“Maybe You Should Talk to Someone”: The Role of Online Communities on Mental Healthcare

Soham Ghosh
Indian Institute of Management (IIM), Calcutta, sohamg17@iimcal.ac.in

Soumyakanti Chakraborty
Indian Institute of Management, soumyakc@iimcal.ac.in

Megha Sharma
Indian Institute of Management, Calcutta, megha@iimcal.ac.in

Anjana Susarla
Michigan State University, asusarla@broad.msu.edu

"Maybe You Should Talk to Someone": The Role of Online Communities on Mental Healthcare

Short Paper

Soham Ghosh
Indian Institute of Management, Calcutta
D. H. Road, Kolkata, W Bengal 700104
sohamg17@iimcal.ac.in

Soumyakanti Chakraborty
Indian Institute of Management, Calcutta
D. H. Road, Kolkata, W Bengal 700104
soumyakc@iimcal.ac.in

Megha Sharma
Indian Institute of Management, Calcutta
D. H. Road, Kolkata, W Bengal 700104
megha@iimcal.ac.in

Anjana Susarla
Eli Broad College of Business, Michigan State University,
East Lansing, Michigan 48824
asusarla@broad.msu.edu

Abstract

Online Health Communities like YouTube offer mental health patients an alternative channel to learn about mental illnesses, the treatment path to follow, and to share their experiences. For many patients who are reluctant to seek professional help, a video on mental health uploaded by a content creator may serve as a substitute for a counsellor. Our work aims to develop an understanding of the relationship between language formality and social support and provide normative guidelines for content creators on social media platforms. Using two transformer-based deep learning classification models, we determine the degree of language formality or informality present in the content, and three dimensions of social support in the comments. We then utilize propensity score estimation to establish the causal effect of (in)formality on the dimensions of social support for 994 videos and 3,10,157 comments. Our findings indicate that informal speech increases emotional support, leading to better health outcomes.

Keywords: Online Health Communities (OHCs), mental health, deep learning, social media, healthcare analytics

Introduction

Mental health patients experience significant psychological vulnerability during interactions with healthcare providers, which leads to a reluctance to seek treatment through the formal healthcare system and avoidance of counselling sessions (Zhou et al., 2022). Online Health Communities (OHCs) offer an alternative way for these patients to learn about mental health and illnesses, the treatment path to follow, and to express their opinions and share their experiences (Bardhan et al., 2020). Research suggests that people often use social media platforms instead of traditional search engines to manage mental health issues (Aref-Adib et al., 2016). A recent study indicates that college students frequently use YouTube to obtain mental health support (Choi et al., 2021). One of the reasons for the growing popularity of YouTube for chronic conditions like mental health issues is the video format which is easily understandable and actionable (Liu et al., 2020).

While content creators on YouTube are quite conversant with the tools and techniques that capture viewers' attention, the language used in the videos is equally important (McLellan et al., 2022). Careful consideration of the language used is particularly important for patients suffering from mental health issues, as complex, inaccessible language discourages patients who are reluctant to seek formal medical advice (Gullslett et al., 2021). One important aspect of disseminating information on social media platforms like YouTube is the stylistic difference in communication (Zhang et al., 2021). In this context, the *formal-informal* dimension is the "most important dimension of variation between styles" (Heylighen & Dewaele, 1999). In the domain of online healthcare, informal medical language is considered accessible, familiar, and understandable (Xie et al., 2021). The usage of medical terms is also found to affect service quality in the context of e-health (Zhang et al., 2021). We consider formality as it is a composite measure consisting of multiple factors like syntactic simplicity, concreteness, and cohesion, which makes it an accurate representation of communication level. It is crucial that the language used in mental health videos is accessible to the patients and increases attention and participation. As a sizable number of mental health patients are young people who tend to rely on social media platforms for their first point of support, it is important that they are able to relate to the language of the content creators.

A video about mental health on YouTube may effectively serve as a substitute for a counsellor, particularly for patients reluctant to seek professional help. It is thus important to determine whether these videos lead to better mental health outcomes. One measure of the effectiveness of social media content in creating better health outcomes is *social support* (Lin & Kishore, 2021), and indeed, OHCs provide the opportunity for patients to give and receive social support (Yan & Tan, 2014). Social support is "information leading the subject to believe he is cared for and loved, esteemed, and a member of a network of mutual obligations" (Cobb, 1976). Social support is given or received through exchanging information and sharing personal experiences and emotions (Lin & Kishore, 2021). It is known to increase well-being and reduce psychiatric symptoms for individuals suffering from serious mental health issues (McCorkle et al., 2008). Extant research on the impact of social support on health outcomes has found that the dimensions of informational and emotional support positively affect patients' mental health (Yan & Tan, 2014). More recent research posits that informational and experiential support are linked to self-care and informational and emotional care are linked to psychological health (Lin & Kishore, 2021). However, to the best of our knowledge, scholars have not studied the antecedents of social support from the content creator's perspective.

The process by which users seek and interact with medical information online is a complex activity with a high cognitive load (Madathil et al., 2015). The choice of language used by the counsellor plays a key role in the perception of the counsellor as empathetic and warm (Lebowitz & Appelbaum, 2019). Past research indicates that informal language builds connections between conversation partners (Tannen, 1991), which in turn leads to perceptions of empathy and warmth. On the other hand, formal language has been found to be better for presenting information (Tannen, 1991). Our work aims to develop an understanding of the relationship between language formality and social support and provide normative guidelines for content creators on social media platforms like YouTube. To the best of our knowledge, this is the first study on social support at the intersection of OHCs and social media from the content creator's perspective. The content creator, who substitutes for a mental health counsellor, would appreciate guidelines on content creation leading to better mental health outcomes. For example, if an expert wants to create a video focusing on self-care for mental health patients, the focus should be on generating informational and experiential support. However, the literature is largely silent on the aspects of the video that would lead to the generation of such support. We therefore aim to address the gap in understanding how the (in)formality of language used by content creators influences the nature of the social support that viewers give or receive on YouTube.

We may infer the (in)formality of language from the speech the speaker(s) use in the posted videos. We collect 1,945 videos related to mental health from YouTube and all the comments in the comment section where available, for a set of 1,311 videos. We then convert the 1,311 audio clips obtained from the videos from speech to text using a pre-trained transformer-based deep learning model known as Whisper, developed by OpenAI (Radford et al., 2022). We classify the text into formal and informal dimensions using another fine-tuned transformer-based deep learning model based on the BERT (Binary Encoder Representation of Transformers) architecture developed by (Devlin et al., 2018). Further, we classify 5,58,176 comments from 1,311 videos into the three dimensions of social support given or received by commenters using an annotated transformer-based deep learning model known as RoBERTa (a variation of the BERT architecture) (Liu et al., 2019). Finally, we employ a propensity score matching approach with 994 matched videos and 3,10,157 matched comments to assess the causal effect of (in)formal speech on the kinds of social support given or received. We estimate the causal effect of (in)formal language based on the Average Treatment Effect (ATE) obtained for each dimension of social support – informational, experiential, and emotional. We find that videos with more informal speech are likely to generate more comments that seek or provide emotional support. We do not find any significant relationship between (in)formality and experiential or informational support, indicating that viewers share their lived experiences and information regardless of the content creator's communication style. Our preliminary results suggest that content creators need to change their speech patterns based on the needs of viewers of mental health channels. Informal speech may lead to higher sharing of emotional support and is likely to translate into better health outcomes for the viewer (Yan & Tan, 2014).

Theoretical Background

Communication Styles in OHCs

Existing design science research has examined the stylistic difference in language through the lens of the *Language-Action Perspective* (LAP) for organizational communication (Flores & Ludlow, 1980), which facilitates sense-making in IS through the classification of speech acts (Abbasi et al., 2018). We study the operationalization of the speech act classification in the context of OHCs by classifying the content creators' speech acts on YouTube based on the stylistic concept of formality.

Embedded medical information in the form of text is useful in determining health outcomes on OHCs (Xie et al., 2021). The vocabulary used in medical information includes formal and informal medical language, leading to different degrees of understandability and trustworthiness, which may indirectly affect health outcomes. An informal communication style is "common, non-official, familiar, casual, and often colloquial, and contrasts in these senses with formal" (Gretry et al., 2017). Formal language is perceived to be credible, authentic, and demonstrates the expertise of the poster (content creator) (Rennekamp & Witz, 2021), while informal language is considered accessible, familiar, and understandable (Majone, 1997). The effect of language (in)formality is different in various contexts. Informal language in social media increases user engagement (Rennekamp & Witz, 2021). On the other hand, formal language is associated with fewer misunderstandings (Heylighen & Dewaele, 2002). In a study on a healthcare forum, the authors note that informal language leads to more helpful information being present on the OHC (Xie et al., 2021). Informal communication is typically perceived as empathetic and prosocial (Marín-López et al., 2019). This perception is supported by the counselling literature, which posits that empathy and warmth are important factors in successfully counselling a patient's mental health disorder (Lebowitz & Appelbaum, 2019). Counselling styles manifesting these factors increase the number of patients opting to visit a mental health counsellor (Zhou et al., 2022). Informal communication can thus lead to better outcomes for mental health counselling. While there is substantial literature on how linguistic differences in written communication can affect health outcomes, we note a lack of studies dealing with video-based communication in online mental healthcare. The language used by content creators, who function as mental health counsellors in this context, may affect health outcomes via social support by displaying characteristics that viewers of these videos prefer.

Mental Health Communication and Social Support in OHCs

Although past studies have suggested that informational and emotional support are the two most useful types of social support for improving health-related outcomes (Yan & Tan, 2014), a recent study posits that in addition to these two types, experiential support is also an important driver of health outcomes on social media (Lin & Kishore, 2021). We thus consider the three dimensions of social support as a representation of health outcomes in the context of this study. Emotional support plays a major role in improving patients' mental health outcomes and is found to be more important than informational support in such contexts (Yan & Tan, 2014). Experiential support has not been studied as extensively. However, people with lived experiences of mental health issues have been found to be positive contributors to collaborative care (Vojtila et al., 2021). Existing studies on OHCs have used methods for detecting the dimensions of social support based on word embeddings (Mikolov et al., 2013). However, classification models for mental health support detection based on more recent advancements in natural language processing, for example, transformers, are comparatively less. We use deep learning techniques to extract three dimensions of social support from user comments on YouTube videos about mental health. Our contribution to each aspect of the OHC literature is summarized in Table 1 below.

A higher degree of informality in the speaker's communication style indicates the usefulness of the information available on OHCs (Xie et al., 2021), which may affect social support. Changes in the speaker's communication style may thus affect the social support given or received by patients. The impact of communication style has been studied in the organizational behaviour literature, with studies indicating that informal interaction provides a basis for giving or receiving more social support (Fay, 2011). Communication competence moderates the impact of emotional support on psychological quality of life, showing that a variation in communication style impacts social support (Yoo et al., 2014). Variations in social support dimensions have different effects on improving health outcomes (Yan & Tan, 2014). Understanding the changes in the dimensions of social support enables us to better understand the connection between communication styles (formal/informal) and health outcomes. Content creators may observe different levels of support across dimensions, thus impacting the engagement and health outcomes based on those videos. Content creators who use informal language in such settings may be perceived as more empathetic and pro-social will lead to patients giving and receiving more emotional support. Formal language utilization is considered more reliable and authentic on OHCs and other social media platforms. As a result, patients seeking medical information or lived experiences of others dealing with mental health issues may look for videos which indicate a higher degree of authenticity. We thus hypothesize:

H1: Informal communication style of content creators leads to a greater degree of emotional support

H2: Formal communication style of content creators leads to a greater degree of informational support

H3: Formal communication style of content creators leads to a greater degree of experiential support

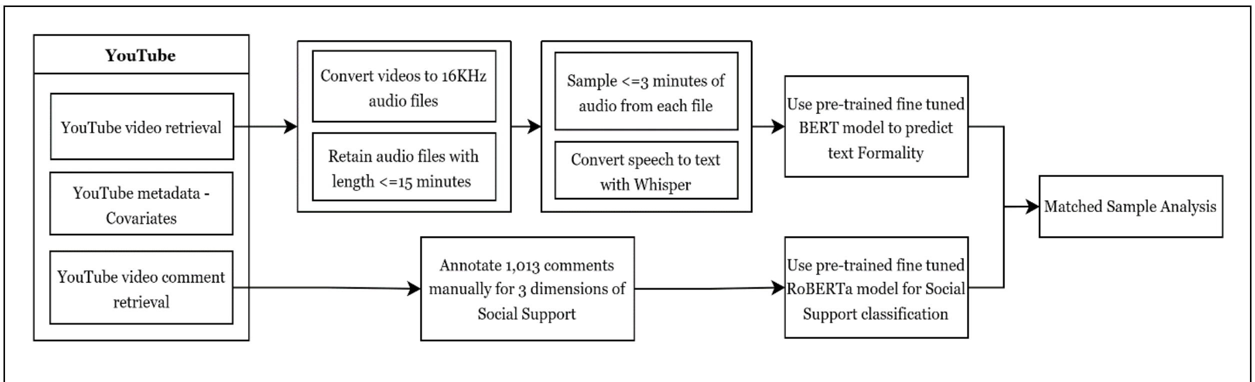

Figure 1: Research Approach

Data and Methods

Data Retrieval and Conversion to Text

We obtain data from the YouTube platform and ensure that the videos collected for the study accurately represent how people look for mental health videos and where support for mental health issues is likely to be found. We select all existing videos from 3 channels listed in the top 10 mental health channels based on popularity on YouTube[1], for a total of 1,945 videos. Content creators on these videos speak about mental health issues, diagnoses, and discuss how to deal with and mitigate such issues and their impact on people's lifestyles. Among the 3 channels, two are run by content creators who are registered clinical counsellors. The third is an organization consisting of licensed counsellors who also function as content creators. These 1,945 videos are downloaded and converted to audio files. We consider short audio files of 15 minutes or less to address the challenges of processing large files. This leads to a set of 1,311 audio files which have comments. We sample 20% of the time duration of each file to obtain a 3-minute (or smaller) sample of each audio file. As most videos have a single speaker, we can infer the presence of (in)formal speech from 3 minutes (or less) of each audio file. The audio files are sampled such that the first 30 seconds, last 30 seconds and 2 minutes centered around the middle are extracted to ensure sampling from different portions

[1] Best Mental Health YouTube Channels (healthcoachinstitute.com)

of the audio file. We use a recent model developed by the OpenAI research group, Whisper (Radford et al., 2022), to convert the sampled audio data to transcript text. We present the research approach in Figure 1.

Covariates – Video Level Measures

The properties of a video uploaded on YouTube may influence the social support given or received for the video. These measures are collected from the video metadata to ensure that confounders do not affect the causal model as per the research approach in Figure 1. The descriptive statistics of the metadata measures are shown along with a brief description of the measures in Table 1.

Variable	Description	Mean	Std. Dev
Age of video (days)	Number of days since the video was posted	2204.83	1077.67
No. of views	Number of views received by the video	110094.16	251683.01
No. of comments	Number of comments received by the video	446.76	970.61
No. of likes	Number of likes received by the video	3475.56	7066.26
Description Length (words)	Length of the description of the video	200.58	79.31
Title Length (words)	Length of the title of the video	9.18	3.59
No. of likes for comments	Number of likes for comments on the video	3293.38	10006.05
No. of comments with replies	Number of comments with replies	152.82	360.89
No. of replies	Number of replies to comments	168.18	436.81
Channel No.	Which channel the video belongs to	2.04	0.41

Table 1: Summary statistics and variable definitions

Formal Language Classification

We then determine the degree of (in)formality present in the text transcripts of the videos. We utilize data from a study on formality detection by (Pavlick & Tetreault, 2016) that analyses linguistic formality, building on multiple annotated corpora of formality in several types of data, specifically from professional emails, sentences from Yahoo Answers, blogs, and community Question&Answer forums, for a total of 11,119 labelled data points. We pre-process and re-train the existing labelled corpus for use as the ground truth of our study using a state-of-the-art model described below using a transfer learning paradigm. BERT (Bidirectional Encoder Representations from Transformers) is a family of models developed by Google (Devlin et al., 2018), pre-trained on language modelling and next-sentence prediction tasks. It outperforms existing models on a wide variety of benchmark datasets. A common approach in the machine learning literature is to use a BERT model variation, pre-trained on a large corpus and fine-tune the same on a downstream NLP task of interest, for example, text classification (Yang et al., 2022).

We proceed along similar lines, using a pre-trained BERT (Base, Cased) model, calculating the loss and accuracy metrics by fine-tuning the model on the annotated data from the previous study with a 20% validation split. We obtain an accuracy of 82.45% on the data. Applying this trained model to the unlabeled text, we obtain a set of (in)formality scores for the 1,311 videos. The scores are converted to 1 for formal text and 0 for informal text. The results are validated by annotating a subset of 100 formality scores by three teaching associates with post-graduate degrees at a reputed business school in India. The associates independently label the video speech as formal or informal, and we obtain an inter-rater reliability score of 0.887. For example, "…*When *** did her study she thought that two thirds of people would have this attachment honestly I think as further research has come it probably not true…*." is classified as formal, while "*And we super fancy under these bright lights and I sweating my ***** off and her husband is filming and I just met him today…*" is classified as informal language. Based on the scores, the number of formal and informal videos is illustrated in Table 2.

Social Support Classification

We also estimate the dimensions of social support given or received by viewers/commentors on these videos. In this study, we examine three categories of social support – informational, experiential, and emotional support. We obtain 5,58,176 comments from the 1,311 videos. Since there are no publicly available models for estimating social support in YouTube comments, we randomly sample 1,013 comments and manually annotate them as seeking or receiving informational, emotional, and experiential support.

		Number of comments (Social Support)		
Video Type	Number of Videos	Emotional	Experiential	Informational
Formal	814	136813	144227	147750
Informal	497	54016	38938	36432
Total	1311	190829	183165	184182

Table 2: Number of formal/informal videos on channels and social support dimensions

We fine-tune the existing RoBERTa (Liu et al., 2019) implementation on the downstream task of classifying text into one of three dimensions of social support. Social support is operationalized by enumerating each dimension of support given/received for a video. Three teaching associates with post-graduate degrees at a reputed business school in India are asked to go through the comments and annotate them as seeking or receiving informational, emotional, and experiential support. The inter-rater reliability is found to be 0.938 for 1013 comments. The 1,013 labelled comments are then passed to the pre-trained RoBERTa classifier, with a 20% validation split, and the model is fine-tuned on these records. We obtain an accuracy of 82.76% on these records. The fine-tuned model is then applied to the larger dataset of comments. We then summarize the number of comments with each dimension of support at the video level, which is our unit of analysis. We thus obtain a total of 5,58,176 classified comments for 1,311 videos. The number of comments of each support type is shown in Table 2. For example, "*…I suppose in this sense I'm an extrovert in that I find I need people to talk to…*" is classified as a comment on emotional support, while "*I think it's important to take responsibility for yourself but that includes asking for help…*" is classified as a comment on informational support. Similarly, "*I spent much of my teen years & my early 20s working on lighthouses, I'm ok with the social distancing…*" is classified as a comment on experiential support.

Propensity Score Matching and Estimation

We are interested in the causal relationship between the (in)formality of the speech in videos and different dimensions of social support, namely, informational, emotional, and experiential support. We consider a video that uses formal speech to be the treatment, and the change in the different dimensions of social support as the treatment effect. In a scenario like uploading videos, observing the counterfactual directly is not possible, as it would require us to randomize the treatment across videos by asking content creators to change their speech pattern and measuring the impact on the different dimensions of social support. We thus measure the actual causal impact using statistical methods for inferring causality from observational data. The endogeneity between types of social support given or received and (in)formality of the video speech is an issue in causal estimation. External factors may affect the content of the video being posted, like the nature of the channel posting the video. We control for covariates based on availability of data on YouTube to minimize the risk of endogeneity in this context. We conduct propensity score matching to model a video's propensity to use speech with a high level of formality. We conduct a matched sample analysis using optimal matching (Hansen & Olsen Klopfer, 2012) with a 2:1 ratio. We report the standardized mean differences (SMD) for each variable in the treatment and control group pre-and post-matching in Table 3. Since we have only three pre-treatment variables (Description length, Title length, and Channel No.), we augment them with seven post-treatment variables which function as surrogates for pre-treatment variables. Augmentation is recommended by (Rosenbaum, 1984) only in cases like ours where surrogates are required, or post-treatment variables show SMDs beyond the threshold even after matching. We note that the absolute value of the SMD pre-matching for most of the variables is greater than 0.20 (Faraone, 2008), indicating an imbalance in the dataset.

Variables	SMD (Pre-matching)	SMD (Post-matching)
Number of Records	1311	994
Age of video	0.956	0.372
No. of views	0.280	0.169
No. of likes	0.328	0.179
Description length	0.416	0.070
Title length	0.325	0.195
No of likes for comments	0.250	0.128
No. of comments with replies	0.356	0.190

No. of replies	0.339	0.177
Channel No.	0.316	0.085

Table 3: Pre- and Post-Matching SMD

The SMD improves on conducting the matching process with a sample of 994 matched videos, where all but two of the variables have an SMD of less than 0.20. The presence of a post-treatment matched variable with an SMD>0.20 further supports the need for the augmentation as suggested by (Rosenbaum, 1984). We proceed to match on propensity scores to determine the effect of (in)formal speech on the dimensions of social support. Table 4 shows the estimated treatment effect of the formal-informal dimension on the three dimensions of social support. We find that formal speech used in videos significantly and negatively impacts emotional support. We also find that experiential and informational support are not significantly impacted by the (in)formality of speech. Videos with more informal speech are thus likely to generate more comments giving/receiving emotional support. This outcome is based on 994 matching samples, which is 75.81% of the data. This is due to the limitation of matching methods (without replacement) based on propensity scores, which discard the unmatched data prior to estimating the treatment effect.

Dimensions of Social Support	Average Treatment Effect	P-value
Emotional	-35.530	$p<0.01$
Informational	6.815	$p>0.05$
Experiential	-0.890	$p>0.1$

Table 4: Estimated Treatment Effect using Propensity Score Matching

Initial Contributions and Future Work

To the best of our knowledge, our study is the first to develop a model to determine the role of the content creators' communication style on the dimensions of social support given or received on YouTube. Our NLP-based approach to studying (in)formality may be useful for researchers conducting data-driven studies on online video-based communication and researchers developing social support detection mechanisms. Researchers can also study the effect of the (in)formality of the language used in workplace settings requiring remote communication through video platforms. A better understanding of the role of content creators' language extends the literature on the role of OHCs in addressing the challenges of mental health. Researchers may also find this causal framework useful for understanding the causal impact of (in)formality on social support in other social media contexts. To highlight the impact of our work, we provide a summary of our contribution to each stream of literature in the context of IS research in Table 5 below.

Stream of Literature	Themes	Our Contribution
Online Communication Style	Understanding cues in communication, role of language in organizations, provides perceptions of empathy and warmth	Classifying speech using (in)formality in the context of video communication related to mental health
OHCs	Social support on OHCs, improves mental and physical health outcomes, improves treatment adherence	Classifying user comments into informational, emotional, and experiential support to understand expert/peer-based support in the domain of online mental healthcare
Causal inference	Causal framework which impacts social support, and by extension, health outcomes	PSM-based framework to understand the role of (in)formality on social support

Table 5: Contributions to Literature

The results can inform researchers and practitioners about the role of the content creator in shaping the type of support these platforms foster. Content creators can treat these results as a guideline for making videos that positively affect health outcomes, thus increasing their reach and popularity. While many content creators uploading such content are healthcare professionals, their credentials are often not visible to prospective viewers of such videos. The perceived lack of formal training in the field of mental health may impede empathic communication between the content creator and viewers. Our work can provide YouTube moderators with the means to improve the creators' credibility by helping to develop guidelines for uploading content on mental health. YouTube moderators can also use this framework to filter videos which adversely affect viewers' mental health and remove them from their platform. This adds an additional layer of safety for psychologically vulnerable viewers. However, while we have focused on the language of the video in our study, emotional valence, length of the video, expertise, and gender impact the number of comments (Munaro et al., 2021; Tsou et al., 2014). Further, our assumptions notwithstanding, there may be unobserved confounding variables which affect the causal estimation process, such as the expertise of the content creator (e.g., experienced content creators may be more effective in communicating on YouTube), which need to be accounted for. We plan to consider these aspects in an expanded version of this study.

The utility of speech in creating actionable results has been studied through the Language-Action Perspective (LAP) in the IS literature (Abbasi et al. 2018). We plan to employ the LAP framework and provide a theoretical grounding for the phenomenon that we have studied in this research. Further, the propensity score matching method has well-known limitations, like reduced sample size (75.81% in this study) and prior functional form assumptions (Steiner & Cook, 2013). More recent causal estimation approaches utilize causal forests and metalearners, which calculate the Conditional Average Treatment Effect (CATE) rather than the ATE in propensity score matching. These approaches are robust to imbalanced data and do not assume prior functional forms. We thus plan to evaluate our data based on these methods to obtain more relevant and robust insights into the role of (in)formality on social support.

References

Abbasi, A., Zhou, Y., Deng, S., & Zhang, P. (2018). Text analytics to support sense-making in social media: a language-action perspective. *MIS Quarterly, 42*(2), 427-464.

Aref-Adib, G., O'Hanlon, P., Fullarton, K., Morant, N., Sommerlad, A., Johnson, S., & Osborn, D. (2016). A qualitative study of online mental health information seeking behaviour by those with psychosis. *BMC Psychiatry, 16*(1), 232.

Bardhan, I., Chen, H., & Karahanna, E. (2020). Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management. *MIS Quarterly, 44(1)*, 185–200.

Choi, B., Kim, H., & Huh-Yoo, J. (2021). Seeking mental health support among college students in video-based social media: Content and statistical analysis of YouTube videos. *JMIR Formative Research, 5*(11), e31944.

Cobb, S. (1976). Social support as a moderator of life stress. *Psychosomatic Medicine, 38(5)*, 300–314.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint* arXiv:1810.04805.

Faraone, S.V. (2008). Interpreting estimates of treatment effects: Implications for managed care. *Pharmacy and Therapeutics, 33*(12), 700.

Fay, M. J. (2011). Informal communication of co-workers: A thematic analysis of messages. *Qualitative Research in Organizations and Management: An International Journal, 6*(3), 212–229.

Gretry, A., Horváth, C., Belei, N., & van Riel, A. C. R. (2017). "Don't pretend to be my friend!" When an informal brand communication style backfires on social media. *Journal of Business Research, 74*, 77–89.

Gullslett, M. K., Kristiansen, E., & Nilsen, E. R. (2021). Therapists' experience of video consultation in specialized mental health services during the COVID-19 pandemic: Qualitative interview study. *JMIR Human Factors, 8*(3), e23150.

Hansen, B. B., & Olsen Klopfer, S. (2012). Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics, 15(3)*, 609–627.

Heylighen, F., & Dewaele, J.M. (1999). Formality of Language: definition, measurement and behavioral determinants. *Interner Bericht, Center "Leo Apostel", Vrije Universiteit Brüssel, 4*.

Heylighen, F., & Dewaele, J.M. (2002). Variation in the contextuality of language: An empirical measure. *Foundations of Science, 7*(3), 293–340.

Lebowitz, M. S., & Appelbaum, P. S. (2019). Biomedical explanations of psychopathology and their implications for attitudes and beliefs about mental disorders. *Annual Review of Clinical Psychology, 15*, 555–557.

Lin, X., & Kishore, R. (2021). Social media-enabled healthcare: A conceptual model of social media affordances, online social support, and health behaviors and outcomes. *Technological Forecasting and Social Change*, *166*, 120574.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pre-training approach. *arXiv preprint* arXiv:1907.11692.

Liu, X., Zhang, B., Susarla, A., & Padman, R. (2020). Go to YouTube and call me in the morning: Use of social media for chronic conditions. *MIS Quarterly, 44*(1)*,* 257-283.

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S., & Gramopadhye, A. K. (2015). Healthcare information on YouTube: A systematic review. *Health Informatics Journal*, *21*(3), 173–194.

Majone, G. (1997). The new European agencies: Regulation by information. *Journal of European Public Policy*, 4(2), 262–275.

Marín-López, I., Zych, I., Monks, C. P., & Ortega-Ruiz, R. (2019). Empathy, morality and social and emotional competencies in interpersonal interactions online. In M. Coetzee (Ed.), *Thriving in Digital Workspaces* (pp. 217–233). Springer International Publishing.

McCorkle, B. H., Rogers, E. S., Dunn, E. C., Lyass, A., & Wan, Y. M. (2008). Increasing social support for individuals with serious mental illness: Evaluating the Compeer model of intentional friendship. *Community Mental Health Journal, 44*(5), 359–366.

McLellan, A., Schmidt-Waselenchuk, K., Duerksen, K., & Woodin, E. (2022). Talking back to mental health stigma: An exploration of YouTube comments on anti-stigma videos. *Computers in Human Behavior, 131*, p. 107214.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint* arXiv:1301.3781.

Munaro, A.C., Hübner Barcelos, R., Francisco Maffezzolli, E.C., Santos Rodrigues, J.P., & Cabrera Paraiso, E. (2021). To engage or not engage? The features of video content on YouTube affecting digital consumer engagement. *Journal of Consumer Behavior, 20*(5), 1336–1352.

Pavlick, E., & Tetreault, J. (2016). An empirical analysis of formality in online communication. *Transactions of the Association for Computational Linguistics, 4*, 61–74.

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. *arXiv preprint* arXiv:2212.04356.

Rennekamp, K. M., & Witz, P. D. (2021). Linguistic formality and audience engagement: Investors' reactions to characteristics of social media disclosures. *Contemporary Accounting Research, 38*(3), 1748-1781.

Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society Series A: Statistics in Society, 147*(5), 656.

Steiner, P. M., & Cook, D. (2013). Matching and propensity scores. In *The Oxford Handbook of Quantitative Methods (Vol 1): Foundations*. (pp. 237-259). Oxford University Press.

Tannen, D. (1991). *You Just Don't Understand: Women and Men in Conversation* (First Edition). New York, NY, US: Ballantine Books, p.352.

Tsou A., Thelwall M., Mongeon P., & Sugimoto C.R. (2014). A community of curious souls: An analysis of commenting behavior on TED talks videos. *PLoS ONE*, *9*(4), e93609.

Vojtila, L., Ashfaq, I., Ampofo, A., Dawson, D., & Selby, P. (2021). Engaging a person with lived experience of mental illness in a collaborative care model feasibility study. *Research Involvement and Engagement, 7*(1), 1–8.

Xie, J., Zhang, B., Brown, S., & Zeng, D. (2021). Write like a pro or an amateur? Effect of medical language formality. *ACM Transactions on Management Information Systems, 12*(3), 1-25.

Yan, L., & Tan, Y. (2014). Feeling blue? Go online: An empirical study of social support among patients. *Information Systems Research, 25*(4), 690–709.

Yang, K., Lau, R. Y. K., & Abbasi, A. (2022). Getting personal: A deep learning artifact for text-based measurement of personality. *Information Systems Research, 34*(1), 194-222.

Yoo, W., Namkoong, K., Choi, M., Shah, D. V., Tsang, S., Hong, Y., Aguilar, M., & Gustafson, D. H. (2014). Giving and receiving emotional support online: Communication competence as a moderator of psychosocial benefits for women with breast cancer. *Computers in Human Behavior, 30*, 13–22.

Zhang, J., Zhang, J., Wang, K., & Yan, W. (2021). Should doctors use or avoid medical terms? The influence of medical terms on service quality of e-health. *Electronic Commerce Research*, 1-31.

Zhou, J., Kishore, R., Amo, L., & Ye, C. (2022). Description and demonstration signals as complements and substitutes in an online market for mental health care. *MIS Quarterly, 46*(4), 2055–2084.