THE UNIVERSITY OF CHICAGO

Self-Disclosure in Psychotherapy: The Development of a Novel Behavioral Measurement via Natural Language Analysis

By

Jiachen Yu

June 2023

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science

Faculty Advisor: Harriet de Wit, Ph.D., Professor of Psychiatry and Behavioral Neuroscience

Preceptor: Sabrina Nardin, Ph.D., Assistant Instructional Professor in MACSS

Jon Clindaniel, Ph.D., Assistant Senior Instructional Professor in MACSS

Abstract

While it is widely accepted that clients' level of self-disclosure has important implications for treatment outcomes and therapeutic relationship, it is difficult to measure. Extant assessment tools fail to capture the fluctuating nature of self-disclosing behaviors, their sensitivity to context and vulnerability to subjective biases. This study sought to address this issue by utilizing a quantitative self-disclosure method assessing distinct linguistic features. Forty-eight participants aged 18-35 were recruited to complete two dyadic, 45-minute conversations with a stranger co-participant. The conversation topics they received either induced high or low self-disclosure. Logistic regression models in combination of natural language processing techniques including a word-count approach (LIWC) and a word-embedding approach (BART) were constructed based on participants' conversation content to classify high vs. low-level of self-disclosure. A logistic regression model built upon LIWC categories achieved 89.66% accuracy, and the one using pre-trained BART model retained 48.28% prediction accuracy. Incorporating demographic variables did not affect the model built on LIWC but improved the accuracy to 62.07% for the BART approach. Results provides preliminary evidence in support of algorithms that assess language content to predict high versus low-levels of self-disclosure. Limitations were discussed that future studies might improve model performance by recruiting a larger, more heterogenous participant sample to discuss a wider range of topics.

Introduction

Self-disclosure—revelation of personal information including feelings, thoughts, and experiences (Derlega et al., 1987)—has important socio-psychological implications in psychotherapy (Farber, 2006). Disclosure of traumatic events benefits physical health (Greenberg et al., 1996; Pennebaker et al., 1988, 1990) and facilitates psychological well-being including social competence, self-efficacy and post-traumatic growth (Jourard, 1964; Levi-Belz, 2015, 2016; Pennebaker, 1997). One major function of self-disclosure is relationship development, insofar as it expedites mutual understanding and liking (Altman & Taylor, 1973; Carpenter & Freese, 1979; Laurenceau et al., 1998). In psychotherapy context, clients' willingness to disclose personal experience reflects their feelings of dyadic boundary with therapists (Derlega et al., 1987). Thus, self-disclosure serves as a concrete predictor of both treatment outcomes and therapeutic relationship (Tschuschke et al., 1996).

Measuring self-disclosure is difficult, especially when applied to a psychotherapy setting (Kreiner & Levi-Belz, 2019). One challenge is the fact that self-disclosure is highly sensitive to context. Most assessment scales, including Jourard Self-Disclosure Questionnaire (Jourard, 1971), Distress Disclosure Index (Kahn & Hessling, 2001), Self-Concealment Scale (Larson & Chastain, 1990) and Self-Concealment Scale (Miller et al., 1983) treat self-disclosure as stable personality traits and thus provide limited information for therapists to track their treatment progress. Importantly, previous research has found that situational factors including ambience (Chaikin et al., 1976; Okken et al., 2013), individual characteristics such as age and gender (Cappella, 1981), and relationship with the conversational partner (Collins & Miller, 1994) affect level of self-disclosure. Some measures of self-disclosure such as Self-Disclosure Situations Survey incorporate hypothetical situations for participants to rate their inclination to disclose personal information, but the scenarios provided are generic with limited ability to capture real-time contextual information (Chelune, 1976).

Another problem with measuring self-disclosure is that most measures depend on self-ratings or observer ratings, which are prone to subjective biases (Jourard, 1971; Miller et al., 1983). Participants' self-reported disclosure behaviors are vulnerable to selective bias, recall bias and demand characteristics by virtue of their self-perceptions and desire to match experimenters' expectations (Brusco & Watts, 2015; Sato & Kawahara, 2011). Although observer ratings by independent judges minimize such biases, subjectivity is still an issue, because of personal experience and research content (Forgas, 2011; Mikulincer & Nachshon, 1991). Indeed, previous studies have found low levels of agreement among trained judges (Barak & Gluck-Ofri, 2007).

The discrepancy between actual behaviors and self-perceptions and the fluctuating nature of self-disclosing behaviors points to the need for an objective, fluid index for measurement. Kreiner and Levi-Belz (2019) has proposed a behavioral, situated assessment of speech as indicator of self-disclosure. The approach is built upon the assumption that self-disclosure is embedded in verbal communication, and it has the advantages of considering real-time situated factors. The current study aims at leveraging this approach to construct an objective, real-time measurement of client self-disclosure during therapy sessions.

Conversational Partner Procedure (Aron et al., 1997)

The methodological challenges described above point to a need for objective measures of self-disclosure from a naturalistic therapy setting. As a first step, we conducted a controlled study with healthy adults, to use spoken language as an indicator of self-disclosure. We utilized a Conversational Partner Procedure (Aron et al., 1997) design to induce high vs. low-level of self-disclosure in healthy adults. The procedure adopts a dyadic conversational paradigm between strangers, and controls self-disclosing behaviors by varying the sets of topics that subjects are asked to discuss (see Study Procedure). Previous studies have found that the high disclosure condition consistently elicited more prosocial feelings and higher closeness with study partners (Aron et al., 1997; Sprecher, Treger, & Wondra, 2013; Sprecher, Treger, Wondra, et al., 2013).

Speech Analysis

Because self-disclosure is closely intertwined with verbal communication to convey personal thoughts and experiences, language features provide a promising index of disclosing behaviors (Chelune, 1976; Cozby, 1973; Jourard, 1971; Tausczik & Pennebaker, 2010). Moreover, recent advances in computational linguistic techniques have provided sensitive and accessible toolkits for psycholinguistic analysis. Hence, we conducted an analysis of the linguistic features of participants' speech to provide a quantitative and objective predictor of self-disclosure. We used bag-of-words and word-embedding approaches to detect implicit social and psychological state in conversation.

Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2015)

In the present study, we used LIWC 2015 (*LIWC*, 2015) based on a bags-of-words approach—representing word occurrences in a document—to transform textual data into meaningful language categories as predictors. Previous studies have utilized speech duration and word counts to measure self-disclosure and demonstrated their ability to capture subtle fluctuations in behaviors (Barak & Gluck-Ofri, 2007; Joinson, 2001). Our analysis expanded this generic total word count approach into more fine-grained psychometric properties of language leveraging built-in dictionaries in LIWC program (Pennebaker et al., 2015). LIWC breaks down language into 93 meaningful linguistic categories such as analytical thinking, clout, authenticity and emotional tones (Pennebaker et al., 2015). The program assigns scores to each category based on word count of participants' speech in the dictionary of that category divided by the total word count. One critique of LIWC is that its bag-of-words approach only concerns whether words occur in a document but not where they occur, thus failing to account for the semantic contextual information.

Pre-trained Language Representations (Devlin et al., 2018)

Pre-trained bidirectional transformers have been widely adopted in natural language processing to accomplish a wide range of tasks including sentiment analysis, question answering, entity recognition and semantic similarity analysis. This approach captures semantic information with a neural network architecture and transforms textual data into representations of word vectors that are machine-readable. It takes into account the context (a window of words, or tokens, before and after the target word) of the sentences in which they are used. In addition to psychometric properties generated by LIWC, these pre-trained models can also accomplish data mining built upon user-defined metrics. In this study, we applied a pre-trained BART model to conduct zero-shot classification to calculate the probabilities of participants' speech belonging to emotions and insights—two aspects of disclosure content (Lewis et al., 2019; Shim et al., 2011). The BART model leverages a bidirectional (BERT-like) encoder and an autoregressive (GPT-like) decoder (Lewis et al., 2019), and is pre-trained on the Multi-Genre Natural Language Inference dataset that includes 433,000 sentence pairs annotated with textual entailment information (Williams et al., 2018).

The Current Study

In this study, we sought to build a situated, objective measure of self-disclosure by examining natural speech of healthy adults under experimenter-controlled high and low levels of self-disclosure. We used LIWC and BART to reconstruct participants' conversations into machine-understandable numerical data and analyzed their language use utilizing logistic regression modeling. We further incorporated demographic variables to improve model performance based on extant literature that self-disclosing behaviors vary depending on individual characteristics (Cappella, 1981). The project is an exploratory analysis to determine the sensitivity of these methods to detect linguistic differences in natural conversations with different levels of self-disclosure.

Methods

Study Procedure

We recruited healthy individuals (N=48) aged 18-35 via flyers around campus and online advertisements. Exclusion criteria include less than a high school diploma, lack of fluency in English, current psychiatric medications, or recent severe psychiatric symptoms. All participants completed two experimental sessions, scheduled at least two days apart, in which they engaged in a conversational Partner Procedure (Aron et al., 1997), involving either high or low levels of self-disclosing conversations with different partners. The high disclosure condition involved conversation topics that were increasingly 'deep' and personal over the 45-minute session whereas the low disclosure condition involved topics that were 'shallow.' See Supplemental Materials for a complete set of questions administered to participants.

In each experimental session, participants were randomly paired with a different same-sex, stranger co-participant to engage in a 45-minute conversation divided in three 15-minute segments. In the shallow condition, the topics given to discuss were shallow in all three 15-minute segments of the dialogue, whereas in the deep condition the topics started as shallow but became more personal across the three 15 min segments. Participants took turns to ask and answer the questions and were allowed to skip questions if they preferred not to answer. They were asked to move on to the next segment at the end of each 15-minute period even if they did not finish all questions. We only employed the last 15-minute of speech in each condition for further analysis because this was the time when the experimental conditions differed. Participants completed self-reported questionnaires before and after each session, and saliva samples were obtained for measures of oxytocin (not reported here). The present analysis examined the participants' natural speech with regards to its psychometric properties. Participants provided written, informed consent prior to the study, and the study procedure was approved by the Institutional Review Board of the University of Chicago.

Data Preprocessing and Feature Selection

We transcribed participants' conversations through automated transcription services with 85% accuracy reported (*Happy Scribe*, n.d.). Transcriptions were double-checked by three independent coders to ensure the participant identification number was labeled correctly to match the text. The level of agreement among coders achieved 95.41% (*SD* = 0.05). All disputes were resolved by a third coder. Every subject's speech was extracted from the conversations and analyzed independently. To minimize biases induced by the sets of questions we administered, we removed all words in the question sets for analysis. The prepared files (the transcriptions of the last 15-minute speech) then went through LIWC programs to generate psychometric predictors. Twenty psychometric word categories were removed because they are irrelevant to our research topic (e.g., filler words) or predominantly scored zero. All categories except Total Word Count are on the scale of 0 to 100, calculated as the percentage of words spoken in the target category dictionary compared to the total number of words spoken. We log-transformed Total Word Count to scale this variable in a lower range. Considering the large numbers of possible predictor features, we conducted dimensionality reduction via a singular vector decomposition (SVD) and non-negative matrix factorization (NMF) to reduce predictors into five manifolds. We examined explained variance ratio of each method to index how representative the resulting components were of the original data. The number of components chosen for each methodology was determined by the characteristics of the data, with the goal of achieving at least 85% of explained variance ratio.

Independently, we ran the transcriptions through a transformer encoder-decoder model to perform a zero-shot classification task that gauge participants' type of disclosure. To be more specific, the model calculated the probabilities that the participants' speech belonged to emotional disclosure or insightful disclosure.

Model Training and Evaluation

We conducted model training and performance evaluation on predictors generated via LIWC and BART, separately. The training set included the whole dataset and testing set consisted of random split (40% of the dataset) of ten iterations of the data. We built a logistic regression classifier on the training data and conducted a five-fold iteration to search for the optimal parameters for the algorithms that yielded the best prediction within the training set. The model then predicted self-disclosure outcomes in the testing data and results were compared to the actual condition. Model performance was indexed by its prediction accuracy, a confusion matrix and the receiver operating characteristic (ROC) curve. Effect sizes and *p*-value at 0.05 level were examined to assess each component's predictor ability. All codes were written in Python 3.7 (Van Rossum & Drake, 2009)

Supplemental Analysis

The main analysis involves a between-subject analysis that incorporates only participants' conversations as predictors (partner speech excluded). To address potential biases with this approach, we conducted three supplemental analyses. First, we examined whether individual characteristics influenced our model performance by adding demographic variables including participant sex, age, and race as predictors, and followed the same procedure of model training and model evaluation. Second, our main analysis failed to recognize clustering in the data induced by the study design that incorporated dyadic groups. We accounted for random effects of groups by running a mixed effect logistic regression model. Predictor robustness was evaluated at *p*-value level of .05.

Results

Descriptives

Nine participants were excluded for analysis owing to recording device malfunctioning, lack of informed consent to use the recordings, and failure to comply with the study procedure. The final sample included 39 participants (19 females and 20 males; see Figure 1) and a resulting 71 conversation files. The average age was 23.92 (*SD* = 3.26). The sample consisted of a predominantly, 66% White population (N = 26), 13% Asian (N = 5), 5% Black or African American (N = 2), 8% Mixed Race (N = 3) and 8% Unknown Race (N = 3).

Before removing the stop words, the average word count for all participants was 1013.44 (*SD* = 293.03). The mean word count for the deep conversations was 914.36 (*SD* = 301.52), and the mean word count for the small conversations was 1076.03 (*SD* = 274.17). Figure 2 displays the most frequent words spoken by participants in the two different conditions. A larger size of the word in the cloud indicates a higher frequency that it occurs in the conversation. As shown in the word clouds, participants tended to use affirmative words (e.g., “yeah”, “okay”) and cognitive processing words (e.g., “think”, “feel”, “know”) in both conditions. When subjects were in the small talk condition, they tended to use words that were more time oriented, whereas when they were in the deep talk condition they focused more on social elements, such as talking about friends.

LIWC Results

The LIWC 2015 program generated a total of 93 psychometric word categories, and 73 were retained for further analysis. Correlations between each variable was calculated and, according to the heatmap illustrating all correlation coefficients (see Figure 3), a weak relationship was implied in all variables. The results indicate low levels of homogeneity within data and therefore supports the need for a higher explained variance ratio from dimension reduction techniques. After iterating from one to seven chosen as the number of components for SVD, the number four was picked as it generated the optimal summed explained variance ratio of 0.90. See Table 1 for the resulting four components. The explained variance ratio for the four is as follows: 0.19, 0.38, 0.25, and 0.09. According to the SVD results, participants' conversations are mostly characterized by authenticity, emotional tone, functional words (“I”, more specifically), clout, analytical thinking, and pronouns. As the summed explained ratio for NMF remained low for the numbers of components that are below ten, we decided to exclude NMF for analysis and used SVD only.

Overall, the predicting accuracy to classify high vs. low-level of self-disclosure was 89.66%. The area under the curve for the ROC curve was .97 (see Figure 4). Logistic regression model indicated that Component 2, $b = .141$, $p = .000$, *95%*CI = [.076, .206], and Component 3, $b = .065$, $p = .011$, *95%*CI = [.015, .116], were significant in classifying self-disclosure conditions. Component 1 and Component 4 were insignificant, with p-value at .634 and at .132.

However, after accounting for the random effects of participant group, Component 2, $b = .822$, $p = .117$, and Component 3, $b = .714$, $p = .153$, became insignificant in classifying self-disclosure.

After adding demographic variables including sex, age, and race to the logistic regression model, the prediction accuracy achieved 89.66%, and the area under the curve for ROC was .96. Component 2, $b = .209$, $p = .001$, *95%*CI = [.087, .329], Component 3, $b = .121$, $p = .006$, *95%*CI = [.035, .206], and Component 4, $b = .166$, $p = .036$, *95%*CI = [.011, .320], were significant. None of the three demographic variables (age, $b = -.165$, $p = .284$, *95%*CI = [-.466, .137], sex, $b = -2.274$, $p = .057$, *95%*CI = [-4.617, .069], racial group, *p*-values ranging from .336-.899) were significant in classifying self-disclosure. Similarly, none of the components or demographic variables were significant in classifying high vs. low-level of self-disclosure after accounting for the random effects of the groups that participants belonged to.

BART Results

Pre-trained BART models were used to generate emotional disclosure and insightful disclosure score. The mean emotional disclosure was .839 ($SD = .190$) in the deep talk condition and .785 ($SD = .158$) in the small talk condition. The average insightful disclosure score for the deep talk condition was .880 ($SD = .122$), and the average for the small condition was .820 ($SD = .163$). See Figure 5 for a distribution for emotional and insightful disclosure score in both experimental conditions.

Overall, the predicting score of the two types of disclosure in classifying high vs. low level of self-disclosure achieved an accuracy of 48.28%, and the area under curve for the ROC curve was 0.442 (see Figure 6). Neither emotional disclosure, $b = -.215$, $p = .906$, *95%*CI = [-3.783, 3.353], nor insightful disclosure, $b = .124$, $p = .944$, *95%*CI = [-3.312, 3.560], was significant in classifying self-disclosure. Likewise, neither variable was significant when the random effects of participant group were accounted for.

With demographic variables added to the logistic regression model, the prediction accuracy was 62.07%, and the area under curve for ROC was .755 (see Figure 6). Neither emotional disclosure, b = 2.179, p = .360, *95%*CI = [-2.484, 6.841], nor insightful disclosure, b = 1.603, p = .579, *95%*CI = [-4.064, 7.269], was significant in predicting high vs. low-level of self-disclosure. None of the demographic variables was significant in predicting self-disclosure (age, b = .001, p = .994, *95%*CI = [-.175, .176], sex, b = -.573, p = .296, *95%*CI = [-1.646, .501], racial group, p-values ranging from .224-.728). Results remained the same when we accounted for the random effects of participant group.

Discussions

The current study aimed at building a situated, objective assessment of self-disclosure leveraging natural language processing techniques and machine learning models. Results provided preliminary evidence that language analysis has potential in distinguishing high versus low self-disclosure in a dyadic conversation. A logistic regression model built upon psycholinguistic word categories achieved 89.66% accuracy (area under curve of ROC = .97) in classifying high vs. low-level of self-disclosure. Results were robust both with and without demographic predictors incorporated in the model.

The current study pinpoints several potentially valuable characteristics of language use when individuals engage in different levels of disclosing behaviors, shedding light on possible predictors to be incorporated in future models in attempt to predict self-disclosure. Logistic regression analysis reveals that Component 2 and Component 3 (see Table 1), two components constructed from SVD that comprise primarily of clout, emotional tone, authenticity and function words, were significant in classifying high vs. low-level of self-disclosure. Such results indicate that participants in the deep condition tend to speak more powerfully, less authentically, with less emotions, and less structured. The findings are consistent with the hypothesis in Kreiner & Levi-Belz‘s study (2019) that emotion words and pronouns (a subcategory of function words) are indicative of disclosing of traumatic events, and expand their proposed indicators to other linguistic categories including clout and authenticity.

Notably, the word-embedding approach (BART) was less sensitive than the LIWC results. Machine learning models built on predictors generated by BART model yielded results that were worse than random (area under curve of ROC <.5). Possible explanations might be that the zero-shot classification performed badly in labeling emotional disclosure and insightful disclosure because the model was originally trained on texts of shorter sentences. The larger chunks of texts in our study might bring more noise in the data and thus influencing the prediction results. On the other hand, the poor performance might stem from the fact that disclosure is intrinsically a complicated construct shaped by various perspectives. The use of emotional disclosure and insightful disclosure may oversimplify self-disclosing speech and therefore lead to ambiguous results.

Previous studies have found that individual characteristics such as age and gender affected self-disclosing behaviors (Cappella, 1981). In contrary to the hypothesis that inclusion of individual characteristics would improve the prediction accuracy, the performance of our logistic regression model built on LIWC results was similar with and without demographic predictors. The results might be attributable to the low diversity of demographic variables in our sample to make the predictors robust (see Figure 1). A larger, more heterogeneous participant sample is needed in future studies to test the effects of including individual characteristics in predicting self-disclosure. Notably, the accuracy of a logistic regression model with BART approach increased from 48.28% to 62.07% when demographic variables including age sex, and race were incorporated in the model. Considering the poor performance of the original model, the improvement after including demographic variables might stem from a more complex model compared to the original one (i.e., increased number of predictors included) rather than from the robustness of demographic as predictors.

Several limitations should be noted. First, our study involved a design of dyadic natural conversations between strangers, and adopted a between-group analytical strategy to build machine learning models. This approach fails to account for within-group variance and is thus potentially subject to bias. Indeed, when we used a mixed effects model to control for random effects of the dyadic groups that the participants belonged to, the previously robust predictors (two decomposed psycholinguistic components) became insignificant. Such results suggest that the high prediction accuracy we have might be explained by the characteristics of conversation partner instead of the language index we employed. Second, although our design had the advantage of manipulating self-disclosure in an experimental setting, it nonetheless confined participants' conversations to certain topics. It is not known how the sets of questions that we gave participants introduced biases in the analysis. We conducted a supplemental sensitivity analysis to run the whole pipeline on participants' conversations without the stop words (i.e., words in the conversation topics) removed. The model performance in the LIWC approach dropped from 89.66% to 75.86%, indicating that the intrinsic psycholinguistic content of the participants' speech was less influenced by the conversation topics in predicting self-disclosure. On the other hand, the accuracy in the BART approach increased from 48.28% to 65.52%, suggesting that participants' emotional self-disclosure and insightful disclosure were affected by the discussion topics in a way that differentiated high vs. low self-disclosure. To address the potential bias caused by conversation topics, future studies can possibly collect data using randomly generated topics from a wider range of discussion materials.

Third, our study had a relatively small sample size (N of data points = 71), which constrained us to employ only a simple, parametric logistic regression model to classify self-disclosing behaviors. In pursuit of a more accurate model, future studies should collect significantly more data points that can allow direct comparison of different machine learning models (e.g., random forest, support vector machine) and feature selection techniques. A larger sample size can also benefit if future study wants to predict more fine-grained, continuous levels of self-disclosure.

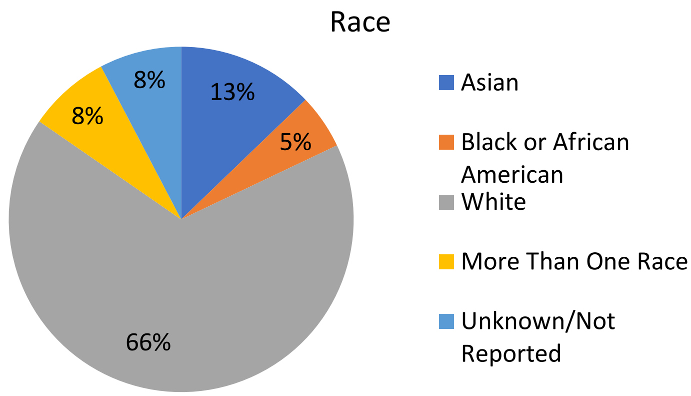
Conclusions

While it is widely accepted that clients' level of self-disclosure has important implications for treatment outcomes and therapeutic relationship, it is difficult to measure. Extant assessment tools fail to capture the fluctuating nature of self-disclosing behaviors, their sensitivity to context and vulnerability to subjective biases. This study sought to address this issue by utilizing a quantitative self-disclosure method assessing distinct linguistic features. Machine learning models in combination of natural language processing techniques including a word-count approach (LIWC) and a word-embedding approach (BART) revealed good prediction accuracy in classifying of high vs. low-level of self-disclosure. Applications of such algorithms can provide informative feedback for therapists to track their clients' disclosing behaviors and fine-tune their therapeutic techniques accordingly.

Acknowledgement

My deep gratitude to Dr. Harriet de Wit for guiding me throughout the whole process of preparing this thesis and for giving me inspirations all the time. Many thanks to Dr. Guillermo A. Cecchi for his valuable feedbacks on my analytical plan, and to Dr. Hanna Molla for her contributions to this study and her kind support and help.

Appendix



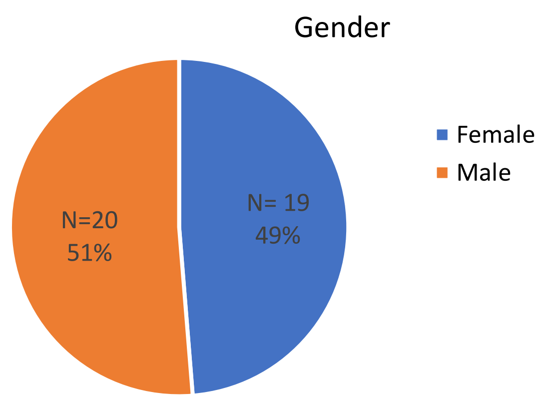


Figure 1. Pie chart depicting participant race (above) and participant gender (below) distribution

Figure 2. Word Cloud of aggregated participant conversations in small talk condition (left) and deep talk condition (right)

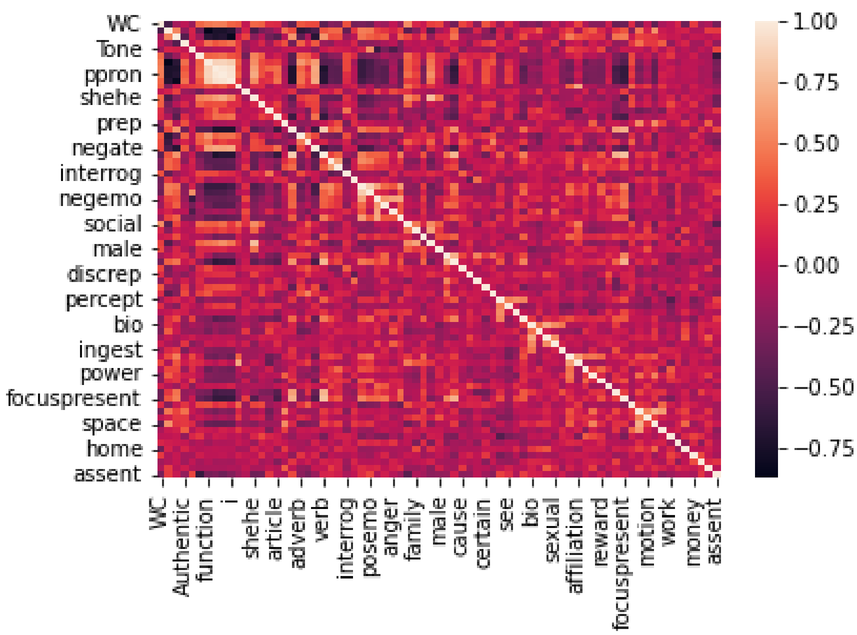


Figure 3. A heatmap of correlations between 73 word categories generated from LIWC

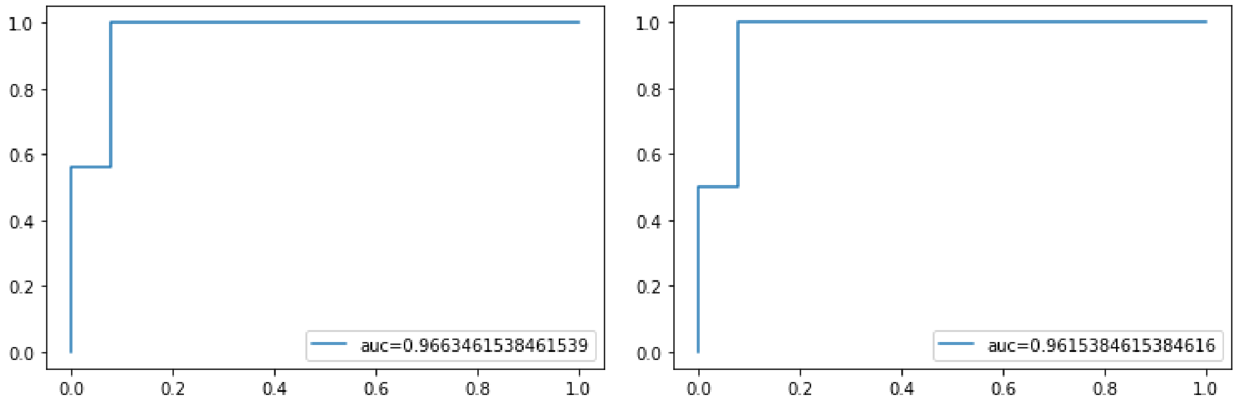


Figure 4. ROC curve for LIWC components as predictors generated (left) and with demographic variables added (right)

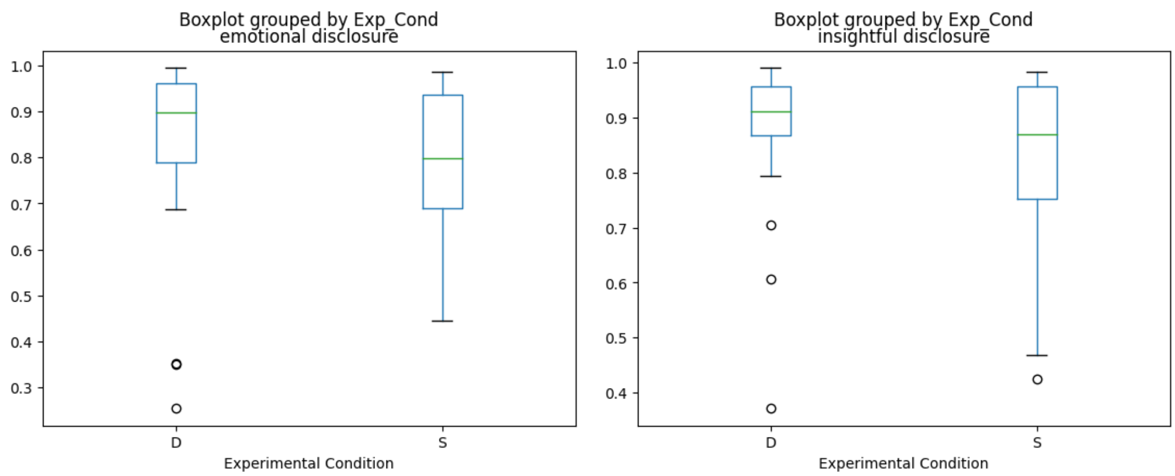


Figure 5. Boxplot depicting the emotional disclosure score (left) and insightful disclosure score (right) in deep talk and small conditions

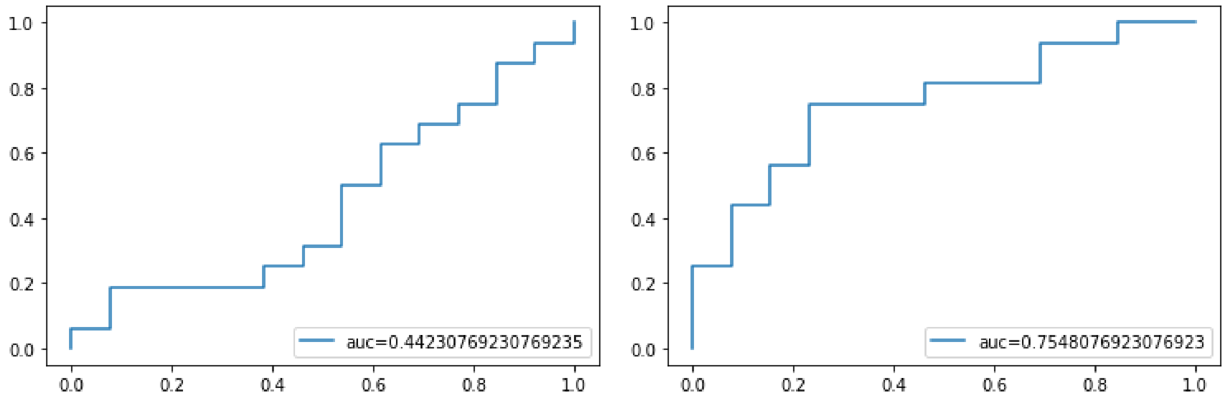


Figure 6. ROC curve for emotional and insightful disclosure as predictors (left) with demographic variables added (right)

Component 1		Component 2		Component 3		Component 4	
Authentic	0.598	Clout	0.656	Tone	-0.771	Analytic	0.701
Tone	0.587	Authentic	-0.447	Clout	0.401	Clout	-0.437
Function	0.294	Analytic	0.356	Authentic	0.367	Function	-0.298
Clout	0.288	Function	-0.239	Analytic	0.176	Authentic	0.287
Analytic	0.128	Tone	0.222	Verb	0.122	Pronoun	-0.180
Verb	0.123	Pronoun	-0.209	Function	0.106	Social	-0.165
Cogproc	0.122	Ppron	-0.190	Cogproc	0.105	Ppron	-0.162
Pronoun	0.109	I	-0.164	Relative	0.093	Conj	-0.110
Relative	0.106	Verb	0.087	Focuspresent	0.072	Space	0.075
Focuspresent	0.078	Focuspresent	0.058	Negemo	0.058	Relative	0.073

Table 1. Word Category Components and their carried weights generated by SVD

References

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships* (pp. viii, 212). Holt, Rinehart & Winston.

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The Experimental Generation of Interpersonal Closeness: A Procedure and Some Preliminary Findings. *Personality and Social Psychology Bulletin*, *23*(4), 363–377. https://doi.org/10.1177/0146167297234003

Barak, A., & Gluck-Ofri, O. (2007). Degree and Reciprocity of Self-Disclosure in Online Forums. *CyberPsychology & Behavior*, *10*(3), 407–417. https://doi.org/10.1089/cpb.2006.9938

Brusco, N. K., & Watts, J. J. (2015). Empirical evidence of recall bias for primary health care visits. *BMC Health Services Research*, *15*(1), 381. https://doi.org/10.1186/s12913-015-1039-1

Cappella, J. N. (1981). Mutual influence in expressive behavior: Adult–adult and infant–adult dyadic interaction. *Psychological Bulletin*, *89*, 101–132. https://doi.org/10.1037/0033-2909.89.1.101

Carpenter, J. C., & Freese, J. J. (1979). Three Aspects of Self-Disclosure as They Relate to Quality of Adjustment. *Journal of Personality Assessment*, *43*(1), 78–85. https://doi.org/10.1207/s15327752jpa4301_11

Chaikin, A. L., Derlega, V. J., & Miller, S. J. (1976). Effects of room environment on self-disclosure in a counseling analogue. *Journal of Counseling Psychology*, *23*, 479–481. https://doi.org/10.1037/0022-0167.23.5.479

Chelune, G. J. (1976). *Self-Disclosure Situations Survey: A New Approach to Measuring Self-Disclosure*. Order Department, American Psychological Association, 1200 Seventeenth Street, N.

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, *116*, 457–475. https://doi.org/10.1037/0033-2909.116.3.457

Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin*, *73*, 79–91. https://doi.org/10.1037/h0033950

Derlega, V. J., Margulis, S. T., & Winstead, B. A. (1987). A Social-Psychological Analysis of Self-Disclosure in Psychotherapy. *Journal of Social and Clinical Psychology*, *5*(2), 205–215.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. https://doi.org/10.48550/arXiv.1810.04805

Farber, B. A. (2006). *Self-disclosure in Psychotherapy*. Guilford Press.

Forgas, J. P. (2011). Affective influences on self-disclosure: Mood effects on the intimacy and reciprocity of disclosing personal information. *Journal of Personality and Social Psychology*, *100*, 449–461. https://doi.org/10.1037/a0021129

Greenberg, M. A., Wortman, C. B., & Stone, A. A. (1996). Emotional expression and physical health: Revising traumatic memories or fostering self-regulation? *Journal of Personality and Social Psychology*, *71*, 588–602. https://doi.org/10.1037/0022-3514.71.3.588

Happy Scribe. (n.d.). Happy Scribe. Retrieved October 26, 2022, from https://www.happyscribe.com/

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, *31*(2), 177–192. https://doi.org/10.1002/ejsp.36

Jourard, S. M. (1964). *The Transparent Self: Self-disclosure and Well-being*. Van Nostrand.

Jourard, S. M. (1971). *Self-disclosure: An Experimental Analysis of the Transparent Self*. Wiley-Interscience.

Kahn, J. H., & Hessling, R. M. (2001). Measuring the Tendency to Conceal Versus Disclose Psychological Distress. *Journal of Social and Clinical Psychology*, *20*(1), 41–65. https://doi.org/10.1521/jscp.20.1.41.22254

Kreiner, H., & Levi-Belz, Y. (2019). Self-Disclosure Here and Now: Combining Retrospective Perceived Assessment With Dynamic Behavioral Measures. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00558

Larson, D. G., & Chastain, R. L. (1990). Self-Concealment: Conceptualization, Measurement, and Health Implications. *Journal of Social and Clinical Psychology*, *9*(4), 439–455. https://doi.org/10.1521/jscp.1990.9.4.439

Laurenceau, J.-P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*, *74*, 1238–1251. https://doi.org/10.1037/0022-3514.74.5.1238

Levi-Belz, Y. (2015). Stress-Related Growth among Suicide Survivors: The Role of Interpersonal and Cognitive Factors. *Archives of Suicide Research*, *19*(3), 305–320. https://doi.org/10.1080/13811118.2014.957452

Levi-Belz, Y. (2016). To share or not to share? The contribution of self-disclosure to stress-related growth among suicide survivors. *Death Studies*, *40*(7), 405–413. https://doi.org/10.1080/07481187.2016.1160164

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). *BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension* (arXiv:1910.13461). arXiv. https://doi.org/10.48550/arXiv.1910.13461

LIWC. (2015). Pennebaker Conglomerates, Inc. https://www.liwc.app/

Mikulincer, M., & Nachshon, O. (1991). Attachment styles and patterns of self-disclosure. *Journal of Personality and Social Psychology, 61*, 321–331. https://doi.org/10.1037/0022-3514.61.2.321

Miller, L. C., Berg, J. H., & Archer, R. L. (1983). Openers: Individuals who elicit intimate self-disclosure. *Journal of Personality and Social Psychology*, *44*, 1234–1244. https://doi.org/10.1037/0022-3514.44.6.1234

Okken, V., van Rompay, T., & Pruyn, A. (2013). Room to Move: On Spatial Constraints and Self-Disclosure During Intimate Conversations. *Environment and Behavior*, *45*(6), 737–760. https://doi.org/10.1177/0013916512444780

Pennebaker, J. W. (1997). Writing about Emotional Experiences as a Therapeutic Process. *Psychological Science, 8*(3), 162–166.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The Development and Psychometric Properties of LIWC2015*. https://repositories.lib.utexas.edu/handle/2152/31333

Pennebaker, J. W., Colder, M., & Sharp, L. K. (1990). Accelerating the coping process. *Journal of Personality and Social Psychology*, *58*, 528–537. https://doi.org/10.1037/0022-3514.58.3.528

Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and immune function: Health implications for psychotherapy. *Journal of Consulting and Clinical Psychology*, *56*, 239–245. https://doi.org/10.1037/0022-006X.56.2.239

Sato, H., & Kawahara, J. (2011). Selective bias in retrospective self-reports of negative mood states. *Anxiety, Stress, & Coping*, *24*(4), 359–367. https://doi.org/10.1080/10615806.2010.543132

Shim, M., Cappella, J. N., & Han, J. Y. (2011). How Does Insightful and Emotional Disclosure Bring Potential Health Benefits?: Study Based on Online Support Groups for Women with Breast Cancer. *The Journal of Communication*, *61*(3), 432–464. https://doi.org/10.1111/j.1460-2466.2011.01555.x

Sprecher, S., Treger, S., & Wondra, J. D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. *Journal of Social and Personal Relationships*, *30*(4), 497–514. https://doi.org/10.1177/0265407512459033

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology, 49*(5), 860–866. https://doi.org/10.1016/j.jesp.2013.03.017

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Tschuschke, V., Mackenzie, K. R., Haaser, B., & Janke, G. (1996). Self-disclosure, feedback, and outcome in long-term inpatient psychotherapy groups. *The Journal of Psychotherapy Practice and Research*, *5*(1), 35–44.

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Williams, A., Nangia, N., & Bowman, S. R. (2018). *A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference* (arXiv:1704.05426). arXiv. https://doi.org/10.48550/arXiv.1704.05426

Supplemental Materials

Deep Talk Question Set 1

1. Given the choice of anyone in the world, whom would you want as a dinner guest?
2. Would you like to be famous? In what way?
3. Before making a telephone call, do you ever rehearse what you are going to say? Why?
4. What would constitute a “perfect” day for you?
5. When did you last sing to yourself? To someone else?
6. If you were able to live to the age of 90 and retain either the mind or body of a 30-year-old for the last 60 years of your life, which would you want?
7. Do you have a secret hunch about how you will die?
8. Name three things you and your partner appear to have in common.
9. For what in your life do you feel most grateful?
10. If you could change anything about the way you were raised, what would it be?
11. Take 4 minutes and tell your partner your life story in as much detail as possible.
12. If you could wake up tomorrow having gained any one quality or ability, what would it be?

Deep Question Set 2

1. If a crystal ball could tell you the truth about yourself, your life, the future, or anything else, what would you want to know?
2. Is there something that you've dreamed of doing for a long time? Why haven't you done it?
3. What is the greatest accomplishment of your life?
4. What do you value most in a friendship?
5. What is your most treasured memory?
6. What is your most terrible memory?
7. If you knew that in one year you would die suddenly, would you change anything about the way you are now living? Why?
8. What does friendship mean to you?
9. What roles do love and affection play in your life?
10. Alternate sharing something you consider a positive characteristic of your partner. Share a total of 5 items.
11. How close and warm is your family? Do you feel your childhood was happier than most other people's?
12. How do you feel about your relationship with your mother?

Deep Question Set 3

1. Make 3 true "we" statements each. For instance, "We are both in this room feeling…"
2. Complete this sentence: "I wish I had someone with whom I could share…"
3. If you were going to become a close friend with your partner, please share what would be important for him or her to know.
4. Tell your partner what you like about them; be very honest this time, saying things that you might not say to someone you've just met.
5. Share with your partner an embarrassing moment in your life.
6. When did you last cry in front of another person? By yourself?
7. Tell your partner something that you like about them already.
8. What, if anything, is too serious to be joked about?
9. If you were to die this evening with no opportunity to communicate with anyone, what would you most regret not having told someone? Why haven't you told them yet?
10. Your house, containing everything you own, catches fire. After saving your loved ones and pets, you have time to safely make a final dash to save any one item. What would it be? Why?
11. Of all the people in your family, whose death would you find most disturbing? Why?
12. Share a personal problem and ask your partner's advice on how he or she might handle it. Also, ask your partner to reflect back to you how you seem to be feeling about the problem you have chosen.

Shallow Talk Question Set 1

1. When was the last time you walked for more than an hour? Describe where you went and what you saw.
2. What was the best gift you ever received and why?
3. If you had to move from [location] where would you go, and what would you miss about [location]?
4. How did you celebrate last Halloween?
5. Do you read a newspaper often and which do you prefer? Why?
6. What is a good number of people to have in a student household and why?
7. If you could invent a new flavor of ice cream, what would it be?
8. What is the best restaurant you've been to in the last month that your partner hasn't been to? Tell your partner about it.
9. Describe the last pet you owned.
10. What is your favorite holiday? Why?
11. Tell your partner the funniest thing that ever happened to you when you were a small child.
12. What gifts did you receive on your last birthday?

Shallow Talk Question Set 2

1. Describe the last time you went to the zoo.
2. Tell the names and ages of your family members, include grandparents, aunts and uncles, and where they were born (to the extent you know this information).
3. One of you say a word, the next say a word that starts with the last letter of the word just said. Do this until you have said 50 words. Any words will do - you aren't making a sentence.
4. Do you like to get up early or stay up late? Is there anything funny that has resulted from this?
5. Where are you from? Name all of the places you've lived.
6. What is your favorite class at [school] so far? Why?
7. What did you do this summer?
8. What gifts did you receive last Christmas / Hanukkah / Kwanza?
9. Who is your favorite actor of your own gender? Describe a favorite scene in which this person has acted.
10. What was your impression of [school] the first time you ever came here?
11. What is the best TV show you've seen in the last month that your partner hasn't seen? Tell your partner about it.
12. What is your favorite holiday? Why?

Shallow Talk Question Set 3

1. Where did you go to high school? What was your high school like?
2. What is the best book you've read in the last three months that your partner hasn't read? Tell your partner about it.
3. What foreign country would you like to visit? What attracts you to this place?
4. Do you prefer digital watches and clocks or the kind with hands? Why?
5. Describe your mother's best friend.
6. What are the advantages and disadvantages of artificial Christmas trees?
7. How often do you get your hair cut? Where do you go? Have you ever had a really bad haircut experience?
8. Did you have a class pet when you were in elementary school? Do you remember the pet's name?
9. Do you think left-handed people are more creative than right-handed people?
10. What is the last concert you saw? How many of that band's albums do you own? Had you seen them before? Where?
11. Do you subscribe to any magazines? Which ones? What have you subscribed to in the past?
12. Were you ever in a school play? What was your role? What was the plot of the play? Did anything funny ever happen when you were on stage?