

Original Paper

Using Digital Phenotypes to Identify Individuals With Alexithymia in Posttraumatic Stress Disorder: Cross-Sectional Study

Tomas Meaney¹, BSc, MSc, PhD; Vijay Yadav², PhD; Isaac Galatzer-Levy², PhD; Richard Bryant¹, PhD

¹School of Psychology, Faculty of Science, University of New South Wales, Sydney, Australia

²Department of Psychiatry, Grossman School of Medicine, New York University, New York City, NY, United States

Corresponding Author:

Tomas Meaney, BSc, MSc, PhD

School of Psychology

Faculty of Science

University of New South Wales

Mathews Building, 11th Fl.

F23 Library Walk

Sydney, 2033

Australia

Phone: 61 2 9385 3600

Email: z5345526@ad.unsw.edu.au

Abstract

Background: Alexithymia, defined as difficulty identifying and describing one's emotions, has been identified as a transdiagnostic emotional process that impacts the course, severity, and treatment outcomes of psychiatric conditions such as posttraumatic stress disorder (PTSD). As such, alexithymia is an important process to accurately measure and identify in clinical contexts. However, research identifying the association between the experience of alexithymia and psychopathology has been limited by an overreliance on self-report scales, which have restricted use for measuring constructs that involve deficits in self-awareness, such as alexithymia. Hence, more suitable and effective methods of measuring and identifying those experiencing alexithymia in clinical samples are needed.

Objective: In this cross-sectional study, we aimed to determine if facial, vocal, and language phenotypes extracted from 1-minute recordings of war veterans with PTSD describing a traumatic event could be used to identify those experiencing alexithymia.

Methods: A total of 96 participants were included in this cross-sectional study. Specialized software was used to extract facial, vocal, and language features from the recordings. These features were then integrated into machine learning (extreme gradient boosting [XGBoost]) classification models that were trained and tested within a 5-fold nested cross-validation pipeline for their capacity to classify veterans scoring above the cutoff for alexithymia on the Toronto Alexithymia Scale-20.

Results: The best performing XGBoost classification model trained in the nested cross-validation pipeline was able to classify those experiencing alexithymia with a good level of accuracy (average F_1 -score=0.78, SD 0.07; average area under the curve score=0.87, SD 0.12). Consistent with theoretical models and past research into phenotypes of alexithymia, language, vocal, and facial features all contributed to the accuracy of the XGBoost classification model.

Conclusions: These findings indicate that facial, vocal, and language phenotypes incorporated in machine learning models could represent a promising alternative to identifying individuals with PTSD who are experiencing alexithymia. The further validation and use of this approach could facilitate more tailored and effective allocation of treatment resources to individuals experiencing alexithymia in clinical settings.

(JMIR Ment Health 2025;12:e83575) doi: [10.2196/83575](https://doi.org/10.2196/83575)

KEYWORDS

mental health; digital health; alexithymia; digital phenotyping; machine learning; posttraumatic stress disorder; veterans

Introduction

Alexithymia is defined as difficulty identifying and describing one's own emotional states, in conjunction with externally focused attention [1,2]. In the attention-appraisal model of alexithymia, those with alexithymia are considered to have difficulties in attending to (due to their externally focused attention) and appraising (due to their difficulty identifying emotions) already occurring emotional responses to stimuli [3]. These issues with attending to and appraising emotional responses mean that it is difficult to subsequently describe them. The 3 core difficulties of people with alexithymia include internal orientation of attention, identifying and describing one's emotional states, have been demonstrated to load directly onto the alexithymia construct and are consistent with the subscales of the primary alexithymia self-report measure, the Toronto Alexithymia Scale-20 (TAS-20) [4].

Alexithymia has been identified as a transdiagnostic risk factor for a range of psychiatric disorders [5-10]. A meta-analysis on the emotion processes relevant in schizophrenia found a large Hedge *g* effect size (1.05) for the association between alexithymia and schizophrenia [10]. A similarly large effect size has been found for the association between alexithymia and posttraumatic stress disorder (PTSD) [11]. Alexithymia has been conceptualized as an important mechanism in exacerbating PTSD symptoms and diminishing treatment response, given its strong association with emotional avoidance and its inhibition of the emotion processing required for gold-standard treatments, such as prolonged exposure therapy, to be effective [12]. This is consistent with findings that alexithymia following a traumatic event is predictive of the development of PTSD [13] and has a substantial influence on outcomes for PTSD interventions [14,15]. Accordingly, as with other transdiagnostic mechanisms of psychological distress, identifying alexithymia as it occurs in clinical settings, such that tailored treatment responses can be used, is important for ameliorating its impacts [12,16,17].

However, findings from most of these studies on the impacts of alexithymia are limited by their overreliance on self-report scales such as the TAS-20 [4]. This is problematic not only because self-report measures are prone to response biases [18,19] but also because alexithymia involves deficits in self-awareness that may impact the accuracy of self-report measures [20]. As such, there is a need for alternate and more construct-appropriate approaches to measure and identify alexithymia, particularly in clinical populations, in which it impacts symptom severity and treatment response.

The identification of alexithymia through individuals' use of language is a construct-relevant approach that has been used in several studies. One such study of individuals with varying levels of alexithymia on the TAS-20 that analyzed their expressive writing samples using the Linguistic Inquiry and Word Count (LIWC) software (version 22; University of Texas) [21] found that those who scored higher on the TAS-20 used fewer words expressing affectivity, sadness, and future perspective [22]. Another study found that those scoring higher on the TAS-20 produced fewer emotion words and a less diverse range of emotion words yet did not have a general vocabulary

deficit relative to low scorers [23]. A systematic review and meta-analysis of 29 empirical studies of language capacity in those with alexithymia found a modest association between language deficits (eg, emotion language use) and alexithymia [24], suggesting that language is not the only relevant expressive measure of alexithymia. This is consistent with findings that participants with alexithymia also demonstrated lower (or the same) physiological reactivity (heart rate, skin conductance, and facial electromyography) to negative stimuli, while reporting subjectively worse experiences than nonalexithymics [25-27]. Of particular relevance to PTSD, a distinction between high subjective distress and low arousal (heart rate) was found in the responses of those with alexithymia to fear imagery [27] and in the subjective report of emotional distress in individuals with PTSD who are alexithymic [28].

To enhance the measurement and identification of this multifaceted, clinically consequential construct of alexithymia, research could benefit from using facial, vocal, and linguistic features of emotional response. This approach is supported by a previous study showing that these features can be used in conjunction with machine learning (ML) models to identify those experiencing psychopathology following traumatic injury [29]. Facial, vocal, and linguistic features were extracted from recordings of participants' responses to questions about their trauma. These features were integrated into an ML neural network model to predict provisional PTSD diagnoses made 1 month after the traumatic injury and variance in PTSD symptom severity. The models achieved an average accuracy score of 0.90 to classify PTSD, based on the contribution of linguistic, vocal, and facial features. As PTSD symptom severity has been associated with alexithymia in several studies [30,31], these features could be shared for individuals with PTSD who score above the cut-off for alexithymia. The possible consistency between the distinctive digital phenotypes of PTSD symptom severity and alexithymia is supported by the relevance of language sentiment and facial expressivity differences for distinguishing individuals with both alexithymia and PTSD [22,27].

This study aimed to estimate the capacity of an ML classification model, built with digital phenotype variables extracted from recordings of war veterans with probable PTSD (hereafter referred to as PTSD) in which they describe traumatic incidents they experienced, to accurately classify individuals with alexithymia. On the basis of the reviewed research, we hypothesized that veterans with PTSD with alexithymia could be classified with a good degree of estimated accuracy, which is what we found. We also hypothesized that language variables would be the most important variables for the estimated capacity of the classification model to classify individuals as alexithymic, given their association with alexithymia in past studies [22,23]. However, in line with the attention-appraisal model of alexithymia and past research demonstrating the different channels through which alexithymia can manifest, we hypothesized that vocal, facial, and linguistic variables would all contribute to the estimated capacity of the best-performing model to make classifications of alexithymia.

Methods

Participants

Participants for this study were 101 veterans of the Australian Defence Force who were recruited via the Trialfacts health research platform. Five participants were excluded due to missing questionnaire responses, leaving 96 participants. The inclusion criteria for this study were being a former member of the Australian Defence Force, having experienced a traumatic event, and scoring above 33 on the Posttraumatic Stress Disorder Checklist for *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (PCL-5). Most of the sample were men (78/96, 81%), and the mean age of participants was 52.38 (SD 11.80) years. The size of the sample was determined based on the requirements of the extreme gradient boosting (XGBoost) classification models to identify those scoring as alexithymic on the TAS-20. Our sample size was larger than those in previous studies examining language features distinctive of individuals high on alexithymia [22] and using an ML classification approach to identify individuals with PTSD [29].

Stimuli

Before the experimental session, participants were asked to confirm if they felt comfortable discussing the traumatic event that had been affecting them the most. Participants were then asked to “think for a moment about a traumatic event you have been through” and to “tell me about this memory in detail... let yourself really try to get into this memory and how it made you feel.” Their responses were recorded for 1 minute.

Measures

Posttraumatic Stress Disorder

The PCL-5 [32] is a 20-item self-report measure of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, symptoms of PTSD. It was used to determine if the veteran participants met criteria for probable PTSD, based on their self-report of PTSD symptoms related to traumatic events they experienced during their military service. Recruited individuals who served in the military were deemed to have probable PTSD if they scored above the cut-off of 33 on the PCL-5 [33]. The PCL-5 has been found to have high internal consistency (Cronbach α from 0.83 to 0.98) and convergent validity (correlations with other PTSD measures of value up to 0.93), indicating it has strong psychometric properties [33,34].

Depression

The Beck Depression Inventory-second edition (BDI-II) [35] is a 21-item self-report measure that was used to index the intensity of depression symptoms in the participants of this study. This measure was used to determine if differences between those who were alexithymic or not alexithymic could be explained by levels of depression, given the overlapping presentation features of alexithymia and depression [36,37]. The BDI-II has been shown to have strong convergent and criterion validity as well as high internal consistency (Cronbach α of 0.9) and reliability [38].

Digital Phenotypes

Participants’ recorded descriptions of their traumatic experiences were processed using the OpenWillis Python (Python Software Foundation) library [39] and LIWC-22 [21] software.

Facial indicators are based on the facial action coding system [40]. This coding system measures the intensity of activity in both individual and groups of muscles in the face (designated by particular facial action units) that have been found to relate to particular emotional experiences, such as the 6 primary emotions of happiness, sadness, surprise, fear, disgust, and anger. OpenWillis uses DeepFace to measure framewise intensity of facial action coding system units on a range of -1 =expressivity of that emotion below baseline and 1 =expressivity of that emotion above baseline to produce facial emotion expressivity scores. DeepFace has been found to have 97% accuracy in correctly identifying the facial landmarks of faces it has been previously trained on [41] and 94% accuracy in identifying human emotions [42]. OpenWillis also uses MediaPipe [43] to measure the frame-by-frame coordinates of 468 unique facial landmarks using its Face Mesh model. From this, it produces a measure of the mean frame-to-frame movement occurring at these coordinates across the length of the video (producing OpenWillis variables such as “Upper face expressivity”; see [Multimedia Appendix 1](#) for a glossary of *OpenWillis* terms). MediaPipe was used as the building block for feature analysis in 1 study that was able to achieve 97% accuracy in correctly detecting human emotion [44].

OpenWillis uses Parselmouth to measure vocal variables [45]. Parselmouth is a Python implementation of the Praat software (University of Amsterdam) library [46]. Measured vocal variables include mean fundamental frequency, deviation in fundamental frequency, loudness, jitter, and shimmer of participants’ vocal production. Parselmouth also measures the percentage of frames without vocal content and the median duration of silences. In more recent updates, it has been able to examine more specific vocal features, such as cepstral peak prominence (CPP) and the mean, variance, and SD in mel-frequency cepstral coefficients. Praat software has been found to have good convergent validity (with other vocal software tools) and reliability in correctly identifying vocal features [47].

For language analysis, OpenWillis uses WhisperX to convert audio into text, which has a word error rate of 9.7%, outperforming previous speech-to-text models [48]. It uses the natural language processing Valence Aware Dictionary and Sentiment Reasoner (Massachusetts Institute of Technology) software [49] to analyze the extracted text in terms of language sentiment using a rule- or lexicon-based algorithm that produces mean scores from -1 =negative sentiment to 1 =positive sentiment. OpenWillis further measures the interaction between speech sentiment and first-person pronoun use (“first person language sentiment”). OpenWillis also uses LexicalDiversity [50], which is a natural language processing tool that measures lexical diversity in terms of moving average type-token ratio, which refers to the ratio of tokens (words) to the different types of words used in windows of 10 words at a time, that are then averaged across the whole segment of speech.

The LIWC library [51], which is the basis for the LIWC-22 software [21], is designed to process text files by counting the words in the text and calculating the percentage of words that correspond to each of the subdictionaries of LIWC (eg, the word “cry” would contribute to increasing the score of the subdictionaries of “emotion,” “affect,” and “verbs”). LIWC provides scores for each of its dictionaries, such as “power word use” and “word use related to feeling” [21]. Previous LIWC software has been found to have higher convergent and discriminant validity with other measures of emotion, such as self-report and rater coding, than competing text analysis software [52].

Affect Scale

Participants were asked to rate how they felt while describing the traumatic event on a 100-point verbal analogue scale (−100=means extremely negative, 0=neutral, and 100=extremely positive).

Alexithymia

The TAS-20 [4] was used to index the participant’s level of alexithymia. It has 20 items that each have a 5-point Likert rating scale (1=totally disagree and 5=totally agree) with items such as “I often don’t know why I am angry.” It has three factors: (1) difficulty describing feeling, (2) difficulty recognizing feeling, and (3) externally oriented thinking. The TAS-20 has an established cut-off score of 61 out of 100, over which an individual is deemed to have “alexithymia” [4]. A review of the TAS-20 measure determined that it has good factor validity, reliability, and internal consistency [53].

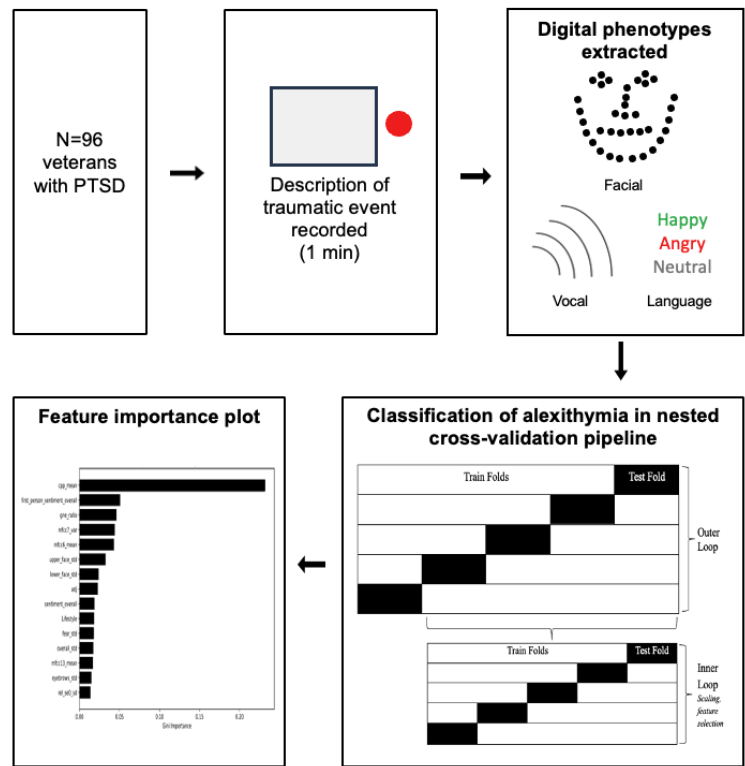
Procedure

The study was conducted via Zoom for all participants because we recruited veterans from across Australia. Participants initially completed informed consent, then completed the PCL-5, the TAS-20, and a range of demographic measures via Qualtrics. Participants were informed that they would be asked to describe a traumatic event they had experienced in detail. These descriptions were audio-visually recorded using the Apple QuickTime app and the record function in Zoom (Zoom Communications, Inc). Each video recording had a frame rate of 60 frames per second. After providing the description, participants were asked to rate how they felt while talking about the traumatic event on the Analogue Affect Scale. After completing this process, they were debriefed on the nature and purpose of the study.

Data Analysis

Participants were classified in terms of alexithymic status according to scores above or below the threshold of 61 on the TAS-20. This categorization resulted in 64 participants (n=53, 83% men; n=11, 17% women) being classified as alexithymic, and 32 participants (n=25, 78% men; n=7, 22% women) as nonalexithymic. XGBoost classification models were then built using the digital phenotype variables extracted from participant descriptions of traumatic events they have experienced. XGBoost classification models have been found to be both efficient and accurate in making classifications of features extracted from recordings of individuals with psychiatric conditions in past research [54]. In this study, these models were built using the scikit-learn package [55] in Python to attempt to classify individuals who scored in the “alexithymia” range on the TAS-20. The hyperparameters were set at the default values for the XGBoost classification algorithm in scikit-learn. Feature selection using the Recursive Feature Elimination method was completed within the inner folds in a 5-fold nested cross-validation pipeline to reduce the bias involved with doing feature selection on the whole sample and subsequently, the possibility of overfitting [56]. The number of features that contributed most to maximizing the precision score were selected and retained in a “best model.” The estimated classification performance of this model was evaluated across the outer 5 folds based on average scores for the precision, recall, F_1 -score, and area under the curve (AUC) metrics. Precision measures the rate at which the model is correctly identifying individuals as being in the “alexithymia” group by the number of times it is making the classification of “alexithymia.” Recall refers to the rate at which the model correctly identifies every individual with “alexithymia.” F_1 -scores are based on the harmonic mean of the precision and recall scores. The F_1 -score was used instead of the standard accuracy score in scikit-learn, as it has been found to be a more robust and subsequently appropriate measure when the two groups to be classified are imbalanced [57]. The AUC provides an indication of the probability that the model will rank each individual scoring above the cut-off for “alexithymia” as having a higher probability of being alexithymic than nonalexithymic. The average AUC scores derived from the receiver operating characteristic curve were used as they have been shown to be suitable for assessing classification with imbalanced datasets [58]. Variable feature importance was scored based on the average decrease in Gini impurity across all decision trees in the best-performing XGBoost model within the inner 5 folds of the nested cross-validation. The experimental process is represented in Figure 1.

Figure 1. Experimental process. PTSD: posttraumatic stress disorder.



Ethical Considerations

All procedures for the study were approved by the University of New South Wales Human Research Ethics Committee (HC230175). All methods were performed in accordance with the relevant guidelines and regulations. Informed consent was obtained from all participants via an online form explaining what was involved in the study before signing. The privacy and confidentiality of participants’ data have been stringently maintained, with no identifying details or features being presented or included in the write up of the study. Participants were given a gift voucher of Aus \$100 for participating in the study.

Results

Participant Characteristics

A chi-square test for participants’ sex at birth in the alexithymia and nonalexithymia groups found no significant group differences. One-way ANOVAs indicated no significant difference between those in the alexithymia and nonalexithymia groups for age, PCL-5, and BDI-II scores (Multimedia Appendix 1). This suggests that differences in digital phenotypes between those who were alexithymic and nonalexithymic were not driven by any of these examined covariates. Summary statistics for these variables are shown in Table 1.

Table 1. Participant characteristics.

Measures	Not alexithymic (n=32), mean (SD)	Alexithymic (n=64), mean (SD)
Age (y)	52.94 (11.83)	52.09 (11.91)
PCL-5 ^a	48.38 (14.21)	48.05 (13.44)
BDI-II ^b	30.00 (11.66)	32.44 (9.96)

^aPCL-5: Posttraumatic Stress Disorder Checklist for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

^bBDI-II: Beck Depression Inventory-second edition.

Nested Cross-Validation Results

The best-performing XGBoost classification model for determining whether participants scored above the cut-off for alexithymia had an average precision of 0.71, an average recall of 0.87, an average F_1 -score of 0.78, and an average AUC of 0.87 (Figure 2) across the 5 outer testing folds, with a model that used 148 features (performance metrics across each fold

are displayed in Table 2). The high recall score indicates that the XGBoost classification model was performing well at correctly identifying those individuals who scored above the threshold for alexithymia (“true positives”). However, in terms of the accuracy of all the alexithymia classifications it made, it was not performing as well, with 29% of those classifications being made incorrectly (“false positives”). These rates are illustrated in Figure 3, which is a confusion matrix depicting

the predictions of the best performing classification model across the 5 outer folds relative to the true alexithymia labels. These predictions produced an average overall accuracy (F_1 -score) of 0.78 (SD 0.07). The average AUC of 0.87 (SD 0.12) suggests that the model was performing well at assigning a higher probability that an individual scoring above the cut-off for “alexithymia” was alexithymic across each of the outer folds. The receiver operating characteristic curve depicting AUC across the outer folds of the nested cross-validation pipeline for this model is displayed in Figure 3.

Figure 4 displays the feature importance plot for the digital phenotype variables that were most important for classifying

individuals as alexithymic based on their Gini importance scores. Language, facial, and vocal variables were important to the capacity of the XGBoost model to classify alexithymia, with “word use related to feeling” emerging as the predictor with the largest Gini importance score. Other language (such as “first person language sentiment” and “language sentiment”), facial (such as “mean mouth openness”), and vocal (eg, “mel frequency cepstral 10 variance”) variables were also important to the classification capacity of the XGBoost models. The importance of language, facial, and vocal variables highlights the value of taking this multimodal approach to identifying a construct such as alexithymia, which has a distinctive presentation across multiple domains.

Figure 2. Receiver operating characteristic (ROC) curve, depicting the area under the curve (AUC) accuracy for alexithymia classification.

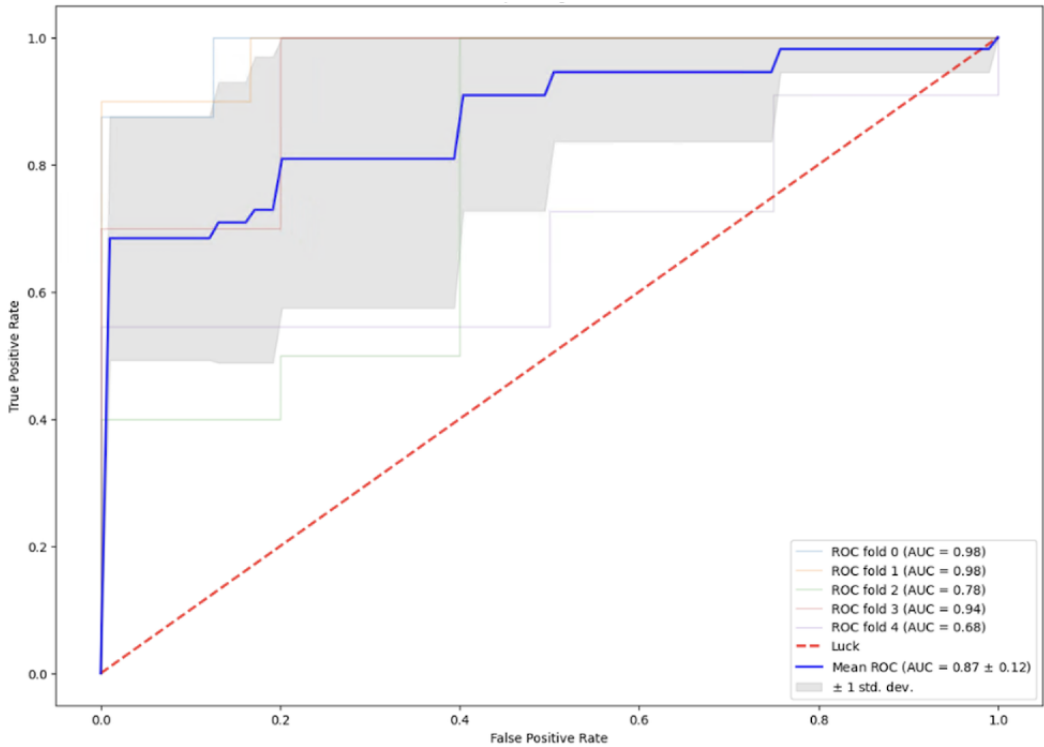


Table 2. Performance metrics for the extreme gradient boosting classification model across the 5 outer folds of the nested cross-validation pipeline.^a

Fold	Precision	Recall	F_1 -score accuracy
Fold 1	0.82	0.93	0.88
Fold 2	0.75	0.92	0.83
Fold 3	0.53	0.90	0.67
Fold 4	0.71	0.77	0.74
Fold 5	0.73	0.85	0.79

^aThe average across folds (SD) were as follows: precision, mean 0.71, SD 0.10; recall, mean 0.87, SD 0.06; F_1 -score accuracy, mean 0.78, SD 0.07.

Figure 3. Confusion matrix for the classification of alexithymia or nonalexithymia by the extreme gradient boosting (XGBoost) model.

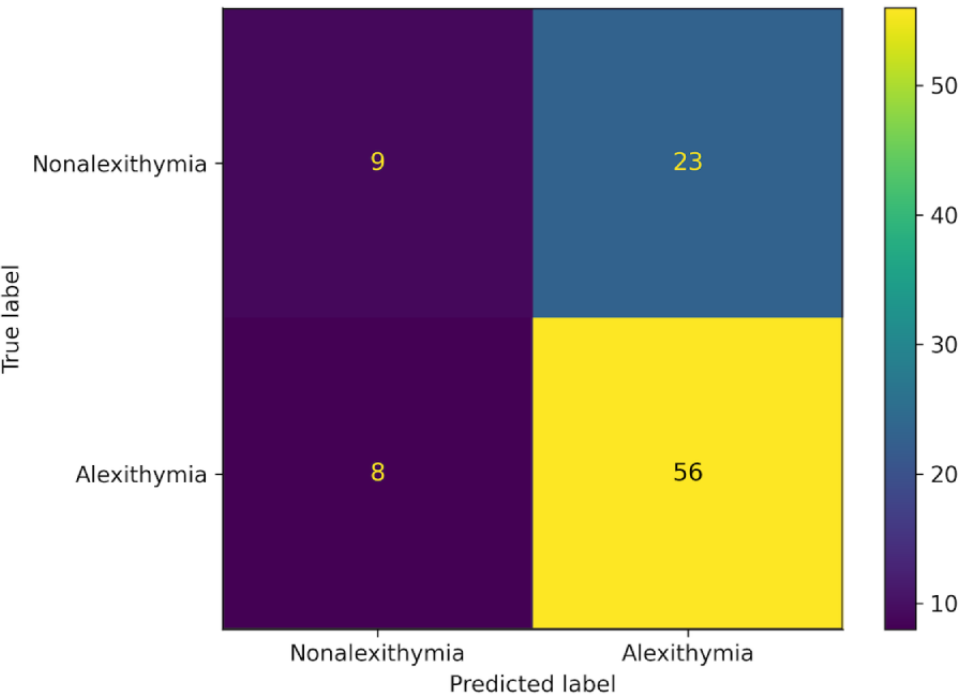
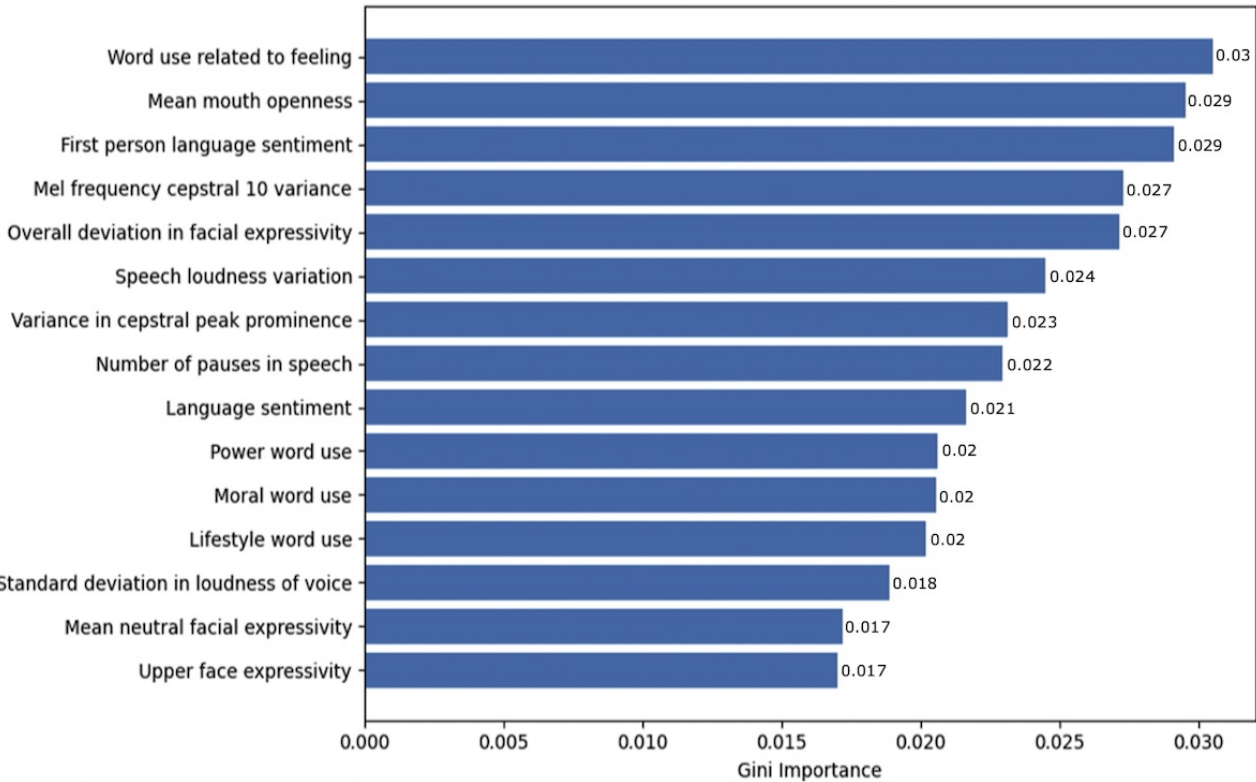


Figure 4. Feature importance plot depicting the most important features for the capacity of the extreme gradient boosting (XGBoost) model to accurately classify alexithymia.



Discussion

Principal Findings

This study examines the estimated capacity of an XGBoost classification model, built with digital phenotype variables extracted from recordings of war veterans with PTSD describing

traumatic events they had experienced, to accurately classify those veterans with alexithymia. These models were built and evaluated in a nested cross-validation pipeline to minimize the impact of bias [56,59].

In line with our hypothesis, the XGBoost classification model tuned and built within the nested cross-validation pipeline

demonstrated a level of accuracy and performance that indicated it could be used for classifying alexithymia in PTSD. Regarding the high recall score, with individuals who scored above the threshold for alexithymia, the model was estimated to be able to accurately classify these individuals as having alexithymia 87% of the time. The precision of the model, measuring how many of the classifications of alexithymia were accurate, was much lower, with only 71% of those classifications being accurate, suggesting the model may have been making too many “alexithymia” classifications. The XGBoost classification model had an average overall F_1 -score of 0.78, which is lower than the average F_1 -score achieved by the classification model for PTSD built using a similar approach [29]. This could be explained by this prior study attempting to classify trauma survivors with and without PTSD [29], whereas the present study focused only on identifying a subgroup of those with PTSD (those with alexithymia). The average AUC was 0.87, which is close to that identified for PTSD classification in a previous study and is considered model performance that suggests it has “considerable” clinical use [60]. However, this result needs to be interpreted with caution given its lack of stability, with the model only achieving an AUC of 0.55 in one of the folds. Overall, the performance of this XGBoost classification model suggests that such a model built with multiple digital phenotypes could be useful for identifying alexithymia in PTSD. This model must also be tested and validated on an independent sample of veterans with PTSD that was used in the model training process.

Digital Phenotypes of Alexithymia in PTSD

This study was the first to examine multiple digital phenotypes in the context of alexithymia in PTSD, and in doing so found that not only language variables but also facial and vocal features were important for the estimated classification of alexithymia. Mean mouth openness was the most relevant facial feature that contributed to classification performance. This may show that differences in how much individuals spoke, as demonstrated by the openness of their mouths, were a factor in the accurate classification of alexithymia. The most relevant vocal feature was the mel-frequency cepstral coefficient variables (mel-frequency cepstral 10 variance) and variance in CPP. The importance of variance in CPP, which is a measure of voice pathology [61,62], to the estimated capacity of the XGBoost model to classify individuals as alexithymic is consistent with past findings linking alexithymia with experiences of voice pathology using other measures [63,64]. The contribution of these facial and vocal features to the estimated classification of alexithymia expands and enhances the understanding of the expressions of emotional experience that could be relevant to this construct.

As hypothesized, language variables were important to the estimated capacity of the XGBoost model to classify individuals as alexithymic. The language variables that had the highest Gini importance scores were associated with the use of feeling words, sentiment of language, and first-person pronoun use. This aligns with foundational theoretical understandings of alexithymia as a deficit in the description of experiences that are associated with feelings and emotional sentiment [3]. It also supports

previous findings that those who score higher on the TAS-20 display differences in their expression of language sentiment [22,23]. In terms of first-person pronoun use, this aligns with theoretical understandings of alexithymia involving differences in the focus placed on oneself [2] and past findings that it is associated with differences in personal pronoun use [65]. This consistency between important predictors in an XGBoost classification model and expectations based on the research domain knowledge about that construct from theoretical models and past findings is an important indication of validity for ML models [66]. However, given that it was not only language variables alone that were important for the estimated classification of alexithymia but also vocal and facial variables, this aligns more closely with the attention-appraisal model's understanding of alexithymia as a multifaceted construct [3] than that of the language hypothesis of alexithymia [67].

Limitations

This study had several limitations. The size of the sample ($N=96$) is modest for an ML model such as XGBoost. The higher number of features (148 features) used in the best performing XGBoost classification model relative to the number of participants in the sample can increase the possibility of overfitting and deleteriously impact the stability of the model [68,69]. However, the use of nested cross-validation has been shown to minimize the impact of overfitting even in small samples, such that the models developed are replicated well in independent test sets [59,70]. Another limitation is that there was a large difference in the size of the groups to be classified (those scoring above or below the cut-off for “alexithymia” on the TAS-20), with more veterans in the sample scoring above the cut-off for “alexithymia.” This imbalance in groups impacts the capacity of the XGBoost classification model to be accurately evaluated. In the case of imbalanced classes, the classification of majority classes tends to be more accurate than that of the minority classes [71,72]. This phenomenon likely contributed to the much higher recall score found for this model relative to the precision score. However, this imbalance also reflects the generally higher incidence of alexithymia in veteran populations with PTSD, and trying to adjust these imbalances through cost-sensitive learning that incorporates oversampling or undersampling has been shown to have substantial limitations such as increasing overfitting [72-74]. The mostly male sample limits the generalizability of the digital phenotype findings to other PTSD populations (eg, civilians and women) and may have also contributed to the imbalance of groups on either side of the “alexithymia” cut-off, given that there is a small effect of sex on the TAS-20, with men generally scoring higher [75]. However, this higher proportion of men is representative of the defense force veteran sample used in this study [76]. Furthermore, this population is one in which alexithymia has a substantial impact on [77], which emphasizes the importance of improving our capacity to identify those who are alexithymic in this population.

Conclusions

Overall, this study suggests that facial, vocal, and language indicators could be used in the identification of veterans with PTSD who are experiencing alexithymia. We emphasize that

the model requires further validation in independent samples, but the findings represent an important first step and attest to the merits of continued research in this area. Particularly considering the limitations of self-report measures of alexithymia, this paradigm has the potential to advance research paradigms and the assessment of alexithymia in clinical settings. These advances could ultimately contribute to alexithymia being more easily identified in psychiatric contexts, leading to the allocation of more tailored and effective treatment resources

for addressing the specific challenges associated with alexithymia. Future research involving the implementation of this approach in clinical settings is required to examine its feasibility, efficiency, and ease of integration into clinical assessment, as well as its accuracy in identifying alexithymia. The improved identification of alexithymia in PTSD would be an important step in ameliorating the specific impacts that alexithymia has on the course and treatment of psychiatric conditions such as PTSD [12].

Data Availability

The deidentified OpenWillis variable datasets analyzed during this study are available in the OSF repository [78]. The raw video recording data analyzed during this study are not publicly available due to privacy and ethical requirements.

Authors' Contributions

TM, VY, IGL, and RB conceived the study. TM, VY, IGL, and RB devised the analysis plan. TM and VY conducted analyses and data curation. TM and RB provided validation and visualization. TM wrote the original draft. TM and RB reviewed and edited the final manuscript. RB acquired funding. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Glossary of relevant digital phenotype variables.

[DOCX File, 15 KB-Multimedia Appendix 1]

References

1. Sifneos PE. The prevalence of 'alexithymic' characteristics in psychosomatic patients. *Psychother Psychosom*. 1973;22(2):255-262. [doi: [10.1159/000286529](https://doi.org/10.1159/000286529)] [Medline: [4770536](https://pubmed.ncbi.nlm.nih.gov/4770536/)]
2. Taylor GJ, Bagby RM, Parker JD. Disorders of Affect Regulation: Alexithymia in Medical and Psychiatric Illness. Cambridge, UK. Cambridge University Press; 1997.
3. Preece D, Becerra R, Allan A, Robinson K, Dandy J. Establishing the theoretical components of alexithymia via factor analysis: introduction and validation of the attention-appraisal model of alexithymia. *Pers Individ Differ*. Dec 2017;119:341-352. [doi: [10.1016/j.paid.2017.08.003](https://doi.org/10.1016/j.paid.2017.08.003)]
4. Taylor GJ, Bagby R, Parker JD. The 20-Item Toronto Alexithymia Scale. IV. Reliability and factorial validity in different languages and cultures. *J Psychosom Res*. Sep 2003;55(3):277-283. [doi: [10.1016/s0022-3999\(02\)00601-3](https://doi.org/10.1016/s0022-3999(02)00601-3)] [Medline: [12932803](https://pubmed.ncbi.nlm.nih.gov/12932803/)]
5. Bankier B, Aigner M, Bach M. Alexithymia in DSM-IV disorder: comparative evaluation of somatoform disorder, panic disorder, obsessive-compulsive disorder, and depression. *Psychosomatics*. May 2001;42(3):235-240. [doi: [10.1176/appi.psy.42.3.235](https://doi.org/10.1176/appi.psy.42.3.235)] [Medline: [11351112](https://pubmed.ncbi.nlm.nih.gov/11351112/)]
6. Preece DA, Mehta A, Becerra R, Chen W, Allan A, Robinson K, et al. Why is alexithymia a risk factor for affective disorder symptoms? The role of emotion regulation. *J Affect Disord*. Jan 01, 2022;296:337-341. [doi: [10.1016/j.jad.2021.09.085](https://doi.org/10.1016/j.jad.2021.09.085)] [Medline: [34606815](https://pubmed.ncbi.nlm.nih.gov/34606815/)]
7. Fukunishi I, Sasaki K, Chishima Y, Anze M, Saijo M. Emotional disturbances in trauma patients during the rehabilitation phase: studies of posttraumatic stress disorder and alexithymia. *Gen Hosp Psychiatry*. Mar 1996;18(2):121-127. [doi: [10.1016/0163-8343\(95\)00121-2](https://doi.org/10.1016/0163-8343(95)00121-2)] [Medline: [8833582](https://pubmed.ncbi.nlm.nih.gov/8833582/)]
8. Monson CM, Price JL, Rodriguez BF, Ripley MP, Warner RA. Emotional deficits in military-related PTSD: an investigation of content and process disturbances. *J Trauma Stress*. Jun 2004;17(3):275-279. [doi: [10.1023/B:JOTS.0000029271.58494.05](https://doi.org/10.1023/B:JOTS.0000029271.58494.05)] [Medline: [15253100](https://pubmed.ncbi.nlm.nih.gov/15253100/)]
9. Kubota M, Miyata J, Hirao K, Fujiwara H, Kawada R, Fujimoto S, et al. Alexithymia and regional gray matter alterations in schizophrenia. *Neurosci Res*. Jun 2011;70(2):206-213. [doi: [10.1016/j.neures.2011.01.019](https://doi.org/10.1016/j.neures.2011.01.019)] [Medline: [21300113](https://pubmed.ncbi.nlm.nih.gov/21300113/)]
10. O'Driscoll C, Laing J, Mason O. Cognitive emotion regulation strategies, alexithymia and dissociation in schizophrenia, a review and meta-analysis. *Clin Psychol Rev*. Aug 2014;34(6):482-495. [FREE Full text] [doi: [10.1016/j.cpr.2014.07.002](https://doi.org/10.1016/j.cpr.2014.07.002)] [Medline: [25105273](https://pubmed.ncbi.nlm.nih.gov/25105273/)]
11. Frewen PA, Dozois DJ, Neufeld RW, Lanius RA. Meta-analysis of alexithymia in posttraumatic stress disorder. *J Trauma Stress*. Apr 2008;21(2):243-246. [doi: [10.1002/jts.20320](https://doi.org/10.1002/jts.20320)] [Medline: [18404647](https://pubmed.ncbi.nlm.nih.gov/18404647/)]

12. Putica A. Examining the role of emotion and alexithymia in cognitive behavioural therapy outcomes for posttraumatic stress disorder: clinical implications. *Cogn Behav Ther*. May 23, 2024;17:e15. [FREE Full text] [doi: [10.1017/S1754470X2400014X](https://doi.org/10.1017/S1754470X2400014X)]
13. Zahradnik M, Stewart SH, Marshall GN, Schell TL, Jaycox LH. Anxiety sensitivity and aspects of alexithymia are independently and uniquely associated with posttraumatic distress. *J Trauma Stress*. Apr 19, 2009;22(2):131-138. [FREE Full text] [doi: [10.1002/jts.20397](https://doi.org/10.1002/jts.20397)] [Medline: [19301252](https://pubmed.ncbi.nlm.nih.gov/19301252/)]
14. Kosten TR, Krystal JH, Giller EL, Frank J, Dan E. Alexithymia as a predictor of treatment response in post-traumatic stress disorder. *J Trauma Stress*. Oct 19, 1992;5(4):563-573. [doi: [10.1007/BF00979225](https://doi.org/10.1007/BF00979225)]
15. Zorzella KP, Muller RT, Cribbie RA, Bambragh V, Classen CC. The role of alexithymia in trauma therapy outcomes: examining improvements in PTSD, dissociation, and interpersonal problems. *Psychol Trauma*. Jan 2020;12(1):20-28. [doi: [10.1037/tra0000433](https://doi.org/10.1037/tra0000433)] [Medline: [30688506](https://pubmed.ncbi.nlm.nih.gov/30688506/)]
16. Nunes da Silva A. Developing emotional skills and the therapeutic alliance in clients with alexithymia: intervention guidelines. *Psychopathology*. 2021;54(6):282-290. [doi: [10.1159/000519786](https://doi.org/10.1159/000519786)] [Medline: [34749373](https://pubmed.ncbi.nlm.nih.gov/34749373/)]
17. Cameron K, Ogrodniczuk J, Hadjipavlou G. Changes in alexithymia following psychological intervention: a review. *Harv Rev Psychiatry*. 2014;22(3):162-178. [doi: [10.1097/HRP.0000000000000036](https://doi.org/10.1097/HRP.0000000000000036)] [Medline: [24736520](https://pubmed.ncbi.nlm.nih.gov/24736520/)]
18. King MF, Bruner GC. Social desirability bias: a neglected aspect of validity testing. *Psychol Mark*. Feb 2000;17(2):79-103. [doi: [10.1002/\(SICI\)1520-6793\(200002\)17:2<79::AID-MAR2>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1520-6793(200002)17:2<79::AID-MAR2>3.0.CO;2-0)]
19. van de Mortel TF. Faking it: social desirability response bias in self-report research. *Aust J Adv Nurs*. 2008;25(4):40-48. [FREE Full text] [doi: [10.37464/anmf](https://doi.org/10.37464/anmf)]
20. Lane RD, Ahern GL, Schwartz GE, Kaszniak AW. Is alexithymia the emotional equivalent of blindsight? *Biol Psychiatry*. Nov 01, 1997;42(9):834-844. [doi: [10.1016/s0006-3223\(97\)00050-4](https://doi.org/10.1016/s0006-3223(97)00050-4)] [Medline: [9347133](https://pubmed.ncbi.nlm.nih.gov/9347133/)]
21. Boyd RL, Ashokkumar A, Seraj S, Pennebaker JW. The development and psychometric properties of LIWC-22. University of Texas at Austin. 2022. URL: <https://www.liwc.app/static/documents/LIWC-22%20Manual%20-%20Development%20and%20Psychometrics.pdf> [accessed 2025-10-29]
22. Renzi A, Mariani R, Di Trani M, Tambelli R. Giving words to emotions: the use of linguistic analysis to explore the role of alexithymia in an expressive writing intervention. *Res Psychother*. Sep 07, 2020;23(2):452. [FREE Full text] [doi: [10.4081/ripppo.2020.452](https://doi.org/10.4081/ripppo.2020.452)] [Medline: [33024722](https://pubmed.ncbi.nlm.nih.gov/33024722/)]
23. Wotschack C, Klann-Delius G. Alexithymia and the conceptualization of emotions: a study of language use and semantic knowledge. *J Res Pers*. Oct 2013;47(5):514-523. [doi: [10.1016/j.jrp.2013.01.011](https://doi.org/10.1016/j.jrp.2013.01.011)]
24. Lee KS, Murphy J, Catmur C, Bird G, Hobson H. Furthering the language hypothesis of alexithymia: an integrated review and meta-analysis. *Neurosci Biobehav Rev*. Oct 2022;141:104864. [FREE Full text] [doi: [10.1016/j.neubiorev.2022.104864](https://doi.org/10.1016/j.neubiorev.2022.104864)] [Medline: [36087760](https://pubmed.ncbi.nlm.nih.gov/36087760/)]
25. Connelly M, Denney DR. Regulation of emotions during experimental stress in alexithymia. *J Psychosom Res*. Jun 2007;62(6):649-656. [doi: [10.1016/j.jpsychores.2006.12.008](https://doi.org/10.1016/j.jpsychores.2006.12.008)] [Medline: [17540222](https://pubmed.ncbi.nlm.nih.gov/17540222/)]
26. Kleiman A, Kramer KA, Wegener I, Koch AS, Geiser F, Imbierowicz K, et al. Psychophysiological decoupling in alexithymic pain disorder patients. *Psychiatry Res*. Mar 30, 2016;237:316-322. [doi: [10.1016/j.psychres.2016.01.021](https://doi.org/10.1016/j.psychres.2016.01.021)] [Medline: [26804974](https://pubmed.ncbi.nlm.nih.gov/26804974/)]
27. Peasley-Miklus CE, Panayiotou G, Vrana SR. Alexithymia predicts arousal-based processing deficits and discordance between emotion response systems during emotional imagery. *Emotion*. Mar 2016;16(2):164-174. [doi: [10.1037/emo0000086](https://doi.org/10.1037/emo0000086)] [Medline: [26461248](https://pubmed.ncbi.nlm.nih.gov/26461248/)]
28. Putica A, O'Donnell ML, Felmingham KL, Van Dam NT. Emotion response discordance among trauma-exposed adults: the impact of alexithymia. *Psychol Med*. Sep 2023;53(12):5442-5448. [FREE Full text] [doi: [10.1017/S0033291722002586](https://doi.org/10.1017/S0033291722002586)] [Medline: [35975360](https://pubmed.ncbi.nlm.nih.gov/35975360/)]
29. Schultebraucks K, Yadav V, Shalev AY, Bonanno GA, Galatzer-Levy IR. Deep learning-based classification of posttraumatic stress disorder and depression following trauma utilizing visual and auditory markers of arousal and mood. *Psychol Med*. Apr 03, 2022;52(5):957-967. [doi: [10.1017/S0033291720002718](https://doi.org/10.1017/S0033291720002718)] [Medline: [32744201](https://pubmed.ncbi.nlm.nih.gov/32744201/)]
30. Edwards ER. Posttraumatic stress and alexithymia: a meta-analysis of presentation and severity. *Psychol Trauma*. Oct 2022;14(7):1192-1200. [doi: [10.1037/tra0000539](https://doi.org/10.1037/tra0000539)] [Medline: [31841012](https://pubmed.ncbi.nlm.nih.gov/31841012/)]
31. Putica A, Van Dam NT, Felmingham KL, O'Donnell ML. Alexithymia and treatment response for prolonged exposure therapy: an evaluation of outcomes and mechanisms. *Psychotherapy (Chic)*. Mar 2024;61(1):44-54. [doi: [10.1037/pst0000512](https://doi.org/10.1037/pst0000512)] [Medline: [37917810](https://pubmed.ncbi.nlm.nih.gov/37917810/)]
32. Weathers FW, Bovin MJ, Lee DJ, Sloan DM, Schnurr PP, Kaloupek DG, et al. The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans. *Psychol Assess*. Mar 2018;30(3):383-395. [FREE Full text] [doi: [10.1037/pas0000486](https://doi.org/10.1037/pas0000486)] [Medline: [28493729](https://pubmed.ncbi.nlm.nih.gov/28493729/)]
33. McDonald SD, Calhoun PS. The diagnostic accuracy of the PTSD checklist: a critical review. *Clin Psychol Rev*. Dec 2010;30(8):976-987. [doi: [10.1016/j.cpr.2010.06.012](https://doi.org/10.1016/j.cpr.2010.06.012)] [Medline: [20705376](https://pubmed.ncbi.nlm.nih.gov/20705376/)]
34. Wilkins KC, Lang AJ, Norman SB. Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. *Depress Anxiety*. Jul 2011;28(7):596-606. [FREE Full text] [doi: [10.1002/da.20837](https://doi.org/10.1002/da.20837)] [Medline: [21681864](https://pubmed.ncbi.nlm.nih.gov/21681864/)]

35. Beck AT, Steer RA, Brown G. Manual for the Beck Depression Inventory-II. San Antonio, TX. Psychological Corporation; 1996.
36. Sagar R, Talwar S, Desai G, Chaturvedi SK. Relationship between alexithymia and depression: a narrative review. *Indian J Psychiatry*. 2021;63(2):127-133. [FREE Full text] [doi: [10.4103/psychiatry.IndianJPsychiatry_738_19](https://doi.org/10.4103/psychiatry.IndianJPsychiatry_738_19)] [Medline: [34194055](https://pubmed.ncbi.nlm.nih.gov/34194055/)]
37. Hemming L, Haddock G, Shaw J, Pratt D. Alexithymia and its associations with depression, suicidality, and aggression: an overview of the literature. *Front Psychiatry*. Apr 11, 2019;10:203. [FREE Full text] [doi: [10.3389/fpsyt.2019.00203](https://doi.org/10.3389/fpsyt.2019.00203)] [Medline: [31031655](https://pubmed.ncbi.nlm.nih.gov/31031655/)]
38. Wang Y, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Braz J Psychiatry*. 2013;35(4):416-431. [FREE Full text] [doi: [10.1590/1516-4446-2012-1048](https://doi.org/10.1590/1516-4446-2012-1048)] [Medline: [24402217](https://pubmed.ncbi.nlm.nih.gov/24402217/)]
39. Worthington M, Efstathiadis G, Yadav V, Abbas A. 172. OpenWillis: an open-source python library for digital health measurement. *Biol Psychiatry*. May 2024;95(10):S169-S170. [doi: [10.1016/j.biopsych.2024.02.407](https://doi.org/10.1016/j.biopsych.2024.02.407)]
40. Ekman P. Facial Action Coding System : Investigator's Guide. Palo Alto, CA. Consulting Psychologists Press; 1978.
41. Taigman Y, Yang M, Ranzato M, Wolf L. DeepFace: closing the gap to human-level performance in face verification. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014. Presented at: CVPR 2014; June 23-28, 2014; Columbus, OH. URL: <https://ieeexplore.ieee.org/document/6909616> [doi: [10.1109/cvpr.2014.220](https://doi.org/10.1109/cvpr.2014.220)]
42. Venkatesan R, Shirley S, Selvarathi M, Jebaseeli TJ. Human emotion detection using DeepFace and artificial intelligence. *Eng Proc*. 2023;59(1):37. [FREE Full text] [doi: [10.3390/engproc2023059037](https://doi.org/10.3390/engproc2023059037)]
43. Lugaresi C, Tang J, Nash H, McClanahan C, Uboweja E, Hays M, et al. MediaPipe: a framework for perceiving and processing reality. In: *Proceedings of the Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition*. 2019. Presented at: CVPR 2019; June 17, 2019; Long Beach, CA. URL: <https://www.semanticscholar.org/paper/MediaPipe%3A-A-Framework-for-Perceiving-and-Reality-Lugaresi-Tang/1cd227fb3dacda18ee94d08b04fcb1d5b9afb351>
44. Siam AI, Soliman NF, Algarni AD, Abd El-Samie FE, Sedik A. Deploying machine learning techniques for human emotion detection. *Comput Intell Neurosci*. Feb 2, 2022;2022:8032673. [FREE Full text] [doi: [10.1155/2022/8032673](https://doi.org/10.1155/2022/8032673)] [Medline: [35154306](https://pubmed.ncbi.nlm.nih.gov/35154306/)]
45. Jadoul Y, Thompson B, de Boer B. Introducing Parselmouth: a Python interface to Praat. *J Phon*. Nov 2018;71:1-15. [doi: [10.1016/j.wocn.2018.07.001](https://doi.org/10.1016/j.wocn.2018.07.001)]
46. Boersma P, Weenink DJM. Praat, a system for doing phonetics by computer. Research Gate. URL: https://www.researchgate.net/publication/208032992_PRAAT_a_system_for_doing_phonetics_by_computer [accessed 2025-11-04]
47. Burris C, Vorperian HK, Fourakis M, Kent RD, Bolt DM. Quantitative and descriptive comparison of four acoustic analysis systems: vowel measurements. *J Speech Lang Hear Res*. Mar 2014;57(1):26-45. [doi: [10.1044/1092-4388\(2013/12-0103\)](https://doi.org/10.1044/1092-4388(2013/12-0103))] [Medline: [24687465](https://pubmed.ncbi.nlm.nih.gov/24687465/)]
48. Bain M, Huh J, Han T, Zisserman A. WhisperX: time-accurate speech transcription of long-form audio. ArXiv.. Preprint posted online on March 1, 2023. [FREE Full text] [doi: [10.21437/interspeech.2023-78](https://doi.org/10.21437/interspeech.2023-78)]
49. Hutto C, Gilbert E. VADER: a parsimonious rule-based model for sentiment analysis of social media text. *Proc Int AAAI Conf Web Soc Media*. 2014;8(1):216-225. [FREE Full text] [doi: [10.1609/icwsm.v8i1.14550](https://doi.org/10.1609/icwsm.v8i1.14550)]
50. Shen L. LexicalRichness: a small module to compute textual lexical richness. GitHub. 2022. URL: <https://github.com/LSYS/lexicalrichness> [accessed 2025-10-29]
51. Pennebaker JW, Booth RJ, Francis ME. Linguistic inquiry and word count: LIWC2007. LIWC.net. 2007. URL: http://www.gruberpeplab.com/teaching/psych231_fall2013/documents/231_Pennebaker2007.pdf [accessed 2024-11-08]
52. Bantum EO, Owen JE. Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychol Assess*. Mar 2009;21(1):79-88. [doi: [10.1037/a0014643](https://doi.org/10.1037/a0014643)] [Medline: [19290768](https://pubmed.ncbi.nlm.nih.gov/19290768/)]
53. Bagby RM, Parker JD, Taylor GJ. Twenty-five years with the 20-item Toronto Alexithymia Scale. *J Psychosom Res*. Apr 2020;131:109940. [doi: [10.1016/j.jpsychores.2020.109940](https://doi.org/10.1016/j.jpsychores.2020.109940)] [Medline: [32007790](https://pubmed.ncbi.nlm.nih.gov/32007790/)]
54. Olah J, Wong WL, Chaudhry AU, Mena O, Tang SX. Detecting schizophrenia, bipolar disorder, psychosis vulnerability and major depressive disorder from 5 minutes of online-collected speech. *Transl Psychiatry*. Jul 12, 2025;15(1):241. [FREE Full text] [doi: [10.1038/s41398-025-03433-0](https://doi.org/10.1038/s41398-025-03433-0)] [Medline: [40651952](https://pubmed.ncbi.nlm.nih.gov/40651952/)]
55. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2012. URL: <https://www.jmlr.org/papers/v12/pedregosa11a.html> [accessed 2025-11-04]
56. Zhong S, Chen Y, Liu S. Facial expression recognition using local feature selection and the extended nearest neighbor algorithm. In: *Proceedings of the Seventh International Symposium on Computational Intelligence and Design*. 2014. Presented at: ISCID 2014; December 13-14, 2014; Hangzhou, China. URL: <https://ieeexplore.ieee.org/document/7064202> [doi: [10.1109/iscid.2014.108](https://doi.org/10.1109/iscid.2014.108)]
57. Branco P, Torgo L, Ribeiro RP. A survey of predictive modeling on imbalanced domains. *ACM Comput Surv*. Aug 13, 2016;49(2):1-50. [doi: [10.1145/2907070](https://doi.org/10.1145/2907070)]
58. Richardson E, Trevizani R, Greenbaum JA, Carter H, Nielsen M, Peters B. The receiver operating characteristic curve accurately assesses imbalanced datasets. *Patterns (N Y)*. Jun 14, 2024;5(6):100994. [FREE Full text] [doi: [10.1016/j.patter.2024.100994](https://doi.org/10.1016/j.patter.2024.100994)] [Medline: [39005487](https://pubmed.ncbi.nlm.nih.gov/39005487/)]

59. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PLoS One*. 2019;14(11):e0224365. [FREE Full text] [doi: [10.1371/journal.pone.0224365](https://doi.org/10.1371/journal.pone.0224365)] [Medline: [31697686](https://pubmed.ncbi.nlm.nih.gov/31697686/)]
60. Çorbacioğlu Ş, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: a guide to interpreting the area under the curve value. *Turk J Emerg Med*. 2023;23(4):195-198. [FREE Full text] [doi: [10.4103/tjem.tjem_182_23](https://doi.org/10.4103/tjem.tjem_182_23)] [Medline: [38024184](https://pubmed.ncbi.nlm.nih.gov/38024184/)]
61. Godino-Llorente JI, Osma-Ruiz V, Sáenz-Lechón N, Gómez-Vilda P, Blanco-Velasco M, Cruz-Roldán F. The effectiveness of the glottal to noise excitation ratio for the screening of voice disorders. *J Voice*. Jan 2010;24(1):47-56. [doi: [10.1016/j.jvoice.2008.04.006](https://doi.org/10.1016/j.jvoice.2008.04.006)] [Medline: [19135854](https://pubmed.ncbi.nlm.nih.gov/19135854/)]
62. Maryn Y, Roy N, De Bodt M, Van Cauwenberge P, Corthals P. Acoustic measurement of overall voice quality: a meta-analysis. *J Acoust Soc Am*. Nov 2009;126(5):2619-2634. [doi: [10.1121/1.3224706](https://doi.org/10.1121/1.3224706)] [Medline: [19894840](https://pubmed.ncbi.nlm.nih.gov/19894840/)]
63. Baker J, Oates JM, Leeson E, Woodford H, Bond MJ. Patterns of emotional expression and responses to health and illness in women with functional voice disorders (MTVD) and a comparison group. *J Voice*. Nov 2014;28(6):762-769. [doi: [10.1016/j.jvoice.2014.03.005](https://doi.org/10.1016/j.jvoice.2014.03.005)] [Medline: [24962231](https://pubmed.ncbi.nlm.nih.gov/24962231/)]
64. Deary IJ, Wilson JA, Carding PN, Mackenzie K. The dysphonic voice heard by me, you and it: differential associations with personality and psychological distress. *Clin Otolaryngol Allied Sci*. Aug 2003;28(4):374-378. [doi: [10.1046/j.1365-2273.2003.00730.x](https://doi.org/10.1046/j.1365-2273.2003.00730.x)] [Medline: [12871257](https://pubmed.ncbi.nlm.nih.gov/12871257/)]
65. Edwards ER, Shivaji S, Micek A, Wupperman P. Distinguishing alexithymia and emotion differentiation conceptualizations through linguistic analysis. *Pers Individ Differ*. Apr 2020;157:109801. [doi: [10.1016/j.paid.2019.109801](https://doi.org/10.1016/j.paid.2019.109801)]
66. Cava WL, Bauer C, Moore JH, Pendergrass SA. Interpretation of machine learning predictions for patient outcomes in electronic health records. *AMIA Annu Symp Proc*. 2019;2019:572-581. [FREE Full text] [Medline: [32308851](https://pubmed.ncbi.nlm.nih.gov/32308851/)]
67. Hobson H, Brewer R, Catmur C, Bird G. The role of language in alexithymia: moving towards a multiroute model of alexithymia. *Emot Rev*. May 23, 2019;11(3):247-261. [doi: [10.1177/1754073919838528](https://doi.org/10.1177/1754073919838528)]
68. Zhu JJ, Yang M, Ren ZJ. Machine learning in environmental research: common pitfalls and best practices. *Environ Sci Technol*. Nov 21, 2023;57(46):17671-17689. [doi: [10.1021/acs.est.3c00026](https://doi.org/10.1021/acs.est.3c00026)] [Medline: [37384597](https://pubmed.ncbi.nlm.nih.gov/37384597/)]
69. Rajput D, Wang WJ, Chen CC. Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*. Mar 14, 2023;24(1):48. [FREE Full text] [doi: [10.1186/s12859-023-05156-9](https://doi.org/10.1186/s12859-023-05156-9)] [Medline: [36788550](https://pubmed.ncbi.nlm.nih.gov/36788550/)]
70. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*. Mar 23, 2006;7(1):91. [FREE Full text] [doi: [10.1186/1471-2105-7-91](https://doi.org/10.1186/1471-2105-7-91)] [Medline: [16504092](https://pubmed.ncbi.nlm.nih.gov/16504092/)]
71. Japkowicz N, Stephen S. The class imbalance problem: a systematic study. *Int Data Anal*. 2002;6(5):429-449. [doi: [10.3233/ida-2002-6504](https://doi.org/10.3233/ida-2002-6504)]
72. Weiss GM, Provost F. Learning when training data are costly: the effect of class distribution on tree induction. *J Artif Intell Res*. Oct 01, 2003;19(1):315-354. [FREE Full text]
73. Welvaars K, Oosterhoff J, van den Bekerom MP, Doornberg J, van Haarst EP, OLVG Urology Consortium, and the Machine Learning Consortium. Implications of resampling data to address the class imbalance problem (IRCIP): an evaluation of impact on performance between classification algorithms in medical data. *JAMIA Open*. Jul 2023;6(2):o0ad033. [FREE Full text] [doi: [10.1093/jamiaopen/o0ad033](https://doi.org/10.1093/jamiaopen/o0ad033)] [Medline: [37266187](https://pubmed.ncbi.nlm.nih.gov/37266187/)]
74. Alkhawaldeh IM, Albalkhi I, Naswhan AJ. Challenges and limitations of synthetic minority oversampling techniques in machine learning. *World J Methodol*. Dec 20, 2023;13(5):373-378. [FREE Full text] [doi: [10.5662/wjm.v13.i5.373](https://doi.org/10.5662/wjm.v13.i5.373)] [Medline: [38229946](https://pubmed.ncbi.nlm.nih.gov/38229946/)]
75. Mendia J, Zumeta LN, Cusi O, Pascual A, Alonso-Arbiol I, Díaz V, et al. Gender differences in alexithymia: insights from an updated meta-analysis. *Pers Individ Differ*. Sep 2024;227:112710. [doi: [10.1016/j.paid.2024.112710](https://doi.org/10.1016/j.paid.2024.112710)]
76. Service with the Australian Defence Force: census. Australian Bureau of Statistics. 2021. URL: <https://www.abs.gov.au/statistics/people/people-and-communities/service-australian-defence-force-census/latest-release> [accessed 2025-10-29]
77. Becirovic E, Avdibegovic E, Softic R, Mirkovic-Hajdukovic M, Becirovic A. Alexithymia in war veterans with post-traumatic stress disorder. *Eur Psychiatr*. Mar 23, 2020;41(S1):S720. [FREE Full text] [doi: [10.1016/j.eurpsy.2017.01.1301](https://doi.org/10.1016/j.eurpsy.2017.01.1301)]
78. Meaney T. Digital phenotypes of alexithymia in posttraumatic stress disorder. *OSF*. 2025. URL: https://osf.io/9kw4j/overview?view_only=89ff5416184547bb87028a01a6c5c2f7 [accessed 2025-11-04]

Abbreviations

AUC: area under the curve

BDI-II: Beck Depression Inventory-second edition

CPP: cepstral peak prominence

LIWC: Linguistic Inquiry and Word Count

ML: machine learning

PCL-5: Posttraumatic Stress Disorder Checklist for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

PTSD: posttraumatic stress disorder

TAS-20: Toronto Alexithymia Scale-20

XGBoost: extreme gradient boosting

Edited by J Torous; submitted 05.Sep.2025; peer-reviewed by A Putica, KW Tay; comments to author 23.Oct.2025; revised version received 26.Oct.2025; accepted 27.Oct.2025; published 13.Nov.2025

Please cite as:

Meaney T, Yadav V, Galatzer-Levy I, Bryant R

Using Digital Phenotypes to Identify Individuals With Alexithymia in Posttraumatic Stress Disorder: Cross-Sectional Study

JMIR Ment Health 2025;12:e83575

URL: <https://mental.jmir.org/2025/1/e83575>

doi: [10.2196/83575](https://doi.org/10.2196/83575)

PMID:

©Tomas Meaney, Vijay Yadav, Isaac Galatzer-Levy, Richard Bryant. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 13.Nov.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.