Depression screening using a non-verbal self-association task: A machine-learning based pilot study

Depression screening using a non-verbal self-association task: A machine-learning based pilot study

Yang S. Liu[1], Yipeng Song[1], Naomi A. Lee[2], Daniel M. Bennett[2], Katherine S Button[3], Andrew Greenshaw[1], Bo Cao[1], Jie Sui[2]

1 Department of Psychiatry, University of Alberta, Edmonton Alberta, Canada

2 School of Psychology, University of Aberdeen, Aberdeen, Scotland

3 Department of Psychology, University of Bath, Bath, England

Background. Effective screening is important to combat the raising burden of depression and opens a critical time window for early intervention. Clinical use of non-verbal depression screening is nascent, yet a promising and viable candidate to supplement verbal screening. Differential self- and emotion-processing in depression patients were previously reported by non-verbal behavioural assessments, corroborated by neuroimaging findings of distinct neuroanatomical markers, Thus non-verbal validated brain-behavior based self- emotion- related assessment data reflect physiological differences and may support individual level screening of depression.

Methods. In this pilot study (n = 84) we collected two longitudinal sessions of behavioural assessment data in a laboratory setting. Depression was assessed using Beck Depression Inventory II (BDI-II), to explore optimal screening methods with machine-learning, and to establish the validity of adapting a novel behavioural assessment focusing on self and emotions for depression screening.

Results. The best machine-learning model achieved high performance in depression screening, 10-Fold cross-validation (CV) Area Under the receiver operating characteristic Curve (AUC) of 0.90 and balanced accuracy of 0.81, using a Gradient Boosting algorithm. Prospective prediction using a model trained with session 1 data to predict session 2 depression status achieved a 10-Fold CV AUC of 0.77 and balanced accuracy of 0.66. We also identified interpretable behavioural signatures for depression patients based on the best model.

Conclusion. The study supports the utility of using behavioural data as a viable and cost-effective solution for depression screening, with a potential wide range of applications in clinical settings. Keywords: Depression, self, machine-learning, matching technique, sensitive objective measurement.

Depression screening using a non-verbal self-association task: A machine-learning based pilot study

Depression is highly prevalent and will become the most burdensome disease world wide by 2030 (World Health Organization, 2008). Early identification and intervention could be an effective first step for harm reduction and prevention (Hall and Reynolds, 2014; James et al., 2018). Currently, popular depression screening tools often rely heavily on screening questions, asked by the clinicians directly, or administered as a self-reported questionnaire. While this mode of screening usually works well, the utility is impacted by subject's cognitive functions (e.g., aphasia), subjective biases (e.g., interpreting questions and Likert scale items differently), and personal agenda of the patients (e.g., manufacturing symptoms or exaggerating symptoms of depression or trying to avoid detection). The impact of these limitations may be cumulative if the same screening tools are used repeatedly on the same patient. Though depression screening may use multiple information sources, here we broadly refer to screening methods that utilize patients' verbal responses as verbal screening methods, and methods that do not utilize patients' verbal responses as non-verbal screening methods. Given the limitations of existing assessments, non-verbal alternatives of depression screening are desirable, and may complement and facilitate screening accuracy of verbal assessments. Although not widely adopted in clinical settings, various non-verbal depression screening methods have already been explored, including facial perception behavioural assessment (Guekht et al., 2013), eye gaze pattern analysis (Alghowinem et al., 2015), facial visual features and voice analysis (Pampouchidou et al., 2019; Scibelli et al., 2018), and there's a plethora of research on identifying depression using neuroimaging data and machine-learning (ML) (Janssen et al., 2018). A common issue with the non-verbal depression screening methods such as EEG and MRI are the high cost (CostHelper, 2021; Sahu et al., 2020)

with poor interpretability (Schnack, 2019). Comparing with non-verbal screening methods relying on EEG and MRI data, costs are less prohibitive for methods relying on video and audio features of the patients, yet costs on additional hardware (e.g., video camera, microphone) and technology-based services (e.g., secure storage and handling of patient's video and audio data) and resource intensive computational pipelines to process video/audio data may prohibit patient access (Pampouchidou et al., 2019). Thus, an effective, yet low-cost alternative non-verbal screening method for depression screening is desirable

Recently, the concept of self-related processing has been proposed as a potential marker for mental illnesses (Sui et al., 2021; Sui and Gu, 2017). Also, self-related processing may have utility in identifying neurodevelopmental disorders, where different behavioural signatures were observed in autistic people compared to non-autistic subjects (Moseley et al., 2021). Bias in self-related processing and emotion processing has already been linked to real-life behaviours (e.g., habits; Verplanken and Sui, 2019), and outcomes (e.g., loneliness, anxiety, and depression; Feldborg et al., 2021; Hobbs et al., 2021). People who are depressed are more sensitive to negative affect and show a negativity bias when interpreting ambiguous stimuli (Sui et al., 2016). In addition, a bias towards perceiving ambiguous emotions as sad may change behaviour, which may elicit negative reactions from others and thus sustain these biases in people with depression (Stolte et al., 2017). Further, the findings of differential behavioural patterns between depressed and non depressed people on the self-referential task are corroborated by neuroimaging findings of district underlying neuroanatomical structure of the self (e.g., the medial prefrontal cortex, the posterior cingulate cortex and dorsolateral prefrontal cortex) (Delaveau et al., 2015; Freton et al., 2014; Nejad et al., 2019; Northoff, 2014) and emotional processing (Delaveau et al., 2011; Enneking et al., 2020; Lemogne et al., 2011; Northoff et al., 2009). This suggests nuances in

self- and emotion- related behavioural patterns reflect physiological differences and may better differentiate depression than using simple response patterns (e.g., mean response time and accuracy). Further, the self- and emotion behavioural assessments are easy to access and has low adaptability barriers, demonstrated by prior successful applications on patients with brain lesions (Sui and Humphreys, 2017) and children between 6 to 10 years old (Maire et al., 2020). The earlier behavioural and neuroimaging evidence, and evidence for ease of access probed us to ask whether accurate detection of depression is achievable using the validated brain-behaviour based self- and emotional- assessments (Sui et al., 2013; Yankouskaya and Sui, 2021). No prior studies have directly tested whether the non-verbal behavioural signatures from self-related and emotional processing could be used directly for screening of depression, and whether the behavioural signatures are meaningful and interpretable.

The adaptation of behavioural assessments for depression screening requires predictive modeling, building models capable of making accurate predictions on subjects' depression status. Traditional statistical models are mainly used for identifying associations between the explanatory variables and the outcome (e.g., How behavioural assessments are associated with depression?), and are not optimized for making individual level predictions. On the other hand, ML-based approaches are excellent for building generalizable predictive models for such a purpose, at the expense of computational complexity (Bzdok et al., 2018). The first aim of the current study was to optimize the model's predictive accuracy of identifying depression, rather than exploring group-level risk factors. Thus, we choose an ML approach rather than traditional multivariate statistical models for prediction.

The second aim was to adopt a new procedure for examining social biases (based on self and emotion) and the underlying neural architecture (Humphreys and Sui, 2016; Stolte et al.,

2017; Sui et al., 2012) for depression screening. We used two non-verbal behavioural tasks: a self perceptual matching task, in which geometric shapes were arbitrarily assigned to the self, a friend, or a stranger, and an emotion matching task, in which shapes are associated with happy, neutral or sad emotion. Participants then judged whether shape-label pairs were as originally shown or re-paired. Response times and accuracy of different trial types were measured and used as behavioural features. Beck Depression Inventory II (BDI-II) (Beck et al., 1996) was used to identify depression. We collected longitudinal data from two separate within-subject sessions and evaluated performance of various models, including linear logistic regression model with or without Lasso type regularization and the nonlinear Gradient Boosting (GB) model (Friedman, 2001) on within session (model 1 and 2), and across-session (model 3) depression classification. We also tested prospective prediction of depression between sessions, and further identified behavioural signatures from this novel approach.

Methods

Behavioural Experiment

Participants

University students were recruited through volunteer sampling. Participants were screened using the BDI-II to ensure balanced levels of depression from no depression, mild depression (BDI-II: 14-19), moderate to severe depression (29-63). Eighty-four subjects participated in this two-session study. All participants had normal or corrected-to-normal vision. Informed consent was obtained from all participants prior to the experiment according to procedures approved by the University of Bath's ethics committee.

Depression symptoms over the last two weeks was assessed using the BDI-II (Beck et al., 1996). This measure consists of 21 questions associated with depression; each question has an integer score with a range of 0 to 3. A total score is calculated by summing individual question

scores. Higher total score indicate more severe depressive symptoms, where a score of 0 to 13 suggest minimal depression symptoms (Beck et al., 1996). Participants with BDI-II score <= 13 were considered as the control group, and participants with a BDI-II score > 13 were considered as the depression group. BDI-II was selected based on previous research, that it can achieve a high sensitivity to screen for clinical depression, reaching a high balanced accuracy, an average of sensitivity and specificity, of 0.82 (von Glischinski et al., 2019). The participants were assessed in two different sessions, which took place one week apart. In session 1, the depression group included 59 subjects, and the control included 25 subjects (Table 1). In session 2, the depression group included 54 subjects and the control group included 29 (Table 1).

--- Insert Figure 1 Here ---

The self-association was measured using a perceptual matching task, where three geometric shapes (triangle, circle, and square) were respectively assigned to three personal labels representing the participant, their named best friend (whom the participant were asked to name), and a stranger (the participant was asked to select a name that was familiar but not held by anyone they knew) (Sui et al., 2012). The assignment of shapes to the different people was counterbalanced across individuals. After the associative instruction, participants immediately underwent a shape-label matching task to judge whether a shape-label pair matched. A shape with 3.5° × 3.5° of visual angle appeared above a white central fixation cross with 0.8° × 0.8°of visual angle. One of three labels (‘You’, ‘Friend, or ‘Stranger’ [covering 1.76°/2.52° × 1.76° of visual angle]) was displayed below the fixation cross. The task was to decide whether the shape-label pair was the same as initially shown or whether the shape and label were mismatched

(Figure 1). All stimuli were displayed on a grey background. E-prime software (version 2.0) (Schneider et al., 2002) was used to present the stimuli and to record responses. The experiment was run on a PC with a 22-in monitor (1920 × 1080 pixels) at 60 Hz.

The emotion-association measurement was identical to that in the self-association measurement, except that the three different geometric shapes were used (e.g., diamond, hexagon, pentagon) and were respectively assigned to drawings of faces, one with a happy expression, one with a neutral expression, and the third with a sad expression.

Procedure

Two testing sessions (session 1 and session 2) took place one week apart. In each session, participants completed the BDI-II first, following by the self-perceptual matching task, and then emotion perceptual matching task. In each session, the order of self and emotion tasks was counterbalanced across participants.

In the perceptual matching task, participants were first asked to name one of their best friends. To measure the familiarity with their best friend, participants were instructed to rate how long they have known each other (years), how familiar with a 1-7 scale (1 = not familiar at all, 7 = highly familiar), how often they meet with a 1-7 scale (1 = not frequent at all, 7 = very frequent), and how often (daily, weekly, monthly, yearly). Participants then selected a sex-matched stranger from a common name list not corresponding to anyone they knew. In the self-matching task, participants were instructed to associate three shapes with the self, their best friend, and the stranger, and three shapes with happy, neutral, or sad line drawings in the emotion task respectively. There were no images of stimuli displayed during the instruction stage. After the instruction, participants were asked to judge whether a simultaneously presented shape and label pair matched. Each trial started with a central fixation cross for 500 ms, followed by a

shape-label pair at the centre of the screen for 100 ms in the self task and for 150 ms in the emotion task. Half the pairings of the shape and label/line drawing conformed to the instruction and were responded to as match trials; on the remaining trials the shapes and labels/line drawings were re-paired to form mismatch trials. For mismatch trials, a shape was paired with the other labels (e.g., self shape with friend and stranger labels in the self task, or happy associated shape with neutral and sad emotions in the emotion task.). The next frame was a 1000 ms blank field. Participants were encouraged to make a response as quickly and accurately as possible within this 1000 ms interval. A feedback message (correct, incorrect, or too slow) was then given in the centre of the screen for 500 ms (see Figure S3). Participants were also informed of their overall accuracy at the end of each block. There were three blocks of 60 trials following 12 practice trials. Thus, there were 30 trials for each match and mismatch condition in each task. Data on reaction time and accuracy in each condition were analysed. For both sessions, all reaction time and accuracy related variables in depression and control groups are presented in supplementary Table S1.

Data processing and ML model development

Data Preparation and Pre-processing

To build machine-learning compatible features, response time and accuracy were aggregated based on the trial types, including all combinations of [response time, accuracy] x [self, friend, stranger, happy, neutral, sad] x [match, mismatch], and we also added the overall ACC of the self-task and emotion task to control for baseline performance, resulting in a total of 26 predicting behavioural features. Then, these 26 behavioural features coupled with sex and age features were used to classify whether the subject was in the depression group or in the control group. Subjects with over 50% missing values of the behavioural features or without

measurements on BDI were excluded from the analysis (one in session 1, and six in session 2). The median imputation method was applied to impute missing values. Also, the sex feature was coded with male as numeric 0 and female as numeric 1. Following the data pre-processing step, a total of 83 subjects (58 in the depression group and 25 in the control) were included in the session 1 data and 78 subjects (49 in the depression group and 29 in the control) were included in session 2 data. Three of the 25 control subjects in session 1 got BDI >13 in session 2, while seven of the 58 depression subjects in session 1 got BDI<=13 in session 2.

ML model development

We applied the nonlinear Gradient Boosting (GB) model (Friedman, 2001) for classifying individuals in the depression group from those in the control group. In the GB model, multiple shallow decision trees, recursively partitioning feature space into a small number of regions with the purpose to assign subjects in different classes to different regions, are combined sequentially to achieve optimal classification performance (Friedman, 2001). We selected the GB model as it is capable to model complex nonlinear and interaction effects between features and outcome, and has shown promising performance in previous research (Bracher-Smith et al., 2020; Parikh et al., 2019; Walss-Bass et al., 2018). See (Friedman, 2001) for detailed explanation about the GB model. Because tree-based models are insensitive to the scales of the features, no further transformation was applied. The standard logistic linear regression model was used as the baseline model for model comparison. Furthermore, to demonstrate the benefit of nonlinear GB model compared to linear ML models, we also compared one of the most commonly used linear ML models, logistic linear regression model with lasso type regularization, in which the regularization was added on the coefficients to mitigate the impact of multi-collinearity. The

details about these models and the associated model selection procedures are presented in the supplementary material.

The hyperparameters of the GB model, the number of trees and the max depth of the base decision tree, were selected using the Bayesian Optimization algorithm (Snoek et al., 2012). Because the performance measured on the subjects used to fit the model tends to overestimate the model's performance on the unseen subjects, we used 10-Fold cross-validation (CV) (Wong, 2015) coupled with the metric, logistic loss, to measure the model's generalizability to unseen subjects during model selection. In a 10-Fold CV, the subjects are split into 10 folds and each fold will be left out as test set for model evaluation while other folders are training set for fitting model. Logistic loss was computed based on the model's predictions on the left-out test sets during the 10-Fold CV process. After model selection, 10-Fold CV AUC, which is invariant to the imbalanced labels of the outcome and represents the probability that the model will rank the subjects in depression group higher than the controls, was used as the main metric to evaluate the selected model. In addition, the selected model was further evaluated with respect to other commonly used metrics, i.e., sensitivity, specificity, balanced accuracy, positive and negative predictive values. Sensitivity refers to the probability that the subjects in the depressed group are correctly identified as depressed subjects; specificity is the probability that the control subjects are correctly identified as control; balanced accuracy is the average of sensitivity and specificity; positive predictive value is the probability that subjects predicted to have depression by the model actually have depression; negative predictive value is the probability that subjects predicted to not have depression by the model actually not have depression.

The variance of the evaluating metrics was mainly estimated through repeated 10-Fold CV, repeatedly applying 10-Fold CV on the selected models, or bootstrapping when applying

repeated 10-Fold CV is not feasible. See supplementary materials for more details on model evaluation methods.

Modelling Procedure

To evaluate model's consistency and generalizability, within session, between session classification of depression were explored. Models were built for the following three different classification problems, with-session classification using behaviour features at session 1, age and sex to predict depression status at session 1 (model 1), using behaviour features at session 2, age and sex to predict depression status at session 2 (model 2), and across-session classification using behaviour features at two sessions, age and sex to predict depression status at session 2 (model 3). We also evaluated model 1's between-session performance in predicting the prospective depression status at session 2. In addition, based on the same modeling procedure, we evaluated a model based on mean response times and mean accuracy only for a baseline prediction performance comparison (See Table S2 for response time and accuracy summary). Also, we further evaluated models focusing on the self task only, or emotion task only. The models were selected in the same way described above except only self task or emotion task related variables coupled with age and sex were used as the predictors.

Variable Importance

The contribution of the behavioural features in the selected GB model on predicting the depression status can be measured by variable importance (Friedman, 2001). We applied a resampling procedure (Efron, 1992) to approximate the variance of the estimated variable importance. Specifically, the GB model was selected and fitted on the resampled data (randomly sampling 90% subjects from the original data without replacement) for 500 times. Then, the

means of the variable importance measures derived from the selected models and their variance were used to evaluate the impact of the features on the classification of individuals in the depression and the control groups.

To further validate the identified features with high variable importance, we tested if the behavioural measures are significantly different in the control and in the depression group posthoc. Because many behavioural measures are not approximately normally distributed, we applied the nonparametric Kolmogorov–Smirnov test procedure (Berger and Zhou, 2014) to test if the distributions of a feature in control and depression groups are significantly different. False discovery rate (FDR), Benjamini-Hochberg Procedure (Benjamini and Yekutieli, 2001), was used to correct the multiple testing issue. The threshold of FDR was set to 0.1, which indicates at most 10% of the significant findings are false.

Results

Model performance

For all classification models, the nonlinear GB models consistently outperformed the linear counterparts (Figure S1). The GB algorithm achieved 10-Fold CV area under the receiver operating characteristic curve (AUC) of 0.800 ± 0.020 for model 1, 0.822 ± 0.024 for model 2, and 0.896 ± 0.018 for model 3 (Figure S2). Among the linear models, the regularized logistic linear model always outperformed the standard one (Figure S1). The regularized logistic linear model achieved AUC 0.774 ± 0.017 for model 1, 0.734 ± 0.027 for model 2 and 0.795 ± 0.021 for model 3 (Figure S1). In addition, the performance of the selected models improved from session 1 to session 2 and can be further improved when prior records from session 1 were used in conjunction with records from session 2 (Figure S2).

For a baseline performance comparison, GB model with only mean reaction time and mean accuracy achieved AUC 0.648 ± 0.022 in predicting depression status in session 1 and AUC 0.657 ± 0.035 in predicting depression status in session 2, performing worse than model 1 and model 2 ($P < 0.001$, derived from the t-test on repeated 10-Fold AUCs).

In addition, for the prediction of depression status in session 1, self task related features achieve AUC 0.746 ± 0.025 and emotion task related features achieve 0.690 ± 0.029, both perform significantly lower than the best-fitting model using features from both self and emotion tasks ($p < 0.001$). For the prediction of depression status in session 2, self task related features achieve AUC 0.779 ± 0.019 and emotion task related features achieve AUC 0.797 ± 0.025, both perform significantly lower than model 1 and model 2 ($p < 0.001$) using all features from both the self- and emotion tasks.

---Insert Table 1 here---

Furthermore, for prospective predictions, where model trained at session 1 is applied directly to explanatory variables and outcome label at session 2, model 1 performed adequately on prediction the depression status at session 2 (AUC 0.774 ± 0.047). The selected GB models were also evaluated with respect to other metrics: sensitivity, specificity, balanced accuracy, and positive and negative predictive values. The results are shown in Table 2.

---Insert Table 2 here---

Important predicting features

The primary goal of building ML models is to make accurate predictions. Interpreting the important variables contributing to machine-learning models is a challenging task, especially when behavioural features are highly correlated. Here we choose to focus on interpreting overlapping features from the models while briefly commenting on other nonoverlapped top features. However, this does not suggest other features not discussed are unimportant (Table 3). The top 10 important features in the selected GB models are shown in Figure 2 (See Table S1 in the supplementary materials for all feature's description). The four features, which appear consistently as the top 10 contributing features for both model 1 and model 2, are the response time of sad matching trials (rt sad), friend mismatch trials (rt mismatch friend), accuracy of self mismatch trials (acc mismatch self), and self matching trials (acc self). Those features are also among top features contributed to model 3, except for the accuracy of self-matching trials.

---Insert Figure 2 here---

Posthoc analysis of the direction of the effects are shown in Table 3. The direction of effects for shared features including response time of sad matching trials, accuracy of self-mismatch trials and accuracy of self matching trials are consistent between session 1 and session 2. Compared to the control group, the response time of sad matching trials are slower in the depression group, the accuracy of self mismatch trials is more accurate in the depression group, and the self matching trials are less accurate in the depression group in both sessions, even though this nominal difference failed to meet statistical significance in session 1. Other top contributing features identified in all 3 models generally follow the pattern that the depression group is nominally more accurate, but slower than controls in both sessions. The cross session

behavioural performance echoes a general learning or practice effect, where session 2 performances are generally enhanced. Accuracy of self-matching trials and response time for friend mismatching trials are the only two exceptions where posthoc comparisons between depression and control group are statistically significant but in the opposite direction. Note that both response time and accuracy features have a theoretical ceiling, thus differences between groups using those measures may be reduced when performance increases in session 2.

---Insert Table 3 here---

Discussion and conclusion

In this interdisciplinary pilot study that integrates both behavioural experiments and ML analysis, we successfully demonstrated that behavioural data from non-verbal self and emotional processing assessments can be applied to identify depression status for individual subjects. Using models trained on two separate sessions of data, we show the ML models can predict depression, with a cross-validated AUC of 0.800, and 0.822, for session 1 and 2, respectively, outperforming models based on simple response patterns and models based on the self- and emotion- assessing behavoural task alone. We also report the model performance can be further enhanced if behavioural data across session is provided to the model, with a cross-validated AUC of 0.896. The best performing model achieved a balanced accuracy of 0.808, with a positive predictive value of 0.846 and negative predictive value of 0.834 (Table 2). The positive predictive values of the GB models are consistently above 0.8 for the three models. We also reported an adequate prospective prediction performance using a model built with session 1 data to predict session 2 performance, achieving an AUC of 0.774 and a balanced accuracy of 0.646. Further, we

identified the core features as behavioural signatures that contribute to the model's prediction performance.

Comparing with verbal screening, non-verbal based depression screening is more inclusive, and is often inconspicuous, thus better suited for routine screening. We showed the self and emotional processing assessment, with a machine-learning based algorithm, can successfully capture behavioural signatures of depressed individuals in risk screening, and make reliable positive screening predictions regardless of prior exposure to the assessment tool. The prediction performance is further enhanced by capturing the differential change in behavioural pattern over a course of subsequent assessment. In both routine risk screening and clinical practices, multi-session assessment could be easily arranged to facilitate screening accuracy, e.g., subject complete the task at home while booking an appointment, and complete the same task at the time of appointment. With the ubiquity of access to smartphones, tablets and personal computers, the cost of behavioural assessment is relatively low compared to models based on specialty neuroimaging devices, and could be more patient accessible and less computational demanding comparing to models relying on video/audio hardware. This non-verbal screening approach could also be incorporated in the existing clinical depression screening pipeline and can provide additional sources of information to facilitate clinician's diagnostic decisions, as well as tracking for treatment efficacy and symptom remission.

Our current results support a promising approach to leverage a non-verbal self- and emotional- behavioural assessment to depression screening, showing adequate prospective prediction performance and good cross-validation performance when subjects are more familiar with the task and previous test session's information is made available, warranting larger scale future studies on this approach.

The machine-learning algorithm is adapted to optimize for classification performance, and is agnostic of theoretical constructs of the features, thus may only rely on a sufficient set of features. In this study, subjects were less familiar with the task in the first session, performing with a lower overall accuracy and slower response time in session 1 compared session 2. Due to the ceiling effects in behavioural measurements, the performance difference between the depression group and control may be reduced as overall performance increase, resembling performance by session 2-way interactions (Figure S3). Despite this limitation, we identified four important behavioural signatures that are consistent across session 1 and session 2, with significant differences between groups in post hoc analysis.

Earlier findings of behavioural and neuroimaging studies converged on supporting neuro-anatomical differences underpinning depressed and non depressed people, leading to key behavioural differences characterized as increased self-focus, body focus, decreased environmental focus, attributing negative emotion to oneself, and increased self-processing or rumination, for depressed patients (Northoff, 2014). In the current study the response time of the sad matching trials are consistently slower in the non-clinical depression group than controls, and this pattern is complemented by an increased accuracy for sad matching trials. This behavioural signature suggests depressed subjects are biased towards processing negative emotions, thus may have difficulty disengaging from negative stimuli (Northoff, 2014). The response time of mismatching trials for friends is inconsistent with the accuracy pattern for mismatching trials for friends in session 1; however, in session 2 the depression group has a faster response time and more accurate responses for those trials that may reflect practice effect over time. This differential behavioural pattern would lead to reduced performance for model 1 on prospective prediction of depression in session 2, but may contribute to model 3's superior performance.

The accuracy of self-mismatching trials is higher in the depressed group than control, complemented by a slower response time, where the accuracy for self-matching trials is showing the opposite pattern. This line of results suggest depressed subjects can negate the mismatching trials more easily compared with controls in general, and there's a heightened focus on the self for depressed subjects, which may lead to other construct (e.g., a shape) more difficult to associate with the self (Northoff, 2014; Northoff and Hayes, 2011; Sheline et al., 2009).

Future research needs to target test-retest consistency, as the current study suggests differential speed accuracy trade offs of the depressed and control subjects. This trade off may have been caused by differential task learning rates of the depressed subjects and partially explained by a ceiling effect in behavioural measurement data. Also, the current study was conducted on a sample of university students without a comprehensive clinical screening including checking recent medication history, future studies need to test on a more diverse sample and screen for other sources influencing depression symptoms, to further establish generalizability. Another area of focus is to upscale the sample size to further validate the current findings. Further, depression status was determined by BDI-II alone, where the cut-off score based classification of depression may cause label changes if the subject's score fluctuates around the cut-off point. Although BDI-II is effective in identifying clinical depression (von Glischinski et al., 2019), additional outcome measurements such as clinical diagnoses and other depression screening tools can be implemented in addition to BDI-II scale, to facilitate the accuracy of depression identification. Moreover, the current experimental procedure of the behavioural assessments should be viewed as a proof of concept instead of a final product. From the patient's perspective the behavioural assessment could be more cumbersome compared to other non-verbal screening methods, but this may no longer be the case with further evidence-

based methodology refinements. For instance, self task related features alone achieved reasonable depression screening performance (AUC = 0.746), suggesting significant reduction in screening time by cutting down emotion assessment can be achieved at a minor cost of accuracy. Future research should focus on refining the procedure to provide a more streamlined patient experience without compromising screening efficacy.

This study provides direct support for the utility of the validated brain-behaviour self- and emotional matching tasks on classifying depression, demonstrating a viable and cost-effective solution to provide routine depression screening, with a potential wide range of applications in clinical settings.

Ethics approval and consent to participate

All participants signed informed consent to participate. The study is approved by University of Bath's ethics committee.

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests

Funding

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program (BC), Alberta Innovates (BC), Mental Health Foundation (BC), MITACS Accelerate program (YL. BC), Simon & Martina Sochatsky Fund for Mental Health (BC), the Alberta Synergies in Alzheimer's and Related Disorders (SynAD) program (YL, BC) and University of Alberta Hospital Foundation (BC). The funding sources had no impact on the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Authors' contributions

YL: Conceptualization, Methodology, Formal Analysis, Writing- Original draft preparation. YS: Conceptualization, Methodology, Data Curation, Software, Formal Analysis, Writing- Original draft preparation. NL: Investigation, Writing – Review & Editing. DB: Writing – Review & Editing. KB: Investigation, Methodology, Writing – Review & Editing AG: Supervision, Writing – Review & Editing. BC: Conceptualization, Methodology, Supervision,

Resources, Writing – Review & Editing. JS: Conceptualization, Methodology, Supervision, Resources, Writing – Review & Editing.

Figures

Figure 1.

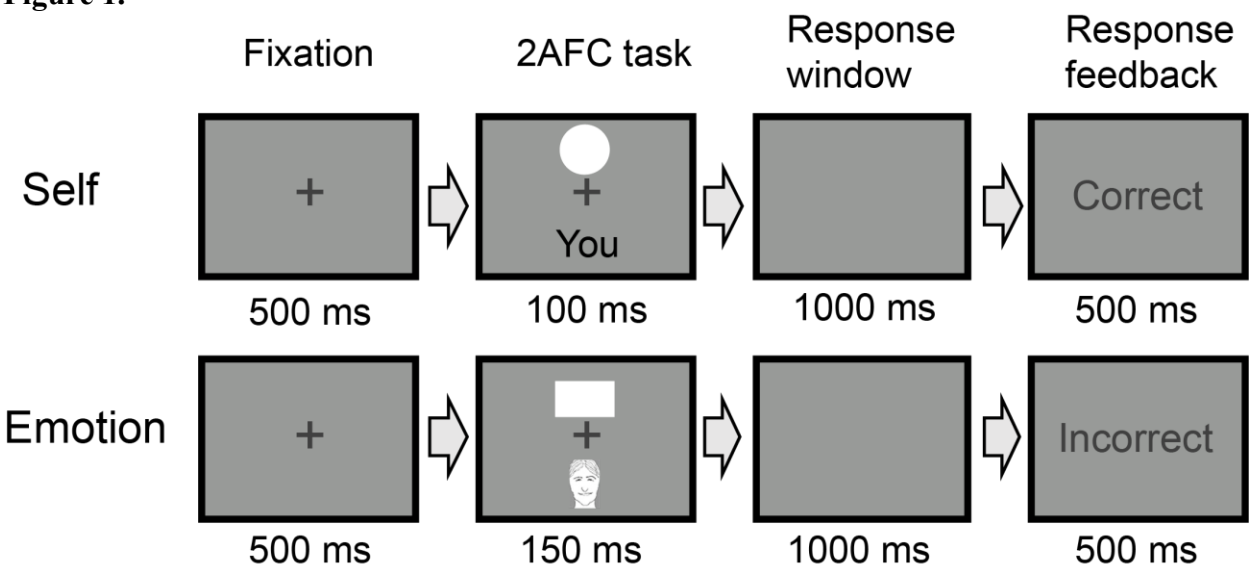
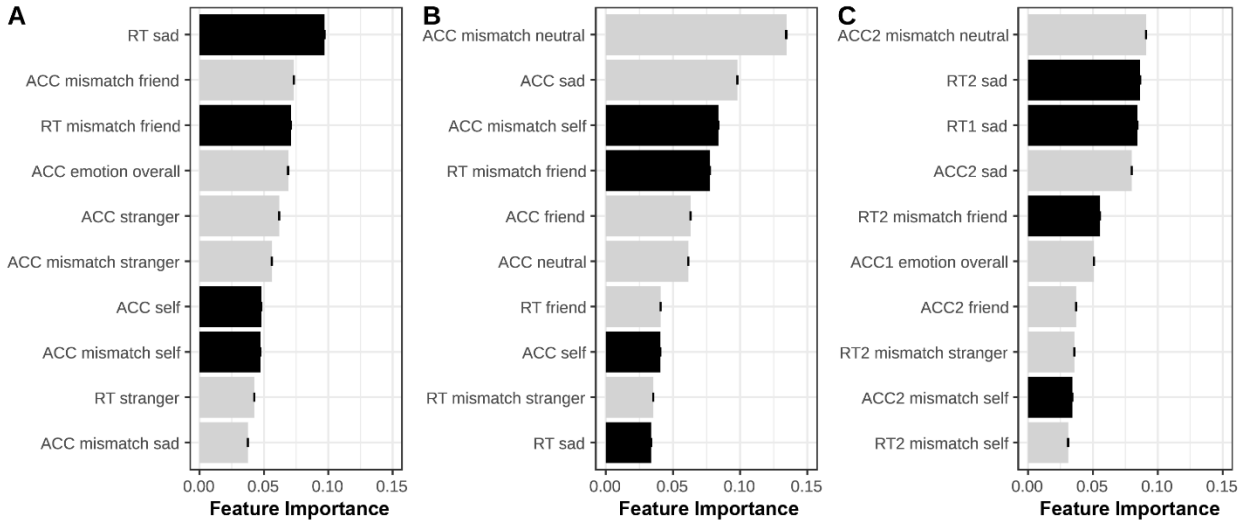


Figure Captions

Examples of The Stimuli and Trial Procedure in The Matching Task.

Note. For the 2-Alternative Forced Choice (2AFC) task, possible trial types are Match and Mismatch. For the response feedback, possible feedback are Correct, Incorrect and Too slow.

Figure 2



The Variable Importance of The Top Features in The Three Selected GB Models

Note. A: the variable importance of the top 10 features derived from the GB models in model 1; B: the variable importance of the top 10 features derived from the GB models in model 2; C: the variable importance of the top 10 features derived from the GB model in model 3. Bars filled with black color indicates the corresponding variables are selected as top features in both task 1 and task 2.

Tables

Table 1

Demographic Data of Participants

Session	Variable	Depression group (BDI >13)	Control group (BDI ≤13)
Time 1	Age, mean (range, SD)	20.60 (18-32, 2.25)	20.60 (19-22, 0.87)
	BDI-II, male/total, n	20/59	9/25
	Mild (14-19)	7/15	
	Moderate (20-28)	11/25	
	Severe (29-63)	2/19	
	BDI-II, mean (SD)	26.10 (9.76)	5.04 (3.42)
	Mild (14-19)	17.30 (2.06)	
	Moderate (20-28)	22.40 (2.52)	
	severe (29-63)	37.80 (8.26)	
Time 2	Age, mean (range, SD)	20.70 (18-32, 2.26)	20.50 (18-23, 1.12)
	BDI-II, male/total, n	18/54	11/29
	Mild (14-19)	8/12	
	Moderate (20-28)	9/31	
	Severe (29-63)	1/11	
	BDI-II, mean (SD)	24.90 (10.40)	5.38 (3.57)
	Mild (14-19)	16.70 (2.06)	
	Moderate (20-28)	21.80 (1.78)	

severe (29-63)	42.60 (9.97)

Table 2

The Core Metrics for The Selected GB Models

Metric	Model 1		Model 2		Model 3		Model 1 to 2	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sensitivity	0.821	0.033	0.890	0.027	0.925	0.019	0.899	0.070
Specificity	0.544	0.052	0.622	0.057	0.692	0.061	0.393	0.153
Balanced Accuracy	0.682	0.028	0.756	0.030	0.808	0.035	0.646	0.053
PPV	0.807	0.017	0.800	0.024	0.846	0.027	0.738	0.038
NPV	0.569	0.049	0.772	0.044	0.834	0.041	0.695	0.100

Note. Model 1 to 2: Model 1 on session 2 data. The mean and standard deviation are calculated based on 30 times repeated 10-Fold CV for Model 1, Model 2, Model 3, and based on 30 times bootstrap for applying Model 1 on session 2 data. PPV: Positive predictive value; NPV: Negative predictive value.

Table 3
Summary of Top Ranked Features

Top 10 Features	Session 1			Session 2		
	$M_{Depresion}$	$M_{Control}$	p	$M_{Depresion}$	$M_{Control}$	p
a) Task 1 and 2 common						
rt sad	833	679	<0.000***	707	633	0.028*
rt mismatch friend	798	703	0.002**	734	770	0.007**
acc mismatch self	0.812	0.736	0.027*	0.861	0.784	0.028*
acc self	0.797	0.837	0.282	0.783	0.926	0.050*
b) Task 1						
acc mismatch friend	0.826	0.709	<0.000***	0.841	0.777	0.014*
acc emotion overall	0.775	0.672	0.013*	0.843	0.787	0.203
acc stranger	0.769	0.623	0.009**	0.896	0.862	0.225
acc mismatch stranger	0.848	0.763	0.012*	0.872	0.821	0.010**
rt stranger	772	696	0.158	736	693	0.174
acc mismatch sad	0.802	0.667	0.003**	0.889	0.793	0.233
c) Task 2						
acc mismatch neutral	0.799	0.673	0.008**	0.875	0.764	0.034*
acc sad	0.773	0.645	0.046*	0.870	0.750	0.009**
acc friend	0.708	0.713	0.986	0.794	0.855	0.077
acc neutral	0.696	0.630	0.065.	0.784	0.763	0.342
rt friend	744	678	0.333	746	676	0.280
rt mismatch stranger	798	717	0.069.	775	739	0.010**

d) Task 3 only						
rt mismatch self	785	745	0.244·	776	758	0.021*

*Note. p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. P-values computed from Kolmogorov–Smirnov test. M denotes mean. The units for response time and accuracy are in milliseconds and proportion correct, respectively. A list of top ranked features a) shared by task 1 and task 2. b) appeared only in task 1, c) appeared only in task 2, and d) appeared only in task 3.*

References

Alghowinem, S., Goecke, R., Cohn, J.F., Wagner, M., Parker, G., Breakspear, M., 2015. Cross-cultural detection of depression from nonverbal behaviour. 2015 11th IEEE Int. Conf. Work. Autom. Face Gesture Recognition, FG 2015. https://doi.org/10.1109/FG.2015.7163113

Beck, A.T., Steer, R.A., Ball, R., Ranieri, W.F., 1996. Comparison of Beck depression inventories -IA and -II in psychiatric outpatients. J. Pers. Assess. 67, 588–597. https://doi.org/10.1207/s15327752jpa6703_13

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29, 1165–1188. https://doi.org/10.1214/aos/1013699998

Berger, V.W., Zhou, Y., 2014. Kolmogorov–Smirnov Test: Overview, in: Wiley StatsRef: Statistics Reference Online. Wiley. https://doi.org/10.1002/9781118445112.stat06558

Bracher-Smith, M., Crawford, K., Escott-Price, V., 2020. Machine learning for genetic prediction of psychiatric disorders: a systematic review. Mol. Psychiatry 2020 261 26, 70–79. https://doi.org/10.1038/s41380-020-0825-2

Bzdok, D., Altman, N., Krzywinski, M., 2018. Statistics versus machine learning. Nat. Methods 15, 233–234. https://doi.org/10.1038/nmeth.4642

CostHelper, 2021. Cost of an EEG - 2021 Healthcare Costs [WWW Document]. unpublished. URL https://health.costhelper.com/eeg.html (accessed 11.23.21).

Delaveau, P., Jabourian, M., Lemogne, C., Allaïli, N., Choucha, W., Girault, N., Lehericy, S., Laredo, J., Fossati, P., 2015. Antidepressant short-term and long-term brain effects during self-referential processing in major depression. Elsevier 247, 17–24. https://doi.org/10.1016/j.pscychresns.2015.11.007ï

Delaveau, P., Jabourian, M., Lemogne, C., Guionnet, S., Bergouignan, L., Fossati, P., 2011. Brain effects of antidepressants in major depression: A meta-analysis of emotional processing studies. J. Affect. Disord. 130, 66–74. https://doi.org/10.1016/J.JAD.2010.09.032

Efron, B., 1992. Bootstrap Methods: Another Look at the Jackknife. https://doi.org/10.1007/978-1-4612-4380-9_41

Enneking, V., Leehr, E.J., Dannlowski, U., Redlich, R., 2020. Brain structural effects of treatments for depression and biomarkers of response: a systematic review of neuroimaging studies. Psychol. Med. 50, 187–209. https://doi.org/10.1017/S0033291719003660

Feldborg, M., Lee, N.A., Hung, K., Peng, K., Sui, J., 2021. Perceiving the Self and Emotions with an Anxious Mind: Evidence from an Implicit Perceptual Task. Int. J. Environ. Res. Public Heal. 2021, Vol. 18, Page 12096 18, 12096. https://doi.org/10.3390/IJERPH182212096

Freton, M., Lemogne, C., Delaveau, P., Guionnet, S., Wright, E., Wiernik, E., Bertasi, E., Fossati, P., 2014. The dark side of self-focus: Brain activity during self-focus in low and high brooders. Soc. Cogn. Affect. Neurosci. 9, 1808–1813. https://doi.org/10.1093/scan/nst178

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 29, 1189–1232. https://doi.org/10.1214/aos/1013203451

Guekht, A., Polunina, A., Sidnyaeva, E., Gudkova, A., Golukhova, E., 2013. Non-Verbal Approach to Screening Diagnostics of Depression and Anxiety. Res. Neurol. An Int. J. 2013, 1–16. https://doi.org/10.5171/2013.791259

Hall, C.A., Reynolds, C.F., 2014. Late-life depression in the primary care setting: Challenges,

collaborative care, and prevention. Maturitas 79, 147–152.

Hobbs, C., Sui, J., Kessler, D., Munafò, M.R., Button, K.S., 2021. Self-processing in relation to emotion and reward processing in depression. Psychol. Med. https://doi.org/10.1017/S0033291721003597

Humphreys, G.W., Sui, J., 2016. Attentional control and the self: The Self-Attention Network (SAN). Cogn. Neurosci. https://doi.org/10.1080/17588928.2015.1044427

James, S.L., Abate, D., Abate, K.H., Abay, S.M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R.S., Abebe, Z., Abera, S.F., Abil, O.Z., Abraha, H.N., Abu-Raddad, L.J., Abu-Rmeileh, N.M.E., Accrombessi, M.M.K., Acharya, D., Acharya, P., Ackerman, I.N., Adamu, A.A., Adebayo, O.M., Adekanmbi, V., Adetokunboh, O.O., Adib, M.G., Adsuar, J.C., Afanvi, K.A., Afarideh, M., Afshin, A., Agarwal, G., Agesa, K.M., Aggarwal, R., Aghayan, S.A., Agrawal, S., Ahmadi, A., Ahmadi, M., Ahmadieh, H., Ahmed, M.B., Aichour, A.N., Aichour, I., Aichour, M.T.E., Akinyemiju, T., Akseer, N., Al-Aly, Z., Al-Eyadhy, A., Al-Mekhlafi, H.M., Al-Raddadi, R.M., Alahdab, F., Alam, K., Alam, T., Alashi, A., Alavian, S.M., Alene, K.A., Alijanzadeh, M., Alizadeh-Navaei, R., Aljunid, S.M., Alkerwi, A., Alla, F., Allebeck, P., Alouani, M.M.L., Altirkawi, K., Alvis-Guzman, N., Amare, A.T., Aminde, L.N., Ammar, W., Amoako, Y.A., Anber, N.H., Andrei, C.L., Androudi, S., Animut, M.D., Anjomshoa, M., Ansha, M.G., Antonio, C.A.T., Anwari, P., Arabloo, J., Arauz, A., Aremu, O., Ariani, F., Armoon, B., Ärnlöv, J., Arora, A., Artaman, A., Aryal, K.K., Asayesh, H., Asghar, R.J., Ataro, Z., Atre, S.R., Ausloos, M., Avila-Burgos, L., Avokpaho, E.F.G.A., Awasthi, A., Ayala Quintanilla, B.P., Ayer, R., Azzopardi, P.S., Babazadeh, A., Badali, H., Badawi, A., Bali, A.G., Ballesteros, K.E., Ballew, S.H., Banach, M., Banoub, J.A.M., Banstola, A.,

Barac, A., Barboza, M.A., Barker-Collo, S.L., Bärnighausen, T.W., Barrero, L.H., Baune, B.T., Bazargan-Hejazi, S., Bedi, N., Beghi, E., Behzadifar, Masoud, Behzadifar, Meysam, Béjot, Y., Belachew, A.B., Belay, Y.A., Bell, M.L., Bello, A.K., Bensenor, I.M., Bernabe, E., Bernstein, R.S., Beuran, M., Beyranvand, T., Bhala, N., Bhattarai, S., Bhaumik, S., Bhutta, Z.A., Biadgo, B., Bijani, A., Bikbov, B., Bilano, V., Bililign, N., Bin Sayeed, M.S., Bisanzio, D., Blacker, B.F., Blyth, F.M., Bou-Orm, I.R., Boufous, S., Bourne, R., Brady, O.J., Brainin, M., Brant, L.C., Brazinova, A., Breitborde, N.J.K., Brenner, H., Briant, P.S., Briggs, A.M., Briko, A.N., Britton, G., Brugha, T., Buchbinder, R., Busse, R., Butt, Z.A., Cahuana-Hurtado, L., Cano, J., Cárdenas, R., Carrero, J.J., Carter, A., Carvalho, F., Castañeda-Orjuela, C.A., Castillo Rivas, J., Castro, F., Catalá-López, F., Cercy, K.M., Cerin, E., Chaiah, Y., Chang, A.R., Chang, H.Y., Chang, J.C., Charlson, F.J., Chattopadhyay, A., Chattu, V.K., Chaturvedi, P., Chiang, P.P.C., Chin, K.L., Chitheer, A., Choi, J.Y.J., Chowdhury, R., Christensen, H., Christopher, D.J., Cicuttini, F.M., Ciobanu, L.G., Cirillo, M., Claro, R.M., Collado-Mateo, D., Cooper, C., Coresh, J., Cortesi, P.A., Cortinovis, M., Costa, M., Cousin, E., Criqui, M.H., Cromwell, E.A., Cross, M., Crump, J.A., Dadi, A.F., Dandona, L., Dandona, R., Dargan, P.I., Daryani, A., Das Gupta, Rajat, Das Neves, J., Dasa, T.T., Davey, G., Davis, A.C., Davitoiu, D.V., De Courten, B., De La Hoz, F.P., De Leo, D., De Neve, J.W., Degefa, M.G., Degenhardt, L., Deiparine, S., Dellavalle, R.P., Demoz, G.T., Deribe, K., Dervenis, N., Des Jarlais, D.C., Dessie, G.A., Dey, S., Dharmaratne, S.D., Dinberu, M.T., Dirac, M.A., Djalalinia, S., Doan, L., Dokova, K., Doku, D.T., Dorsey, E.R., Doyle, K.E., Driscoll, T.R., Dubey, M., Dubljanin, E., Duken, E.E., Duncan, B.B., Duraes, A.R., Ebrahimi, H., Ebrahimpour, S., Echko, M.M., Edvardsson, D., Effiong, A., Ehrlich, J.R., El Bcheraoui, C., El Sayed Zaki, M., El-Khatib,

Z., Elkout, H., Elyazar, I.R.F., Enayati, A., Endries, A.Y., Er, B., Erskine, H.E., Eshrati, B., Eskandarieh, S., Esteghamati, A., Esteghamati, S., Fakhim, H., Fallah Omrani, V., Faramarzi, M., Fareed, M., Farhadi, F., Farid, T.A., Farinha, C.S.E., Farioli, A., Faro, A., Farvid, M.S., Farzadfar, F., Feigin, V.L., Fentahun, N., Fereshtehnejad, S.M., Fernandes, E., Fernandes, J.C., Ferrari, A.J., Feyissa, G.T., Filip, I., Fischer, F., Fitzmaurice, C., Foigt, N.A., Foreman, K.J., Fox, J., Frank, T.D., Fukumoto, T., Fullman, N., Fürst, T., Furtado, J.M., Futran, N.D., Gall, S., Ganji, M., Gankpe, F.G., Garcia-Basteiro, A.L., Gardner, W.M., Gebre, A.K., Gebremedhin, A.T., Gebremichael, T.G., Gelano, T.F., Geleijnse, J.M., Genova-Maleras, R., Geramo, Y.C.D., Gething, P.W., Gezae, K.E., Ghadiri, K., Ghasemi Falavarjani, K., Ghasemi-Kasman, M., Ghimire, M., Ghosh, R., Ghoshal, A.G., Giampaoli, S., Gill, P.S., Gill, T.K., Ginawi, I.A., Giussani, G., Gnedovskaya, E. V., Goldberg, E.M., Goli, S., Gómez-Dantés, H., Gona, P.N., Gopalani, S.V., Gorman, T.M., Goulart, A.C., Goulart, B.N.G., Grada, A., Grams, M.E., Grosso, G., Gugnani, H.C., Guo, Y., Gupta, P.C., Gupta, Rahul, Gupta, Rajeev, Gupta, T., Gyawali, B., Haagsma, J.A., Hachinski, V., Hafezi-Nejad, N., Haghparast Bidgoli, H., Hagos, T.B., Hailu, G.B., Haj-Mirzaian, Arvin, Haj-Mirzaian, Arya, Hamadeh, R.R., Hamidi, S., Handal, A.J., Hankey, G.J., Hao, Y., Harb, H.L., Harikrishnan, S., Haro, J.M., Hasan, M., Hassankhani, H., Hassen, H.Y., Havmoeller, R., Hawley, C.N., Hay, R.J., Hay, S.I., Hedayatizadeh-Omran, A., Heibati, B., Hendrie, D., Henok, A., Herteliu, C., Heydarpour, S., Hibstu, D.T., Hoang, H.T., Hoek, H.W., Hoffman, H.J., Hole, M.K., Homaie Rad, E., Hoogar, P., Hosgood, H.D., Hosseini, S.M., Hosseinzadeh, M., Hostiuc, M., Hostiuc, S., Hotez, P.J., Hoy, D.G., Hsairi, M., Htet, A.S., Hu, G., Huang, J.J., Huynh, C.K., Iburg, K.M., Ikeda, C.T., Ileanu, B., Ilesanmi, O.S., Iqbal, U., Irvani, S.S.N., Irvine, C.M.S., Mohammed, Sheikh, Islam, S., Islami, F., Jacobsen, K.H.,

Jahangiry, L., Jahanmehr, N., Jain, S.K., Jakovljevic, M., Javanbakht, M., Jayatilleke, A.U., Jeemon, P., Jha, R.P., Jha, V., Ji, J.S., Johnson, C.O., Jonas, J.B., Jozwiak, J.J., Jungari, S.B., Jürisson, M., Kabir, Z., Kadel, R., Kahsay, A., Kalani, R., Kanchan, T., Karami, M., Karami Matin, B., Karch, A., Karema, C., Karimi, N., Karimi, S.M., Kasaeian, A., Kassa, D.H., Kassa, G.M., Kassa, T.D., Kassebaum, N.J., Katikireddi, S.V., Kawakami, N., Kazemi Karyani, A., Keighobadi, M.M., Keiyoro, P.N., Kemmer, L., Kemp, G.R., Kengne, A.P., Keren, A., Khader, Y.S., Khafaei, B., Khafaie, M.A., Khajavi, A., Khalil, I.A., Khan, E.A., Khan, M.S., Khan, M.A., Khang, Y.H., Khazaei, M., Khoja, A.T., Khosravi, A., Khosravi, M.H., Kiadaliri, A.A., Kiirithio, D.N., Kim, C. Il, Kim, D., Kim, P., Kim, Y.E., Kim, Y.J., Kimokoti, R.W., Kinfu, Y., Kisa, A., Kissimova-Skarbek, K., Kivimäki, M., Knudsen, A.K.S., Kocarnik, J.M., Kochhar, S., Kokubo, Y., Kolola, T., Kopec, J.A., Kosen, S., Kotsakis, G.A., Koul, P.A., Koyanagi, A., Kravchenko, M.A., Krishan, K., Krohn, K.J., Kuate Defo, B., Kucuk Bicer, B., Kumar, G.A., Kumar, M., Kyu, H.H., Lad, D.P., Lad, S.D., Lafranconi, A., Lalloo, R., Lallukka, T., Lami, F.H., Lansingh, V.C., Latifi, A., Lau, K.M.M., Lazarus, J. V., Leasher, J.L., Ledesma, J.R., Lee, P.H., Leigh, J., Leung, J., Levi, M., Lewycka, S., Li, S., Li, Y., Liao, Y., Liben, M.L., Lim, L.L., Lim, S.S., Liu, S., Lodha, R., Looker, K.J., Lopez, A.D., Lorkowski, S., Lotufo, P.A., Low, N., Lozano, R., Lucas, T.C.D., Lucchesi, L.R., Lunevicius, R., Lyons, R.A., Ma, S., Macarayan, E.R.K., Mackay, M.T., Madotto, F., Magdy Abd El Razek, H., Magdy Abd El Razek, M., Maghavani, D.P., Mahotra, N.B., Mai, H.T., Majdan, M., Majdzadeh, R., Majeed, A., Malekzadeh, R., Malta, D.C., Mamun, A.A., Manda, A.L., Manguerra, H., Manhertz, T., Mansournia, M.A., Mantovani, L.G., Mapoma, C.C., Maravilla, J.C., Marcenes, W., Marks, A., Martins-Melo, F.R., Martopullo, I., März, W., Marzan, M.B., Mashamba-Thompson, T.P., Massenburg,

B.B., Mathur, M.R., Matsushita, K., Maulik, P.K., Mazidi, M., McAlinden, C., McGrath, J.J., McKee, M., Mehndiratta, M.M., Mehrotra, R., Mehta, K.M., Mehta, V., Mejia-Rodriguez, F., Mekonen, T., Melese, A., Melku, M., Meltzer, M., Memiah, P.T.N., Memish, Z.A., Mendoza, W., Mengistu, D.T., Mengistu, G., Mensah, G.A., Mereta, S.T., Meretoja, A., Meretoja, T.J., Mestrovic, T., Mezerji, N.M.G., Miazgowski, B., Miazgowski, T., Millear, A.I., Miller, T.R., Miltz, B., Mini, G.K., Mirarefin, M., Mirrakhimov, E.M., Misganaw, A.T., Mitchell, P.B., Mitiku, H., Moazen, B., Mohajer, B., Mohammad, K.A., Mohammadifard, N., Mohammadnia-Afrouzi, M., Mohammed, M.A., Mohammed, Shafiu, Mohebi, F., Moitra, M., Mokdad, A.H., Molokhia, M., Monasta, L., Moodley, Y., Moosazadeh, M., Moradi, G., Moradi-Lakeh, M., Moradinazar, M., Moraga, P., Morawska, L., Moreno Velásquez, I., Morgado-Da-Costa, J., Morrison, S.D., Moschos, M.M., Mousavi, S.M., Mruts, K.B., Muche, A.A., Muchie, K.F., Mueller, U.O., Muhammed, O.S., Mukhopadhyay, S., Muller, K., Mumford, J.E., Murhekar, M., Musa, J., Musa, K.I., Mustafa, G., Nabhan, A.F., Nagata, C., Naghavi, M., Naheed, A., Nahvijou, A., Naik, G., Naik, N., Najafi, F., Naldi, L., Nam, H.S., Nangia, V., Nansseu, J.R., Nascimento, B.R., Natarajan, G., Neamati, N., Negoi, I., Negoi, R.I., Neupane, S., Newton, C.R.J., Ngunjiri, J.W., Nguyen, A.Q., Nguyen, Ha Thu, Nguyen, H.L.T., Nguyen, Huong Thanh, Nguyen, L.H., Nguyen, M., Nguyen, N.B., Nguyen, S.H., Nichols, E., Ningrum, D.N.A., Nixon, M.R., Nolutshungu, N., Nomura, S., Norheim, O.F., Noroozi, M., Norrving, B., Noubiap, J.J., Nouri, H.R., Nourollahpour Shiadeh, M., Nowroozi, M.R., Nsoesie, E.O., Nyasulu, P.S., Odell, C.M., Ofori-Asenso, R., Ogbo, F.A., Oh, I.H., Oladimeji, O., Olagunju, A.T., Olagunju, T.O., Olivares, P.R., Olsen, H.E., Olusanya, B.O., Ong, K.L., Ong, S.K., Oren, E., Ortiz, A., Ota, E., Otstavnov, S.S., øverland, S., Owolabi, M.O., Mahesh, P.A., Pacella,

R., Pakpour, A.H., Pana, A., Panda-Jonas, S., Parisi, A., Park, E.K., Parry, C.D.H., Patel, S., Pati, S., Patil, S.T., Patle, A., Patton, G.C., Paturi, V.R., Paulson, K.R., Pearce, N., Pereira, D.M., Perico, N., Pesudovs, K., Pham, H.Q., Phillips, M.R., Pigott, D.M., Pillay, J.D., Piradov, M.A., Pirsaheb, M., Pishgar, F., Plana-Ripoll, O., Plass, D., Polinder, S., Popova, S., Postma, M.J., Pourshams, A., Poustchi, H., Prabhakaran, D., Prakash, S., Prakash, V., Purcell, C.A., Purwar, M.B., Qorbani, M., Quistberg, D.A., Radfar, A., Rafay, A., Rafiei, A., Rahim, F., Rahimi, K., Rahimi-Movaghar, A., Rahimi-Movaghar, V., Rahman, M., Ur Rahman, M.H., Rahman, M.A., Rahman, S.U., Rai, R.K., Rajati, F., Ram, U., Ranjan, P., Ranta, A., Rao, P.C., Rawaf, D.L., Rawaf, S., Reddy, K.S., Reiner, R.C., Reinig, N., Reitsma, M.B., Remuzzi, G., Renzaho, A.M.N., Resnikoff, S., Rezaei, S., Rezai, M.S., Ribeiro, A.L.P., Robinson, S.R., Roever, L., Ronfani, L., Roshandel, G., Rostami, A., Roth, G.A., Roy, A., Rubagotti, E., Sachdev, P.S., Sadat, N., Saddik, B., Sadeghi, E., Saeedi Moghaddam, S., Safari, H., Safari, Y., Safari-Faramani, R., Safdarian, M., Safi, S., Safiri, S., Sagar, R., Sahebkar, A., Sahraian, M.A., Sajadi, H.S., Salam, N., Salama, J.S., Salamati, P., Saleem, K., Saleem, Z., Salimi, Y., Salomon, J.A., Salvi, S.S., Salz, I., Samy, A.M., Sanabria, J., Sang, Y., Santomauro, D.F., Santos, I.S., Santos, J.V., Santric Milicevic, M.M., Sao Jose, B.P., Sardana, M., Sarker, A.R., Sarrafzadegan, N., Sartorius, B., Sarvi, S., Sathian, B., Satpathy, M., Sawant, A.R., Sawhney, M., Saxena, S., Saylan, M., Schaeffner, E., Schmidt, M.I., Schneider, I.J.C., Schöttker, B., Schwebel, D.C., Schwendicke, F., Scott, J.G., Sekerija, M., Sepanlou, S.G., Serván-Mori, E., Seyedmousavi, S., Shabaninejad, H., Shafieesabet, A., Shahbazi, M., Shaheen, A.A., Shaikh, M.A., Shams-Beyranvand, M., Shamsi, M., Shamsizadeh, M., Sharafi, H., Sharafi, K., Sharif, M., Sharif-Alhoseini, M., Sharma, M., Sharma, R., She, J., Sheikh, A., Shi, P., Shibuya, K., Shigematsu, M., Shiri, R.,

Shirkoohi, R., Shishani, K., Shiue, I., Shokraneh, F., Shoman, H., Shrime, M.G., Si, S., Siabani, S., Siddiqi, T.J., Sigfusdottir, I.D., Sigurvinsdottir, R., Silva, J.P., Silveira, D.G.A., Singam, N.S.V., Singh, J.A., Singh, N.P., Singh, V., Sinha, D.N., Skiadaresi, E., Slepak, E.L.N., Sliwa, K., Smith, D.L., Smith, M., Soares Filho, A.M., Sobaih, B.H., Sobhani, S., Sobngwi, E., Soneji, S.S., Soofi, M., Soosaraei, M., Sorensen, R.J.D., Soriano, J.B., Soyiri, I.N., Sposato, L.A., Sreeramareddy, C.T., Srinivasan, V., Stanaway, J.D., Stein, D.J., Steiner, C., Steiner, T.J., Stokes, M.A., Stovner, L.J., Subart, M.L., Sudaryanto, A., Sufiyan, M.B., Sunguya, B.F., Sur, P.J., Sutradhar, I., Sykes, B.L., Sylte, D.O., Tabarés-Seisdedos, R., Tadakamadla, S.K., Tadesse, B.T., Tandon, N., Tassew, S.G., Tavakkoli, M., Taveira, N., Taylor, H.R., Tehrani-Banihashemi, A., Tekalign, T.G., Tekelemedhin, S.W., Tekle, M.G., Temesgen, H., Temsah, M.H., Temsah, O., Terkawi, A.S., Teweldemedhin, M., Thankappan, K.R., Thomas, N., Tilahun, B., To, Q.G., Tonelli, M., Topor-Madry, R., Topouzis, F., Torre, A.E., Tortajada-Girbés, M., Touvier, M., Tovani-Palone, M.R., Towbin, J.A., Tran, B.X., Tran, K.B., Troeger, C.E., Truelsen, T.C., Tsilimbaris, M.K., Tsoi, D., Tudor Car, L., Tuzcu, E.M., Ukwaja, K.N., Ullah, I., Undurraga, E.A., Unutzer, J., Updike, R.L., Usman, M.S., Uthman, O.A., Vaduganathan, M., Vaezi, A., Valdez, P.R., Varughese, S., Vasankari, T.J., Venketasubramanian, N., Villafaina, S., Violante, F.S., Vladimirov, S.K., Vlassov, V., Vollset, S.E., Vosoughi, K., Vujcic, I.S., Wagnew, F.S., Waheed, Y., Waller, S.G., Wang, Y., Wang, Y.P., Weiderpass, E., Weintraub, R.G., Weiss, D.J., Weldegebreal, F., Weldegwergs, K.G., Werdecker, A., West, T.E., Whiteford, H.A., Widecka, J., Wijeratne, T., Wilner, L.B., Wilson, S., Winkler, A.S., Wiyeh, A.B., Wiysonge, C.S., Wolfe, C.D.A., Woolf, A.D., Wu, S., Wu, Y.C., Wyper, G.M.A., Xavier, D., Xu, G., Yadgir, S., Yadollahpour, A., Yahyazadeh Jabbari, S.H., Yamada, T., Yan, L.L.,

Yano, Y., Yaseri, M., Yasin, Y.J., Yeshaneh, A., Yimer, E.M., Yip, P., Yisma, E., Yonemoto, N., Yoon, S.J., Yotebieng, M., Younis, M.Z., Yousefifard, M., Yu, C., Zadnik, V., Zaidi, Z., Zaman, S. Bin, Zamani, M., Zare, Z., Zeleke, A.J., Zenebe, Z.M., Zhang, K., Zhao, Z., Zhou, M., Zodpey, S., Zucker, I., Vos, T., Murray, C.J.L., 2018. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet 392, 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Janssen, R.J., Mourão-Miranda, J., Schnack, H.G., 2018. Making Individual Prognoses in Psychiatry Using Neuroimaging and Machine Learning. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 3, 798–808. https://doi.org/10.1016/J.BPSC.2018.04.004

Lemogne, C., Gorwood, P., Bergouignan, L., Pélissolo, A., Lehéricy, S., Fossati, P., 2011. Negative affectivity, self-referential processing and the cortical midline structures. Soc. Cogn. Affect. Neurosci. 6, 426–433. https://doi.org/10.1093/SCAN/NSQ049

Maire, H., Brochard, R., Zagar, D., 2020. A Developmental Study of the Self-Prioritization Effect in Children Between 6 and 10 Years of Age. Child Dev. 91, 694–704. https://doi.org/10.1111/cdev.13352

Moseley, R.L., Liu, C.H., Gregory, N.J., Smith, P., Baron-Cohen, S., Sui, J., 2021. Levels of Self-representation and Their Sociocognitive Correlates in Late-Diagnosed Autistic Adults. J. Autism Dev. Disord. 1–14. https://doi.org/10.1007/s10803-021-05251-x

Nejad, A.B., Rotgé, J.Y., Valabregue, R., Guérin-Langlois, C., Hoertel, N., Gorwood, P., Dubertret, C., Limosin, F., Fossati, P., Lemogne, C., 2019. Medial prefrontal disengagement during self-focus in formerly depressed patients prone to rumination. J. Affect. Disord. 247,

36–44. https://doi.org/10.1016/J.JAD.2019.01.004

Northoff, G., 2014. How is our self altered in psychiatric disorders? A neurophenomenal approach to psychopathological symptoms. Psychopathology 47, 365–376. https://doi.org/10.1159/000363351

Northoff, G., Hayes, D.J., 2011. Is our self nothing but reward? Biol. Psychiatry. https://doi.org/10.1016/j.biopsych.2010.12.014

Northoff, G., Schneider, F., Rotte, M., Matthiae, C., Tempelmann, C., Wiebking, C., Bermpohl, F., Heinzel, A., Danos, P., Heinze, H.J., Bogerts, B., Walter, M., Panksepp, J., 2009. Differential parametric modulation of self-relatedness and emotions in different brain regions. Hum. Brain Mapp. 30, 369–382. https://doi.org/10.1002/HBM.20510

Pampouchidou, A., Simos, P.G., Marias, K., Meriaudeau, F., Yang, F., Pediaditis, M., Tsiknakis, M., 2019. Automatic Assessment of Depression Based on Visual Cues: A Systematic Review. IEEE Trans. Affect. Comput. 10, 445–470. https://doi.org/10.1109/TAFFC.2017.2724035

Parikh, R.B., Manz, C., Chivers, C., Regli, S.H., Braun, J., Draugelis, M.E., Schuchter, L.M., Shulman, L.N., Navathe, A.S., Patel, M.S., O'Connor, N.R., 2019. Machine Learning Approaches to Predict 6-Month Mortality Among Patients With Cancer. JAMA Netw. Open 2, e1915997–e1915997. https://doi.org/10.1001/JAMANETWORKOPEN.2019.15997

Sahu, A., Vikas, H., Sharma, N., 2020. Life cycle costing of MRI machine at a tertiary care teaching hospital. Indian J. Radiol. Imaging 30, 190–194. https://doi.org/10.4103/ijri.IJRI_54_19

Schnack, H., 2019. Bias, noise, and interpretability in machine learning: From measurements to features, in: Machine Learning: Methods and Applications to Brain Disorders. Academic

Press, pp. 307–328. https://doi.org/10.1016/B978-0-12-815739-8.00017-1

Schneider, W., Eschman, A., Zuccolotto, A., 2002. E-Prime: User's Guide. Reference Guide. Getting Started Guide. Psychology Software Tools, Incorporated.

Scibelli, F., Roffo, G., Tayarani, M., Bartoli, L., De Mattia, G., Esposito, A., Vinciarelli, A., 2018. Depression Speaks: Automatic Discrimination between Depressed and Non-Depressed Speakers Based on Nonverbal Speech Features. ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. 2018-April, 6842–6846. https://doi.org/10.1109/ICASSP.2018.8461858

Sheline, Y.I., Barch, D.M., Price, J.L., Rundle, M.M., Vaishnavi, S.N., Snyder, A.Z., Mintun, M.A., Wang, S., Coalson, R.S., Raichle, M.E., 2009. The default mode network and self-referential processes in depression. Proc. Natl. Acad. Sci. U. S. A. 106, 1942–1947. https://doi.org/10.1073/pnas.0812686106

Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian optimization of machine learning algorithms, in: Advances in Neural Information Processing Systems. pp. 2951–2959.

Stolte, M., Humphreys, G., Yankouskaya, A., Sui, J., 2017. Dissociating biases towards the self and positive emotion. Q. J. Exp. Psychol. 70, 1011–1022. https://doi.org/10.1080/17470218.2015.1101477

Sui, J., Greenshaw, A.J., Macrae, C.N., Cao, B., 2021. Self research: A new pathway to precision psychiatry. J. Affect. Disord. https://doi.org/10.1016/j.jad.2021.06.041

Sui, J., Gu, X., 2017. Self as Object: Emerging Trends in Self Research. Trends Neurosci. https://doi.org/10.1016/j.tins.2017.09.002

Sui, J., He, X., Humphreys, G.W., 2012. Perceptual effects of social salience: Evidence from

self-prioritization effects on perceptual matching. J. Exp. Psychol. Hum. Percept. Perform. 38, 1105–1117. https://doi.org/10.1037/a0029792

Sui, J., Humphreys, G.W., 2017. The self survives extinction: Self-association biases attention in patients with visual extinction. Cortex 95, 248–256. https://doi.org/10.1016/j.cortex.2017.08.006

Sui, J., Ohrling, E., Humphreys, G.W., 2016. Negative mood disrupts self- and reward-biases in perceptual matching. Q. J. Exp. Psychol. 69, 1438–1448. https://doi.org/10.1080/17470218.2015.1122069

Sui, J., Rotshtein, P., Humphreys, G.W., 2013. Coupling social attention to the self forms a network for personal significance. Proc. Natl. Acad. Sci. U. S. A. 110, 7607–7612. https://doi.org/10.1073/PNAS.1221862110

Verplanken, B., Sui, J., 2019. Habit and identity: Behavioral, cognitive, affective, and motivational facets of an integrated self. Front. Psychol. 10, 1504. https://doi.org/10.3389/fpsyg.2019.01504

von Glischinski, M., von Brachel, R., Hirschfeld, G., 2019. How depressed is “depressed”? A systematic review and diagnostic meta-analysis of optimal cut points for the Beck Depression Inventory revised (BDI-II). Qual. Life Res. https://doi.org/10.1007/s11136-018-2050-x

Walss-Bass, C., Suchting, R., Olvera, R.L., Williamson, D.E., 2018. Inflammatory markers as predictors of depression and anxiety in adolescents: Statistical model building with component-wise gradient boosting. J. Affect. Disord. 234, 276–281. https://doi.org/10.1016/J.JAD.2018.03.006

Wong, T.T., 2015. Performance evaluation of classification algorithms by k-fold and leave-one-

out cross validation. Pattern Recognit. 48, 2839–2846.

https://doi.org/10.1016/J.PATCOG.2015.03.009

World Health Organization, 2008. The global burden of disease: 2004 update.

Yankouskaya, A., Sui, J., 2021. Self-Positivity or Self-Negativity as a Function of the Medial Prefrontal Cortex. Brain Sci. 2021, Vol. 11, Page 264 11, 264.

https://doi.org/10.3390/BRAINSCI11020264