above two mental health attributes.

Interpersonal Focus. We will capture self-referentiality and interpersonal awareness through interpersonal pronoun use in the textual social media data, for which we will employ the psycholinguistic lexicon called LIWC, or Linguistic Inquiry and Word Count (Pennebaker, Boyd, Jordan, & & Blackburn, 2015).

Hopelessness. Future orientation indicates a tendency to be hopeful, whereas ruminating over the past, especially over negative experiences and events, is considered less emotionally healthy. The three types of tenses -- past, present, and future will be captured by using LIWC.

Circadian Rhythms. Sleep disturbances are common in individuals at risk of mental health challenges. Extending Co-I De Choudhury's prior work (De Choudhury, Gamon, Counts, & Horvitz, 2013), we will extract the normalized patterns of social media activity over a 24-hour cycle or a 7-day week, by using the activity and engagement data in social media, alongside temporal data. We will calculate the entropy and variance of these diurnal/weekly patterns to understand the extent of disruption.

Social Capital. Previous research has found an inverse relationship between an individual's access to cognitive social capital (social trust, sense of belonging, mutual aid) and mental illness (Cohen & Wills, 1985). We will leverage the following types of measures of an individual's social capital, based on our team's prior work (De Choudhury, Counts, Horvitz, & Hoff, 2014): likes and comments on status updates made by a participant; posts/media and associated likes and comments on timeline posts made by friends on participant's timeline; likes and comments on posts and media with specific friends tagged; and number of friends in which a participant had directed communication.

Social Support. A number of studies have demonstrated that the adequacy of social support is directly related to the severity of psychological symptoms and/or acts as a buffer between distressful events and stress (Cohen & Wills, 1985). These observations have held true in studies of online social platforms as well. Accordingly, drawing on the prior work of the team (Sharma & De Choudhury, 2018), we will leverage pre-trained classifiers that assess a message received by an individual on such a social media platform to be either emotional, or informational, or neither.

Social Proclivity and Orientation. Activities on social media sites can serve as a way to understand an individual's social orientation and social functioning, such as whether they seek out for social interaction, events, or activities. To capture this attribute in our participant pool, we will capture social interactions in all of the social media platforms based on the network data. We will assess the volume and frequency of pairwise interpersonal interactions, as well as measure social network attributes like network size, density, betweenness, closeness and eigenvector centrality of the participant, network constraint, embeddedness with social ties, clustering coefficient, and a measure of the strength and reciprocity of social ties of the participants across social media platform.

Beyond these, to allow data-driven exploration of the social media content, for each post, we will extract a variety of linguistic attributes that will serve as features in the ensuing machine learning models. By employing natural language analytic techniques on the social media posts; we will extract *n*-grams (*n*=1, 2, 3), psycholinguistic attributes based on LIWC, sentiment based on the Stanford CoreNLP toolkit (Manning, et al., 2014), word and document embeddings (such as Word2Vec (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013) and Doc2Vec (Wang, Tang, Aggarwal, & Liu, 2016)) that represent latent semantics in text, and narrative frames (Swayamdipta, Thomson, Dyer, & Smith, 2017) that would allow us to capture everyday language and the rich variation in writing styles of social media users.

Machine Learning Modeling

Having extracted these attributes, we will now use machine learning to explore to what extent these attributes can help predict the self-reported responses of the participants in the various surveys/questionnaires. To do so, given the large number of attributes, we will first perform appropriate tests of collinearity. Then we will employ various feature selection and dimensionality reduction techniques, tested in our team's prior work (Saha, et al., 2019), in order to prevent overfitting and ensure model stability and robustness. These include: a) Selecting Features on Coefficient of Variation: First, we will reduce the feature space on the basis of explained variance using the measure of coefficient of variation, that essentially quantifies the ratio of standard deviation to the mean for each feature. b) Selecting Features on Pairwise Correlations: Correlated features typically affect or distort machine learning prediction models by potentially yielding unstable solutions or masking the interactions between significant features. With a suitably chosen threshold absolute value, we will drop those features that are highly correlated with another feature. c) Transforming Features using Principal Component Analysis: This method will reduce the dimensions in the feature space by transforming features into orthogonal or principal components.

We will begin with simpler, more interpretable supervised machine learning techniques such as Naive Bayes, Random Forest, and Support Vector Machine/Regression to train and test classifiers and regression models, based on the type of survey responses and type/granularity of prediction desired. We will then progressively explore the efficacy and suitability of more sophisticated models: a) Gradient boost technique on an ensemble of decision tree classifiers/regressors (Dietterich, 2000); and b) Deep neural networks where we will use the multi-layered perceptron (MLP) technique that works in a feed-forward fashion (no cycles) with multiple internal layers (Haykin, S. 1994). For all of these models, we will employ bootstrapping and *K*-best univariate statistical scoring models using mutual information, to obtain the relative importance among features, and establish their statistical significance using ANOVA and selection probability by applying LASSO shrinkage (Tibshirani, 1996).

Model Evaluation

For all of the above supervised machine learning models, to assess concurrent validity of the inferences/predictions, we will use measures of model fit (e.g., the Widely Applicable Information Criterion, Hosmer-Lemeshow goodness of fit) and statistical power (e.g., effect size). Then, we will examine the internal validity of the inferences based on their error rate, bias, and inferential ability (e.g., accuracy, root mean squared error). Sensitivity, specificity, positive predictive value, and negative predictive value will be calculated for various cut-offs using a Receiver Operating Characteristic (ROC) analysis and observing the area under the curve. We will also use *k*-fold cross validation for parameter tuning, which will allow us to assess performance on a held-out sample. We will examine improvements in more sophisticated models over simpler models using the log likelihood ratio test and other measures of incremental validity.

Significance and Expected Impact

Regular community climate and needs assessments, informative public messaging and surveillance, and reciprocal discussions such as town halls, can help communities to develop appropriate and effective plans, communications messaging, training, and resources; and foster an informed and vigilant re-entry to reasonable daily activities that continues to keep public health and safety at the forefront of this process. This study will: 1) provide rapid data to better understand the range of immediate, short-term, and potentially long-term impacts of a pandemic on healthcare and first responder community members, and 2) demonstrate the use of integrating knowledge from online interview data and open source social media to drive real-time impact analysis that will inform time-sensitive community interventions and training. Based on the findings from this study, additional studies can be developed in order to maximize understanding of these subject areas.

APPENDIX: References

Bagroy, S., Kumaraguru, P., & De Choudhury, M. (2017). A social media based index of mental well-being in college campuses. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1634–1646). ACM.

Bedford, J., Enria, D., Giesecke, J., Heymann, D. L., Ihekweazu, C., & Kobinger, G. (2020). COVID-19: towards controlling of a pandemic. *Lancet, 395*, 1015-1018. Retrieved from https://doi.org/10.1016/50140-6736(20)30673-5

Birnbaum, M. L., Ernala, S. K., Rizvi, A., Arenare, E., Van Meter, A., De Choudhury, M., & Kane, J. M. (2019). Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient- contributed digital data from facebook. *NPJ Schizophrenia, 5*(1), 1-9.

Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: A new approach to personality in a digital world. *Current Opinions in Behavioral Science*, 63-68.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The Lancet, 395*, 912-920. Retrieved from www.thelancet.com

Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: psychometric properties of the English version. *Behaviour Research and Therapy, 40*, 931–945.

Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the Brief COPE. *International Journal of Behavioral Medicine, 4*(1), 92-100.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing Coping Strategies: A Theoretically Based Approach. *Journal of Personality and Social Psychology, 56*(2), 67-283.

Center for Disease Control. (2020). *Coronavirus Disease 2019 (COVID-19): Stress and Coping.* Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological bulletin, 98*(2), 310.

Conway III, L. G., Woodard, S. R., & Zubrod, A. (2020). Social Psychological Measurements of COVID-19: Coronavirus Perceived Threat, Government Response, Impacts, and Experiences Questionnaires [preprint].

Coyne, J. C. (1991). Social factors and psychopathology: Stress, social support, and coping processes. *Annual review of psychology, 42*(1), 401-425.

Cramer, P. (1996). *Storytelling, Narrative, and the Thematic Apperception Test.* New York, NY: Guilford Press.

Crooks, C. L. (2020, manuscript in preparation). Influence and Vulnerability in the Information Environment: Implications for Cyber-Enabled Information Operations and National Security. In M. (. Kosal, *Innovate for Future Threats: Disruptive Innovation Efforts and Uses of the Technology Environment by State and Non-State Actors.*

Crooks, C. L., Muchlinski, D., Cross, E., Jurczyk, K., Martin, S., & Srivastava, R. (2020, manuscript in preparation). *Qualitative analysis of social response to COVID-19 outbreak through Twitter observational study.*

De Choudhury, M., & De, S. (2014). Mental health discourse on reddit: Self-disclosure, social support, and anonymity. *Eighth international AAAI conference on weblogs and social media.*

De Choudhury, M., & Kiciman, E. (2017). The language of social support in social media and its effect on suicidal ideation risk. *ICWSM*, 32-41.

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behavior via social media. *Proceedings of the 2013 ACM annual conference on Human factors in computing systems* (pp. 3267–3276). ACM.

De Choudhury, M., Counts, S., Horvitz, E., & Hoff, A. (2014). Characterizing and predicting postpartum depression from facebook data. *In Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing.* ACM.

De Choudhury, M., Gamon, M., & Counts, S. (2012). Happy, nervous or surprised? Classification of human affective states in social media. *In Sixth International AAAI Conference on Weblogs and Social Media.*

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *AAAI Conference on Weblogs and Social Media.*

De Choudhury, M., Morris, M. R., & & White, R. W. (2014). Seeking and sharing health information online: comparing search engines and social media. *Proceedings of the SIGCHI conference on human factors in computing systems*, (pp. 1365-1376).

DeChoudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Humar, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. *Mental Health in Technology Design and Social Media*, 2098-2110.

Del Vicario, M., Bessi, A., Zollo, F., Petronic, F., Scala, A., Caldarelli, G., . . . Quattrociocchi, W. (2016). The spreading of misinformation online. *PNAS, 113*(3), 554-559. doi:doi/10.1073/pnas.1517441113

Dietterich, T. G. (2000). Ensemble methods in machine. *Multiple classifier systems, 1857*, 1-15.

Goldenberg, A., & Finkelstein, J. (2020). *Cyber Swarming, Memetic Warfare and Viral Insurgency: How Domestic Militants Organize on Memes to Incite Violent Insurrection and Terror Against Government and Law Enforcement.* Network Contagion Research Institute.

Grasso, D., Briggs-Gowan, M., Ford, J., & Carter, A. (2020). *The Epidemic – Pandemic Impacts Inventory (EPII).*

Haykin, S. (n.d.). *Neural networks: a comprehensive foundation.* 1994: Prentice Hall PTR.

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., & Ballard, C. e. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *The Lancet*, 1-14. Retrieved from https://doi.org/10.1016/S2215-0366(20)30168-1

IASC Reference Group on Mental Health and Psychosocial Support in Emergency Settings. (2020). Addressing Mental Health and Psychosocial Aspects of COVID-19 Outbreak (Version 1.5).

Kang, L., Ma, S., Chenb, M., Yangb, J., & Wang, Y. e. (2020, in press). Impact on mental health and perceptions of psychological care among medical and nursing staff in Wuhan during the 2019 novel coronavirus disease outbreak: A cross-sectional study. *Brain, Behavior, and Immunity*, xxx-xxx.

Kim, S., & So, J. (2018). How message fatigue toward health messages leads to ineffective persuasive outcomes: Examining the mediating roles of reactance and inattention. *Journal of Health Communication, 23*, 109-116. doi:10.1080/10810730.2017.1414900

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. *Paper presented at the Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations.*

Matheson, K., & Anisman, H. (2003). Systems of Coping Associated with Dysphoria, Anxiety and Depressive Illness: A Multivariate Profile Perspective. *6*(3), 223-34. doi:10.1080/10253890310001594487

McAdams, D. P. (1988). *Power, Intimacy, and the Life Story.* New York, NY: Guilford Press.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 3111–3119.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *In Advances in neural information processing systems, pages* , 3111–3119.

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.

Oxford Anaytica. (2020). *Misinformation will undermine coronavirus responses*. Retrieved from Expert Briefings: https://www.emerald.com/insight/content/doi/10.1108/OXAN-DB250989/full/html

Paul, C., & Matthews, M. (2016). *The Russian "Firehose of Falsehood" Propaganda Model.* RAND.

Peacock, E. J., & Wong, P. T. (1990). The Stress Appraisal Measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine, 6*, 227-236. Retrieved from http://www.drpaulwong.com/wp-content/uploads/2018/03/Stress-Appraisal-Measure-SAM-Peacock-Wong-1990-Scale.pdf

Pennebaker, J. W., Boyd, R. L., Jordan, K., & & Blackburn, K. (2015). The development and psychometric properties of LIWC2015.

Pfefferbaum, B., & North, C. S. (2020). Mental Health and the Covid-19 Pandemic. *The New England Journal of Medicine*, 1-3.

Rosmarin, D. H. (2020, March 20). What's Scarier than the Coronavirus? . *Scientific American: Observations [electronic version]*. Retrieved from https://blogs.scientificamerican.com/observations/whats-scarier-than-the-coronavirus/?utm_source=newsletter&utm_medium=email

Saha, K., Bayraktaroglu, A. E., Campbell, A. T., Chawla, N. V., De Choudhury, M., D'Mello, S. K., & ... & Mark, G. (2019). Social media as a passive sensor in longitudinal studies of human behavior and wellbeing. *In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, (pp. 1-8).

Saha, K., Kim, S. C., Reddy, M. D., Carter, A. J., Sharma, E., Haimson, O. I., & De Choudhury, M. (2019). The language of lgbtq+ minority stress experiences on social media. *Proceedings of the ACM on Human-Computer Interaction, 3(CSCW)*, 1-22.

Saha, K., Reddy, M. D., das Swain, V., Gregg, J. M., Grover, R., Lin, W., . . . Mulukutla, R. (2019). Imputing missing social media data stream in multisensor studies of human behavior. *8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 178–184.

Sharma, E., & De Choudhury, M. (2018). Mental health support and its relationship to linguistic accommo- dation in online communities. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, (pp. 1–13).

Sohrabia, C., Alsafib, Z., O'Neilla, N., Khanb, M., Kerwanc, A., Al-Jabir, A., . . . Aghad, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery, 76*, 71-76.

Substance Abuse and Mental Health Services Administration. (2020). *Coping With Stress During Infectious Disease Outbreaks.* Retrieved from https://store.samhsa.gov/product/Coping-with-Stress-During-Infectious-Disease-Outbreaks/sma14-4885

Substance Abuse and Mental Health Services Administration. (2020). *Tips For Social Distancing, Quarantine, And Isolation During An Infectious Disease Outbreak.* Retrieved from https://store.samhsa.gov/product/tips-survivors-coping-anger-after-disaster-or-other-traumatic-event/pep19-01-01-002

Swayamdipta, S., Thomson, S., Dyer, C., & Smith, N. A. (2017). Frame-semantic parsing with softmax- margin segmental rnns and a syntactic scaffold. *arXiv preprint arXiv:1706.09528*.

Taha, S., Matheson, K., Cronin, T., & Anisman, H. (2014). Intolerance of uncertainty, appraisals, coping, and anxiety: The case of the 2009 H1N1 pandemic. *British Journal of Health Psychology*, *19*, 592–60.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society., Series B (Methodological)*, 267–288.

Torales, J., O'Higgins, M., Castaldelli-Maia, J. M., & Ventriglio, A. (2020). The outbreak of COVID-19 coronavirus and its impact on global mental health. *International Journal of Social Psychiatry*, 1-4. doi:10.1177/0020764020915212

Van Bavel, J. J. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature: Human Behavior*. Retrieved from https://doi.org/10.1038/s41562-020-0884-z

Wang, S., Tang, J., Aggarwal, C., & Liu, H. (2016). Linked document embedding for classification. *Proceedings of the 25th ACM international on conference on information and knowledge management*, (pp. 115-124).